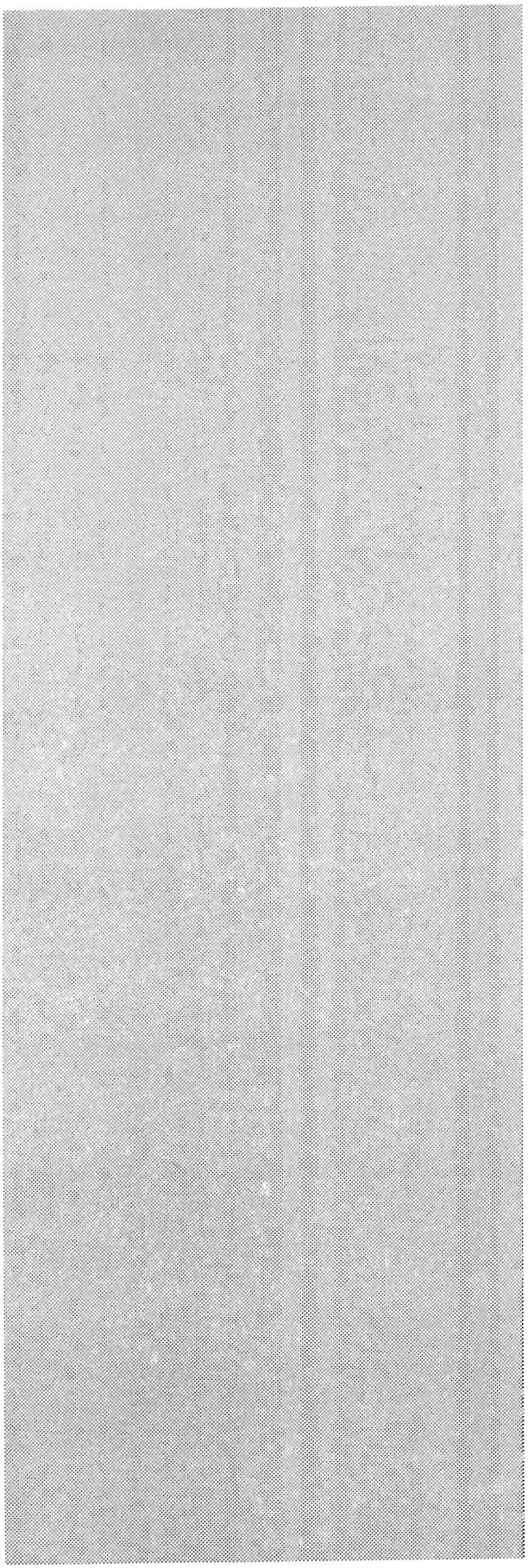


NIVEL
nederlands instituut
voor onderzoek van de
gezondheidszorg
drieharingstraat 6
postbus 1568 3500 bn utrecht
tel. 030 2319946 fax 030 2319290

DUTCH NATIONAL SURVEY OF GENERAL PRACTICE

A PORTFOLIO

September 1995

NIVEL Netherlands institute of primary health care
P.O. Box 1568 3500 BN Utrecht
Telephone ..-31302319946 Fax ..-31302319290

CONTENTS

CONTENTS

CONTENTS

A policy of strengthening primary care has been on the political agenda, for a long time. There has been however a number of obstacles preventing its implementation. One of these being the lack of research into the evidence supporting primary care. This paucity of scientific data was felt at the national level. In consequence, several initiatives were launched; one of these being the construction of a large database based on what actually happens in everyday general practice.

So it was, that nearly ten years ago, NIVEL started with the Dutch National Survey of General Practice. It was one of the largest surveys ever conducted in primary care in the Netherlands, unique in its kind by combining data from general practice registrations with extensive health interviews among the patients. Since then, numerous books, reports and papers have been published based upon the data from this survey, covering a wide range of subjects. The data has been analyzed by many researchers, from various research institutes who have published their results both in Dutch and in English.

This book gives a view of the top of the iceberg of publications from the Dutch National Survey. We have compiled a list of English language publications published in leading scientific journals to give a bird's eye view of the results of this National Survey as these have been published in the international literature. We received the kind permission of many editors or publishers concerned for this purpose for which we are grateful.

We are pleased to learn that the number of publications deriving from the Dutch National Survey is still increasing. Clearly part of the iceberg is still below the surface. It will keep on growing for quiet a while.

Prof. Jozien Bensing
Prof. Jouke van der Zee

directors of NIVEL

THE IMPACT OF A LARGE SURVEY IN A SMALL COUNTRY

It is almost 20 years ago that the first ideas of a national survey of general practice in the Netherlands were brought forward. It took 10 years of preparation before the actual collection of data started in the general practice.
After one year census data from 335.000 persons, data from 386.000 consultations, this population had with 103 general practices (161 GP's/177 practice assistants) and (figure 1) data from a health interview survey among a random sample of the practice population were obtained.

Figure 1

The original aim of the Dutch National Survey had been described as follows: "to obtain at national scale an insight in the presentation of complaints and health problems to general practice, in the actions of the GP's, related to these problems and in factors who influence the presentation of health problems as well as the differing reaction of GP on presented health problems" (Foets et al., 1986). Altogether the following issues have played a role during our work:

* organization and implementation of data collection, data processing and data analysis of the initial and follow-up study's.
* advise on study design and datacollection for a number of related surveys (eg Almere-survey, Urk-study, ROME-project, Aletta study) and on development of international data sets in primary care (eg AIM programme/EEC, Telematics in Primary Health Care/EEC, surveys in Catalonia, Spain, Hungary and Flandres).
* cmmunication of results through scientific journals and mass media.

THE IMPACT

Now ten years after the start of the pilot phase of this largest PHC-research project in the Netherlands it may be interesting to review what the impact has been of all our efforts. Two indicators may help us in this exercise:

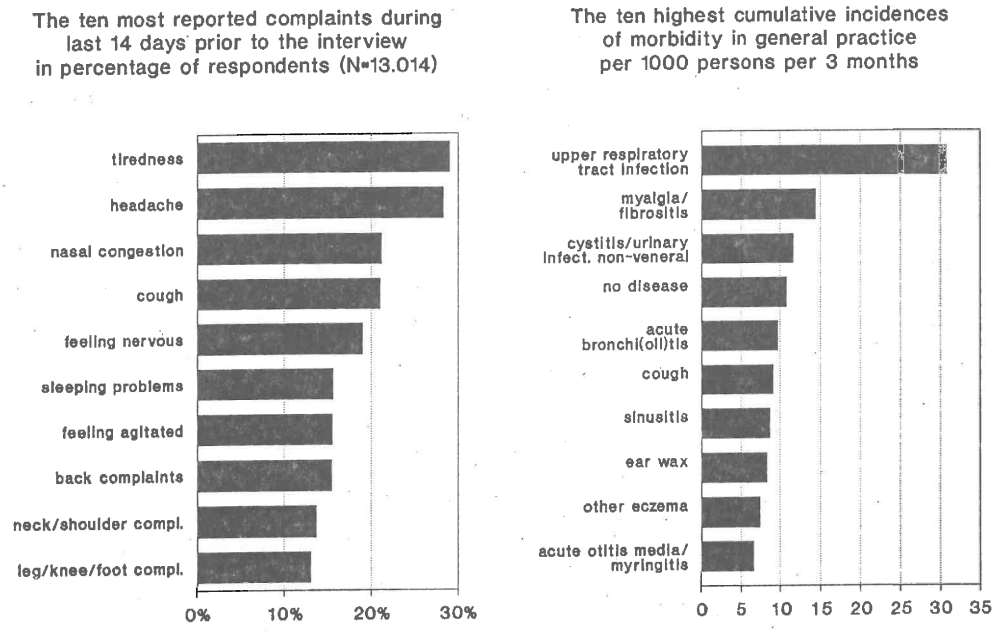
* the number of scientific pulications and citations (eg articles in peer review journals, citation and use of data for different purposes). The national survey produced up to now 40 scientific reports, contributed to 10 theses and led to 40 international and 65 national published articles.
* the influence of the obtained information for health policy. This is an difficult assessment with limited reliability and therefore not carried out quantitatively but qualitatively in this exercise. The survey was supportive for developing health policies at national and local level. It's data have been used in
  - the most important national policy documents: financial review of care, public health forecast and scenario research;
  - the problem solving of certain professional issues: the position of GPs in deprived areas in inner cities, the position of the practice assistant and the debate on the implementation of preventive programmes through general practice;
  - the development of NHG-standards, providing the designers with background information;
  - the further development of certain concepts like co-morbidity, prescribed daily doses.

In this booklet we have made a compilation of publications in different international journals, organized through the major determinants of health services research: demand, utilization and supply of care.
In addition an overview is given of all reports, all articles in national journals and a list of request for data.

1. Demand for care

The Dutch National Survey of General Practice has explicitly paid of attention towards demand of general practice care and its determinants. In the design of the survey the prominent place of the health interview survey (+ health diary), the census data and the morbidity registration has led to a large number of analyses and publications, both methodological and technical.
The iceberg phenomenon has become clearly visible. The top ten of complaints in the population differs in quantity and content completely from the top ten of health problems, presented to GPs (figure 2., see next page).

Figure 2



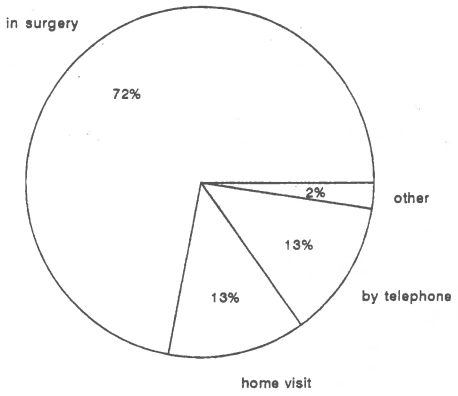
Yet access to general practice is considered good by patients, with minor critiques on the availability of service during evening/night and weekends in some areas. Major determinants of demand of care are morbidity, age, gender and social class.
The outcome may not be surprising, but it has been described on a national scale for the first time.
Further important (and published) issues are:

* epidemiology of specific diseases;
* the phenomena of comorbidity, described in different fora;
* the detection of psychological problems in general practice;
* the relation of social networks and health;
* healthy life expectancy and the effect of chronic disease;
* life style factors (incl. smoking, alcohol);
* gender differences in health;
* social health inequalities.

2. Utilization of care

The largest component of the Dutch National Survey of General Practice consists of a registration of data from all consultations with general practice in 3 months (N=386.000) of which the marjority was done in surgery (figure 3, see next page).

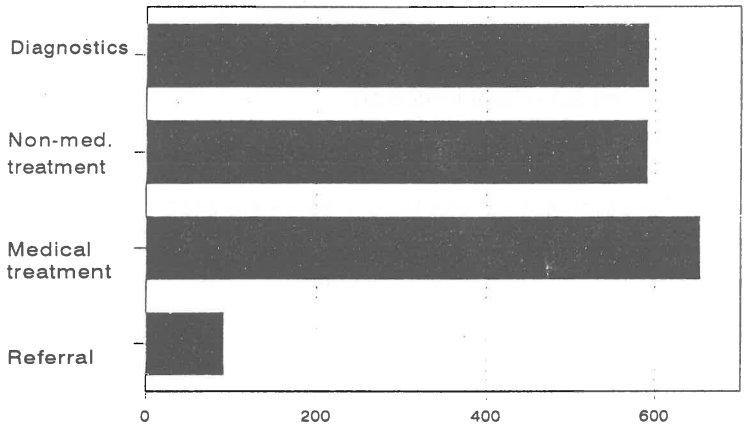
Figure 3



Type of consultation in general practice

Besides a description of the process of care in a large number of written in reports Dutch, three areas got specific attention: workload, quality of care and cost.
Descriptive studies on different interventions (physical examination/laboratory and X-ray investigation/educational activities/ minor survey/prescription/referrals) and follow-up arrangements have yielded widely used data (figure 4).

Figure 4: Practice activities per 1000 consultations



Further descriptive studies, requested by different national committees (eg Steering Committee Future Scenarios/Biesheuvel-committee) and national research (eg Future public health status) contributed to a higher status of the survey. The management of a wide range of diseases was subject of a large number of studies and publications, frequently in conjunction with University Departments of General Practice (eg respiratory infections/skin infections/childhood diseases/reumatoid diseases/headache and migrain/gastro-intestinal diseases/neurological diseases).

Studies related to workload have looked into the number of contacts with general practice and determining factors; the relation of workload and quality of care is topic of a more specific project. Out of hours care was topic of a separate analysis.
Quality of care received a lot of attention in a large number of publications. Medical technical skills of GPs were considered in a project, whereby the implementation of a number of NHG-standards of care were reviewed. This has resulted in a report including 13 articles. Adverse effects of prescribed medicines are studied in a separate programme. Communication skills were frequently considered in sidelines of analyses and in a number of articles related to women health care. Practice management was subject in analyses around the position of the practice-assistant, preventive care and repeat prescription.
Although not a direct objective of study, costs of care received attention in relation to laboratory- and X-rays investigation, prescription of medicine and GP-income studies.

3 Supply of care

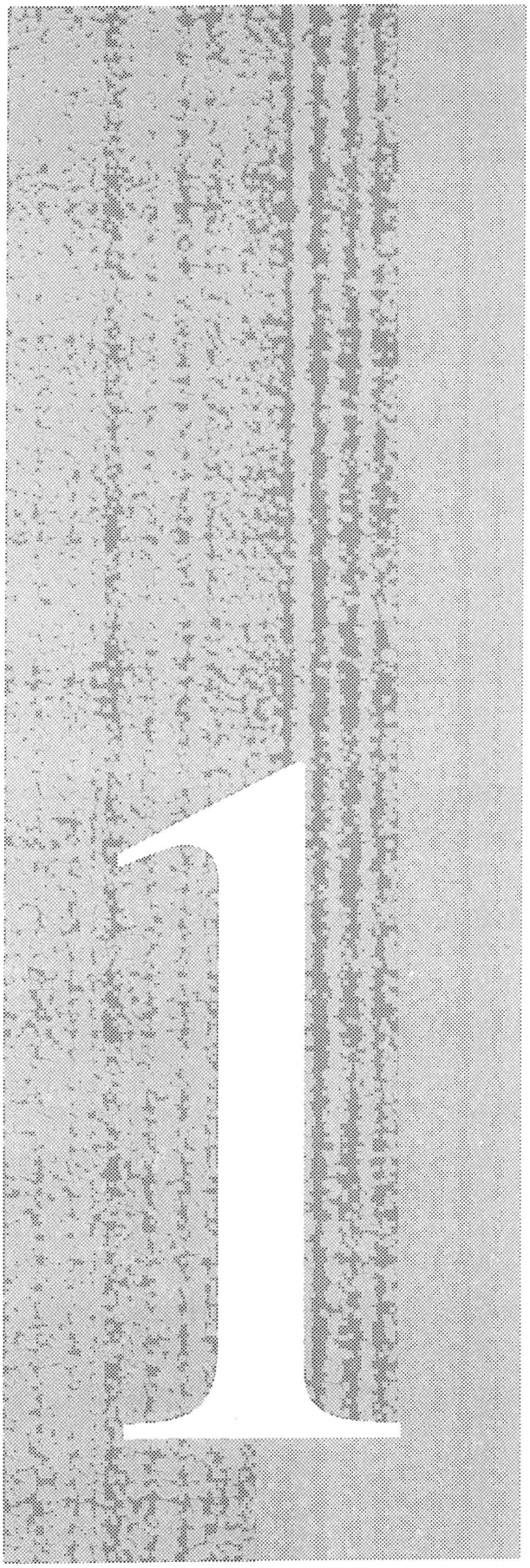
This subject was not a specific area of study, but the social and professional characteristics of both the GP and the practice-assistant were often considered in analyses and publications, related to the demand and utilization of care. Opinions and attitudes of GPs were used in analyses and publications around the role of general practice with regard to occupational health, screening programmes and preventive care in general, quality of care and the influence of GP in secondary care. The important subject of interdoctor-variation received attention in relation to the management of different diseases (e.g. migrain, acne vulgaris, gastro-intestinal diseases) and around most interventions (e.g. diagnostic procedures, patient education, preventive care, prescription of medicines, referrals to medical specialists and other PHC-disciplins and follow-up). In relation to workload the characteristics of GP were often used.

4 Conclusion

The Dutch National Survey has led to a comprehensive database, containing a lot of relevant information to be used in daily practice, scientific research and health policy. The database has been consulted by hundreds of health professionals, researchers and policymakers for different purposes. Surprisingly the database is still widely used by people in- and outside NIVEL.
One of the major successes of the survey is the good cooperation we could establish with other research bodies like the university departments of general practice, public health and other disciplins, national (RIVM, TNO, NEI, SCP, IOO) and other (AS/tri, IPSO facto) research institutions, professional bodies (LHV/NHG) and insurance companies, industrial research (GLAXO), patient organisations (Wound care society, migrain foundation), mass media (Boerderij, Nieuwe Revu, Trouw, Volkskrant, Elsevier) and many individuals.
The value of this survey can never be underestimated, only for the fact that it may serve as baseline for future research on changes in Dutch health care.

THE IMPACT

The researchers involved with carrying out datacollection, dataprocessing and data-analysis of the Dutch National Survey have gained unique experience, as to the possibilities (successes and shortcomings) of this type of survey. After all it has been a learning and rewarding social experiment.



1

DEMAND FOR CARE

AN ORIENTATION TOWARD HELP-SEEKING FOR EMOTIONAL PROBLEMS

M. A. R. TIJHUIS, L. PETERS and M. FOETS

Netherlands Institute of Primary Health Care (NIVEL), P.O. Box 1568, 3500 BN Utrecht, The Netherlands

Abstract—In recent years, many researchers tried to explain the social selection in use of mental health care services. A modest role is attributed to the orientation toward help-seeking. This article studies this orientation. Our research-population consisted of 10,171 Dutch persons, aged 18 and older. Analysis showed that most people are prone to seek help for one or more emotional problems. People who are more prone to seek help are younger, have had more education and have a higher family income. They have more often acquaintances working in mental health care. People who are more prone to seek help do not see chance as the locus of control of health. These people are less dependent on their GP for common disorders and are more open about mental health matters. The results of discriminant analysis are not satisfactory, but when we try to distinguish the groups of people who are and who are not willing to seek help, we see that the best discriminating factor is their help-seeking attitude for common disorders. People who have high expectations from the GP for common disorders, clearly do have a preference to seek help for the emotional problems. The groups of people who are more willing to seek help from the GP compared to mental health professionals cannot be distinguished by these expectations. Here the level of education discriminates fairly well: people who are more prone to seek help from a GP have a lower educational level. Future research should be focussed on the testing of a theoretical model that explains the orientation toward help-seeking for emotional problems and selection in help-seeking with longitudinal data.

Key words—help-seeking orientation, emotional problems

INTRODUCTION

Many psychological problems are never brought to the attention of specific professional providers. Goldberg *et al.* [1] assume that most patients with psychological problems seek help at the primary care level (although in many cases not with well articulated psychological problems, but with somatic reasons for a visit), and that only a small proportion of them enters the specialized mental health care system. Within the group that reaches specialized mental health care, inequalities in age, socio-economic status and the like can be observed. Important factors in this respect are due to the health care system. For instance, in countries like the U.K. and the Netherlands, the specialized mental health care agencies are accessible, formally, only after referral by the general practitioner. The GP might have preferences in his referral policy for certain professionals and kinds of patients. A second important system factor might be the costs, attached to specialized mental health care, which may constitute serious barriers for some groups of patients.

Another possible explanation for differences between patients in help-seeking behaviour might be found on the part of the patient. Different orientations toward seeking specific professional help for psychological problems have proven to be important determinants of differences in actual help-seeking [2–12]. We think, therefore, that it is useful to explore the relationship between explanatory variables and this orientation. As Leaf *et al.* [13, p. 276] stated, "Knowledge of (..) correlates will allow for better targeting of future inventions to reduce barriers to the use of mental health services."

We return to our starting-point. The first filter in the well-known Goldberg and Huxley model reflects the help-seeking behaviour. Only a small proportion of the people in the community that feel some amount of stress (stage 1), presents this to the GP (stage 2). Kessler *et al.* [6] tried to explain this filtering process in terms of gender, amount of problems and orientation toward help-seeking ('propensity to seek care for psychological problems' as they call it). Here, we'll extend the model to other explanatory variables and concentrate on the relationship of individual characteristics and the orientation toward help-seeking (see Fig. 1). The variables related to an orientation toward help-seeking will be grouped into three categories: (1) personality characteristics, (2) demographic characteristics and (3) network characteristics.

Personality characteristics. One of the correlates with help-seeking attitudes is locus of control [14]. Respondents with an external locus of control showed a negative attitude toward help-seeking. A second characteristic of the personality that correlates significantly with a help-seeking orientation is authoritarianism. Highly authoritarian students and female respondents [14, 15] tend to express negative orientations. Fischer *et al.* [14] suggested that the ability to disclose oneself to another could be related to a help-seeking orientation. Subsequently they construct their scale for help-seeking orientation, which included several items about interpersonal openness. We feel (as they do [14, p. 88]) that this is not a dimension of

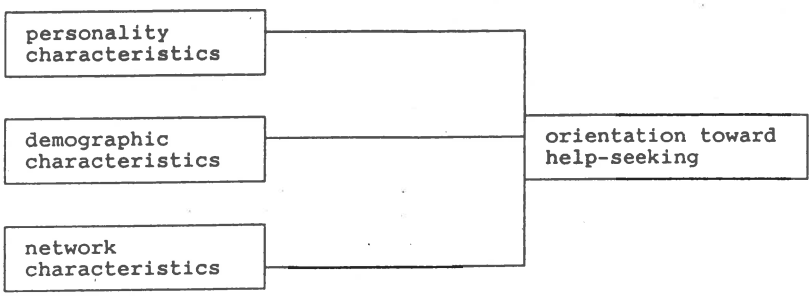


Fig. 1. Model posited to explain an orientation toward help-seeking.

help-seeking orientation, but merely one of the correlating personality characteristics. In the Fischer study, other personality characteristics showed significant correlations for men but not for women. This was the case for social desirability and trust in professionals and institutions for mental health.

Demographic characteristics. Gender, age, socio-professional level, education and income are shown to affect a help-seeking orientation [6, 13, 15–17]. The findings for age are somewhat contradictory but female respondents with a higher education, socio-professional level and income have a more positive help-seeking orientation.

Network characteristics. As Barbot [15] pointed out, women who knew somebody who had had psychotherapy held more favourable orientations. People who had prior contact with psychotherapists were also more positively oriented toward help-seeking for emotional problems [17, 18]. Other characteristics of the network (like number of friends, meeting frequency) have not been included in the analysis for help-seeking orientation. However, these characteristics do have an influence on the use of mental health services [19]. Respondents with more friends and higher meeting frequency were less likely to use mental health services. Sherbourne [19] states that "The more social resources available to a person, the less likely that person is to use mental health services." It would be useful to examine whether these characteristics affect the help-seeking orientation.

RESEARCH QUESTION

As stated earlier, we will explore the model set up by Kessler *et al.* [6]. We have restricted ourselves to the relationship between individual characteristics and the orientation toward help-seeking. Looking at earlier studies in this respect [6, 13–15, 17, 20], we see that only some variables are used to explain an orientation toward help-seeking. In this article we'll try to study most of the variables that are expected to be related to this orientation. The primary goal, therefore, is to examine the relationships between relevant personality characteristics, demographic and network characteristics and an orientation toward help-seeking.

This goal can be translated in two questions, relevant for this article. These are: (1) Is it possible to distinguish the people who are willing to seek help for emotional problems from those who are not, on the ground of certain individual characteristics? and (2) When we take a closer look at the people who are willing to seek help: Is it possible to distinguish the people who are willing to seek help from the GP from those who are willing to seek help from mental health care professionals (MHP) using individual characteristics?

Based on the earlier studies and on the available data, the following individual characteristics are selected. The *personality characteristics* chosen for analysis are locus of control, expectations from GP with respect to common disorders, and openness about mental health matters. We assume that people with an internal locus of control and who are open about mental health matters will be more oriented toward help-seeking. People who do expect a lot from the GP for common disorders are supposed to be highly dependent on the GP and will also be oriented toward help-seeking.

The *demographic characteristics* included in the analyses are gender, age, education, income and marital status. We assume that female, younger persons with a higher education and higher income will be more prone to seek help. Marital status has not been studied before in relationship with a help-seeking orientation. Research on social selectivity in seeking help for psychological problems [3] showed that unmarried people are more likely to apply for help at a counseling centre. In accordance with this finding, we assume that single persons will be more prone to seek help.

We have chosen three *network characteristics*, i.e. the number of close friends, whether or not people have acquaintances working in mental health care, and whether or not people have had prior contacts with mental health care professionals. As people with a lot of friends receive enough social support, we assume that people with more friends will not be oriented toward help-seeking. People with acquaintances working in mental health care will be more oriented toward help-seeking. People with prior contacts with mental health care professionals are assumed to be more willing to seek help for emotional problems.

As we do not control our analysis for number of problems or morbidity, we cannot say anything about the effect of the health-status of respondents on their orientation or vice versa. We'll return to this point in our discussion.

METHOD

Data were collected as part of the National Study of Morbidity and Interventions in General Practice in

Table 1. Demographic characteristics of the sample ($N = 10,171$)

	N	%		N	%
Gender:			Marital status:		
Male	4969	48.9	Not married	2316	22.8
Female	5202	51.1	Married	6824	67.1
			Widowed	634	6.2
Education:			Divorced	397	3.9
Highest	1082	11.1	Age:		
High	1958	20.1	18–24	1451	14.3
Low	4054	41.5	25–44	4556	44.8
Lowest	2668	27.3	45–64	2751	27.0
(mv = 409)			65+	1413	13.9

the Netherlands [21]. In 1987, interviews were conducted among a random sample of 100 persons per general practitioner who took part in the study. The participating GPs form a random, non-proportional stratified sample of GPs in the Netherlands ($N = 161$). The interview consisted of 6 components, i.e. indicators of morbidity, health behaviour and use of health services, demographic variables, indicators of health-endangering habits, attitudes and beliefs, network characteristics and social support, and life events. The response rate for the interviews was 76.7% ($N = 13,067$). We selected all respondents aged 18 and over for analysis ($N = 10,171$).

Sample characteristics

The demographic characteristics of the sample are presented in Table 1. As indicated by the table women comprise somewhat over half of the sample, about one-third of the respondents have had only lowest education, over 40% is middle-aged (between 25 and 44) and two-thirds of the respondents are married. Differences between the Dutch population and the population under study, concerning gender and age, are very small according to Foets *et al.* [22].

The orientation toward help-seeking

The questionnaires the respondents filled in during the interview, included a list about help-seeking orientation. This list combines 5 emotional problems with the orientation toward help-seeking. The respondents indicated to which (if any) of the professionals people should go for each of the problems. Possible answers were: general practitioner, social work, mental health services, and seeking no help at all. The answers of respondents are given in Table 2.

Because we want to distinguish the groups of people who are and who are not prone to seek help, we conducted Guttman scale statistics. For the purpose of Guttman scale analysis, the answers are recoded to 1 = will seek help, and 0 = will not seek help for each problem in the list (Table 3). The fifth item ('Someone is afraid of using an elevator. Who could he/she turn to?') scored badly; the H_i coefficient [23] for this item was 0.16. We left this item out. Loevingers' H_{ij} between items varied from 0.51 to 0.68. H_is, Loevingers' H_{ij}s and H are high enough to draw the conclusion that we are dealing with a (Guttman) scale.

We determined scores for all respondents on this scale. To do so, we totalled the responses per respondent for each category. If, for example, a person answered that people should go to a GP for each of the problems in the list, he or she scores 4 times on 'help-seeking'. We assume that this score reflects what the respondents think they should do themselves.

As indicated by Table 4, nearly all respondents would turn to a professional for at least the first problem in the list (depression). Nearly 70% of all respondents would even turn to a professional for problems with raising a 9-year-old boy.

Individual characteristics

The demographic variables in this study are gender, age, marital status, household income and education. Education indicates the completed level of education.

Network characteristics of the respondents are the number of close friends, whether or not the respondents have acquaintances working in mental health care, and whether or not the respondents have had prior contacts with mental health care professionals.

The three personality characteristics chosen for this study are locus of control, expectations from GP for common disorders and openness about mental health matters. The locus of control scale we used [24] is a translated version of the Wallston [25] scale for health locus of control. This scale measures to what extent the respondent thinks his illness or health is determined by himself, the GP or by chance (Cronbach's alpha 0.76, 0.80 and 0.71 respectively).

Van de Lisdonk [26] developed a scale for 'expectations from GP for common disorders', named 'Nijmeegse Verwachtingen Lijst'. It is a list of common disorders which do not require medical

Table 2. Preference for professionals per emotional problem, in percentages

Item		Preference				
		Nobody	The GP	Social work	Mental health care	N
1.	Someone is having great difficulties raising a 9-year-old boy. Who could give him/her some advice about this?	13.7	40.4	24.6	21.3	10,030
2	Serious problems have arisen in a 3-year-old marriage. Who would be the best person for this couple to turn to?	16.3	27.6	36.1	20.0	10,006
3.	Someone is feeling very lonely. Who could he/she turn to?	17.9	27.1	36.8	18.3	9991
4.	Someone has been depressed for months. Who could he/she turn to?	5.4	61.9	6.5	26.1	10,019
5.	Someone is afraid of using an elevator. Who could he/she turn to?	26.5	33.1	3.0	37.3	9994

Table 3. Guttman scale statistics for help-seeking orientation

Item		H_i coefficient
(4)	Someone has been depressed for months. Who could he/she turn to?	0.64
(2)	Serious problems have arisen in a 3-year-old marriage. Who would be the best person for this couple to turn to?	0.39
(3)	Someone is feeling very lonely. Who could he/she turn to?	0.37
(1)	Someone is having great difficulties raising a 9-year-old boy. Who could give him/her some advice about this?	0.34
H		0.41

attention. The respondents are asked whether or not they would visit a GP for these disorders. The items in our questionnaire are a selection of 12 (out of 23 in Van de Lisdonk's study), all loading on one factor, resembling the dimensions of visiting the doctor for diagnostic or therapeutic reasons or to recover more quickly (Cronbach's alpha 0.91).

Respondents were asked to answer three questions about openness toward mental health care matters; whether people should hide the fact that they are in treatment for mental health problems, whether people should talk about existing problems (reverse-coded), and whether people should take medicines to solve the problems. These items were combined in one measure for openness by totalling negative responses (Cronbach's alpha 0.80).

RESULTS

Table 5 presents the mean score of respondents scoring extreme (the score of 0 or 1 compared to 3 or 4) on the orientation toward help-seeking per explanatory variable. The table also shows the significance of calculated t-test statistics.

It can be seen that there are significant differences between the groups of peope who are and who are not prone to seek help for personal problems. Of course, we must keep in mind that we are dealing with a very large sample here. Therefore, we only consider differences greater than 10% of importance. For our results this approximates the statistical significance with $P \leqslant 0.001$.

Most findings for respondents tending not to seek help at all for the problems in the list are in concordance with our expectations. This group is usually older, widowed and they more often have a lower level of education and income. They have less often acquaintances working in mental health care and have less often had prior contacts with mental health care professionals. The results for the number of friends, internal locus of control and GP as locus of control are not significant. Their score on chance as the locus of control is usually higher. This group is more often not open about mental health matters and does not expect much from GP for common disorders.

Next we will try to distinguish the two groups by discriminant analysis (Table 6).

The strongest discriminating factors here, are the extent to which people have expectations from the GP for common disorders and age. People who have higher expectations in this respect and people who are younger belong more often to the group of people who are willing to seek help. This group (respondents who are prone to seek help) can be well classified

Table 4. Willingness to seek help for emotional problems, in percentages ($N = 9880$)

Score	Help-seeking
0×	2.6
1×	3.4
2×	7.6
3×	17.3
4×	69.1
Mean	3.5
SD	1.0

Table 5. Mean score per variable for people who are more willing to seek help ($N = 8536$) vs people who not ($N = 593$)

Variables	Yes	No
Gender (1 = male, 2 = female)	1.5	1.4*
Age (18–97)	42.2	51.4**
Marital status:		
1 = not married	0.2	0.1**
1 = married	0.7	0.7
1 = widowed	0.1	0.2**
Education, 1 (low)–9 (high)	2.9	2.3**
Income, 1 (low)–16 (high)	8.3	6.9**
Number of friends (0–95)	6.5	6.4
Acquaintances in mental health care (1 = yes)	0.2	0.1**
Prior contact with MHP (1 = yes)	0.1	0.0**
Internal locus of control, 1 (low)–31 (high)	16.0	16.0
GP as locus of control, 1 (low)–31 (high)	13.0	13.4
Chance as locus of control, 1 (low)–31 (high)	14.6	16.1**
Common disorder expectations from GP, 1 (low)–49 (high)	24.8	22.6**
Openness, 1 (low)–18 (high)	13.7	12.8**

Significance calculated from t-test statistics:
*$P \leqslant 0.05$, **$P \leqslant 0.001$.

Table 6. Weight of explanatory variables on a discriminant function with differences between extreme low and high scorers on the orientation to seek help ($N = 9880$), being maximalized

Variables	Weights
Gender (1 = male)	0.26
Age (18–97)	−0.46
Marital status:	
1 = not married	0.07
1 = married	0.07
1 = widowed	−0.15
Education, low (1)–high (9)	0.29
Income, low (1)–high (16)	0.17
Friends (0–95)	−0.01
Acquaintances working in mental health care (1 = yes)	0.12
Prior contact with MHP (1 = yes)	0.10
Internal locus of control, low (1)–high (31)	−0.06
GP as locus of control, low (1)–high (31)	−0.02
Chance as locus of control, low (1)–high (31)	−0.38
Expectations from GP for common disorders, low (1)–high (49)	0.75
Openness, low (1)–high (18)	0.23
Eigenvalue	0.04
Wilks' lambda	0.96
% Group members predicted in the right group:	
Low score: respondents not prone to seek help	1.9
High score: respondents prone to seek help	99.9
All respondents	95.7

Table 7. Guttman scale statistics for help-seeking orientation towards the GP or towards mental health care professionals ($N = 9468$)

Item		H_i coefficient
(4)	Someone has been depressed for months. Who could he/she turn to?	0.57
(2)	Serious problems have arisen in a 3-year-old marriage. Who would be the best person for this couple to turn to?	0.44
(3)	Someone is feeling very lonely. Who could he/she turn to?	0.43
(1)	Someone is having great difficulties raising a 9-year-old boy. Who could give him/her some advice about this?	0.36
H		0.43

according to the statistics, but the characteristics of people who are not prone to seek help cannot be distinguished well.

Finding that the differences between the groups of people who are and people who are not prone to seek help for personal problems are not strong, one wonders what differences could exist between people who are prone to seek help for personal problems at the GP's office in contradiction with mental health care professionals. To answer this question we selected all respondents who are prone to seek help for at least one of the problems mentioned. Since we are dealing with a Guttman scale this means that we selected all persons who would seek help for being depressed ($N = 9468$). Recoding the answers of four items to (0) seeking help at the GP's office, and (1) seeking help from mental health professionals (social work plus mental health care), we see that here also a (Guttman) scale exists (Table 7).

Loevingers' H_{ij} between items ranges from 0.53 to 0.71. H_is, Loevingers' H_{ij}s and H are high enough to draw the conclusion that we are dealing with a (Guttman) scale again. Scores for each respondent are shown in Table 8. A score of 3 means that respondents are prone to seek help for difficulties raising a 9-year-old boy at the GP's office and, given the Guttman scale, are prone to seek help at the office of a mental health care provider for the three other problems in our list. Similarly, a score of 4 means that respondents are not prone to seek help from the GP for these problems, but are willing to go to mental health care professionals.

More people are willing to go to a mental health care professional (22.6%) for all four personal problems mentioned than there are people willing to go to the GP (17.9%). Table 9 shows differences between the groups of people with a preference for the GP vs mental health care professionals (the score of 0 or 1 compared to 3 or 4).

Table 8. Willingness to seek help for emotional problems at the office of a mental health care professional (MHP) for respondents willing to seek help for at least one of the problems, in percentages ($N = 9468$)

Score	MHP
0×	17.9
1×	12.8
2×	20.3
3×	26.4
4×	22.6
Mean	2.2
SD	1.4

Table 9. Mean score per variable for people who are more willing to seek help from the GP ($N = 2907$) vs from mental health care professionals (MHP) ($N = 4639$)

Variables	GP	MHP
Gender (1 = male, 2 = female)	1.6	1.5**
Age (18–97)	45.4	40.7**
Marital status:		
1 = not married	0.2	0.3**
1 = married	0.7	0.6**
1 = widowed	0.1	0.0**
Education, 1 (low)–9 (high)	2.5	3.1**
Income, 1 (low)–16 (high)	7.7	8.5**
Number of friends (0–95)	6.9	6.4*
Acquaintances in mental health care (1 = yes)	0.1	0.2**
Prior contact with MHP (1 = yes)	0.0	0.1*
Internal locus of control, 1 (low)–31 (high)	16.1	16.0
GP as locus of control, 1 (low)–31 (high)	13.8	12.6**
Chance as locus of control, 1 (low) 31 (high)	15.5	14.2**
Common disorder expectations from GP, 1 (low)–49 (high)	25.8	24.2**
Openness, 1 (low)–18 (high)	13.3	13.9**

Significance calculated from t-test statistics:
*$P \leqslant 0.05$, **$P \leqslant 0.001$.

Nearly all individual characteristics differ between the two groups. People who have a greater tendency to seek help from mental health care professionals compared to people who have a greater tendency to seek help from a GP are more often male, younger and not married. These people have a higher educational level and a higher income than people who are prone to seek help from the GP. They also have more often acquaintances working in mental health care. Finally, these people see the GP or chance as their locus of control less often, they have lower expectancies from the GP for common disorders and are more open about mental health matters. In Table 10 we show the results of discriminant analysis, used to distinguish the two groups more pronounced.

As can be seen, education distinguishes the two groups at best. Still, the results of the analysis are not satisfactory at all. The percentage group members that can be classified in the group 1 and Wilks' lambda are not as high as they should be to make an accurate distinction between the two groups.

DISCUSSION

Some of the variables we used in our analyses (i.e. marital status, number of friends, whether or not the respondent has acquaintances working in mental health care, expectations the respondent had from the GP for common disorders, and openness) were not used before in the studies mentioned in the introduction, and can therefore not be compared to results of others. Making the possible comparisons between our results and those of others we see that we can confirm most of the findings.

Table 10. Weight of explanatory variables on a discriminant function with differences between low scorers on seeking help from mental health care professionals ($N = 2907$) and high scorers on seeking help from mental health care professionals ($N = 4639$), being maximalized

Variables	Weights
Gender (1 = male)	−0.12
Age (18–97)	−0.29
Marital status:	
1 = not married	0.32
1 = married	−0.10
1 = widowed	−0.07
Education, low (1)–high (9)	0.52
Income, low (1)–high (16)	0.15
Friends (0–95)	−0.13
Acquaintances working in mental health care (1 = yes)	0.00
Prior contact with MHP (1 = yes)	0.07
Internal locus of control, low (1)–high (31)	−0.03
GP as locus of control, low (1)–high (31)	−0.02
Chance as locus of control, low (1)–high (31)	−0.16
Expectations from GP for common disorders, low (1)–high (49)	0.06
Openness, low (1)–high (18)	0.10
Eigenvalue	0.05
Wilks' lambda	0.96
% Group members predicted in the right group:	
Low score:	
respondents prone to seek help from the GP	6.6
High score:	
respondents prone to seek help from mental health care professionals	96.7
All respondents	67.6

Several authors [13, 14, 16, 17] suggested that women held more positive attitudes toward help-seeking than men. Others [27, 28] did not find a difference in help-seeking orientation between the sexes. The samples of most of these studies consist of students, as they do in the case of Dadfar [27] and Zeldow [28] too. In our case, a representative sample, we found a very small difference in orientation toward help-seeking between the sexes, possibly due to our large sample. Future research in this respect is needed.

In our study, younger people were more willing to seek help than older people, which confirms the findings of Leaf *et al.* [13]. Several authors [13, 15, 17, 20] also indicated the positive relationship between education and income, and the orientation toward help-seeking we found. That people with prior contacts with mental health care appeared to be more willing to seek help in our study, was also found by the majority of authors on this subject [14, 15, 17, 18, 27, 29]. As our measure of locus of control differs from the one used by Fischer *et al.* [14], we cannot compare our results completely. Nevertheless, in both studies 'externals' were less oriented toward help-seeking.

Unfortunately we have not found any publications about a help-seeking attitude towards specific professionals. Frank *et al.* [30] used a similar measure for help-seeking behaviour, but they did not include an orientation toward help-seeking in their analyses. Their main conclusion was that (mental) health plays an important role in the decision to seek care but has little effect on the type of provider chosen. As mentioned earlier, we did not control our analysis for health status.

The results of our discriminant analysis were not satisfactory. Other, more discriminating, individual characteristics have to be found (for instance 'threat to self-esteem', as suggested by Amato *et al.* [31] and by Nadler [32]). One way to start this research is by giving some more theoretical background for the analysis. The work of Kessler *et al.* [6], mentioned earlier, is one step in this direction. In this respect, we have to keep the ultimate goal in mind. This is, we want to explain why certain people do seek help for their emotional problems and others do not. One important variable, of course, is the amount of problems people have. In the case of one of our network variables for instance, there exists extensive literature about the effect of social support on this amount of problems. To explore the relationship of an orientation towards help-seeking, the amount of problems and the actual help-seeking, one should have longitudinal data. As several authors [17, 18, 20, 28] pointed out, people with prior contacts (whether positive or not) with mental health care have a more positive orientation toward help-seeking. We should keep in mind that older people have had more opportunities to use mental health services. So when we try to determine the effect of the orientation on actual help-seeking we should at least have longitudinal data. In the studies about selectivity in the use of mental health care mentioned earlier, there is not much attention for this causality problem. Some researchers had longitudinal data [12] at their disposal and did not have to worry about causality. Others [7, 17] have not determined the causal direction of the relationship between the orientation and the use of mental health care: "Those who are in treatment may use being in treatment as an indicator of a positive orientation" [7, p. 1334], but recognize the problem. We feel that this causality problem should be solved before going on with this subject.

REFERENCES

1. Goldberg D. P. and Huxley D. *Mental Illness in the Community. The Pathway to Psychiatric Care.* Tavistock, London, 1980.
2. Greenley J. R. Social factors, mental illness and psychiatric care: recent advances from a sociological perspective. *Hosp. Commun. Psychiat.* **35**, 813, 1984.
3. Greenley J. R. and Mechanic D. Social selection in seeking help for psychological problems. *J. Hlth soc. Behav.* **17**, 249, 1976.
4. Greenley J. R., Mechanic D. and Cleary P. D. Seeking help for psychologic problems. A replication and extension. *Med. Care* **25**, 1113, 1987.
5. Hosman C. M. H. *Psychosociale problematiek en hulpzoeken. Een sociaal-epidemiologische studie ten behoeve van de preventieve geestelijke gezondheidszorg.* Dissertation Catholic University of Njmegen, Swets & Zeitlinger, Lisse, 1983.
6. Kessler R. C., Reuter J. A. and Greenley J. R. Sex differences in the use of psychiatric outpatient facilities. *Soc. Forces* **58**, 557, 1979. Gove W. and Swafford M. Sex differences in the propensity to seek psychiatric treatment: prevailing folk beliefs and misused log-linear analysis—Comment on Kessler *et al. Soc. Forces* **59**, 1281, 1981. Kessler R. C. Reply to Gove and Swafford. *Soc. Forces* **59**, 1291, 1981.

7. Leaf P. J., Livingston B. M., Tischler G. L., Weissman M. M., Holzer III C. E. and Myers J. K. Contact with health professionals from the treatment of psychiatric and emotional problems. *Med. Care* **23,** 1322, 1985.
8. Leaf P. J., Livingston B. M. and Tischler G. L. The differential effect of attitudes on the use of mental health services. *Soc. Psychiat.* **21,** 198, 1986.
9. Leaf P. J. and Livingston B. M. Gender differences in the use of mental health-related services: a re-examination. *J. Hlth soc. Behav.* **28,** 171, 1987(a).
10. Mechanic D., Cleary P. D. and Greenley J. R. Distress syndromes, illness behavior, access to care and medical utilization in a defined population. *Med. Care* **20,** 361, 1982.
11. Robbins J. M. Lay attribution of personal problems and psychological help-seeking. *Soc. Psychiat.* **16,** 1, 1981.
12. Tessler R., Mechanic D. and Dimond M. The effect of psychological distress on physician utilization: a prospective study. *J. Hlth soc. Behav.* **17,** 353, 1976.
13. Leaf P. J., Livingston B. M., Tischler G. L. and Holzer III C. E. The relationship between demographic factors and attitudes toward mental health services. *J. Commun. Psychol.* **15,** 275, 1987(b).
14. Fischer E. H. and Turner J. LeB. Orientations to seeking professional help: development and research utility of an attitude scale. *J. consult. clin. Psychol.* **35,** 79, 1970.
15. Barbot F. M. de. A cross-cultural study of attitudes toward seeking professional psychotherapeutic help. *Dissert. Abstr. Int.* **38,** 1875, 1977.
16. Sanchez A. R. and Atkinson D. R. Mexican-American cultural commitment, preference for counselor ethnicity, and willingness to use counseling. *J. Counsel. Psychol.* **30,** 215, 1983.
17. Surgenor L. J. Attitudes toward seeking professional psychological help. *N. Z. J. Psychol.* **14,** 27, 1985.
18. Kligfeld M. Medical student attitude toward seeking professional psychological help. *Dissert. Abstr. Int.* **39,** 4143A, 1979.
19. Sherbourne C. D. The role of social support and life stress events in use of mental health services. *Soc. Sci. Med.* **27,** 1393, 1988.
20. Fischer E. H. and Cohen S. L. Demographic correlates of attitude toward seeking professional psychological help. *J. consult. clin. Psychol.* **39,** 70, 1972.
21. Foets M., Velden J. van der and Zee J. van der. *Morbidity and interventions in general practice. A cross-national survey in the Netherlands. Study design.* Netherlands Institute of Primary Health Care (NIVEL), Utrecht, 1986.
22. Foets M. and Velden J. van der. *Methoden van onderzoek. Basisrapport van de Nationale Studie van Ziekten en Verrichtingen in de Huisartspraktijk.* NIVEL, Utrecht, 1990 (in press).
23. Swanborn P. G. *Schaaltechnieken. Theorie en praktijk van acht eenvoudige procedures.* Boom, Meppel, Amsterdam, 1982.
24. Halfens R. J. G. *Locus of control: de beheersingsorientatie in relatie tot ziekte- en gezondheidsgedrag.* Dissertation State University of Masstricht, 1985.
25. Wallston K. A., Wallston B. S. and DeVellis R. Development of the multidimensional health locus of control (MHLC) scales. *Hlth Educ. Monogr.* **6,** 160, 1978.
26. Lisdonk E. H. van de. *Ervaren en aangeboden morbiditeit in de huisartspraktijk. Een onderzoek met dagboeken.* Dissertation Catholic University of Nijmegen, 1985.
27. Dadfar S. and Friedlander M. L. Differential attitudes of international students toward seeking professional psychological help. *J. Counsel. Psychol.* **29,** 335, 1982.
28. Zeldow P. B. and Greenberg R. P. Attitudes toward women and orientation to seeking professional psychological help. *J. clin. Psychol.* **35,** 473, 1979.
29. Cash T. F., Kehr J. and Salzbach R. F. Help-seeking attitudes and perceptions of counselor behavior. *J. Counsel. Psychol.* **25,** 264, 1978.
30. Frank R. G. and Kamlet M. S. Determining provider choice for the treatment of mental disorder: the role of health and mental health status. *Hlth Serv. Res.* **24,** 83, 1989.
31. Amato P. R. and Saunders J. The perceived dimensions of help-seeking episodes. *Soc. Psychol. Q.* **48,** 130, 1985.
32. Nadler A. Determinants of help seeking behaviour: the effect of helpers similarity, task centrality and recipient's self esteem. *Eur. J. Soc. Psychol.* **17,** 57, 1987.

Family Practice

The Importance of the GHQ in General Practice

P F M VERHAAK, H J WENNINK AND M A R TIJHUIS

Verhaak P F M, Wennink H J and Tijhuis M A R. The importance of the GHQ in general practice. *Family Practice* 1990; **7**: 319–324.

The relationship between General Health Questionnaire (GHQ) score and complaints presented at the general practitioners office was examined, and showed that the correlation between them is not as high as might be expected. Many patients who present psychosocial problems to their GP appear to have a low GHQ score; many patients with a high GHQ score exclusively present somatic complaints, which are also assessed by the GP as being purely somatic. Implications of the results are discussed.

INTRODUCTION AND PRESENTATION OF THE QUESTION

Most people with mental disorders [defined as the presence of symptoms and/or signs which originate from or are sustained by emotional conflicts and/or stressing conditions. (cf Ref 10)] are not treated by the health care services.[1,2] This is often put down to the fact that a GP does not always identify the true nature of the complaint. As a solution, the use of screening instruments in primary care has been suggested, and the General Health Questionnaire (GHQ) has been frequently mentioned.

The GHQ is used as a first stage screening instrument for the identification of mental disorders,[1,3-10] and measures the likelihood of a person being classified a psychiatric 'case'. Given a standardized clinical interview (eg the Present State Examination) as a 'gold standard', the GHQ has proved to be a specific, sensitive instrument[6,10-14] and has been used successfully in clinical and general populations.[4] Hodiamont and Veling[15] came to the conclusion that the GHQ is an efficient questionnaire for estimating the prevalence of serious psychiatric problems in a specific population. Tarnopolsky *et al.*[12] concluded that 'in the region of 15–30% high GHQ scores, where the results of most surveys are expected to fall, the GHQ predicts with reasonable closeness the proportion of psychiatric disorders'.

These characteristics may be sufficient to make a reliable estimation of the probability that a patient is a psychiatric case, but it remains questionable whether screening for mental disorders in primary care is useful at all. The GHQ measures the prevalence of psychiatric complaints in the population: a GP reacts primarily to the demand for assistance at a specific moment in time. We are dealing, therefore, with the difference between epidemiological and clinical medicine, between screening and diagnosing. A patient who requests an analgesic for influenza, may also be known to the doctors as a diabetic, or as a person suffering from severe strain, but at the moment of the encounter it is the influenza which is presented. The GHQ detects 'psychiatric cases' regardless of the patient's immediate needs. The GP on the other hand, is primarily guided by those needs.

Moreover, the GHQ represents a specialist point of view, whereas the GP has a generalistic point of view. The difference between the two is sometimes characterized as: 'treat unless' as opposed to 'do not treat unless'. This difference can clearly be seen in the complaints which psychiatrists raise about GPs failing to identify serious psychiatric problems (compare the studies cited above), but it also explains the GPs complaint that the psychiatric point of view does not fit in with their way of looking at things.[16,17] This also raises the question of whether, in principle, GPs should follow up matters which they suspect, but which are not presented to them by the patient. In other words, one must consider whether it is useful or desirable for a GP to transfer to screening and the active detection of psychiatric problems.

To decide about the feasibility of the GHQ as an instrument for screening in general practice, we need to know how GHQ-cases present themselves in general practice. A second question is to what degree patients who present with mental problems in general practice, or those diagnosed by the GP as having psychologically induced complaints are detected by the GHQ.

Such a description might contribute to an insight in the practical relevance of screening for general practice.

Netherlands Institute for Primary Health Care (NIVEL), PO Box 1568–3500 BN Utrecht, The Netherlands.

METHOD

The basic material used to answer these questions consisted, on the one hand, of the GHQ-30 and, on the other hand, of data concerning the presented complaints, and the assessment of the situation by the GP. These data were collected within the framework of a Morbidity and Intervention Survey in General Practice, conducted by the NIVEL. For this study a representative sample survey of Dutch general practices ($n = 160$) was selected and all contacts with patients were registered over a period of 3 months. In addition 100 patients per practice were interviewed; this interview included filling in of GHQ-30 (for further details of this survey see Foets, Van der Velden and Van der Zee[18]). Using this data as a basis it was possible to compare the GHQ scores and the nature of the complaint as recorded in the practice notes.

Having taken the term GHQ score to be a known entity, we must now turn to the way in which the presentation of the problem and the assessment was determined in the morbidity study mentioned. This took place on the basis of the complaint as expressed by the patient and on the basis of the subsequent assessment recorded by the doctor. These variables were registered by the GP during each consultation on a so-called 'reason for encounter form'. Two complaint aspects were assessed on this form.

Reason for encounter

We are dealing here with the complaint presented in the words of the patient. For our analysis we distinguish two groups: complaints which are partly of a psychosocial nature (eg feeling depressed, fear, cannot sleep, relationship problems), and complaints whereby this is not the case (eg cough, stomach ache). The complaint was recorded verbatim by the GP and later coded with the aid of the so-called ICPC coding.[19] Complaints from chapter 'P' (psychological) and 'Z' (social) were considered 'psychosocial' and others somatic by presentation.

Assessment by the doctor

Due to the fact that the ICPC coding system is restricted to a single axis classification in organ systems, the extent to which psychosocial problems are present (eg anxiety, underlying vague complaints; stress causing gastritis) cannot be expressed within this system. This is why each complaint was also assessed by the GP on a five-point scale ranging from entirely somatic to entirely mental. In this way complaints presented as somatic can also be assessed as contributory factors to eventual mental problems.

These two aspects make it possible to distinguish three relevant categories within the scale of problems presented to the GP: complaints which are presented as somatic, which the doctor assessed as being entirely somatic (SS), mental complaints, both in terms of presentation as in terms of assessment by the doctor (PP) and the cases, in which the patient expresses somatic complaints, but which the doctor thinks to be caused by factors other than the purely somatic (SP). The fourth combination, mental complaint in presentation but assessed by the doctor as purely somatic, hardly occurred; therefore we have not included it here.

The interviews in which the GHQs were taken were all carried out in a short space of time, whilst the encounters with the GPs took place over a period of 3 months. For this reason we selected for our analysis all the complaints from patients (who completed a GHQ) who in the week of the GHQ administration consulted their GP. The contacts in this week were subsequently aggregated to patient level, so that each patient only counts once (17% of the patients in that week had either consulted their GP more than once or had presented more than one complaint during an encounter).

RESULTS

In this week 'GHQ patients' had 799 contacts. These contacts originated from 627 patients, of which 151 had one or more values missing, which led to their figures being excluded. The analysis on the week-file takes in 476 patients. The GHQ-30 was completed by all the patients involved in the research taken all together ($n = 8747$) and by the patients from the week's sample survey; the results are shown in Table 1.

Using a cut-off point between GHQ score 4 and 5, as recommended by Goldberg,[3] 20% of the attenders were found to be a 'case'; 12% of the random sample from the population was above this threshold. The GHQ score in the group we selected is in general higher than in the whole study: this can be explained by the fact that in our selection only the patients who consulted their GP in the same period are included.

The psychosocial problems in the group we selected were characterized by the words in which the patient presented the complaint, that is to say as a 'psychosocial problem' or as a 'somatic complaint', and also by the assessment which was made by the doctor, that is to say 'purely somatic' or 'partly of a psychosocial nature'.

In a cross-tabulation the two dichotomies can be seen set against one another (Table 2).

Due to the fact that the data from the table above are

TABLE 1 *GHQ sum scores, in the total sample survey of the National Study and in the week's sample survey*

GHQ sum score	National study ($n = 8747$)	Week sample ($n = 476$)
0	60%	48%
1	14%	17%
2	6%	9%
3	4%	3%
4	3%	4%
5	2%	2%
6	1%	2%
7	1%	2%
8	1%	1%
9	1%	1%
10	6%	11%

TABLE 2 *Patient's complaint (RFE) and its assessment by the GP (n = 476)*

	Somatic RFE	Psychosocial RFE	Total
Somatic assessment	294 (62%)	5 (1%)	299 (63%)
Psychosocial assessment	125 (26%)	52 (11%)	177 (37%)
Total	419 (88%)	57 (12%)	476 (100%)

based on aggregated data, this is not a record of the absolute number of encounters, but data about patients who visited their GP within the period of a week. This implies that a number of patients consulted their GPs more than once. Sixty-two per cent of all patients presented a somatic complaint which the doctor also assessed as somatic; 11% of the patients presented psychosocial complaints, which the doctor also estimated to be psychosocial. In 26% of the patients comorbidity of a psychological nature was involved, at least according to the judgement of the GP. It is worth noting that practically the same relationship between the patient's presentation and GP's assessment was found in the total file from which the week was selected.

The question which is central to this analysis is, of course, how the GHQ compares proportionately to the encounter data. The relationships between the GHQ and the patients' presentations of complaints, and between the GHQ and the assessment of the doctor, have been examined consecutively. Table 3 illustrates the situation which emerges if the patients' presentations of complaints are split into somatic on the one side and psychosocial on the other, and a cut-off point for the GHQ is established at 4/5.

In the same way we can compare the GP's assessment with that of GHQ (Table 4).

Patients with psychosocial complaints have, in a majority of cases, a low GHQ score, both when based on RFE and on the assessment of the doctor. Only 39% (22/57) of the patients who present psychosocial problems and 28% (49/128) of those in whom GP assesses that psychosocial aspects played a part in their complaints, were actually GHQ cases, if we take the cut-off point to be 4. Figure 1 shows comparable percentages obtained by increasing or decreasing the cut-off point for GHQ. Using the chosen cut-off point, the majority of the so called GHQ cases present purely somatic complaints, and the assessment of the GP in the majority of these cases is purely somatic. At this point it should be pointed out that the GHQ cases showed a higher proportion of psychosocial assessments by GPs than did the number of patients presenting psychosocial complaints. Most patients who do not present psychosocial problems during the consultation or whose complaints were assessed as somatic fall beneath the cut-off point for GHQ fall (Figure 2).

We shall round the analysis off by briefly considering the content of complaints.

PP-complaints* from patients with a low GHQ score (termed PP-low), differ from PP-complaints from patients with a high GHQ score (PP-high) in that the psychosocial complaint is not always expressed as the first symptom. PP-cases with a high GHQ-score always have a mental symptom as the most important one. Among the 'PP-low' cases there were half as many mental complaints mentioned as the first symptom (including requests for psychotropic drugs) as among the 'PP-high' encounters. In addition, slightly more patients with a high GHQ score appeared to present complaints which resulted in a psychiatric diagnosis; this does not mean, however, that within the 'GHQ-low' group no psychiatric diagnosis at all has been established.

However, in other respects no differences were found: both 'PP-low' and 'PP-high' encounters highlighted complaints about housing and financial problems.

Which patients present an SS complaint at the consultation, but belong nevertheless to the category of 'psychiatric cases' according to the GHQ? Among the 'SS-high' contacts (in other words: somatic complaints, somatic assessment and a high GHQ score), respiratory disorders and gynecological complaints were over-represented when compared to the 'SS-low' contacts; complaints concerning the limbs and skin were rather under-represented. The complaints that one might expect, like neurovegetative complaints, and complaints related to neurology did not occur in relatively larger numbers within the SS-high group compared with SS low patients.

PP stands for mental complaint combined with a mental assessment. SS stands for a somatic complaint combined with a somatic assessment. SP stands for a somatic complaint combined with a mental assessment.

TABLE 3 *Relationship between RFE (patient's complaint) and GHQ (n = 476)*

	Somatic RFE	Psychosocial RFE	Total
GHQ ≤ 4	341 (72%)	35 (7%)	376 (79%)
GHQ ≤ 4	78 (16%)	22 (5%)	100 (21%)
Total	419 (88%)	57 (12%)	476 (100%)

TABLE 4 *Relationship between the assessment of the GP and GHQ (n = 476)*

	Somatic assessment	Psychosocial assessment	Total
GHQ ≤ 4	248 (52%)	128 (27%)	376 (79%)
GHQ > 4	51 (11%)	49 (10%)	100 (21%)
Total	299 (63%)	177 (13%)	476 (100%)

We mentioned the fact that some adjustments were made to the original material. This was done in order to prevent people who frequently visit their GP to be over-represented and to forestall the occurrence of differences in the intervals between GHQ administration and encounter registration. Each patient was counted once, and only encounters in the week of the interview were selected. Bearing this in mind, we can state that the same tendencies, outlined above, are to be found in the total file (considering all the encounters of all patients in a 3-month period); these measures do not improve if we take a cross-section of a day from the interview period instead of a week.

DISCUSSION

Presentation of complaints and assessments by the GP were compared with the GHQ score of the patients involved; 20% of the patients who visited their GP in a one week period were probable psychiatric cases. This figure is rather low, compared with other studies within general practice. Wilmink[10] for example reports 46% GHQ cases within a random general practice population, and Von Korff *et al.*[20] arrived at a figure of 40%.

Our percentage still exceeds the number of 'cases' detected in a random population sample. The latter result is comparable to that of Finlay-Jones and Burvill,[21] who reported from an Australian morbidity survey that 1 in 60 'GHQ cases' consulted their GP, compared to 1 in 130 non-cases. Goldberg *et al.* [22] also confirmed this finding. 'Those attending a general practitioner are shown to be more psychiatrically disturbed than a random sample of the practice population.' Likewise, Berwick *et al.*[23] report that those who score high on the GHQ are more likely to enter the medical care system.

Twelve per cent of the patients presenting during the study period presented with psychological complaints. However 37% of all complaints were considered to be not entirely somatic by the GP. Both figures are not uncommon in similar studies. Verhaak and Wennink[24] report 9% psychosocial complaints (patients' own words) and 47% non-somatic assessments (by the GP) in a sample of 30 GPs who were slightly more psychologically minded than the sample in this study.

Wilmink[10] reports that 12% patients presented with psychosocial complaints, while 26% of patients were identified as a 'mental health problem' by the GP. His positive definition of a mental health problem can be considered as somewhat stricter than our negatively formulated definition of 'not entirely somatic'.

This study was concerned with the relationship between the GHQ on the one hand and complaints/assessments on the other. This relationship proved to be rather weak in two senses: most psychosocially presented complaints/psychosocial assessments were not GHQ cases, and most GHQ cases presented somatic complaints, assessed as being purely somatic. As Wilmink undertook the same kind of analyses, we may compare our results with his. Most of Wilmink's GHQ cases (58%) were assessed by GPs as having no mental health problem. Most mental health problems (defined by the GP) were detected by the GHQ (74%). The first result fits with the data presented in Table 4, the second does not. We must keep in mind that Wilmink found a higher proportion of GHQ cases and a lower proportion of positive assessments by the GP.

What are the consequences of these results for the evaluation of the GHQ as a screening instrument for general practice? Let us first take the GHQ as the

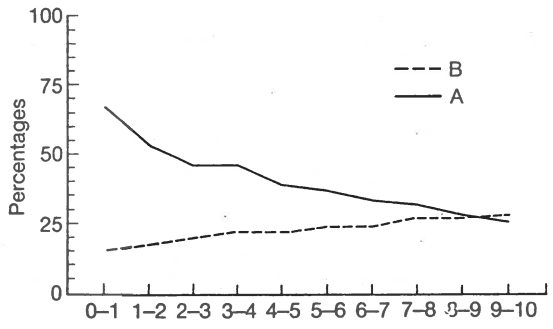


FIGURE 1 *Percentage psychosocial RFE's who are also GHQ cases (A), against percentage of GHQ cases who present psychosocial RFEs (B), with variable cut-off points*

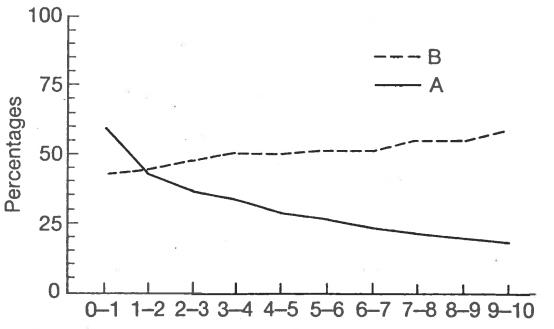


FIGURE 2 *Percentage non-somatic GP assessments who are also GHQ cases (A) against percentage GHQ cases where the GP assessment is non-somatic (B), with variable cut-off points*

departure point for our reflections. We are dealing here with a well-validated instrument, and we may therefore conclude that many patients who consult their GP have a psychiatric problem which they do not reveal to their GP: secondly, the GP, given a somatic presentation of complaints, sees no reason to suspect a psychiatric problem. If we look closely at the sort of complaints involved, then it becomes clear that it is difficult for the GP to find any indicators from the complaint presented by the patient. In actual fact, the somatic complaints made by patients with a high GHQ score and by those with a low GHQ score are strikingly similar. An interesting question to tackle here is the extent to which interweaving of mental and somatic problems (comorbidity) influences the proceedings. Earlier research[24] has shown that in patients with comorbidity the GP tends to follow up the somatic aspects. In this respect it may be worthwhile to point to the nature of the GPs population, which is known to contain a large number of patients with chronic somatic complaints. Van der Zee[25] pointed to the fact that these patients score very high on the 'VOEG', an instrument which bears some similarity to the GHQ, on which a high score indicates not psychiatric problems but chronic discomfort. In the case of the GHQ a high score might also be indicative of somatic indisposition.

Other investigations within our research group[26] have indicated that it is possible to successfully teach a GP a Rogerian approach. The GPs performance proved to be restricted however to an attitude of listening and counselling; they failed to detect a possible psychiatric background to a somatic presentation. Furthermore, we should like to point out that, generally speaking, GPs tend not to make use of their patient's attendance at their surgery to carry out screening activities. Research into preventive activities carried out by GPs has shown that the average GP took the actual request made by the patient as the starting point for his treatment.[25] It seems that GPs have certain reservations about the detection of diseases when there is no hint of a complaint actually in existence.

Finally, it should be pointed out that an essential condition for the execution of a screening program is the availability of a valid intervention program. Authors like Schnabel[27] and Gerards[28] point out that in many cases there is no reason to believe that early detection (without recognition by the patient) would prevent the problems becoming any more serious. If one accepts this viewpoint, it would seem to be undesirable to make the suspicion of psychiatric problems more explicit at an early stage. Caution is called for, because it appears there is a risk of 'mental fixation'[29] when problems are defined as being psychiatric when they have not been presented as such. This might be an explanation for GPs (correct) reluctance to search for problems which the patient is not apt to present spontaneously.

In summary, hidden psychiatric problems do not manifest themselves very conspicuously; a GP lacks the specialist requirements to uncover them and he does not have the intention to do so, when the patient is reluctant to present them. It is doubtful if early detection in those cases would be very sensible.

If we take the GP's psychosocial caseload as a point of departure we are arrested by a striking phenomenon. The GP is in fact confronted with quite a number of patients with psychosocial problems, both explicit and non-formulated, but the majority of these patients are not classifed as psychiatric cases by the GHQ. One possible explanation for this might be the insensitivity of the GHQ for the measurement of borderline psychological problems. The fact that GPs are mostly confronted with milder psychological problems underlines the findings in the 'Nijmegen-Region Project', namely that psychosocial (indicating psychological problems of a relatively mild character, contrary to the psychological problems which belong to psychiatry) problems occur six times as often as psychiatric problems.[30]

How practical is the use of the GHQ as a screening device, given the evidence brought forward? In the present primary care situation, where a large group of undetected mental disorders exists (along with large groups of undetected internal, neurological, and circulatory diseases), it seems only attainable for a GP to wait for more conclusive evidence; this might be constituted by pertaining use of medical care, by accumulating evidence during next visits, etc. After all, one of the other conclusions from Wilmink's study was that the GHQ did not significantly improve the GP's prediction of cases, determined by the Present State Examination. Perhaps the combination of intuition and time (not the duration of a consultation, but the lifelong relationship with a patient in the hands of a GP), is at least as good as a screening device. Further research should be directed to a better understanding of the course of mental problems. It might perhaps in the future be possible to design screening instruments, aimed at reaching a better diagnosis when intervention might be successful, instead of the aim of detecting mental distress without looking at the possibilities of treatment.

REFERENCES

[1] Goldberg D P, Huxley P. *Mental illness in the community: the pathway to psychiatric care.* London: Tavistock, 1980.

[2] Schulberg H C, Burns B J. Mental disorders in primary care. *Gen Hosp Psychiat* 1988; **10:** 79–87.

[3] Goldberg D. *The detection of psychiatric illness by questionnaire.* London: Oxford University Press, 1972.

[4] Goldberg D. Screening for disease. *Lancet* 1974; **ii:** 1245–1248.

[5] Goldberg D. Detection and assessment of emotional disorders in a primary care setting. *Int J Mental Health* 1979; **2:** 30–48.

[6] Goldberg D, Bridges K. Screening for psychiatric illness in general practice: the general practitioner versus the

screening questionnaire. *J R Coll Gen Pract* 1987; **1:** 15–18.

[7] Giel R. Onderzoek naar psychische stoornissen. Een epidemiologische beschouwing. *Tijdschrift Soc Gezondheids* 1983;**6:** 198–203.

[8] Wright A F, Perini A F. Hidden psychiatric illness: use of the general health questionnaire in general practice. *J R Coll Gen Pract* 1987; **4:** 164–167.

[9] Whewell P J, Gore V A, Leach C. Training general practitioners to improve their recognition of emotional disturbance in the consultation. *J R Coll Gen Pract* 1988; **6:** 259–262.

[10] Wilmink F M. Patient, physician, psychiatrist. Assessment of mental health problems in primary care. Groningen: Dissertation, 1989.

[11] Goldberg D. Identifying psychiatric illness among general medical patients. *B Med J* 1985; **291:** 161–162.

[12] Tarnopolsky A, Hand D J, Mclean E K, Roberts H, Wiggins R D. Validity and uses of a screening questionnaire (GHQ) in the community. *Br J Psychiatr* 1979; **134:** 508–515.

[13] Tennant C. The General Health Questionnaire: a valid index of psychological impairment in Australian populations. *Med J Aust* 1977; **9:** 392–394.

[14] Wilkinson M J B, Barczak P. Psychiatric screening in general practice: comparison of the general health questionnaire and the hospital anxiety depression scale. *J R Coll Gen Pract* 1988; **38:** 311–313.

[15] Hodiamont P P G, Veling S H J. Een model voor het bepalen van psychiatrische prevalentie: de relatie GHQ-PSE. *Tijdsch Psychiatr* 1984; **8:** 592–607.

[16] Broom A K, Kat B J B. Would more mental illness services help general practitioners manage their difficult patients? *J R Coll Gen Pract* 1981; **5:** 303–307.

[17] Huygen F J A. A family physician's concept of mental health. In: Visser G J, Bensing J M, Gersons B P R, Maoz B, Margolis C Z. *Mental Health and primary care. Dutch and Israelian experience.* Utrecht: NIVEL, 1986.

[18] Foets M, Velden J van der, Zee J van der. Een nationale studie van Ziekten en verrichtingen in de huisartspraktijk. Projectvoorselt. Utrecht: NIVEL, 1986.

[19] Lamberts H, Wood M (eds). *ICPC, International Classification of Primary Care.* Oxford: Oxford University Press, 1987.

[20] Von Korf M, Shapiro S, Burke J D, *et al.* Anxiety and depression in a primary care clinic. *Arch Gen Psychiatr* 1987; **44:** 152–156.

[21] Finlay-Jones R A, Burvill P W. Contrasting demographic patterns of minor psychiatric morbidity in general practice and the community. *Psychol Med* 1978; **8:** 455–456.

[22] Goldberg D, Kay C, Thompson L. Psychiatric morbidity in general practice and the community. *Psychol Med* 1976; **6:** 565–569.

[23] Berwick D M, Budman S, Damico-White J, Feldstein M, Klerman G L. Assessment of psychological morbidity in primary care: explorations with the general health questionnaire. *J Chronic Dis* 1987; **40 supp:** 71S–79S.

[24] Verhaak P F M, Wennink H J. What do doctors do with psychosocial complaints. Psychiat Med; In Press.

[25] Zee J van der, Verhaak P F M. Preventie, een kwestie van routine. *Med Contact* 1989; **3:** 75–78.

[26] Bensing J M, Sluijs E M. Evaluation of an interview training course for general practitioners. *Social Sci Med* 1985; **20:** 737–744.

[27] Schnabel P. Het eigene van de geestelijke gezondheidszorg. In: Geestelijke Volksgezondheid, Ministerie van WVC, 1984.

[28] Gerards F M. Van kwaad tot erger? De geldigheid van de continuiteitshypothese in de psychosociale preventie. *Maandblad Geestelijke Volksgezondheid* 1985; **3:** 243–258.

[29] Grol R P T M (ed). Huisarts en somatische fixatie. Utrecht: Bohn, Scheltema en Holkema, 1983.

[30] Furer J W, Tax B. Somatische klachten, psychiatrische symptomen en psychosociale problemen. Eindrapportage van het Regioproject geneeskunde, Katholieke Universiteit. Nijmegen: Instituut voor sociale Psychiatry, 1987.

PSYCHOSOCIAL PROBLEMS IN PRIMARY CARE: SOME RESULTS FROM THE DUTCH NATIONAL STUDY OF MORBIDITY AND INTERVENTIONS IN GENERAL PRACTICE

PETER F. M. VERHAAK and MARJA A. R. TIJHUIS

Netherlands Institute of Primary Care, P.O. Box 1568, 3500 BN Utrecht, The Netherlands

Abstract—According to standardized screening instruments, mental distress is a common phenomenon among many patients who visit their general practitioner. However, a number of patients who seem to be in need of mental help do not put forward such a demand for help, whereas other patients who express psychosocial problems to their GP are not considered to be in need, according to a standardized measure. In this paper, a distinction has been made between the objectified needs of the patient as expressed by a standardized assessment, and the demands of the patient, expressed by the Reason for Encounter, stated during their visit at the GP. Results of a follow-up study of two cohorts of patients have been presented: one cohort presented during a 3 month period at least one articulated demand for psychosocial help, a second cohort presented at least one somatic complaint, considered by the GP as being psychological by character, without presenting any psychosocial complaint in that period. Objective needs for mental help of patients in both cohorts were assessed by means of the General Health Questionnaire. During one year all consultations of these two cohorts were registered. The following questions have been put forward: what demands for help have been put forward by the patients, what treatment have these patients got, and what has been the course of the problems during one year of patients with different needs and demands. From the results the following conclusions may be drawn: many patients with a probable mental illness (according to their objective need) present only physical symptoms. The severity of their distress however appears to be less than that of patients with a probable mental illness who do express their psychological distress overtly. More therapeutic effort is directed at psychological symptoms than at somatic symptoms, assessed as being mainly psychosocial by character. Mental health referrals have been made almost exclusively with the former group. Within both groups, most energy is devoted to patients who are really in need, according to the General Health Questionnaire. The majority of the patients with mental health problems (be they overtly presented or not) did not present psychosocial or psychosomatic complaints anymore after a 6 month period. Recovery is higher for patients with psychological symptoms; within each group recovery is higher for patients with a low GHQ-score.

Key words—general practice, mental illness, somatization, treatment of psychological complaints

INTRODUCTION

The prevalence of mental disorders within General Practice has been thoroughly investigated over the last twenty years [1–8]. An important conclusion is that most patients with psychiatric disorders consult their general practitioner regularly but not many of them are recognized by the GP as having mental disorders. Many psychological problems remain untreated, or are treated purely as somatic disorders. If a mental disorder is treated at all, it is done by the GP himself in most cases. Only a relatively small number of them are referred to specialized care.

At first sight, one cannot but conclude that GPs are doing a poor job. They have a low rate of detection and withhold specialized treatment from those patients whose mental disorders have been recognized. Such a conclusion might be drawn too quickly, due to too schematic an approach to the problem at hand. Most of the epidemiological studies mentioned above assessed patients' clinical status by means of a standardized questionnaire or interview at a particular moment in time and compared it with GP's assessment; such an approach gives a random picture in time, neglecting the past history of doctor and patient and also neglecting possible co-morbidity. In earlier publications we have reported that many patients who were probably cases of mental illness, in clinical terms, nevertheless visited their GP with demands for help for physical problems [9]. Videotaped consultations in combination with a questionnaire, completed by the GP, showed that, although mental problems were not discussed during a particular contact, the GP was aware of such problems in the majority of cases [10].

In our opinion, screening tests provide a picture of the objectified need, irrespective of the articulated demands of the patient at that particular moment, whereas a General Practitioner assesses the reason for one particular visit (i.e. the demand) and not the general mental health status of the patient. The question remains however, as to what the implications of such a situation are for the treatment and course of this heterogeneous conglomerate of explicit and hidden mental needs, combined with physical demands.

It has been argued, especially by general practitioners, that in many cases the mental need, which accompanies physical problems might be self-limiting. Explicitation of psychological aspects and treatment of them might be counterproductive [11]. In contradistinction, the possibility of worsening the case by not recognizing these aspects has also been put forward [12, 13].

In this discussion, the former arguments are mainly at the level of demand, whereas the latter point to the objectified need of the patient, regardless of the explicit requests, advanced by the patient. A combination of both points of view might be worthwhile. Consequently, a one year follow-up of patients whose problems were considered by the GP as psychological in nature, was conducted. Some of these patients were probably cases of mental illness, according to standardized measurements, and others were not. Some expressed their need explicitly as psychological complaints, others were somatizing according to the GP. The goal of this study was to describe the treatment and natural course of psychological problems in general practice, in terms of needs as well as demands.

The research questions are:

(1) What demands for help have been put forward by patients, assessed by the GP as being probably mentally ill, with or without an articulated demand for mental help and with or without a probable need for mental help?
(2) What treatment has been given by the GP to patients with various combinations of need and demand?
(3) What is the course of the complaints of these various types of patients and how is it influenced by treatment probable need and explicitation of the psychological character of the problem?

METHOD

Data was collected within the framework of the Dutch National Study of Morbidity and Intervention in General Practice, conducted by the Netherlands Institute of Primary Care (NIVEL). For this study a representative sample survey of Dutch general practices ($N = 103$) was selected and all contacts with patients were registered over a period of three months. The data included the reason for visit, diagnosis, treatment and possibly referral. Reason for visit and diagnosis were classified according to the International Classification of Primary Care (ICPC) [14]. Nine of these practices cooperated in a longitudinal follow-up study. This provided us with the opportunity to select two cohorts of patients based on their illness behaviour during a three month period: one cohort of patients who presented psychosocial reasons for encounter (classified within chapters Psyche or Social of the ICPC) during the 3 month registration of the National Survey and one cohort of patients who presented somatic reasons for encounter (all other chapters of the ICPC), assessed by the GP as being mainly psychosocial,* without presenting explicit psychological or social reasons for the encounter. All contacts with the patients from those two cohorts ($N = 397$ resp. 411) were registered during another nine months on the same registration form as that of the National Study Registration, including reason for visit, diagnosis, treatment and referrals. In addition, all patients completed questionnaires including the General Health Questionnaire (GHQ-30) and the Biographical Problem Inventory.

The GHQ is used as a first stage screening instrument for the identification of mental disorders [1, 2, 8, 15, 16]. It measures the likelihood of a person being classified as a psychiatric 'case'. Given a standardized clinical interview (like for example the Prsent State Examination) as a 'golden standard', the GHQ has proved to be a specific, sensitive instrument [1, 8, 17].

The Biographical Problem Inventory is an adaptation of the Mooney Problem Checklist [18], it is an inventory of problematic experiences (emotional, material, social). Its summation is an indication of the number of problem situations a person has recently experienced.

These questionnaires were completed directly after selection of the cohorts and again one year later, when the follow-up registration had been concluded.

Twenty-one GPs in 9 practices recorded contacts of the two cohorts during a period of one year. During the first 3 months of the National Study, each contact had to be recorded by the GP. For this purpose, a research assistant was present in the practice daily to collect and check the registration forms. After selection of the cohorts, the patients could be recognized on the GP's register. After each consultation with such a patient a registration-form was completed, indicating reason for encounter, diagnosis by the GP, treatment, diagnostics, referral and several assessments by the GP. Every two weeks a research-assistant visited the practice, and collected and checked the forms. It is reasonable to assume that during the first three months the registration covered all contacts completely. As registration was less routine during the follow-up, omissions in this period of nine month might be expected. A reliability study revealed that on average 83% of all contacts with the GP during the follow-up period had resulted in a completed form.

The patients in the two cohorts could not be expected to constitute a random sample. Their symptoms had been assessed at least once as being merely psychosocial in nature. Patients in cohort I (those who expressed their psychosocial complaints explicitly) were more often male and are older on average than patients in cohort II (those who presented

*These complaints will be called 'psychosomatic' in the rest of this paper.

Table 1. Reasons for encounter presented by patients who present explicit psychosocial problems (cohort 1) and by patients who only present somatic problems, assessed by the GP as being mainly psychosocial by nature (cohort 2). Both cohorts divided in GHQ-cases and non-GHQ-cases

	Cohort 1		Cohort 2		Average patients
	GHQ+	GHQ−	GHQ+	GHQ−	
N:	175	216	130	279	
per 100 patients/3-month period					
Complaints	371	322	254	267	208
Psychological complaints	181	152	—	—	14
Psychosomatic complaints	61	84	132	120	43
Somatic complaints	130	136	122	147	150

psychosomatic complaints). The distribution of the sexes in cohort I resembled the average distribution, in cohort II women were overrepresented. In both cohorts patients between 25 and 65 years were overrepresented. They did not differ from one another in this respect. As a consequence, married and divorced people are overrepresented in both groups and widowed and single persons are underrepresented, compared with the average practice populations.

The frequency of doctor–patient contacts in the first cohort during the 3 month selection is significantly higher than the mean number of contacts of the second cohort. Both cohorts have visited their GP more frequently than average patients. There was no difference between the two cohorts in the degree of seriousness of their complaints, as estimated by the GPs.

RESULTS

What proportion of these patients was a probable case of mental illness, according to the GHQ and what demands for help were advanced by them during the 3 months in which the patients were selected?

About 45% of the respondents who put forward explicit psychosocial complaints (cohort 1) scored above the threshold of 4/5 of the GHQ-30. Thirty two percent of the patients who presented psychosomatic complaints were probable cases of mental illness according to the GHQ.

In both cohorts, patients who scored above the threshold of the GHQ had higher frequencies in all respects than patients who scored below the threshold.

From Table 1, it may also be concluded that patients with psychological complaints (cohort 1) do not only limit themselves to explicitly stated psychological complaints. Somatic and psychosomatic complaints, are presented too.

Patients from cohort 2 could, by definition, not present explicit psychosocial complaints. They present however about as many somatic complaints, which are also considered purely somatic by the GP as an average patient.

Because of the selection criteria, the figures for psychological and psychosomatic complaints in both cohorts are not comparable with the figures for average patients. Patients from both cohorts can be distinguished, however, by a high average frequency of contacts, compared with average patients. In this respect, cohort 1 also exceeds cohort 2.

The next question to be answered concerns the treatment of the kind of problems distinguished. In Table 2 treatment of explicitly psychological complaints and somatic complaints, assessed as being mainly psychological is compared, split up by GHQ-caseness.

Complaints that are overtly presented as being psychological get more treatment than complaints which are judged by the GP as being psychological but which are presented in a somatic way. This cannot be attributed to the higher average GHQ-score of the former group; patients with psychological complaints, who are below threshold on the GHQ, receive more treatment than patients with psychosomatic complaints and high GHQ-scores.

Table 2. Treatment of explicitly presented psychosocial problem and of somatizations with a psychosocial assessment by the GP, divided by GHQ-caseness

	Explicitly psychosocial		Somatization with psychosocial assessment	
	GHQ+	GHQ−	GHQ+	GHQ−
N[1]:	216	262	223	363
Treatment				
Counseling by GP	83%*[3]	76%*	64%**	53%
Referral to mental health specialist	14%	11%*	4%	3%
Consultation by GP of mental health specialist	8%	4%*	—	—
Mean duration of visit (min)	13**	11*	10	10

*Difference between treatment of explicit psychosocial complaints and treatment of psychosomatic complaints is significant (t-test $P < 0.05$).

**Difference beween treatment of GHQ+ and GHQ− cases within one problem-category is significant (t-test $P < 0.05$).

[1] *N* differs from preceding tables because patients from cohort 1, presenting psychosomatic symptoms and patients from cohort 2, presenting psychological symptoms have been added to the respective groups.

Table 3. Proportion of patients that did not present any explicit or hidden psychosocial problems in the second half of the year

	Somatization with psychosocial assessment		Explicitly psychosocial	
	GHQ+	GHQ−	GHQ+	GHQ−
N:	130	279	175	216
No psychosocial or psychosomatic problem in 2nd half year	67%	79%	55%	67%
At least one psychosocial or psychosomatic problem in 2nd half year	33%	21%	45%	33%

There is one and a half times as much counselling for psychological symptoms than for psychosomatic symptoms; referrals to and consultation of mental health specialists almost always concern psychological symptoms, and the average duration of a consultation is 20% longer when purely psychological symptoms are concerned.

Within each group, patients who are to be considered as probable cases of mental illness, according to their GHQ-score, are more often treated by the GP himself, as can be concluded from the proportion of patients who get counselling by the GP and the average amount of time spent with them.

What are the consequences of differences in presentation of the—presumably—psychosocial-complaint and consequent differences in treatment?

In Table 3 the proportion of patients who continued to present psychological and psychosomatic symptoms is shown for the two cohorts, subdivided in GHQ-cases and non-GHQ cases.

Initially all patients in both cohorts presented explicit psychosocial or psychosomatic complaints. After 6 months a majority in all groups had recovered, in that sense that they no longer visited their GPs with either somatic manifestations of psychological ill-health (according to the GP) or overt presentation of psychological complaints.

However, 45% of the GHQ-cases and 33% of the non-GHQ-cases who presented psychological complaints at the time of selection, still presented complaints of a psychological origin in the second half year of the follow-up.

Of those patients who initially presented psychosomatic complaints, 33% of the GHQ-cases and 21% of the non-GHQ cases still presented psychosomatic or psychosocial complaints in the last six months.

It can be seen in Table 3 that the changes of recovery are better for non-GHQ cases than for GHQ-cases. Actually, the odds-ratio is 1.66 in the case of psychosocial complaints, and even 1.85 in the case of psychosomatic complaints.

The conclusion that illness behaviour associated with mental problems, be they overtly manifest or hidden behind a somatic presentation, largely tends to disappear after some time would appear to be justified. As the probability of serious mental illness rises, the chances of a reduction in illness behaviour decreases.

To what degree can this decrease be attributed to interventions undertaken by the general practitioner or mental health specialists? See Table 4.

In this respect the results are disappointing. In fact, we were not able to trace any effect at all: patients with psychological complaints with whom the GP had counselling sessions or who had been referred to mental health agencies tended to show about the same rate of recovery as patients for whom this had not been done. There were no differences at all in psychosomatic complaints.

Patients who did receive counselling or a referral were also considered to be more serious cases (cf. Table 2). After controlling for GHQ-caseness, we may conclude that non-cases have benefited slightly more from treatment of psychological complaints (the effect is above level of significance) while for the more severe cases no difference at all between treatment and non-treatment groups can be found.

DISCUSSION AND CONCLUSION

Regarding the relationship between standardized needs and the demands put forward by the patient, the following points may be mentioned.

Within a limited period, many patients with a probable mental illness, in terms of their score on the GHQ, present only physical symptoms. We should

Table 4. Proportion of patients without psychosocial respectively psychosomatic symptoms in the 2nd half year with or without counselling resp. referral

	Counselling		Referral	
	No	Yes	No	Yes
Psychosocial complaints:				
GHQ+:	62% ($N = 78$)	62% ($N = 138$)	62% ($N = 191$)	64% ($N = 25$)
GHQ−:	69% ($N = 111$)	76% ($N = 151$)	72% ($N = 233$)	76% ($N = 29$)
Somatizations				
GHQ+:	73% ($N = 122$)	75% ($N = 101$)	75% ($N = 209$)	64% ($N = 14$)
GHQ−:	79% ($N = 209$)	79% ($N = 154$)	79% ($N = 343$)	80% ($N = 20$)

[1] *N* differs from preceding tables because patients from cohort 1, presenting psychosomatic symptoms and patients from cohort 2, presenting psychological symptoms have been added to the respective groups.

keep in mind in this respect that only cases which were recognized by the GP have been taken into consideration. It is a common finding in general practice that many probable cases of mental illness only present physical complaints [8, 19].

A relatively large number of patients, who are not probable cases of mental illness, nevertheless present psychosocial or psychosomatic complaints to their general practitioner. Though an objectified need has not been established with them, their demands are such, that they surpass the somatizing patients who actually are a probable case of mental illness, in the the attention they require from their GP. Besides, patients who are known to have psychological complaints, visit their GP with purely physical symptoms too.

In summary: patients with an explicit psychological demand exceed patients with somatic reasons for encounter in degree of distress and in the actual appeal they do to their GP.

Within each group, patients with a higher objectified need present more symptoms than patients who are according to the GHQ not in need of mental help.

The result that somatizing patients in general are less psychologically distressed than patients presenting with psychological problems has been reported earlier by Bridges *et al.* [18] and Wright [20]. Wright also confirms our result that patients who present psychological problems visit their GP more frequently with longer consultations than patients who restrict themselves to somatic complaints.

The same conclusion may be drawn as regards treatment: more therapeutic effort is directed at psychosocial problems which are presented as such than at somatic complaints, assessed by the GP as being mainly psychosocial by character. Mental health referrals in particular have been made almost exclusively with the former group of complaints. Within each group (patients who present overtly psychosocial distress and somatizers), most energy is apparently devoted to those patients whose complaints are most serious, as measured by the GHQ. But again, objectified need comes after the subjective demand of the patient. Kessler *et al.* [21] concluded also that GHQ and psychiatric reason for visit were significantly correlated with receiving any mental health treatment, but attached more weight to GHQ-score.

Ormel *et al.* [22] conclude that psychological reasons for encounter, besides severity, recency of onset and psychiatric co-morbidity, contribute to GP's case recognition as well as mental health treatment and outcome. Earlier analysis of videotaped consultations [23] also revealed that in about 90% of the cases a psychologically stated request for help resulted in mental health treatment, whereas reasons for encounter, assessed by the GP as being 'physical complaints with psychosocial background' only resulted in mental health treatment in 50% of the cases.

In general more than half of the persons with mental health problems (be they overtly presented or not) do not present psychosocial or psychosomatic complaints at the end of a 6 month period. A higher proportion of the patients with initial psychosomatic complaints recover than patients with explicit psychosocial complaints. Again, within both groups, the objectified need as indicated by the GHQ introduces a second ranking.

Mann *et al.* [24] report about 50% recovery for a sample of 100 neurotic men and women. The only significant variable, assessed at the beginning, to affect this outcome was the overall severity of the psychiatric illness, assessed by a psychiatrist. Their worse outcome, compared with the outcome reported in Table 4, might be due to a stricter definition of 'caseness'. The effect of severity is not contradictory to our findings: patients with psychosomatic problems tend to recover more than patients with psychological problems (the former as a whole being considered as less severe than the latter [18]) and patients with high GHQ have a poorer rate than patients with a low GHQ.

A significant finding in this respect is the lack of proof of any treatment effect, even after controlling for severity. Ormel *et al.* [21] who found similar results (especially regarding counselling and referral) hypothesize that the less effective GPs are in their interventions, the more mental health treatment they will provide.

In the introduction we distinguished the demand for help, representing the subjective experience of the patient, and the objectified need, as expressed by standardized instruments. We may conclude from our findings that such a need relatively often results in a request for help; such a request is a good guarantee that actual treatment will be offered. If a patient sticks to somatizations, the odds that the symptoms disappear are nevertheless higher than for psychological symptoms.

We cannot but conclude from patients' illness behaviour that a lot of patients do not put forward any demands for help at the end of a year, although a lot of them keep a high score on GHQ. However, the common sense theory that most psychological distress disappears as time goes by, cannot be considered as being corroborated. Too many patients continued to suffer from their psychological complaints. As our results do not indicate that specialists in mental health care are more successful than general practitioners, it does not seem fair to put the blame for this situation on the side of the GP. On the contrary, though general practice should do better still, it should be pointed out once more, that most energy was spent on those patients who 'objectively' needed it most. The task in specialists' hands is perhaps to keep on trying to provide general practice with better tools for those areas where yet more is to be gained.

REFERENCES

1. Goldberg D. P. and Bridges K. Screening for psychiatric illness in general practice: the general practitioner versus the screening questionnaire. *J. R. Coll. Gen. Pract.* **37,** (1), 15, 1987.
2. Goldberg D. P. and Huxley P. *Mental Illness in the Community, the Pathway to Psychiatric Care.* Tavistock, London, 1980.
3. Lamberts H. and Hartman B. Psychische en sociale problemen in de huisartspraktijk. *Huisarts en Wetenschap* **25,** 333, 1982.
4. Sartorius N., Goldberg D., Girlamo G. de, Costa e Silva J. A., Lecrubier Y. and Wittchen H-U. (Eds) *Psychological Disorders in General Medical Settings.* Hogrefe & Huber, Toronto, 1990.
5. Shepherd M., Cooper B., Brown A. C. and Kalton G. *Psychiatric Illness in General Practice,* 2nd edn. Oxford University Press, Oxford, 1981.
6. Shepherd M., Wilkinson G. and Williams P. (Eds) *Mental Illness in Primary Care Settings.* Tavistock, London, 1986.
7. Wilkinson G. *Mental Health Practices in Primary Care Settings. An Annotated Bibliography.* Tavistock, London, 1985.
8. Wilmink F. M. Patient, physician, psychiatrist. Assessment of mental health problems in primary care. Dissertation State University of Groningen, 1989.
9. Verhaak P. F. M., Wennink H. J. and Tijhuis M. A. R. The importance of the GHQ in general practice. *Family Practice* **7,** (4), 319, 1990.
10. Bensing J. M. and Beerendonk P. Psychosociale problemen in de huisartspraktijk. (Psychosocial problems in general practice). *Maandblad Geestelijke Volksgezondheid* **45,** (6), 595, 1990.
11. Huygen F. J. A. A family physician's concept of mental health and mental health services. In *Mental Health and Primary Care* (Edited by Visser G. J., Bensing J. M., Gersons B. P. R., Maoz B. and Margolis C. Z.). Dutch and Israeli experience. NIVEL, Utrecht, 1986.
12. Goldberg D. Reasons for misdiagnosis. In *Psychological Disorders in General Medical Settings* (Edited by Sartorius N., Goldberg D., Girlamo G. de, Costa e Silva J. A., Lebrubier Y. and Wittchen H-U.). Hogrefe & Huber, Toronto, 1990.
13. Ormel J. and Giel R. Medical effects of nonrecognition of affective disorders in primary care. In *Psychological Disorders in General Medical Settings* (Edited by Sartorius N., Goldberg D., Girlamo G. de, Costa e Silva J. A., Lecrubier Y. and Wittchen H-U.). Hogrefe & Huber, Toronto, 1990.
14. Lamberts H. and Wood M. (Eds) *ICPC, International Classification of Primary Care.* Oxford University Press, Oxford, 1987.
15. Goldberg D. P. *The Detection of Psychiatric Illness by Questionnaire.* Oxford University Press, London, 1972.
16. Wright A. F. and Perini A. F. Hidden psychiatric illness: use of the General Health Questionnaire in general practice. *J. R. Coll. Gen. Pract.* **37,** (4), 164, 1987.
17. Wilkinson M. J. B. and Barczak P. Psychiatric screening in general practice: comparison of the general health questionnaire and the hospital anxiety scale. *J. R. Coll. Gen. Pract.* **38,** 311, 1988.
18. Furer J. W. and Tax B. (Eds) Somatische klachten, psychiatrische symptomen en psychosociale problemen: Prevalentie, overlappingen en verbanden met demografische variabelen. Eindrapport van het Regioprojekt Nijmegen, deel 1, 1987.
19. Bridges K., Goldberg D., Evans B. and Sharpe T. Determinants of somatization in primary care. *Psychol. Med.* **21,** 473, 1991.
20. Wright A. F. A study of the presentation of somatic symptoms in general practice by patients with psychiatric disturbance. *Br. J. gen. Pract.* **40,** (11), 459, 1990.
21. Kessler L. G., Amick B. C. and Thompson J. Factors influencing the diagnosis of mental disorder among primary care patients. *Med. Care* **23,** (1), 50, 1985.
22. Ormel J., Brink W. vd., Koeter M., Giel R., Meer K. vd., Willige G. vd and Wilmink F. W. Recognition, management and outcome of psychological disorders in primary care: a naturalistic follow-up study. *Psychol. Med.* **20,** 909, 1990.
23. Verhaak P. F. M. and Wennink H. J. What does a doctor do with psychosocial problems in primary care? *Int. J. Psychiat. Med.* **20,** (2), 151, 1990.
24. Mann A. H., Jenkins R. and Belsey E. The twelve month outcome of patients with neurotic illness in general practice. *Psychol. Med.* **11,** 535, 1981.

Alcohol consumption and general health status in the Dutch national survey of general practice.

An analysis of the position of abstainers.

Marleen Foets* and Margriet E. van Baar*

Introduction

The relation between health status and lifestyle factors such as smoking, physical exercise, dietary habits and alcohol consumption is the subject of much research.

The association between alcohol use and mortality often has been reported to be U-shaped: moderate drinkers would be at a lower risk than heavy drinkers and abstainers[1].

Most evidence for a non-linear, U-shaped relationship occurred in the case of coronary heart disease (CHD). In many ecological, case control and prospective studies, an inverse relationship has been found between moderate alcohol use and coronary heart disease, while excessive alcohol use often seems to increase the risk of CHD[2]. In his review of prospective studies on alcohol and mortality, Shaper[1] also concluded that there is a U-shaped association between alcohol consumption and cardiovascular disease, coronary heart disease and total mortality.

Several biological mechanisms can explain this U-shaped relationship. A recent review of the different physiological effects of alcohol relevant to CHD-pathogenesis was given by Veenstra[3]. The most widespread explanation for the U-shaped association with cardiovascular morbidity and mortality is the U-shaped association between alcohol use and the level of high density lipoprotein cholesterol (HDL-C) and blood pressure. There seems to be a potential benefit on the HDL-cholesterol level up to two glasses per day. This effect seems to disappear at a higher consumption level, especially through the rise of blood pressure[4].

In the case of indicators of perceived health

Summary

The aim of this study was to investigate the relationship between alcohol consumption and perceived health status. Special attention was given to the position of the abstainers. The assumed relatively low health status of abstainers was examined by testing two research questions:

- Is the health status of ex-drinkers relatively bad compared with the health status of life-long abstainers?
- Are there still health differences between abstainers and alcohol consumers, after adjustment for other lifestyle habits and socio-economic status?

Results are based on a population based health survey from a sample of 10,728 persons, aged 15 years and older. Included data are age, sex, perceived health indicators, self-reported alcohol use, measures for smoking habits, exercise, overweight and socio-economic status.

Alcohol consumption and perceived health status were significantly associated especially between 25-64 years of age. Abstainers and to a lesser extent the excessive drinkers reported a lower health status.

In analysing the position of the abstainers, firstly a distinction was made between ex-drinkers and life-long abstainers. Both groups hardly differed in their perceived health status. The possibility that ex-drinkers contribute to the relative low health status of abstainers was not confirmed.

Secondly, after controlling for other lifestyle habits and socio-economic status, alcohol consumption remained positively associated with a low perceived health status. Therefore, the supposition concerning the fade-out of health differences between abstainers and alcohol consumers, after controlling for other lifestyle habits and socio-economic status, was also not confirmed.

An important restriction of this study lies in the cross-sectional design. Further longitudinal research is required to clarify the position of abstainers.

Keywords: Alcohol drinking, health status, the Netherlands

* *Netherlands Institute of Primary Health Care (NIVEL), Utrecht, the Netherlands*

Address for correspondence and reprint requests: Dr. M. Foets, Netherlands Institute of Primary Health Care (NIVEL), Posthus 1568, 3500 BN Utrecht, The Netherlands

Int. J. Health Sciences 1993, Vol. 4, no. 2, pp. 75-84

status, research results are less consistent. In a review article Colsher and Wallace[2] reported inconsistent results for indicators of general morbidity, varying from a symptom checklist to hospitalisation. Drenthen et al[5] found that abstainers and excessive drinkers rate their own health worse than moderate drinkers and score higher on the neuroticism scale. Also more psychiatric cases were found in these groups. The number of psychosocial problems was higher with a higher drinking level. These relationships between alcohol consumption and health status were hardly found in the younger age group. Mechanisms to explain these associations are more difficult to give.

However, many studies into the relationship between health status and alcohol consumption are subject to critical comments the interpretation of the results is much debated: often the question is raised wether the association is an artefact[1,4].

Firstly, the position of the abstainers remains unclear. Indeed, it is often impossible to distinguish the former drinkers from the life-long abstainers. Current abstainers may be former heavy drinkers. They may be abstainers because of poor health status, including health damaged by previous drinking[1]. In the case of life- long abstainers, reasons for abstinence can be diverse: motivations may be found e.g. in religious beliefs or in a healthier lifestyle. Following this hypothesis, life-long abstainers may have a relatively good health status, compared to ex-drinkers. The U-shaped association between the use of alcohol and ill health may be due to the relatively bad health status of former drinkers. Therefore the importance of the so far very limited research on the heterogeneity within the group of abstainers is emphasized[1,4,5]. Recently an investigation by Jackson et al[6], in which it was possible to differentiate former drinkers from never-drinkers, showed no evidence that the U-shaped relationship was due to the migration of former drinkers to the abstainer-group.

Secondly, biological mechanisms may not be the only possible explanation for the association between health status and the use of alcohol: behavioural and social factors may influence the relationship as well. The association between alcohol consumption and ill health may well be attributable to other factors such as other lifestyle habits, social status and personality. Marmot[7] already suggested the possibility that teetotallers may have an increased risk of coronary heart disease for reasons such as personality and diet.

Blaxter[8] emphasized that the effects of healthy habits should not be considered in isolation, because few people's lifestyles are totally healthy or unhealthy. Therefore, ideally speaking information on different sorts of health behaviour has to be obtained.

Another determinant of health status receiving much attention is socio-economic status. There is accumulating evidence pointing at the health risk accompanying low socio-economic status. Moreover, the higher morbidity and mortality rate among abstainers may reflect their lower socio-economic status[4].

The question whether the relatively bad health status of non-drinkers is the result of a less healthy lifestyle or of a lower socio-economic status remains unanswered.

In this article we focus on the position of abstainers. The assumed relatively low health status of abstainers was examined by testing the following questions:

- Is the health status of former drinkers relatively low compared to the health status of life-long abstainers?
- Are there still differences in health status between abstainers and alcohol consumers, after adjustment for other lifestyle habits and socio-economic status?

Prior to answering the research questions, the alcohol consumption in the study population is described and related to the perceived health status.

Methods

Study population

This analysis is based on data of the Dutch National Survey of General Practice. In this study a random sample of 161 general practitioners participated. A health interview was carried out in a random sample of 100 registered patients in the participating practices. The response rate amounted to 76.6%. Persons younger then 15 years were withdrawn form this analysis (N=10,728). The study has been described in detail by Foets et al[9] and by Foets and Sixma[10].

Selected variables

This analysis was based on information collected in the health interview study on health status, socio-demographic characteristics and lifestyle habits.

As indicators of *perceived health status* were selected:

- the number of acute complaints during the two weeks before the interview took place
- the number of chronic conditions (Van den Bos[11])
- the score on the general health questionnaire (GHQ), a psychiatric screening questionnaire (30 items)[12]
- the score on a questionnaire, concerning 7 life areas designed to detect psychosocial problems[13]

The *socio-demographic variables* included are: age, sex, educational status and professional status. Educational and professional status are used as indicators of *socio-economic status*. Professional status is based on a prestige scale from Ultee and Sixma ranging from 0 to 100[14].

The following *lifestyle habits* were included as dichotomic variables (yes/no) in the analysis: smoking habits and physical exercise. Overweight was based on the Body Mass Index (BMI) (in kg/m2): the threshold for overweight was defined at a BMI of 25 kg/m2, according to the classification of Garrow[15].

The self-reported *alcohol consumption* in this study included several measures:

- the absolute alcohol consumption, differentiating current drinkers from non-drinkers. These groups were based on a question whether or not alcoholic beverages had been consumed in a period of six months before the interview. Respondents were classified as current alcohol-abstainers if they had reported no use of alcoholic beverages in the past 6 months.
- current abstainers were distinguished in a group of ex-drinkers and a group of life-long abstainers. Current abstainers were asked how long it was since they last used alcoholic beverages; those reporting no alcohol history were classified as life-long abstainers; the others were the ex-drinkers.
- usual frequency of alcohol drinking during the past six months, based on a question listing seven options ranging from seldom to every day.
- usual quantity, based on a question on the number of glasses daily, weekly or monthly consumed.

On the basis of the information on usual frequency and on usual quantity a quantity-frequency-index (QF-index) was computed for the current drinkers. This index is comparable to the one constructed by Lemmens[16].

Methods of analysis

The distribution of usual alcohol quantity was positively skewed. Therefore a logarithmic transformation was undertaken, resulting in a geometric mean, instead of an untransformed arithmic mean.

Socio-demographic characteristics are potential confounders of the relation between alcohol consumption and health status. On the basis of a preparatory analysis, age and sex were used as control variables, being the most strongly related variables to both alcohol consumption (Cramers V $\leq$ .32) and health indicators ($\eta \leq .13$).

Because of the expected interactions between sex, age and alcohol consumption, the statistical method used to analyse the relation between alcohol consumption and health status (the first research question) was analysis of variance.

For the analysis of the relative importance of lifestyle and socio-economic status for perceived health status (the second question), regression analysis was used. In a preliminary analysis the correlations between the independent variables were low ($r \leq .15$), with some exceptions (alcohol & smoking r=.29, alcohol & education r=-.25 and profession & education r=.26). All analyses were performed using SPSS software[17].

Results

Alcohol consumption in the study population

Table 1 shows the distribution of the reported alcohol consumption for both men and women. The QF-index was used to indicate the alcohol consumption of the current drinkers. Women reported the highest prevalence of abstinence and consumed less alcohol compared with men.

Alcohol consumption varied with age. The highest level of consumption and a relatively low prevalence of abstainers (6.8% for men, 22.9% for women) was reported in the category 25-44 years. In the older age- categories (45-64 and 65 years and older) the number of abstainers was higher for

Table 1. Alcohol consumption by sex and age (in %) (N=10787).

Age	15-24	25-44	45-64	65+	total
Men (N=5260)					
Life-long abstainer	15.7	3.2	5.5	9.9	7.0
Ex-drinker	4.8	3.6	10.6	15.2	7.0
Minimal drinker	19.8	17.2	20.8	22.1	19.2
Mild drinker	46.4	43.4	34.1	29.4	40.0
Moderate drinker	10.0	20.6	18.9	17.1	17.7
Excessive drinker	3.2	12.0	10.1	6.4	9.1
Women (N=5468)					
Life-long abstainer	23.6	13.0	20.5	34.2	20.0
Ex-drinker	7.9	9.9	14.8	19.3	12.1
Minimal drinker	31.1	29.2	26.9	25.7	28.5
Mild drinker	34.7	37.4	28.4	16.3	31.6
Moderate drinker	2.4	9.0	7.5	3.9	6.6
Excessive drinker	0.3	1.6	1.9	0.6	1.3

both men and women. The prevalence of abstinence in the highest age-category was 25.1% for men and 53.5% for women.

Within the group of abstainers, 62% of the females and 50% of the males could be classified as life-long abstainers. However, there were large differences between the different sex/age groups. As expected, the relative number of life-long abstainers was especially high in the 15-24 age-groups, and for women, also in the 65+ age-group. On the contrary, the relative number of life-long abstainers was relatively low in males of 45 and older.

The former alcohol consumption level reported by the ex-drinkers was low compared with the consumption level of the actual drinkers. The mean reported former alcohol consumption was respectively 2.1 glasses per week in the ex-drinkers and 4.0 glasses per week in the current drinkers (geometric means), after standardisation for age and sex.

Alcohol consumption and perceived health status

Table 2 shows the results of the analysis of variance of alcohol consumption on perceived health status.

When looking at main effects, most indicators of perceived health status were, as expected, significantly associated with age and sex. Three of the health indicators, the number of acute complaints, of psychosocial problems and the GHQ-score, were significantly associated with alcohol consumption.

In interaction with age, the alcohol consumption was significantly associated with all indicators of perceived health status. In interaction with sex, alcohol consumption was significantly associated only with psychosocial problems.

Figure 1 gives a graphical representation of the relationship between alcohol consumption and health status within each age-sex group. The associations between perceived health status and alcohol consumption occurred particularly in the age groups 25-44 and 45-64. There were hardly significant differences in perceived health status in the youngest and the oldest age-groups.

Table 2. Perceived health status by alcohol consumption, age and sex (ANOVA) (N=10.728)

Source of variance	Df	F-value Acute compl.	Chronic conditions	Psychosoc. problems	GHQ-score
Main effects					
alcohol consumption	4	1.8	0.4	6.1 **	3.4 **
age	1	13.0 **	503.7 **	8.6 *	1.6
sex	1	71.5 **	9.5 **	25.1 **	19.0
Interaction effects					
alcohol consumption + age	4	6.9 **	4.1 **	5.1 **	6.7 **
alcohol consumption + sex	4	0.9	1.5	5.2 88	1.4
age + sex	1	6.4 *	0.7	5.2 *	1.1
Residual variance		17.1	1.5	4.2	16.1
df		10712	10712	10511	10234

* p<.05 ** p<.01

Figure 1. Perceived health status and alcohol consumption according to age and sex

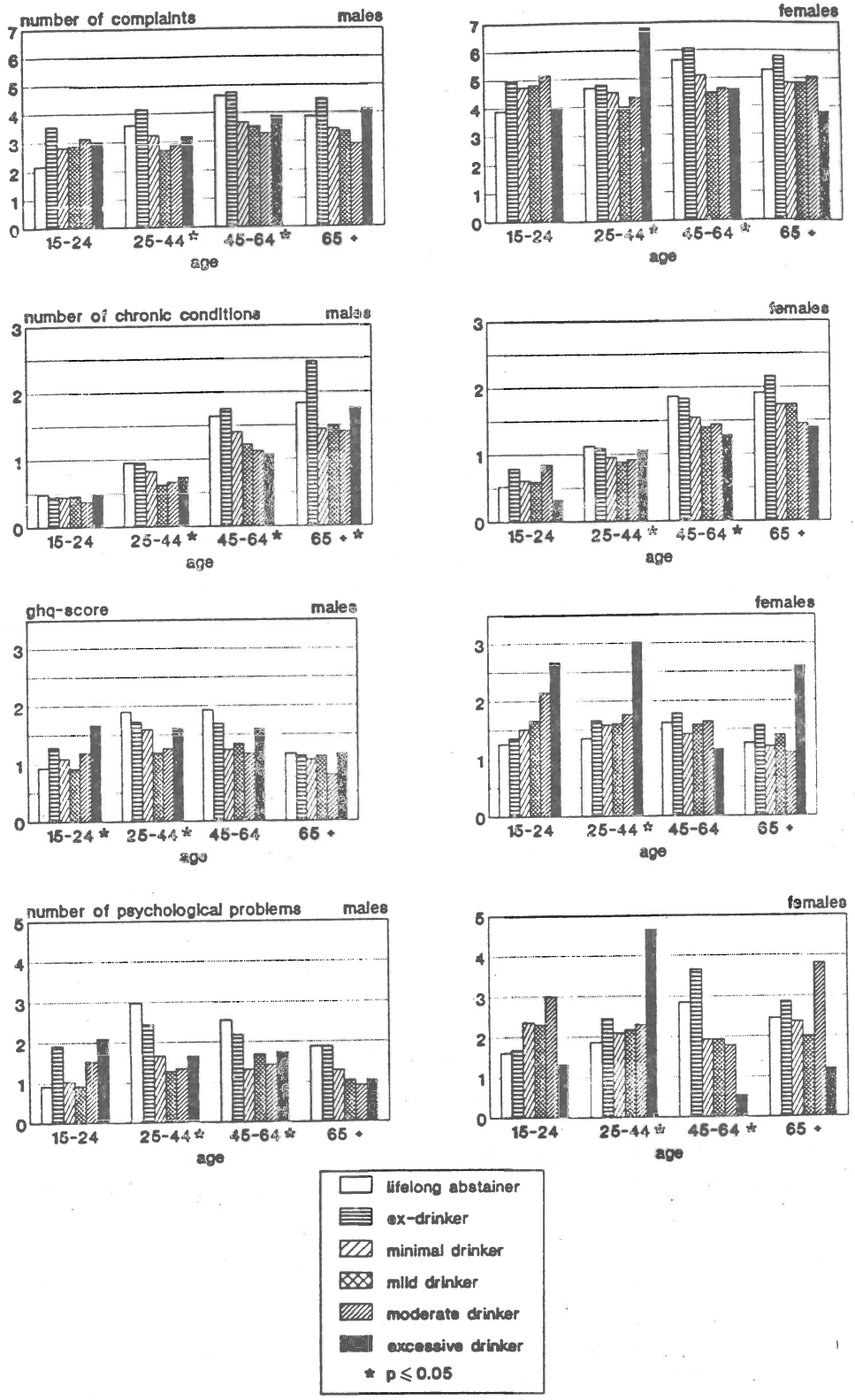


Table 3. Perceived health status by type of abstinence (ex-drinker vs. life-long abstainer), age and sex (ANOVA) (N=2489).

Source of variance	Df	F-value Acute compl.	Chronic conditions	Psychosoc. problems	GHQ-score
Main effects					
type of abstinence	1	3.9 *	0.3	3.7	1.0
age	1	28.7 **	329.1 **	1.4	10.0 **
sex	1	10.4 **	4.3 *	0.7	0.3
Interaction					
type of abstinence + age	1	0.8	2.0	2.0	0.1
type of abstinence + sex	1	0.3	0.1	1.0	1.0
age + sex	1	0.8	2.9	0.6	0.1
Residual variance		23.5	2.1	5.0	21.8
df		2482	2482	2413	2341

* p<.05 ** p<.01

The general pattern is as follows. In the age/sex groups of 25-44 years and of 45-64 years the association between alcohol consumption and perceived health status was significant in all cases but two. In these age-groups, and especially in the male population, minimal, mild and moderate drinkers reported the best perceived health. Abstainers in general reported most health problems. This again was especially the case in the male population. In the female population this was only the case for the 45-64 age-group. In the age group between 25 and 44 excessive drinkers in general reported most health problems: the number of complaints, of psychosocial problems and the GHQ-score are dramatically high here.

question 1: type of abstinence and perceived health status

One possible explanation for differences in health status between abstainers and alcohol consumers is the relatively bad health status of former drinkers as compared to life-long abstainers.

In an analysis of variance (Table 3) the relations between type of abstinence, age, sex and perceived health status were examined. The importance of type of abstinence was relatively low compared to the importance of age.

The main effect of type of abstinence was significant (p<.05) only for the number of acute complaints: ex-drinkers reported more complaints then life-long abstainers.

These results are illustrated by means of the age and sex adjusted scores for perceived health status for both ex-drinkers and life-long abstainers (Table 4).

The ex-drinkers reported a lower health status compared with life-long abstainers. However, after correction for age, sex and their interaction most differences were not significant. The difference in acute complaints was the only exception (p=.048).

question 2: Alcohol consumption, other lifestyle habits, socio-economic status and perceived health status

Other lifestyle habits and socio-economic status possibly contribute to differences in health status between abstainers and alcohol consumers. Therefore, the abstainers and alcohol consumers were compared for lifestyle habits and socio-economic status, after standardisation for age and sex.

A higher percentage of abstainers had scores that point to moderate or serious overweight according to the Body mass index (7.9% vs. 5.1%). They also reported less exercise in an active way (25.4% vs.

Table 4. Perceived health by type of abstinence, corrected for age and sex. P-values are computed, corrected for age, sex and their interaction (ANOVA).

Abstinence	ex-drinkers N=1027	life-long abstainers N=1462	p-value
Acute complaints	5.04	4.48	.048*
Chronic conditions	1.46	1.34	.571
Psychosocial problems	1.58	1.37	.053
GHQ-score	2.54	2.12	.330

36.7%). However, abstainers less often reported to smoke (33.5% vs. 42.6%), compared with alcohol consumers. Thus, according to those indicators, abstainers did not show an unambiguous lifestyle with regard to health risks.

There were also differences in educational and professional status, other potential confounders of a relation between alcohol consumption and health status. Abstainers more often had a lower educational level, consisting of primary school or a lower vocational school (69.9% vs. 51.9%). Their professional status was also lower, referring to the values 0 to 49 on the prestige scale from Ultee and Sixma (74.6% vs 61.3%). These findings point to a relative low socio-economic status of abstainers, which may contribute to the differences in perceived health between abstainers and alcohol consumers.

The relative importance of these indicators of lifestyle and socio-economic status for health status was established by a regression analysis (Table 5).

As expected, the feminine sex was associated with a lower perceived health status. A higher age was associated with a lower health status only in the case of acute complaints and chronic conditions.

Lifestyle habits were also often associated with health status. Alcohol consumption was significantly associated with perceived health status, except for the score on psychosocial problems: alcohol consumption was related to a better health status. Smoking and overweight were significantly associated with 3 resp. 2 indicators of perceived health status; both characteristics pointed to a relatively low health status. Exercise was significantly associated with all indicators of perceived health status and was associated with a better health. Concerning the socio-economic status indicators, a higher education was associated with a relatively high number of chronic conditions and of psychosocial problems.

Discussion

We analysed differences in health status between abstainers and alcohol consumers.

The results of our analysis showed a relation between alcohol consumption and perceived health status especially in the age-categories of 25-44 and 45-64 years old. In these age-categories abstainers and, to a lesser extent, the excessive drinkers reported a lower health status. However, for women in the age-group of 45-64 years an inverse relation between alcohol consumption and perceived morbidity was found.

Thus, in general abstainers clearly reported the lowest health status. Therefore as in previous research, there were indications for the existence of a non-linear, U-shaped, association between alcohol consumption and health status.

Until now, support for a U-shaped relation between alcohol consumption and health has especially been found with respect to cardiovascular diseases[2]. In our analysis the perceived health status was object of study. Our results correspond to some degree with those found by Drenthen et al[5]. They also found that for most indicators, the perceived health status of moderate drinkers is better than the health status of abstainers and ex-

Table 5. Health status by age, sex, lifestyle habits and socio-economic status. Multiple regression-analysis, in standardised regression coefficients (beta) (p-value * < .01, ** < .05)

	acute complaints	chronic conditions	psychosocial problems	GHQ-score
Age	0.059 *	0.309 *	-0.006	0.011
Gender (m/w)	0.181 *	0.073 *	0.077 *	0.093 *
Alcohol (no/yes)	-0.032 *	-0.063 *	-0.016	-0.049 *
Smoking (no/yes)	0.077 *	0.000	0.091 *	0.077 *
Overweight	0.045 *	0.045 *	0.005	0.013
Exercise (no/yes)	-0.047 *	-0.038 *	-0.035 *	-0.027 **
Education	0.011	0.027 *	0.067 *	-0.015
Professional status	-0.015	-0.018	-0.004	0.007
R^2	0.05	0.14	0.02	0.02

cessive drinkers. Similarly, they found hardly any relationship in the younger age group. Persons of 65 and older were not included in their research.

In contrast to most findings in previous research, in which the health status of excessive drinkers was worst, in this study in general the health status of abstainers was worst. This again is similar to the results found by Drenthen et al[5]. Perhaps excessive drinkers with a very poor health were not included in this research because of a selectively high non-response in this group[18]. The problem of bias due to selective non-response increases in importance when non-response rates are high. However, compared to most surveys in general populations, the response rate in this survey was, relatively high (nearly 77%). Moreover, a study of Lemmens led to the conclusion that the effect of non-response on estimates of alcohol consumption is small[16].

Abstainers are used as the comparison group in many studies, but their position remains unclear. Differences in health status between abstainers and alcohol consumers may be the result of differences in health status between former drinkers and life-long abstainers. These differences may also be the result of factors such as other lifestyle habits and the socio-economic status. Both possibilities were examined in this study.

First, a distinction was made between ex-drinkers and life-long abstainers. The differences in perceived health status between both groups were hardly significant. The possibility that the ill health of the ex-drinkers contributes to the relatively low perceived health status of abstainers cannot be confirmed. Recently Jackson et al[6] came to a similar conclusion in an investigation into the relationship between alcohol consumption and risk of coronary heart disease: former drinkers had a lower risk of myocardial infarction than never-drinkers, but a similar risk of fatal coronary heart disease. Perhaps in line with this result is our finding that ex-drinkers did not report a high former alcohol consumption level. On the contrary, they reported a rather low consumption level, but this may also reflect the fact that they decreased their drinking gradually.

In a second attempt to clarify the position of the abstainers, other lifestyle habits and socio-economic status were included in the analysis. The consumption of alcohol in general remained positively associated with perceived health status. Thus, the supposition concerning the fade-out of health differences between abstainers and alcohol consumers, after controlling for other lifestyle habits and socio-economic status, is not confirmed either.

In this study we looked at the self-reported perceived health status. The result that abstainers, and to a somewhat lesser degree, excessive drinkers perceive their health as being worse compared to moderate drinkers, could not be explained by our suggestion, pointing in the direction of an artefact. Therefore the question arises what mechanisms can explain these results. Possibly a greater health consciousness contributes to the relatively low perceived health status of abstainers. A great health consciousness can lead to an early signalling of defects and a following low perceived health status[19]. Drenthen et al[5] found some evidence for this possibility: abstainers turned out to pay more attention at health information and hive a hogher score on preventive behaviour. Since excessive drinkers report a lower health status as well, Drenthen et al[5] suggest that both abstainers and excessive drinkers may be characterised by psychological instability; other factors determine whether this results in abstaining or in excessive drinking.

This study is based on the so-called frequency – quantity index, based on a few recall questions in a health interview. Several factors may be threats to the validity of these self-reported alcohol consumption data. Besides the problem of selective non-response, already mentioned, there is the problem of underestimation of alcohol consumption by surveys as a consequence of several factors, of which memory effects is one of the main error sources. Besides unconscious underreporting, respondents may deliberately underreport their alcohol consumption because they experience questions on alcohol as threatening. An important question here is whether underestimation is uniform across all consumption levels. A study of Lemmens indicated that the relationship between the consumption level and underreporting is linear[16]. Therefore we assume that our index does not to a large degree misclassify the respondents.

Lemmens also found some indications suggesting that heavy drinkers do not particularly underreport their consumption.

An important restriction of this study lies in the

cross-sectional character of the analysed data. In this way only a restricted insight can be obtained in the dynamics of alcohol consumption and health status. A more thorough study of these dynamics, especially focused on the change from drinking towards non-drinking behaviour, will help to understand the position of abstainers. It is a fact that, in a cross-sectional design, the reporting of more health problems in life-long abstainers cannot be interpreted as alcohol being health-protective. A relatively bad reported health in life-long abstainers may for example reflect the presence of chronic conditions present early in life which may condition drinking habits.

Further research is needed to clarify the position of the health status of abstainers, in comparison with minimal to moderate drinkers.

Acknowledgments

The Dutch National Survey of General Practice was financed by the Ministry of Welfare, Health and Cultural Affairs.
We thank the participating physicians and their patients as well as the research team which conducted the survey. We are especially grateful to Harald Abrahamse for technical assistance.

References

1 SHAPER AG. Alcohol and mortality: a review of prospective studies. Br J Addiction 1990;85:837-847.
2 COLSHER PL. WALLACE RB. Is modest alcohol consumption better than none at all? An epidemiological assessment. Ann Rev Public Health 1989;10:203-219.
3 VEENSTRA J. Moderate alcohol use and coronary heart disease: a U-shaped curve? World Rev Nutr Dietetics 1991; 65:38-71.
4 CRIQUI MH. The reduction of coronary heart disease with light tot moderate alcohol consumption: effect or artifact? Comment on Shaper's Alcohol and mortality: a review of prospective studies. Br J Addiction 1990;85:854-857.
5 DRENTHEN T. TAX B. TER HEINE E. et al. Alcoholconsumptie en gezondheidsproblemen. T Soc Gezondhz 1987;65:158-164.
6 JACKSON R. SCRAGG R. BEAGLEHOLE R. Alcohol consumption and risk of coronary heart disease. BMJ 1991; 303: 211-216.
7 MARMOT MG. Alcohol and coronary heart disease. Int J Epidemiol 1984;13:160-167.
8 BLAXTER M. Health and lifestyles. London: Routledge, 1990.
9 FOETS M. VAN DER VELDEN J. DE BAKKER D. Dutch National Survey of General Practice. A summary of the study design. Utrecht, NIVEL 1992.
10 FOETS M. SIXMA HJM. Een nationale studie naar ziekten en verrichtingen in de huisartspraktijk. Basisrapport: Gezondheid en gezondheidsgedrag. Utrecht, Nederlands Instituut voor onderzoek van de eerstelijnsgezondheidszorg 1991.
11 VAN DEN BOS GAM. Zorg van en voor chronisch zieken. Utrecht, Bohn, Scheltema en Holkema 1989.
12 GOLDBERG D. The detection of psychiatric illness by questionnaire. Maudsley Monographs no. 21. London. Oxford University Press 1972.
13 FÜRER JW. TAX B. eds. Somatische klachten, psychiatrische symptomen en psychosociale problemen. Eindrapport van het Regioproject Nijmegen, deel 1. Nijmegen: Instituut voor Sociale Geneeskunde. Katholieke Universiteit Nijmegen 1987.
14 SIXMA H. ULTEE W. Een beroepsprestigeschaal voor Nederland in de jaren tachtig. Mens en Maatschappij 1983;4:360-382.
15 GARROW JS. Treat obesity seriously, a clinical manual. London, Churchill Livingstone 1981.
16 LEMMENS PHHM. Measurement and distribution of alcohol consumption. Maastricht, Datawyse 1991.
17 SPSS Inc. SPSS-x user's guide. New York. McGraw Hill 1989.
18 PHILIPSEN H. KNIBBE RA. VAN REEK J. Alcohol consumption in the Netherlands as a social phenomenon. In: TNO. Alcohol, Health and Society. Zeist 1983.
19 MECHANIC D. Medical sociology: some tensions among theory, method and substance. J Health Soc Behaviour 1989;30:147-160.

received	13-01-1993
revision asked by the editor	16-03-1993
accepted	29-06-1993

VALIDITY OF DIAGNOSES OF CHRONIC DISEASES IN GENERAL PRACTICE

THE APPLICATION OF DIAGNOSTIC CRITERIA

F. G. SCHELLEVIS,[1,2]* E. VAN DE LISDONK,[2] J. VAN DER VELDEN,[3] J. TH. M. VAN EIJK[1] and C. VAN WEEL[2]

[1]Department of General Practice and Nursing Home Medicine, Vrije Universiteit, Amsterdam, [2]Department of General Practice, University of Nijegen and [3]Netherlands Institute of Primary Health Care (NIVEL), Utrecht, The Netherlands

(Received in revised form 30 November 1992)

Abstract—Certainty of a diagnosis is not only important for the patient but also for morbidity studies. In the absence of a gold standard, agreement with diagnostic criteria is often the best approach in measuring the certainty of a diagnosis. The agreement with diagnostic criteria has been studied for 5 chronic diseases (hypertension, chronic ischemic heart disease, diabetes mellitus, chronic nonspecific lung disease and osteoarthritis) in 7 general practices with a total practice population of 23,534 persons. Agreement with diagnostic criteria is operationalized into 3 categories. For each chronic disease a diagnostic quality measure per general practitioner is computed. Retrospective data have been collected in the practices on 2295 diseases in 1989 patients. Two-thirds of the diagnoses were made in general practice. The agreement with the diagnostic criteria for the cases diagnosed in general practice is high, ranging from 96% true positive cases in diabetes mellitus to 58% in chronic nonspecific lung disease. The highest rate of false positive cases is 4%. On the level of general practitioners diagnostic qualities vary from 62 to 96% true positive cases for the different diseases. The variation in diagnostic quality between general practitioners is substantial. The prevalence rates for the 5 chronic diseases are lower after adjustment by only including true positive cases. Diagnoses of the 5 chronic diseases recorded in general practice are generally valid with low numbers of false positive cases.

General practice Chronic diseases Diagnostic criteria Diagnostic procedures Validity of diagnoses Prevalence

INTRODUCTION

The management of chronic diseases is considered to be the "very stuff of general practice" [1]. The diagnostic process and long-term care are two major clinical aspects of managing chronic diseases in general practice.

An increasing number of publications deals with various aspects of the long-term care of patients with chronic diseases: the organization of care and the management of chronic diseases [2], standards in view of the quality of care [3], and compliance with therapy and follow-up controls [4]. Much less attention has been given to the process of diagnosing a chronic disease in general practice. Yet the diagnosis of a chronic disease is of great importance: it labels the patient, often for his lifetime, and often has implications for daily life. Those suffering from a chronic disease are at risk of complex or serious illness, and of potentially harmful medi-

*All correspondence should be addressed to: F. G. Schellevis, Department of General Practice and Nursing Home Medicine, v.d. Boechorststraat 7, 1081 BT Amsterdam, The Netherlands.

cal interventions. Therefore, the certainty of the diagnosis of a chronic disease is of crucial importance: false-positive and false-negative diagnoses may have unacceptable consequences for the patient.

Apart from the importance of accurate diagnoses for the patient, certainty of the diagnosis is important for epidemiological research. This is especially the case in countries like the U.K. and The Netherlands where all non-institutionalized persons are registered in a general practice, which allows a valid estimation of the epidemiological denominator. Estimations of morbidity rates often rest upon population surveys with self-report about diseases. In measuring the concurrent validity of these data a comparison is often made with morbidity data from medical sources on the assumption that the latter data have a higher validity [5–7].

The aim of this study was to establish the validity of the diagnoses of chronic diseases on medical records in general practice.

Unfortunately, there is no absolute certainty or "gold standard" for diagnoses, except for some areas where diagnostic criteria are linked to underlying pathologic confirmation. Therefore, agreement with a set of diagnostic criteria is usually the best approach. Diagnostic criteria, often determined by international expert fora, reflect the actual common consensus on the nature and outcome of diagnostic procedures determining the diagnosis. The **International Classification of Health Problems in Primary Care** [8] lists diagnostic criteria to be used in general practice.

The application of such criteria in diagnosing chronic diseases in general practice has not been widely studied. A pilot study performed in Maastricht (The Netherlands) to evaluate the feasibility of application of criteria in daily work in general practice provided encouraging results [9].

We have compared information on applied diagnostic procedures by general practitioners in diagnosing patients with hypertension, chronic ischemic heart disease, diabetes mellitus, chronic nonspecific lung disease, and osteoarthritis of hip and/or knee with the diagnostic criteria of the ICHPPC-2. We also describe the effects of only including true positive cases on the prevalence rates of these chronic diseases.

METHODS

Practices and population

From 103 practices that took part in the "Dutch National survey of morbidity and interventions in general practice" [10] (a non-proportional stratified random sample of all general practitioners in The Netherlands) 8 practices were selected and invited to participate in a follow-up project on chronic diseases. The selection was based on the period of participation in the national survey (third and fourth trimester of the 1-year period of data collection) and the location of the practice (south-east part of the country). The reason for selecting this part of the country was that from this region the Department of General Practice of Nijmegen recruits practices for educational and research purposes. One practice (3 general practitioners (GP)) refused participation for reasons of the expected high workload. Two of the 7 participating practices are single-handed, 4 have 2 GPs and 1 is a group practice with 5 GPs. In 1988, 56% of the Dutch GPs worked in single handed practices, 30% in duo-practices and 14% in group practices or health centers. Three of the participating practices are involved in vocational general practice training, the others have no special relationship with a university department of general practice.

In each practice an age/sex register was compiled with the help of trained students. On 1 January 1988 the 7 practices covered 23,534 people. Table 1 lists some population characteristics for these practices compared with the total population of The Netherlands.

Information on the distribution in the population of the practices of risk factors for chronic diseases, like body weight, blood pressure, cholesterol, nutrition and smoking habits, is not available.

Identification of cases

In each practice the GPs were asked to identify from the practice list all patients with any of the following diseases:

—hypertension
—chronic ischemic heart disease (angina pectoris, previous myocardial infarction, coronary sclerosis) (CIHD)
—diabetes mellitus
—chronic nonspecific lung disease (asthma, chronic bronchitis, emphysema) (CNSLD)
—osteoarthritis of hip and/or knee.

Criteria for identification were: the patient was

currently known by the GP or by notation on the patient's record as having any of the chronic diseases mentioned and the diagnosis was made before 1 January 1988. Identification of patients was a two-step process.

Firstly, identification was made on the occasion of an encounter, a repeat prescription or other administrative reason for one of the chronic diseases during the first 3 months of the study, resulting in 1073 patients. In the second step the GP reviewed systematically the patient records of all patients on the practice list who had not had contact with the practice during the previous 3 months. Another 916 patients were identified in this way, resulting in a total number of 1989 patients with 2295 diseases. These figures imply the presence of more than one of the chronic diseases in 1 out of 6 patients. At the time of the identification of patients the GPs were not aware of the aim of the study, nor had they discussed the diagnostic criteria.

Information on diagnostic procedures

In a questionnaire for each patient and each disease the GPs were asked to supply: the diagnostic procedures used, the date of diagnosis, and the diagnosing physician (GP or specialist). They were asked to consult all available sources of information such as patients' records and archives. In answering the questions the GPs were instructed that any "Yes" had to be based on written information on patient history, physical examination or diagnostic tests. Any available information was considered relevant for this study as the diseases involved are chronic and lasting. In collecting these retrospective data we set no limitation in time.

The questions regarding the diagnostic procedures were derived from the ICHPPC-2 diagnostic criteria as is shown in Table 2.

Validity measure

A measure of validity for each chronic disease was determined by comparing the reported pro-

Table 1. Characteristics of the populations of the 7 study practices ($N = 23,534$) compared to the population of The Netherlands ($N = 14,714,948$) on 1 January 1988; percentages

	Study practices	The Netherlands
Age		
0–4	5.4	6.1
5–14	12.9	12.4
15–24	17.3	16.8
25–44	35.8	31.9
45–64	19.3	20.4
⩾65	9.3	12.5
Sex		
Male	48.8	49.4
Female	51.2	50.6
Health care insurance (income-related)		
Health Care Fund members	68.6	61.0
Privately insured or not insured persons	31.4	39.0
Country of birth		
The Netherlands	97.1	96.0
Turkey/Morocco	0.3	2.0
Other Western countries	0.9	1.2
Other non-Western countries	1.7	0.8
Highest educational level (only persons ⩾ 18: 72.4%)		
No education/primary school	26.2	20.9
Secondary school	61.9	62.8
University	11.9	16.3
Socioeconomic class (*profession*) (missing data: 25.6%)		
Brain work: high/middle	23.8	26.1
Brain work: low class	20.9	24.4
Farmers/independent business	8.3	5.7
Hand work: high/middle	23.9	21.5
Hand work: low class	23.2	22.3

Table 2. ICHPPC-2 Diagnostic criteria and corresponding questions per disease

ICHPPC-2 diagnostic criteria	Questions
Hypertension	
• Blood pressure at two readings > 160/95 mmHg	• Date encounters? • Blood pressure (mmHg)?
Chronic ischemic heart disease	
• Old myocardial infarction history, ECG or X-ray evidence • Chest pain compatible with angina pectoris • ECG evidence of myocardial ischemia or ventricular aneurysm • X-ray evidence of narrowed coronary arteries or ventricular aneurysm	• Diagnosis based on —history (Y/N)? —ECG (Y/N)?
Diabetes mellitus	
• Blood glucose level: —fasting ⩾ 8.0 mmol/l —not fasting ⩾ 11.0 mmol/l • Classic symptoms	• Blood glucose level (mmol/l)? • Condition at the time of taking the sample (fasting, 2 hr after meal, glucose tolerance test, arbitrary occasion)?
CNSLD/Asthma	
• Variable obstruction at pulmonary function test • Wheeze, dry cough, prolonged expiratory phase	• Diagnosis based on —history (Y/N)? —physical examination (Y/N)? —pulmonary function test (Y/N)? —X-ray (Y/N)?
CNSLD/Chronic bronchitis	
• History of cough with purulent sputum • Scattered rales or ronchi on auscultation	• See Asthma
CNSLD/Emphysema	
• X-ray evidence • Obstruction at pulmonary function test • Dyspnea • Shape of chest with reduced breath sounds	• See Asthma
Osteoarthritis hip/knee	
• X-ray evidence • Joint disorder with —irregular swelling —crepitation —stiffness/limited movement —normal laboratory tests —age	• Diagnosis based on —history (Y/N)? —physical examination (Y/N)? —laboratory test (Y/N)? —X-ray (Y/N)?

cedures with the diagnostic criteria. Agreement with the diagnostic criteria is categorized (Table 3):

(a) full agreement (true positive)
(b) partial agreement
(c) no agreement (false positive).

This measure of validity was used for analysis on case level. If no data on the diagnostic criteria were available, the case was designated as "missing". The ICHPPC-2 diagnostic criteria for diabetes mellitus and hypertension list cut-off points of numeric values. For the other diseases the criteria are of a descriptive nature, allowing only a qualitative judgment of the diagnostic validity.

Table 3. Categories of agreement with the diagnostic criteria

Full agreement	Partial agreement
Hypertension	
• DBP ≥ 95 mmHg at two encounters	• DBP > 95 mmHg at one encounter
Chronic ischemic heart disease	
• Diagnosis based on ECG	• Diagnosis based only on history
Diabetes mellitus	
• Blood glucose level: —fasting: ≥8.0 mmol/l or —not fasting: ≥11.0 mmol/l	Not categorized
CNSLD	
• Diagnosis based on pulmonary function test or X-ray	• Diagnosis based only on history and physical examination
Osteoarthritis hip/knee	
• Diagnosis based on laboratory test or X-ray	• Diagnosis based only on history and physical examination

GP's diagnostic quality

Data on diagnostic procedures performed in diagnosing chronic diseases in patients of any one GP have to be considered interdependently: the diagnostic process is likely to be a physician characteristic rather than a patient characteristic. Consequently, the validity measures for the different chronic diseases are aggregated on the GP level [11]. A diagnostic quality measure was computed by dividing the number of true positive cases by the total number of identified cases per GP, expressed in percentages. The influence of characteristics of GP (gender) and practice (practice type, distance between practice office and nearest hospital, and urbanization level of the community in which the practice is located) on this diagnostic quality measure is analyzed by means of subgroup analysis, comparing means in subgroups by univariate analysis and by calculating the statistical significance with 2-tailed probability at the 5% level.

Prevalence

The prevalences reflect point-prevalences: the number of disease cases per 1000 in the population on 1 January 1988. As the information on morbidity was derived from general practice records, the time period for measuring prevalence for each patient was potentially lifelong. The mean duration of the diseases on 1 January 1988 varied from 4 to 9 years.

The prevalences of the chronic diseases are based on the diseases of the identified patients ("unadjusted" prevalence). Adjustment was carried out by only taking into account true positive cases.

RESULTS

The cooperation of the GPs was satisfying, considering the intensive search for data in archives and patient files that had to be made. No data about the applied criteria could be traced in 17% of the diagnoses. In 5% of the diagnoses it is unknown whether it was made by a GP or by a specialist.

Validity of diagnoses

Table 4 shows the agreement of the applied procedures with the ICHPPC-2 diagnostic cri-

Table 4. Agreement of the diagnostic procedures performed with the diagnostic criteria (in percentages) of the cases for each disease—for the cases diagnosed by the GP

	No. of cases (abs)	Agreement Full (%)	Part (%)	No (%)	Missing data (abs)
Hypertension	(719)	85.1	12.6	2.3	(75)
Chronic ischemic heart disease	(194)	67.9	29.9	2.2	(10)
Diabetes mellitus	(172)	96.1	—	3.9	(44)
Chronic non-specific lung disease	(292)	57.6	42.4	0.0	(35)
Osteoarthritis hip/knee	(68)	80.6	19.4	0.0	(6)

Table 5. Mean diagnostic quality (percentage of true positive cases) of the participating GPs ($n = 15$) per chronic disease. Percentages and 95% confidence intervals

	No. of patients (abs)	True positive (%)	95% CI
Hypertension	(644)	85.5	81.0–90.1
Chronic ischemic heart disease	(184)	76.3	58.4–94.3
Diabetes mellitus	(128)	97.6	95.6–99.6
Chronic non-specific lung disease	(257)	62.5	44.3–80.7
Osteoarthritis of hip/knee	(62)	84.0	71.1–96.8

teria for all cases diagnosed by GPs (63% of all diagnoses). The agreement is high in cases of diabetes mellitus (96% true positive cases), hypertension (85%) and osteoarthritis (81% true positive). Partial agreement is a substantial category in CIHD and CNSLD. The category "no-agreement with the diagnostic criteria" (the false positive cases) is highest in diabetes mellitus (4%).

In patients whose diagnosis was made by a specialist (results no shown) data is often missing, especially in hypertension and diabetes mellitus. The specialists' diagnoses in patients with CIHD and CNSLD are more in agreement with the ICHPPC-2 criteria than cases diagnosed by GPs (85 vs 68% and 81 vs 58% respectively). In diabetes mellitus diagnosed by specialists the agreement is lower (82 vs 96% in GPs' diagnoses).

Diagnostic quality

In Table 5 the diagnostic quality measures of the participating GPs are summarized. The GPs diagnosed hypertension, diabetes mellitus, ischemic heart disease and osteoarthritis in more than 70% of the cases in full agreement with the criteria; in chronic lung disease the diagnostic quality is lower. The relatively small confidence intervals in diabetes mellitus and hypertension indicate little variation between the GPs.

Further analysis of the diagnostic quality measure in subgroups defined by GP and practice characteristics shows significant differences ($p < 0.05$) only on the variables 'single-handed and duo practice' vs 'group practice' (GPs from the group practice scored lower in CIHD and CNSLD) and urbanization level of the community served (GPs in suburbanized communities scored lower for CIHD and CNSLD). There was no correlation between the diagnostic quality ratios for the different diseases for the individual GP (data not shown).

Prevalence

Table 6 summarizes the prevalence of the chronic diseases, regardless of whether the diagnosis was made in general practice or by a specialist, in relation to agreement with the diagnostic criteria. The prevalence rates are computed for two age-groups: <65 and $\geqslant 65$ years old. Adjustment by only including true positive cases lowers the prevalence of all chronic diseases.

DISCUSSION

Validity

Diagnoses of chronic diseases made by GPs agree very well with the diagnostic criteria of the ICHPPC-2-Defined. The highest rate of false positive cases is 4% (diabetes mellitus).

In the case of diabetes mellitus, it should be taken into account that the widely used diagnostic criteria changed in 1980 to more stringent ones [12]. In this study 4 out of the 14 false positive cases of diabetes mellitus were diagnosed before 1980.

The number of cases with CIHD and CNSLD diagnosed on the basis of history and physical examination ("partial agreement") is relatively high. This reflects the common diagnostic procedures in general practice in The Netherlands where electrocardiographic and spirometric examinations are not available in general practice. Apparently GPs consider the patient's history, and signs and symptoms at physical examination, to be a sufficient basis for diagnosing these cardiac and respiratory diseases.

For many patients osteoarthritis is a silent disease: only patients with complaints or symptoms consult a GP, and the GP will only report on these cases. Consequently, in this study probably only osteoarthritis patients with complaints were included as being "at risk" for undergoing X-ray examination. This is reflected in the high percentage of true positive cases of osteoarthritis.

The number of missing data is acceptable, considering the retrospective nature of the collected data. Moreover, for the diagnoses made by medical specialists, it should be kept in mind

that patients in The Netherlands will only come under the care of a specialist upon referral by a GP. The medical specialist usually reports his diagnosis to the GP, but does not always include the criteria upon which the diagnosis is based.

Routine general practice care does not necessarily imply detailed documentation of performed diagnostic procedures. As this study is based on recorded evidence the level of agreement with diagnostic criteria is an underestimation of validity. The "partial agreement" category, for example, probably includes a number of true positive cases. This could be verified by performing additional diagnostic procedures. Moreover, we have no information on false negative cases.

We suggest that agreement with international standard diagnostic criteria for general practice is the best way to assess the validity of the diagnosis. The diagnostic criteria of the ICH-PPC-2-Defined seem to be useful in assessing the quality of diagnoses, although difficult to use in qualitative retrospective data. Operationalization of these criteria for use in research needs further elaboration.

Diagnostic quality

The diagnostic quality per chronic disease varies substantially between the GPs. When the results on validity are aggregated from the patient level to the level of the GP the mean percentages do not change very much. The confidence intervals, however, are larger because of the sample size (15 GPs). This is especially the case for CIHD and CNSLD, which is possibly explained by differences between GPs in the use of clinical diagnostic facilities (electrocardiography and spirometry).

We find no support in our data for the assumption that there is a general diagnostic ability of GPs reflected in a correlation between the quality ratios for different diseases, but the number of GPs in our study is too small for definitive conclusions.

Prevalence

Substantial differences in prevalences are found by adjustment for true positive cases only, mainly due to the number of missing cases left out. The age-specific adjusted prevalence rates of the chronic diseases are lower in this

Table 6. Point-prevalence and 95% confidence intervals unadjusted and adjusted (true positive cases only) in people < 65 years (N = 21,349) and ≥65 years (N = 2185) for 5 chronic diseases in the study practices per 1000 patients

	<65 years			
	Unadj.	95% CI	Adj.	95% CI
Hypertension	24.7	22.6–26.8	18.5	16.6–20.3
Chronic ischemic heart disease	10.9	9.5–12.3	7.5	6.4–8.7
Diabetes mellitus	7.3	6.1–8.4	4.3	3.4–5.1
Chronic non-specific lung disease	19.9	18.0–21.0	12.2	10.8–13.7
Osteoarthritis hip/knee	2.2	1.5–2.8	1.7	1.1–2.2
	≥65 years			
	Unadj.	95% CI	Adj.	95% CI
Hypertension	143.2	129.0–158.0	89.7	78.1–102.0
Chronic ischemic heart disease	100.2	87.6–113.0	75.1	64.4–86.9
Diabetes mellitus	60.9	51.2–71.7	33.0	25.8–41.3
Chronic non-specific lung disease	65.4	55.5–76.6	43.5	35.3–52.9
Osteoarthritis hip/knee	38.0	30.3–46.8	29.3	22.6–37.3

study than those in other morbidity surveys in general practice in The Netherlands [13, 14]. On the patient level our sample can be considered representative for the entire population of The Netherlands. On the level of practices the sample is too small to allow generalization.

For chronic lung disease the differences in prevalence in morbidity surveys reflect the differences in definition between The Netherlands and the U.K. [15, 16].

Osteoarthritis of hip and/or knee is probably the most underreported chronic disease in this study, because only patients with complaints are seen. Comparison with other morbidity data is difficult as generally all cases of osteoarthritis are reported without specifying the affected joints.

These discrepancies can be explained by differences between GPs in registration discipline [17], differences in case-finding, differences in applying diagnostic criteria or by real morbidity differences between populations. We have no information about applied diagnostic criteria in these large surveys. We have no reason to suppose that the population of our study practices differs that much from other populations, that the studied diseases really occur less frequently.

CONCLUSIONS

Diagnoses of the 5 chronic diseases recorded in general practice are generally valid with low numbers of false positive cases. The diagnosing physician as a source of variability in the validity of diagnoses should not be ignored. The validity of morbidity data originating from population surveys can well be measured by comparison with GPs' records.

REFERENCES

1. Hasler JC. The very stuff of general practice. **J R Coll Gen Pract** 1985; 35:121–127.
2. Hasler J, Schofield T, Eds. **Continuing Care: the Management of Chronic Disease**. Oxford: Oxford University Press; 1990: 2nd edn.
3. Grol R, Mesker P, Schellevis F, Eds. **Peer Review in General Practice**. Lelystad: Meditekst; 1987.
4. Campen C v, Sluijs EM. Patient compliance. A survey of reviews (1979–1989) [Bibliography]. Utrecht: Netherlands Institute of Primary Health Care; 1989.
5. Kars-Marshall C, Spronk-Boon YW, Pollemans MC. National health interview surveys for health care policy. **Soc Sci Med** 1988; 26: 223–234.
6. Tretli S, Lund-Larsen PG, Foss OP. Reliability of questionnaire information on cardiovascular disease and diabetes: cardiovascular disease study in Finnmark county. **J Epidemiol Commun Health** 1982; 36: 269–273.
7. Midthjell; K. Holmen J, Bjørndal A, Lund-Larsen PG. Is questionnaire information valid in the study of a chronic disease such as diabetes? The Nord-Trøndelag diabetes study. **J Epidemiol Commun Health** 1992; 46: 537–542.
8. Classification Committee of WONCA. **International Classification of Health Problems in Primary Care (ICHPPC)-2-Defined**. Oxford: Oxford University Press; 1983: 3rd edn.
9. Huijnen LGJ, Muyrers PEM, Knottnerus JA. Registreren van diagnosen met behulp van diagnostische criteria. **Huisarts Wet** 1987; 30: 237–239 (English Summary).
10. Foets M, Velden J vd, Bakker D de. **Dutch National Survey of General Practice. Survey Design**. Utrecht: Netherlands Institute of Primary Health Care; 1992.
11. Whiting-O'Keefe QE, Henke C, Simborg DW. Choosing the correct unit of analysis in medical care experiments. **Med Care** 1984; 22: 1101–1114.
12. WHO Expert Committee on Diabetes Mellitus. Technical Report Series No. 646. Geneva: WHO; 1980.
13. Hoogen HJM vd, Huygen FJA, Schellekens JWG, Straat JM, Velden HGM vd. **Morbidity Figures from General Practice. Data from Four General Practices**. Nijmegen: Nijmegen University Department of General Practice; 1985.
14. Lamberts H. **Morbidity in General Practice. Diagnosis Related Information from the Monitoring Project**. Utrecht: Huisartsenpers; 1984.
15. Sluiter HJ, Koëter GH, Monchy JGR de, Postma DS, Vries K de, Orie NGM. The Dutch hypothesis (chronic non-specific lung disease) revisited. **Eur Respir J** 1991; 4: 479–489.
16. Vermeire PA, Pride NB. A "splitting" look at chronic nonspecific lung disease (CNSLD): common features but diverse pathogenesis. **Eur Respir J** 1991; 4: 490–496.
17. Mant D, Tulloch A. Completeness of chronic disease registration in general practice. **Br Med J** 1987; 294: 223–224.

COMORBIDITY OF CHRONIC DISEASES IN GENERAL PRACTICE

F. G. SCHELLEVIS,[1,2,*] J. VAN DER VELDEN,[3] E. VAN DE LISDONK,[2] J. TH. M. VAN EIJK[1] and C. VAN WEEL[2]

[1]Department of General Practice and Nursing Home Medicine, Vrije Universiteit, Amsterdam, [2]Department of General Practice, University of Nijmegen and [3]Netherlands Institute of Primary Health Care (NIVEL), Utrecht, The Netherlands

(Received in revised form 30 November 1992)

Abstract—With the increasing number of elderly people in The Netherlands the prevalence of chronic diseases will rise in the next decades. It is recognized in general practice that many older patients suffer from more than one chronic disease (comorbidity). The aim of this study is to describe the extent of comorbidity for the following diseases: hypertension, chronic ischemic heart disease, diabetes mellitus, chronic nonspecific lung disease, osteoarthritis. In a general practice population of 23,534 persons, 1989 patients have been identified with one or more chronic diseases. Only diseases in agreement with diagnostic criteria were included. In persons of 65 and older 23% suffer from one or more of the chronic diseases under study. Within this group 15% suffer from more than one of the chronic diseases. Osteoarthritis and diabetes mellitus are the diseases with the highest rate of comorbidity. Comorbidity restricts the external validity of results from single-disease intervention studies and complicates the organization of care.

Comorbidity Chronic diseases General practice Prevalence

INTRODUCTION

The morbidity pattern in general practice is well-documented, particularly in The Netherlands and in the U.K., where the fixed practice population allows for population-based description [1–4].

General practice covers its own clinical spectrum as has been demonstrated previously [5]. Chronic diseases are an important feature of this clinical spectrum: hypertension, chronic ischemic heart disease, diabetes mellitus, chronic nonspecific lung disease, and osteoarthritis all have a prevalence above 10 per 1000 and are mostly managed in general practice [3, 4, 6, 7]. In the near future the number of elderly people will increase in The Netherlands. As a consequence, the prevalence of chronic diseases will rise.

These changes in morbidity pattern will influence the daily work in general practice. Standards and guidelines for proper diagnosis, treatment and management of chronic diseases are crucial for maintaining the quality of care. Intervention studies provide an essential basis for adequate treatment and prevention. Most of such studies analyze the effects of intervention on a single disease. General practitioners (GPs), however, recognize that their patients often suffer from more than one chronic disease. As a generalist the GP, alone or in cooperation with the specialist, deals with all diseases of a patient. As the natural course and the therapeutic interventions of one disease will influence the co-existing second (or even third) disease [8], comorbidity diminishes the practical value of

*All correspondence should be addressed to: F.G. Schellevis, Department of General Practice and Nursing Home Medicine, v.d. Boechorststraat 7, 1081 BT Amsterdam, The Netherlands.

single-disease standards for treatment and management, derived from single-disease trials. We found no publications on the frequency of comorbidity in general practice populations. The aim of this study is to describe the extent of comorbidity of chronic diseases in general practice in The Netherlands: how many of the patients are under care for more than one of the following, most common, chronic diseases: hypertension, chronic ischemic heart disease, diabetes mellitus, chronic nonspecific lung disease, and osteoarthritis?

The findings indicate the prevalence of these problems and thus contribute to our insight in disease clustering [9].

METHODS

This study is part of a larger research on the prevalence of chronic diseases in general practice, and of the effects of systematic surveillance on the quality of care.

General practices and population

Seven general practices (15 GPs) participated in the study. The practices covered 23,534 persons at the start of the data collection (1 January 1988). An age/sex register of the practices was completed with the help of trained students. The number of persons of 65 years and over is 3% less than in the entire country (9.3 vs 12.5%—Table 1). For that reason all results are presented for two subpopulations: that of persons below 65 and that of persons of 65 and over. In other characteristics the practice population differs only marginally from the entire population [10].

Table 1. Age and sex distribution of the population of the general practices studied ($N = 23,534$) compared to the population of The Netherlands ($N = 14,714,948$) 1 January 1988

	Practice population (%)	The Netherlands (%)
0–4	5.4	6.1
5–14	12.9	12.4
15–24	17.3	16.8
25–44	35.8	31.9
45–64	19.3	20.4
⩾ 65	9.3	12.5
Male	48.8	49.4
Female	51.2	50.6

Case identification

The participating GPs identified all patients known to them with the following diseases:
—hypertension
—diabetes mellitus
—chronic ischemic heart disease (CIHD) (angina pectoris, previous myocardial infarction, coronary sclerosis)
—chronic nonspecific lung disease (asthma, chronic bronchitis, emphysema (CNSLD)
—osteoarthritis of hip and/or knee.
Identification took place on the occasion of a consultation, a repeat prescription or another administrative reason for visiting the practice during the first 3 months. Finally, the GP reviewed systematically all patient records to identify diseases of patients who were not seen.

This process of identification resulted in a total number of 1989 patients with 2295 diseases (cases).

Application of diagnostic criteria

The GP provided retrospective data of the medical history from the patients' records in relation to the diagnostic procedures applied in diagnosing the chronic disease, regardless of whether the diagnosis was made in general practice or by a medical specialist. These data were compared with the inclusion criteria of the **International Classification of Health Problems in Primary Care** [10, 11]. Only the cases (diseases) meeting these inclusion criteria were used for analysis.

Analysis

Comorbidity of chronic diseases is defined as the "point-prevalent concurrence" of the studied diseases known to the participating GPs. Point-prevalence reflects the number of diseases in the population at 1 January 1988. Comorbidity was analyzed on patient level by means of the multiple response technique in SPSS [12]. Comorbidity is expressed as the number of the studied diseases per patient, the mean number of diseases per patient in each disease category, and the proportion of patients with at least one of the other diseases. Due to the cross-sectional measurement and the method of presentation, patients with comorbidity appear in each of the disease categories that apply to them and are therefore counted more than once (e.g. a patient with diabetes and with hypertension appears in the hypertension group as well as in the diabetes

Table 2. Point-prevalence in persons <65 (N = 21,349) and ≥65 (N = 2185) of 5 chronic diseases per 1000 persons in the general practices studied

	<65		≥65	
	Prevalence	95% CI	Prevalence	95% CI
Hypertension	18.5	16.6–20.3	89.7	78.1–102.0
CIHD	7.5	6.4–8.7	75.1	64.4–86.9
Diabetes mellitus	4.3	3.4–5.1	33.0	25.8–41.3
CNSLD	12.2	10.8–13.7	43.5	35.3–52.9
Osteoarthritis hip/knee	1.7	1.1–2.2	29.3	22.6–37.3

category). Proportions and means are presented with the 95% confidence intervals (CI).

RESULTS

Our definition of comorbidity is strongly connected with the prevalences of the studied diseases. Table 2 lists the point-prevalences on 1 January 1988 of the diseases meeting the inclusion criteria. The prevalence of most diseases is high in persons over 65. Hypertension is the most frequent of the studied diseases. Table 3 shows the distribution of the number of the diseases per patient. In the younger subgroup there are few persons known to have one of the studied diseases, and comorbidity occurs in only 0.3% of these persons. Of the persons over 65 years old, more than 75% are known not to have one of the 5 chronic diseases, but of the older patients who do have one of these diseases, 16% has more than one chronic disease. Tables 4(A) and (B) show the extent of comorbidity in patients with at least one disease for the two age groups. In patients under 65 years old patients with osteoarthritis have the highest rate of comorbidity. The most frequent second disease in these patients is CNSLD (5 of 36 patients). In diabetics under 65 (N = 91) hypertension is the most frequent second disease (15%). In patients of 65 years and older the highest frequency of comorbidity is found in patients with diabetes mellitus. The most frequent second chronic disease in diabetics over 65 is CIHD (22%), followed by hypertension (19%). High rates of comorbidity are also found in patients with osteoarthritis (mostly hypertension and CNSLD) and CIHD (mostly hypertension).

Table 3. Number of studied chronic diseases per person in the population of the general practices studied

Number of chronic diseases	<65 (N = 21,349) (%)	≥65 (N = 2185) (%)
None	95.9	76.9
One	3.8	19.5
Two	0.3	3.2
Three	<0.1	0.3
Four	—	<0.1

DISCUSSION

Prevalence of chronic diseases

This study is based on data obtained from medical records. Generally this leads to an underestimation of the number of cases in the population. Moreover, only cases in agreement with diagnostic criteria were included. the prevalence of hypertension, chronic ischemic heart disease, diabetes mellitus, and chronic nonspecific lung disease is lower than in other Dutch reports from general practice [3, 4]. Compared with data from the U.K., the prevalence of diabetes mellitus is higher, as has been reported by others [13].

Chronic nonspecific lung disease has been identified as an area of diagnostic confusion. Differences in opinion between physicians in the U.K. and The Netherlands exist as to whether asthma, chronic bronchitis and emphysema have the same pathophysiological characteristics, the so-called Dutch hypothesis, which is heavily disputed [14, 15]. As Dutch GPs are familiar with the diagnostic label of chronic nonspecific lung disease, this term is used in our description of comorbidity.

Extent of comorbidity

Comorbidity is a quantitatively important phenomenon in patients over 65 with a chronic disease. Most people over 65 (77%) do not suffer from any of the 5 most common chronic diseases, but within the affected group 16% is known to suffer from at least one other of the 5 chronic diseases studied.

The occurrence by chance of two diseases in one person can be estimated by multiplying the

Table 4 (A). Extent of comorbidity per disease for patients <65, presented as number of chronic diseases per patient group and as fraction of patients with comorbidity per patient group

		Number of chronic diseases per patient		Fraction of patients with comorbidity
	N	Mean	95% CI	% of *N*
Hypertension	394	1.1	1.07–1.13	9.6
CIHD	161	1.2	1.14–1.28	19.9
Diabetes mellitus	91	1.2	1.13–1.33	20.9
CNSLD	250	1.1	1.04–1.11	7.6
Osteoarthritis hip/knee	36	1.3	1.12–1.43	27.8

CIHD = chronic ischemic heart disease; CNSLD = chronic nonspecific lung disease.

Table 4 (B). Extent of comorbidity per disease for patients ≥65, presented as number of chronic diseases per patient group and as fraction of patients with comorbidity per patient group

		Number of chronic diseases per patient		Fraction of patients with comorbidity
	N	Mean	95% CI	% of *N*
Hypertension	196	1.3	1.18–1.32	21.9
CIHD	164	1.3	1.24–1.42	28.0
Diabetes mellitus	72	1.5	1.32–1.62	40.3
CNSLD	93	1.3	1.18–1.42	24.7
Osteoarthritis hip/knee	64	1.4	1.22–1.50	32.8

prevalences of the separate diseases. The observed comorbidity of the 5 diseases under study is significantly higher. Having a chronic disease apparently means being at higher risk to have a second or even third disease.

By including only diagnoses meeting diagnostic criteria and by disregarding false negative diagnoses, we probably underestimate also the extent of comorbidity. Moreover, these figures are related only to the 5 chronic conditions under study. The rate of comorbidity would have been even larger, if additional diseases, like malignant neoplasms, epilepsy and other neurological diseases, stroke, peripheral vascular disease, peptic ulcer disease, had been considered. On the other hand, by estimating the extent of comorbidity in a general practice setting bias due to the Berkson's fallacy cannot be excluded: patients under care for a chronic disease are at higher risk for detection of diseases than persons who do not receive such care.

Comorbidity is partly the result of a common pathophysiological process or of complications in the natural course of a disease, as is the case for diabetes mellitus and cardiovascular disease [16]. In other cases, comorbidity of chronic diseases is accidental and cannot be explained pathophysiologically.

CONSEQUENCES

Persons suffer from more than one chronic disease more frequently than could be expected by chance from the prevalence of the disease in the general population. This is a clinical reality of medical practice with consequences for research and for the organization of daily care.

Research

Optimal patient care should ideally be based on valid results from clinical trials. In intervention studies, however, patients with comorbidity often are excluded in the selection of a study group, e.g. in the well-known therapeutic trials in hypertension [17, 18]. This selection restricts the external validity of the results for excluded patient categories, as has recently been described for the elderly and women [19, 20]. The existence of a second disease complicates the choice of the antihypertensive treatment that is proven to be effective in single disease patients (e.g. diuretics in diabetes, bèta-blockers in lung disease). Strictly speaking, these studies have not proven the effectiveness of antihypertensive treatment in lowering blood pressure and decreasing cardiovascular morbidity and mortality for patients with comorbidity. In intervention studies on treatment of chronic diseases patients

with comorbidity should be included. In analyzing the data, patients with comorbidity can be handled as a subgroup, or adjustment of the results for comorbidity can be carried out.

The specific combinations of chronic diseases need further exploration, in order to gain more insight into patterns of disease clustering and hypothetically common etiology.

Care

Systematic surveillance of patients with chronic diseases is essential in order to provide them with optimal care [7]. Patients with more than one chronic disease are at risk of being included in more than one surveillance scheme. This should be recognized when designing surveillance programs, since it would be counter-productive to have patients visit the practice on various different occasions, as a result of following different schemes for each of their diseases. Careful registration of all diseases is conditional not only for organizational reasons but also for the care to be provided.

Chronic diseases are regarded as "the very stuff of general practice" [6]. Proper management of patients with comorbidity of chronic diseases presents a real challenge to the GP.

REFERENCES

1. Fry J. **Common Diseases.** Lancaster: MTP Press Limited; 1985: 4th edn.
2. Royal College of General Practitioners, Office of Population Censuses and Surveys and Department of Health and Social Security. **Morbidity Statistics from General Practice. Third National Study 1981–1982.** London: HMSO; 1986.
3. Hoogen HJM vd, Huygen FJA, Schellekens JWG, Straat JM, Velden HGM vd, Eds. **Morbidity Figures from General Practice. Data from Four General Practices 1978–1982.** Nijmegen: Nijmegen University Department of General Practice; 1985.
4. Lamberts H. **Morbidity in General Practice.** Utrecht: Huisartsenpers; 1984.
5. Fry J. Perspectives on health care. **Update** 1977; 15: 999.
6. Hasler J. The very stuff of general practice. MacKenzie lecture. **J R Coll Gen Pract** 1985; 35: 121–127.
7. Hasler J, Schofield Th. **Continuing Care. The Management of Chronic Diseases.** Oxford: Oxford University Press; 1990: 2nd edn.
8. Feinstein AR. The pre-therapeutic classification of co-morbidity in chronic disease. **J Chron Dis** 1970; 23: 455–468.
9. Anonymous. Disease clustering: hide or seek? [Editorial]. **Lancet** 1990; 336: 717–718.
10. Schellevis FG, Lisdonk EH vd, Velden J vd, Eijk JThM v, Weel C v. Validity of diagnoses of chronic diseases in general practice. The application of diagnostic criteria. **J Clin Epidemiol** 1993; 46: 461–468.
11. Classification Committee of WONCA. **International Classification of Health Problems in Primary Care (ICHPPC)-2-Defined.** Oxford: Oxford University Press; 1983: 3rd edn.
12. Anonymous. **SPSS Version 4.1.** Chicago: SPSS Inc; 1990.
13. Reenders K, Nobel E de, Hoogen HJM vd, Rutten GEHM, Weel C v. Diabetes and its longterm complications in general practice. A survey in a well defined population. **Fam Pract** 1993; in press.
14. Sluiter HJ, Koëter GH, Monchy JGR de, Postma DS, Vries K de, Orie NGM. The Dutch hypothesis (chronic non-specific lung disease) revisited. **Eur Respir J** 1991; 4: 479–489.
15. Vermeire PA, Pride NB. A "splitting" look at chronic nonspecific lung disease (CNSLD): common features but diverse pathogenesis. **Eur Respir J** 1991; 4: 490–496.
16. Kaplan MH, Feinstein AR. The importance of classifying initial comorbidity in evaluating the outcome of diabetes mellitus. **J Chron Dis** 1974; 27: 387–404.
17. Management Committee. The Australian therapeutic trial in mild hypertension. **Lancet** 1980; i: 1261–1267.
18. Medical Research Council Working Party. MRC-trial of treatment of mild hypertension: principal results. **Br Med J** 1985; 291: 97–104.
19. Gurwitz JH, Col NF, Avorn J. The exclusion of the elderly and women from clinical trials in acute myocardial infarction. **JAMA** 1992; 268: 1417–1422.
20. Wenger NK. Exclusion of the elderly and women from coronary trials. Is their quality of care compromised? [Editorial]. **JAMA** 1992; 268: 1460–1461.

The effect of elimination of selected chronic diseases on disability-free life expectancy: compression or expansion of morbidity? Preliminary results[1]

WJ Nusselder (1,2,3), J van der Velden (1), ME Lenior (2), JLA van Sonsbeek (4), and GAM van den Bos (2)

(1) Netherlands Institute for Primary Health Care (NIVEL)

(2) Institute of Social Medicine, University of Amsterdam

(3) Department of Public Health, Erasmus University Rotterdam

(4) Netherlands Central Bureau of Statistics (CBS)

Abstract

This paper examines whether complete elimination of a specific disease—and its consequences in terms of disability and mortality—leads to absolute or relative compression of morbidity or to relative expansion of morbidity. The Sullivan method is used to combine morbidity data from the Dutch National Survey of General Practice (conducted by the Netherlands Institute for Primary Health Care (NIVEL) in 1987–88) and mortality data collected by the Dutch Central Bureau of Statistics, to obtain disability-free life expectancy (DFLE). The effect of complete elimination of a disease on DFLE is estimated in three steps. Firstly, the cause-elimination life table technique is used to calculate cause-deleted probabilities of dying. Secondly, cause-deleted disability rates are estimated by means of a logistic regression model, taking into account age and co-morbidity. Finally, the cause-deleted disability rates are substituted into the cause-elimination life table to obtain the DFLE after elimination.

The analyses show that elimination of disabling diseases that cause little mortality, like 'arthritis/back complaints' and 'migraine/severe headache', results in absolute compression of morbidity. On the other hand, elimination of chronic diseases with high case-fatality, like cancer, leads to relative expansion of morbidity. The reason for this paradox is that in addition to a gain in DFLE, a considerable increase in total life expectancy takes place, resulting in an increase in life expectancy with disability. We conclude that while complete elimination of a disease and its consequences in terms of disability and mortality results in a gain in DFLE, as a side effect of the mortality decline, life expectancy with disability may increase as well.

---

[1] A cooperative project of the Institute of Social Medicine, University of Amsterdam, the Netherlands Institute of Primary Health Care and the Netherlands Bureau of Statistics.

Introduction

The objective of the present study is to calculate the effect of elimination of a specific disease (and its consequences in terms of mortality and disability) on disability-free life expectancy (DFLE). The central question is whether this elimination results in absolute or relative compression or in relative expansion of morbidity. To answer this question, two outcome-measures are important: firstly, the gain in years without disability and total life expectancy that would result from elimination of the disease, and secondly, the proportion of life without disability: healthy life percentage (HLP).

Absolute compression is defined as a decrease in number of years with disability. Relative compression takes place if the proportion of life with disability declines (HLP increases), while the number of years with disability rises. Relative expansion of morbidity occurs if the proportion of life without disability—HLP—declines, but DFLE increases. The worst possible evolution of health status, 'absolute expansion of morbidity', is characterised by a decline in DFLE (Robine and Mathers 1993). However, the latter is not a possible outcome of elimination of a disease.

Data

The primary data source, on which this study is based, is the Dutch National Survey of General Practice, which was conducted by the NIVEL in 1987–88 (Foets, van der Velden and de Bakker 1992). In this study, a random sample of respondents was taken via a sample of 161 general practitioners (GP). The GPs were selected by a non-proportional stratified sample of the Dutch GP population. This sample was stratified in terms of region, degree of urbanisation and distance from the hospital. A sample of approximately 100 respondents per GP was selected for a Health Interview Survey.

The total sample for the Health Interview Survey included 17,047 respondents. The response percentage was 76%; the net sample consisted of 13,014 persons. The age and sex distribution in the sample deviated slightly from the Dutch population: men and middle-aged respondents (25–44 years) were overrepresented, while women and persons above 65 years of age were underrepresented. In order to correct our results both for deviations related to differential non-response and for deviations due to the stratification procedure, the sample was weighted. Since in the Netherlands practically the entire population is registered with a GP and because the sample is almost identical to the Dutch population in terms of age and gender (Foets, van der Velden and de Bakker 1992) the survey can be regarded as representative for the non-institutional population. Although in the survey no age limits were set, the analysis of this paper deals only with persons aged 16 and over (n = 10,147), because of proxy interviews. Furthermore, the reported disabilities were not assumed to be valid for young children, because no differentiation could be made between having a disability and being not yet able to perform the activities due to their young age.

The NIVEL survey includes information on age, sex, long-term disabilities and chronic diseases for the non-institutional population.[2] With respect to questions concerning long-term disabilities and chronic diseases, a self-administered questionnaire was used.

Having one or more disabilities was measured by means of the OECD indicator. The OECD indicator consists of 16 items referring to the ability to carry out a number of activities which are essential for daily independent functioning (McWhinnie 1981). In this study 11 items were selected dealing with the ability: to bend down and pick up something; to get in and out bed; to dress and undress; to move between rooms; to walk 400 m; to carry an object of 5 kg for 10 m; to read small letters in a newspaper; to recognise someone's face; to have a conversation with another person; to follow a conversation in a group; to go up and down the stairs.[3] Persons were considered to be disabled when they indicated that they needed help from another person, or were unable or had (great) difficulty in carrying out one or more of the selected activities included in the OECD indicator. Using aids, like glasses or a hearing aid, was not considered to be disabled. In the NIVEL survey the long duration of disabilities was emphasised in the introductory text of the question concerning the OECD indicator.

From the checklist of chronic diseases in the NIVEL survey, the following disease-clusters were selected: non-specific lung disease (CNSLD), heart diseases, cancer, diabetes mellitus, arthritis/back complaints, migraine/severe headache and 'other neurological diseases'. For a description of the disease-clusters, see appendix 1. The presence or absence of these diseases was used in the elimination analysis.

The number of persons with long-term institutional disability (e.g. psychiatric hospitals, nursing homes, homes for the elderly, homes for the mentally deficient) as well as their age and sex distribution, was estimated on the basis of both additional administrative data sources and surveys among elderly and people in homes for the elderly (SCP 1991; CBS 1984; CBS Statistiek Bejaardenoorden, several years; CBS, Intramurale Gezondheidszorg, several years; CBS, Maandstatistiek voor de Bevolking, several years; SIG, PIGG, several years; SIG, Jaarboek Verpleeghuizen, several years). Persons living in one of the above-mentioned institutions were considered to have long-term institutional disabilities. Only for persons living in a home for the elderly, an adjustment had to be made to take into account those without any disability. The estimation of the proportion of residents living in homes for the elderly without disabilities was based on data of the CBS survey among the population 55 years and older in 1982 (CBS 1984) and on surveys among persons living in homes for the elderly (CBS, Statistiek Bejaardenoorden, several years) as well as on the SCP AVO survey (1991).

---

[2] Persons living in a home for the elderly were included in the survey. However, they were excluded from the further analysis, because their numbers were too small to allow for distinction in age, sex and disability status.

[3] Two separate questions were asked; they were joined into one disability item for further analysis.

Data on mortality and the population distribution by sex and single year of age were derived from the Dutch Central Bureau of Statistics (CBS 1987, 1992). The number of deaths by sex, five-year age groups and underlying cause of death were also derived from the CBS (CBS, Overledenen naar doodsoorzaak, several years). The selected (underlying) causes of death are summarised in appendix 1.

Methods

Complete life tables covering the period 1982–83 to 1990–91 were calculated on the basis of population and mortality data. The latter were classified by sex and single year of age (according to a cohort-age observational plan, i.e. one age group, two calendar-years). From these complete life tables, abridged life tables (by 5-year age groups) were derived for further analysis. Details of the life table techniques can be found in Namboodiri and Suchindran (1987) and Manton and Stallard (1984). The DFLE was calculated using the Sullivan method (a prevalence rate life table technique) (Sullivan 1971a, 1971b; see also: Mathers 1992; Robine 1992). According to this method the number of person-years per age interval is subdivided into years with and years without disability by multiplying the person-years by age-specific disability rates (prevalence rates).

In the hypothetical situation in which a disease is completely eliminated, people will not have this disease any more, neither will they have any disabilities or die because of this disease. In other words, eliminating a disease leads to a decline in age-specific probabilities of dying, as well as to a decrease in age-specific prevalence rates of disability.

The potential gain in DFLE due to elimination was calculated in several steps. At first, the effect of elimination on the probabilities of dying was estimated using the 'cause-elimination life table' technique under the assumption of independent causes of death (Namboodiri and Suchindran 1987; Tsai, Lee and Harvey 1978; Manton and Stallard 1984).

Secondly, cause-deleted disability rates were estimated. However, for this purpose we could not make use of a standard method or an approach used in earlier studies. The method that was used by Mathers (1992) and Colvez and Blanchet (1983) was not appropriate, because all disabilities indicated by a person with the considered disease were completely eliminated in case the disease was eliminated. From the Dutch Health Interview Surveys (for example the NIVEL survey mentioned above) we cannot infer whether the disabilities were caused by the disease considered, by other chronic diseases (co-morbidity) or the so called 'senescence process'. Application of this method in our study would therefore result in a considerable overestimation of the disability effect. So, in order to obtain an estimation of the 'net effect' of a chronic disease on the prevalence of disabilities, controlled for co-morbidity and age, a logistic regression model was fitted. In this model the probability of having one or more

disabilities was taken as the dependent variable. Age,[4] the selected disease clusters and a category of several chronic diseases were included in the model as independent variables. A separate model was estimated for men and women. The difference between the estimated disability rates before elimination and those after elimination can be attributed to the considered disease. Among those living in an institution, changes in long-term disability rates that might occur due to elimination were not taken into account, because data on age, sex, chronic conditions and disability were not available on an individual level for persons living in institutions.

Finally, the cause-deleted probabilities of dying and the cause-deleted disability rates were combined into (cause-deleted) disability-free life expectancies using the Sullivan method (Mathers 1992). The difference between DFLE before and after elimination is the potential gain in disability-free years due to elimination.

Results

Before elimination, total life expectancy at age 15 was 59.3 years for men and 65.6 years for women, of which respectively 47.7 and 45.6 years were spent without disability (79.9% and 69.6%). The remaining 11.9 and 20.0 years were years with long-term disability.

Elimination of the (underlying) cause of death has a large impact on the probabilities of dying (and consequently on total life expectancy) for heart diseases and cancer. The potential gain in life expectancy due to elimination of heart diseases amounts to 4 years for men and 2.9 years for women (Tables 1a and 1b). For cancer, the gain is 3.8 and 3.3 years respectively. Other disease-clusters show smaller effects; for migraine it was even zero.

The change in disability rates due to elimination of a disease depends on both the net effect of the disease on the probability of having disabilities (odds ratio derived from the B-coefficient in the logistic regression analysis), and on the prevalence of the disease. The decline in disability rates due to elimination differs greatly between the disease-clusters. Elimination of arthritis/back complaints, heart diseases and to a lesser extent CNSLD (among men) has a large effect on the disability rates. On the contrary, the effect of cancer and 'other neurological diseases' is small. For the latter disease, the small effect can be explained by the low prevalence of this disease. In Figures 1a and 1b the two extremes, arthritis/back complaints and cancer, are represented. The results of the logistic regression analysis were not statistically significant for cancer and diabetes mellitus among men.

---

[4] Age was included as a continuous variable using a transformation (i.e. age to the 2.5th power for men and age to the 3rd power for women). The correctness of this transformation was tested by the Box–Tidwell transformation. 'The Box–Tidwell approach adds a term of the form xln(x) to the model. If the coefficient for this variable is significant we have evidence for non-linearity in the logit' (Hosmer and Lemeshow 1989, p.90).

Combining the effects of the decline in both the probabilities of dying and the disability rates in a single measure, the DFLE, informs us about the potential gain that can be obtained by elimination of a disease (and its consequences in terms of mortality and disability). In Tables 1a and 1b the effect of elimination is presented.

Elimination of heart diseases and arthritis/back complaints leads to the greatest gain in DFLE (at age 15). The ranking of impacts of these diseases differs between the two sexes: among men heart diseases have the largest impact (2.5 years compared with 1.9 years for arthritis/back complaints), while among women arthritis/back complaints have the largest effect (2.8 years compared with 1.4 years) (Tables 1a and 1b and Figures 2a and 2b). The impact of migraine/severe headache and diabetes mellitus is small (varying between 0.1 and 0.5 years). CNSLD and cancer take up a middle position.

Since, for some diseases the total life expectancy increases considerably too, it is important to consider the change in LEWD as well. With the exception of heart diseases and cancer (as well as diabetes mellitus among men), the LEWD decreases (Tables 1a and 1b). Therefore, elimination of these diseases results in absolute compression of morbidity. For cancer and heart diseases (and diabetes mellitus among men), the LEWD increases as well, because the increase in DFLE is accompanied by a larger gain in total life expectancy. This results in a decline in HLP: relative expansion

Table 1a: *Change in total life expectancy (LE), disability-free life expectancy (DFLE), life expectancy with disability (LEWD) and healthy life percentage (HLP) due to elimination (men and women, age 15)*

	LE (yr)	DFLE (yr)	LEWD (yr)	HLP (% points)
Men				
CNSLD	0.3	0.7	–0.4	0.7
heart diseases	4.0	2.5	1.6	–1.2
cancer[1]	3.8	1.7	2.1	–2.2
diabetes mellitus[1]	0.2	0.1	0.1	–0.1
arthritis/back complaints	0.0	1.9	–1.9	3.1
migraine/severe headache	0.0	0.4	–0.4	0.6
other neurological diseases	0.2	0.2	–0.0	0.1
Women				
CNSLD	0.1	0.6	–0.5	0.8
heart diseases	2.9	1.4	1.5	–1.0
cancer	3.3	1.1	2.2	–1.7
diabetes mellitus	0.3	0.4	–0.1	0.3
arthritis//back complaints	0.1	2.8	–2.7	4.2
migraines/severe headache	0.0	0.5	–0.5	0.7
other neurological diseases	0.2	0.2	–0.0	0.1

[1] Effect was not significant in the regression analysis.

Table 1b: *Change in total life expectancy (LE), disability-free life expectancy (DFLE), life expectancy with disability (LEWD) and healthy life percentage (HLP) due to elimination (men and women, age 65)*

	LE (yr)	DFLE (yr)	LEWD (yr)	HLP (% points)
Men				
CNSLD	0.3	0.5	–0.2	2.3
heart diseases	3.1	1.5	1.6	+0.0
cancer[1]	2.7	0.9	1.8	–2.3
diabetes mellitus [1]	0.1	0.0	0.1	–0.1
arthritis//back complaints	0.0	0.7	–0.7	5.0
migraine/severe. headache	0.0	0.1	–0.1	0.4
other neurololgical diseases	0.1	0.1	+0.0	0.3
Women				
CNSLD	0.1	0.2	–0.1	1.0
heart diseases	2.7	0.9	1.8	+0.0
cancer	1.9	0.4	1.5	–1.2
diabetes mellitus	0.3	0.3	–0.0	1.0
arthritis//back complaints	0.1	1.0	–1.0	5.3
migraine/severe. headache	0.0	0.1	–0.1	0.4
other neurological diseases	0.1	0.1	+0.0	0.3

[1] Effect was not significant in the regression analysis.

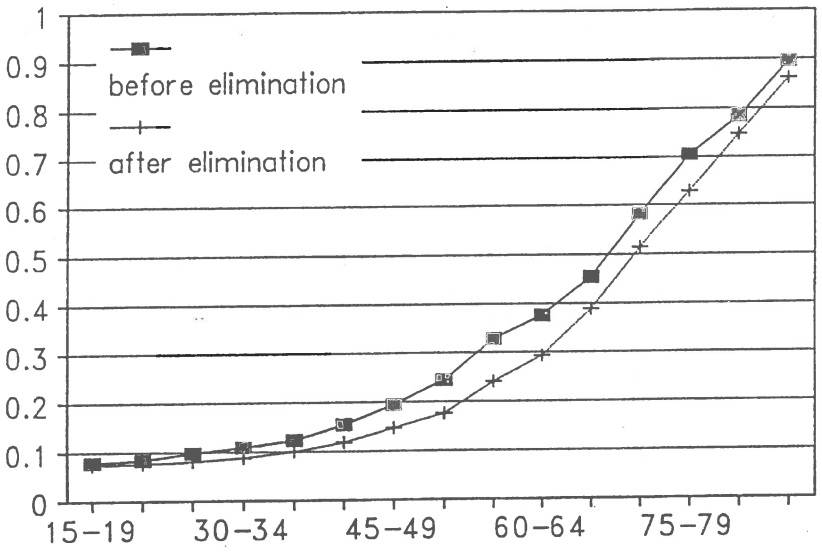


Figure 1a: *Estimated disability rates before and after elimination of arthritis/back complaints (women)*

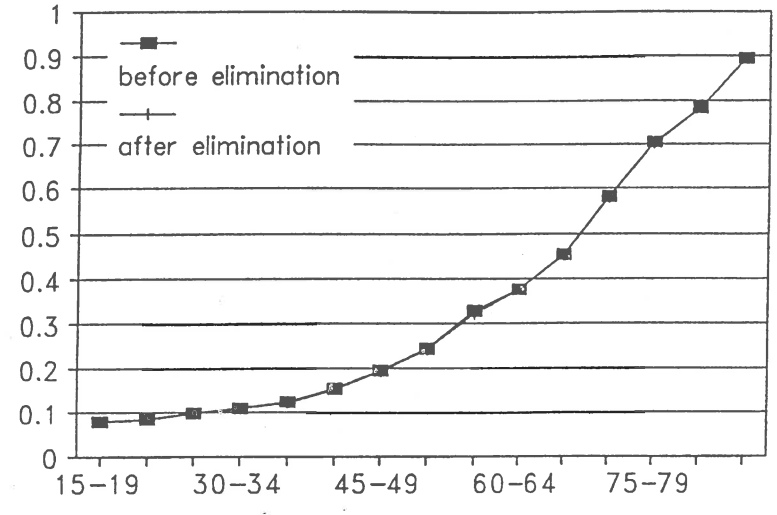


Figure 1b: *Estimated disability rates before and after elimination of cancer (women)*

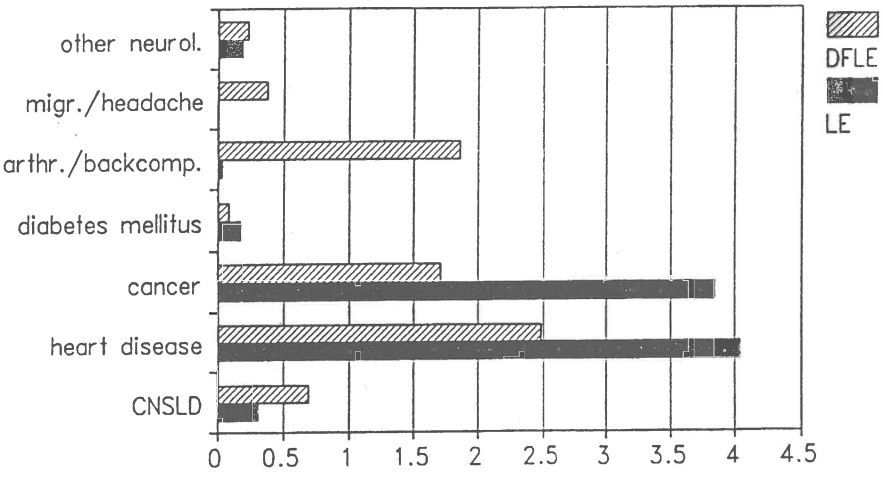


Figure 2a: *Increase in total and DFLE due to elimination of a disease (men, age 15)*

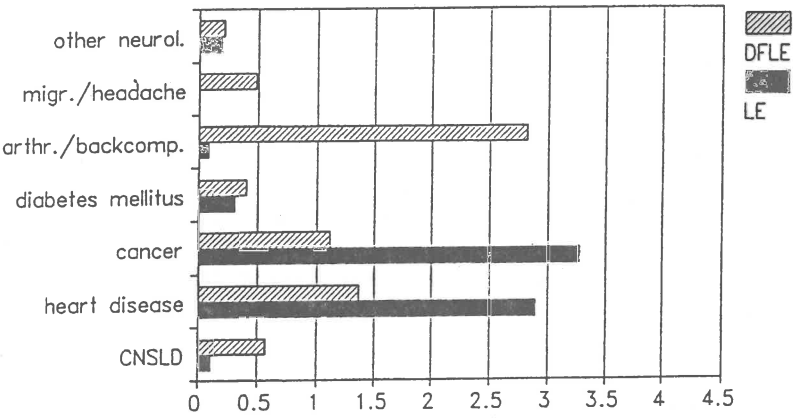


Figure 2b: *Increase in total and DFLE due to elimination of a disease (women, age 15)*

of morbidity takes place. With respect to the effect of elimination above age 65, similar results were found. The only important difference is the constant HLP for heart diseases, so, neither relative compression nor relative expansion of morbidity takes place. Elimination of 'other neurological diseases' has no significant effect on LEWD.

Conclusions

The main objective of this paper was to determine whether complete elimination of a disease leads to relative or absolute compression or to relative expansion of morbidity. Our findings show that elimination of a disabling chronic disease that causes little mortality (e.g. arthritis/back complaints, migraine/severe headache) leads to an increase in DFLE, while the total life expectancy remains more or less constant. Consequently, the LEWD declines and absolute compression of morbidity takes place. However, for migraine/severe headache the gain in disability-free years was small.

For cancer, a disease with high case-fatality and a low prevalence, elimination results not only in a gain in DFLE, but also in an even larger increase in the LEWD. In the years that people are saved from dying of cancer, they experience disabilities due to other causes. So, people live longer, but most of these extra years are with disability. Elimination of cancer therefore results in relative expansion of morbidity.

Elimination of diseases that cause significant disability and mortality, like heart diseases, results in an increase in total life expectancy and DFLE. Depending on the size of the mortality and disability decline, the increase in DFLE is accompanied by an increase in LEWD. With respect to heart diseases, elimination leads to both an increase in DFLE and in LEWD. This results in a decline in HLP at age 15: relative expansion of morbidity occurs. At age 65, the rate of increase in LEWD due to elimination of heart diseases was in equilibrium with that of the total life expectancy: neither relative

compression nor expansion of morbidity takes place. For other diseases that have consequences both in terms of disability and mortality, like CNSLD among men and diabetes mellitus among women, the (small) increase in total life expectancy is accompanied by a larger increase in DFLE. So, LEWD declines and absolute compression of morbidity takes place.

To conclude, complete elimination of a disease and its consequences in terms of disability and mortality results in a gain in DFLE. However, a side effect of the decline in mortality is that more people survive to older ages where the risk of becoming disabled is higher. If this is not compensated by a larger increase in DFLE, LEWD increases and relative expansion of morbidity may occur. This represents an increasing burden for society.

Acknowledgements

This project is supported financially by the Dutch Ministry of Health, Welfare and Cultural Affairs (WVC). The paper is supported by the Netherlands Organization for Scientific Research (NWO).

References

CBS (1984) *De leefsituatie van de Nederlandse bevolking van 55 jaar en ouder 1982*, deel 1A kerncijfers, 's-Gravenhage, Staatsuitgeverij

CBS (1987) *Overlevingstafels naar geslacht en leeftijd, 1981-1985*, 's-Gravenhage, Staatsuitgeverij

CBS (1992) *Overlevingstafels naar geslacht en leeftijd, 1986-1990*, 's-Gravenhage, SDU Uitgeverij

CBS *Intramurale Gezondheidszorg*, several years

CBS *Overledenen naar doodsoorzaak, leeftijd en geslacht in het jaar 19--* , 's-Gravenhage, SDU Uitgeverij, several years

CBS *Statistiek van de Bejaardenoorden*, 's-Gravenhage, SDU Uitgeverij, several years

CBS Leeftijdsopbouw van Nederland, 1 januari, *Mndstat bevolk*, (CBS), several years.

Colvez A, Blanchet M (1983) Potential gains in life expectancy free of disability: a tool for health planning, *Int J Epidem*, vol.12, no.2, pp.224–229

Foets M, van der Velden J, de Bakker D (1992) *Survey-design, Dutch National Survey of General Practice*, NIVEL, Utrecht

Hosmer DW, Lemeshow S (1989) *Applied logistic regression*, John Wiley & Sons, New York

Manton KG, Stallard E (1984) *Recent trends in mortality analysis studies in population*, Academic Press, Orlando

Mathers C (1992) *Estimating gains in health expectancy due to elimination of specified diseases*, Paper presented at the 5th meeting of the International Network on Health Expectancy (REVES), Ottawa, February 1992

McWhinnie JR (1981) Disability assessment in population surveys: results of the OECD common development effort, *Revue Epidémiol Santé*, vol.29, pp.413–419

Namboodiri K, Suchindran CM (1987) *Life table techniques and their applications*, Academic Press, Orlando

Robine J-M, Mathers C (1993) Measuring the compression or expansion of morbidity through changes in health expectancy. In Robine J-M, Mathers CD, Bone MR, Romieu I (eds) *Calculation of health expectancies: harmonization, consensus achieved and future perspectives*, Colloques INSERM, vol.226, John Libbey Eurotext, London

Robine J-M (1992) Disability-free life expectancy. In Robine J-M, Blanchet M, Dowd EJ (eds) *Health expectancy*, OPCS, HMSO, London

SCP (1991) *Aanvullend Voorzieningen Onderzoek* (AVO), Rijswijk

SIG, *Jaarboek Verpleeghuizen*, Utrecht, SIG, several years

SIG, *PIGG*, Utrecht, several years

Sullivan DF (1971a) A single index of mortality and morbidity, *HSMHA Health Reports*, vol.86, no.4, pp.347–254

Sullivan DF (1971b) *Disability components for an index of health*, Public Health Service Publication No. 1000 Series 2, No.42, Government Printing Office, Washington DC

Tsai PT, Lee EU, Hardy RJ (1978) The effect of a reduction in leading causes of death: potential gains in life expectancy, *Am J Public Health*, vol.68, no.10, pp.966–971

Appendix 1

Disease-clusters and related chronic diseases and causes of death

Diseases-clusters	Health Interview Survey (NIVEL)	Cause of death (ICD-9)
chronic non-specific lung disease (CNSLD)	chronic bronchitis; emphysema; asthma	490-493
heart diseases	heart complaints, cardiac failure	390-398; 410-414; rest of 390-459 (with exception of 401-405; 430-438; 440)
cancer	cancer (inc leukaemia)	140-208
diabetes mellitus	diabetes mellitus	250
arthritis/back complaints	backache (slipped disc, sciatica); rheumatism, arthritis; arthrosis	710-719; 720-724; 725-729; 730-739
migraine/severe headache	migraine; severe headache	346
other neurological diseases	Parkinson's disease, multiple sclerosis, epilepsy	330-337; 340-349 (with exception of 346); 350-359

INT'L. J. PSYCHIATRY IN MEDICINE, Vol. 24(2) 157-177, 1994

THE SOMATIZING PATIENT IN GENERAL PRACTICE

P. F. M. VERHAAK

M. A. R. TIJHUIS

Netherlands Institute of Primary Health Care

ABSTRACT

Objective: The exploratory study described in this article followed two groups of patients over a twelve-month period. Subjects were drawn from a pool of patients who had consulted their general practitioner during the three-month selection period. One group consisted of patients who had consulted their general practitioner at least once about a physical complaint that the GP regarded as predominantly psychosocial; these patients did not articulate complaints of an explicitly mental or social nature. The second group was characterized by the fact that its members voiced precisely such mental or social complaints. *Method:* The study investigated the extent to which the two groups (which were comparable in the severity of their complaints) differ with respect to patient characteristics such as the severity of their possible psychological problems, the frequency with which they visited their GPs, and the types of complaints—e.g. mental, psychosomatic and purely physical—they presented. *Results:* It was found that patients in the first group, whose somatic complaints were seen to have a psychosocial basis, are not the dependent types generally mentioned in theories about somatization. In fact, they adopt a more independent attitude to the GP than do patients voicing mental complaints. There are indications that for "somatizing" patients, underlying mental problems are less important than for "psychologizing" patients. *Conclusions:* Both the somatizing patients and the psychologizing patients continued very frequent visits to their GP during the 12-month research period, although chiefly to address physical complaints that the GP also assessed as such.

(*Int'l. J. Psychiatry in Medicine* 24:157-177, 1994)

Key Words: somatizing, psychosomatic, primary care, psychosocial problems

INTRODUCTION

Persons diagnosed as having some form of mental disturbance, as well as those who believe that they have problems of a mental or social nature, make disproportionately greater use of medical facilities, with general practitioners bearing the brunt of this [1]. Estimates of the percentages of mental disturbances among visitors to general practice physicians (from 22% to 35%) [2-5] are significantly higher than the 8 to 15 percent of the general population estimated to have some form of mental disturbance [6-9].

Most patients with mental problems come to their GP with an explicitly physical complaint [8-11]. Often the GP agrees that the request for help concerns a purely physical matter, but GPs also consider some physical complaints to be psychosocial in nature. Such a judgment need not result from a single consultation, but may be based on a lengthier history of interaction with the patient. For example, the number or the intensity of the complaints may be disproportional to demonstrable somatic dysfunctions. The judgment that a complaint is psychosocial is then based on the absence of evidence for a physical explanation, and is consistent with Lipowski's definition of somatization as "a tendency to experience and communicate somatic distress and symptoms unaccounted for by pathological findings, to attribute them to physical illness and to seek medical help for them" [12]. "Somatizing" is thus a purely descriptive concept.

Somatizing is often linked to mental problems [13, 14], in which case it is interpreted as "the articulation of emotional problems and psychosocial stress by way of physical symptomatology" [14]. According to Katon et al. "somatization is a metaphor for personal distress. It as an idiom to obtain help from care-givers and a universally powerful mechanism of obtaining social support and manipulating relationships." In many cases somatizing is also equated with the psychiatric DSM-III diagnosis "somatization disorder" [15]. A distinction is made between "acute somatization" lasting some days or weeks, "subacute somatization" for a period of months, and "chronic somatization" over much longer periods. This distinction is important because in primary care settings somatization disorder—identified as a form of chronic somatization—is much more rare than is the patient who presents with one or two symptoms that are not accounted for by pathological findings. We stress that this study reports on the latter type of patients, few of whom are likely to have a somatization disorder as defined by DSM-III criteria.

Although GPs often suspect a psychosomatic basis in patients whose complaints are difficult to explain physically, they do not always treat the presumed mental component of such cases. While 90 percent of those patients who describe their problems as mental or social receive some form of psychosocial treatment, only 50 percent of patients who arrive with physical complaints that appear to have mental causes receive such attention [16, 17]. Thus many patients who are emotionally disturbed do not express these feelings to their doctors, and present

with physical symptoms that appear to have no physical basis. GPs often do not discuss the probable emotional background of such symptoms with either type of patient. Those who argue that physical complaints for which no explanation can be found are in general an expression of emotional distress advocate a "reattribution process:" the GP must try to give the patient insight into the psychosocial factors that cause or aggravate the symptoms [18]. There are indications that recognized mental complaints have a better prognosis than unrecognized complaints and that GPs must be better trained in recognizing possible mental problems, and in broaching these with the patients, even if the patient does not raise them [16, 19-22].

On the other hand, notably in the world of general practice, psychosocial causes need not always be sought. Thus Huygen has remarked that many complaints—including psychosomatic ones—fade of their own accord, and that the physician who seeks psychosocial causes unasked may prompt undesired side-effects [23]. This does not imply that physicians must proceed narrowly on the basis of patient-described physical complaints and neglect the mental component, only that a reserved attitude is desirable [24].

The present study investigated the prognosis of patients with physical symptoms for which medical explanations are lacking. A group of somatizing[1] patients has been followed for one year. The group has been compared with a group of patients who presented explicitly mental complaints and with a control group of average patients.

The following questions have been put forward:

1. *What characterizes somatizing patients and in what respects do they differ from patients who present explicitly mental complaints (psychologizers) and from average patients?*

   The degree of mental disturbance and psychosocial problems experienced by somatizing and non-somatizing patients and average patients is compared. The extent to which the demographic and personality characteristics of somatizing patients may be distinguished from those of non-somatizing and average patients has also been investigated. It has been assumed that somatizing patients will tend to be older, less well educated and more dependent on their GPs than patients who explicitly discuss their mental problems [13, 14].
2. *What trends can be distinguished in the medical histories of somatizers and psychologizers over a one-year period?*

[1] By somatizing we mean the presenting of physical complaints that the GP considers to be mainly of a psychosocial nature. Physical complaints that the GP considers to be of a psychosocial nature are designated as "psychosomatic complaints." Somatizing patients are patients whose GPs judge their physical complaints to be mainly psychological ones, non-somatizing patients are patients who bring forward explicitly psychological complaints. We shall address them in this article as "somatizers" and "psychologizers" respectively.

The number of consultations during the year after selection and the kind of complaints presented by somatizing and non-somatizing patients are compared. A distinction is made between complaints explicitly formulated as psychosocial; complaints without sufficient evidence which are, according to the GP, somatically formulated (psychosomatic complaints); and somatically formulated complaints the GP agrees are mainly somatic.

To allow a valid comparison, cohorts have initial complaints of the same severity, based on the GP's assessment of the life-threatening or incapacitating impact of the symptoms presented during the selection period. The degree of possible mental illness was evaluated by means of the General Health Questionnaire (GHQ-30) completed by all patients.

We assume that failure to recognize the psychosocial basis of psychosomatic complaints during the selection period will be identifiable by the more frequent presentation of psychosomatic and mental complaints in the follow-up year in comparison with "non-somatizers," assuming that there are no differences in the treatment of the two groups. Therefore, the treatment histories of the follow-up year will be taken into account.

DATA AND METHOD

The patients and practices selected all participated in the Dutch National Study into Morbidity and Services in General Practice [25]. This larger study registered all contacts (n = 335,000) between 103 practices and their registered patients over a three month period. On the basis of this registration, somatizing patients and patients with explicitly mental complaints were selected from nine practices involving sixteen GPs. For an additional nine-month period the number of contacts with the GP's office and the complaints presented there by the two groups were recorded; the patients also completed questionnaires twice during the study period.

The complaints, registered during the registration period were coded with the aid of ICPC.[2] A somatized mental complaint is defined as a somatic complaint (neither P nor Z) that the physician assessed as predominantly psychosocial (scoring 4 or 5 on a five-point scale from somatic to psychosocial). The group of psychologizers explicitly described mental complaints at least once during the selection period. Somatizing patients presented only somatic complaints, which, at least once (during the selection period) were assessed by the GP as psychosocial. Based on a three-month registration six percent of all visitors could be defined as somatizers and twelve percent as psycholizers.

[2] This well-known classification system for general practice (Lamberts and Wood, 1987) places each complaint in one of 17 main categories; for example, P = psychic complaints; Z = social complaints.

The following data was registered for contacts with each patient during the twelve-month period:

1. All complaints presented:
   - ICPC-code;
   - (differential) diagnosis by the GP (coded in ICPC);
   - five-point scale assessment by the GP of the possible psychosocial character of the complaint;
   - five-point scale assessment by the GP of the life-threatening character of the complaint;
   - five-point scale assessment by the GP of the disabling character of the complaint;
2. Treatment:
   - counselling;
   - referral;
   - consultation of specialist;
   - duration of the consultation.

The patients completed questionnaires in the fourth month and during the final quarter of the registration year. The final study sample was compiled on the basis of the response to these questionnaires. Full records are available on 397 patients with explicitly mental complaints and 411 somatizing patients.

Questionnaires recorded demographic data (age, sex, marital status, education), and included the following attitude lists and screening questions:

- a list used to measure the patient's dependence on the GP (on a scale of 12 to 60). Higher scores indicate less dependent attitudes;
- a Dutch adaptation of the list originally developed by Rotter [26] to measure "Health Locus of Control" (specifically, the scale measuring the orientation of internal control);
- a screening list, the General Health Questionnaire (GHQ-30), [27] was used to trace psychosocial problems and to assess the likelihood that a patient suffered from severe psychosocial problems are involved (GHQ scores range from 0 to 30; usually 4 or 5 is used as cut-off), with scores higher than 4 considered to indicate possible psychopathology; and:
- a list aiming to quantify the specific psychosocial problems that patients have faced [28]. Scores can range from zero (no psychosocial problems, no burden) to twenty-two (high burden).

The aforementioned national study, which includes data derived from interviews with some 13,000 respondents, served as a base reference. These respondents answered the same questionnaires used in the present study. Because the data from the contact registration of these respondents are available, it is possible to make two kinds of comparisons: each group may be compared with an average

practice population (all respondents to the national survey) and with the profile of the average person who visited a GP in a comparable three-month period. Data from the first three months, during which the individuals were selected, can be readily compared with the national survey data, because by definition all respondents visited their GP during the period. This data is less useful for comparison with later periods, for which data on the total population (including non-visitors) is more relevant, given that patients need not have visited their GPs in each successive quarter of the study period. Differences between the groups have been statistically tested by means of a unilateral *t*-test or, in the case of nominal variables (for example marital status) using Pearson's Chi-square. Following the recent recommendations of the editorial board of this journal [29], effect sizes are reported, using Cohen's *d* as a measure for effect size. Following Cohen's recommendations, a value of .20 is considered a small effect, .50 indicates a moderate effect and.80 a large effect [30].

RESULTS

Demographic and Personality Characteristics of Somatizing, Psychologizing, and Average Patients

Table 1 lists the demographic and personality characteristics of somatizing patients, patients who explicitly indicated the existence of psychosocial problems and an average group of visitors to GPs.

Somatizing patients are younger and (therefore) less frequently divorced or widowed ($p < 0.05$). Somatizers are less dependent on their GP than are psycholizers. However, somatizing and psychologizing patients could not be distinguished on the basis of gender or control orientation, nor did their educational level differ. Although the reported differences reach the level of significance, the effect size is rather moderate, Cohen's *d* not exceeding .26.

Compared to the average patient in the national study, patients with concealed or unconcealed mental problems are more often between twenty-five and forty-five years old and more frequently divorced; they also have a lower internal control orientation than the average GP visitor, although the effect size in this respect is rather low (Cohen's $d = .12$ respectively .07). Somatizing patients are less dependent on their GP then average visitors ($d = .21$).

Severity of Symptoms of Somatizers and Psychologizers

Comparisons of the course of complaints of somatizing and psychologizing patients requires that the severity of the complaints be comparable. To determine the extent to which the selected patients have complaints that are of similar severity, the life-threatening or disabling potential of their complaints was assessed by their GPs (see Table 2).

Table 1. Demographic and Personality Characteristics of Somatizers and Psychologizers, Both Divided according to GHQ-Scores

	Somatizers	Psychologizers	Average Consulters
N	411	397	4042
Sex			
Male	39	42	43
Female	61	58	57
Marital Status			
Married	67	64	65
Divorced	7	9	4
Widowed	3	7	7
Never married	23[a,b]	20[a,b]	23[b]
Education			
Higher	13	13	11
Secondary	20	20	21
Primary	54	51	47
Unknown	12	16	21
Age (years)			
16-25	14	11	18
25-45	58	49	43
45-65	24	33	29
65 and over	5	7	10
Mean Age	39[a,b]	42[a]	42[b]
Mean Scores for Dependency	39.4[a,b]	37.6[a,b]	36.3[b]
Internal Locus of Control	18.4[b]	18.8[b]	19.3[b]

[a]Statistically significant difference between somatizers and psychologizers.
[b]Statistically significant difference between average consulters and somatizers or psychologizers.

People who explicitly raise psychosocial concerns during their consultation with their GPs visit their GP more often in the quarter in which they first report such complaints ($p < 0.05$; $d = .30$). Similar findings were reported in the UK by Wright, who also divided patients with "clinically significant psychiatric disturbance" into patients with mental symptoms and patients with somatic symptoms without a physical explanation [31].

The "severity" of the complaints, indicated by the GP as life-threatening or disabling potential, is the same for both somatizing and non-somatizing patients. For a number of types of complaints, namely involving the gastrointestinal tract,

Table 2. Average Severity of Symptoms Presented Within the First Three Months of Registration

	Somatizers	Psychologizers
N		
Patients	411	397
Consultations	2.1	2.6[a]
All Complaints		
Life threatening	4.93	4.90
Incapacitating	4.78	4.75
Mental Complaints		
Life threatening	—	4.90
Incapacitating	—	4.76
Psychosomatic Complaints		
Life threatening	4.94	4.87
Incapacitating	4.80	4.71
Somatic Complaints		
Life threatening	4.93	4.91
Incapacitating	4.77	4.74

[a]Statistically significant difference between somatizers and psychologizers ($p < 0.05$; $d = .30$)

nervous system, muscular system and with reference to pregnancy, somatizers consulted their GPs more frequently that did those who presented openly mental difficulties. After adjusting for age and sex, however, the difference with reference to pregnancy disappeared [32].

Psychopathology of Somatizers and Psychologizers

Two measures of the extent to which somatizing patients, patients with mental complaints and average patients differ from each other in psychopathology are available. Table 3 indicates the number of psychosocial areas in which the two groups of patients indicated they were experiencing problems and which percentage of the patients from the two groups had GHQ scores indicating the significant likelihood of a severe mental problem.

This comparison shows that the patients who made explicitly psychosocial complaints have experienced more psychosocial problems and are more likely to be suffering from a serious mental disorder. At first sight, one would say that they are in a "more serious" psychosocial state. On average, both groups had a higher

Table 3. Number of Psychosocial Problem Areas and GHQ-Scores for Somatizers and Psychologizers

	Somatizers	Psychologizers	Average Consulters
N	411	397	4775
Number of Problem Areas			
No problems	23%	17%	45%
1-5	49%	46%	45%
6-8	20%	19%	7%
> 8	8%	18%	3%
Mean	3.0[b,c]	4.0[a,c]	1.6[a,b]
GHQ			
Mean score	4.6[b,c]	7.3[a,c]	2.7[a,b]
Percent > 4	32%	45%	16%

[a]Statistically significant difference with somatizers ($p < 0.05$).
[b]Statistically significant difference with psychologizers ($p < 0.05$).
[c]Statistically significant difference with average consulters ($p < 0.05$).

Effect sizes:	Somatizers-Psychologizers	Somatizers-Average	Psychologizers-Average
Mean *n* of problem areas	.30	.29	.38
Mean GHQ	.34	.18	.34

contact frequency, experienced more problems and had higher GHQ scores than the average visitor to the doctor's office. These differences are all somewhat larger than earlier reported differences: as is reproduced in the note at Table 3, effect-sizes mostly vary between .30 and .40.

To allow a comparison of the characteristics and the course of the complaints of somatizing and psychologizing patients, the severity of the mental problems as assessed by the GHQ must be taken into account. The following analyses subdivides each group of patients into those who, according to the GHQ, have a considerable chance of a serious mental disorder, and those who do not.

The Course of Complaints among Somatizing and Non-Somatizing Patients

To investigate the extent to which somatizing patients continued to voice psychosomatic complaints, a comparison is made between the complaints voiced in the first (selection) quarter and the last quarter of the study.

Figure 1 shows the number of contacts per 100 patients in the first and last quarters. Because patients were selected on the basis of having made at least one contact during the first quarter, *T1* is artificially high and a drop is to be expected. At *T1*, psychologizers had a higher contact frequency than somatizers, and among the former, those with elevated GHQ scores had the highest number of contacts. (Cohen's $d = .41$).

By *T2*, the contact rate for psychologizers with low GHQs had dropped to the level of the average patient (80 contacts per 100 registered patients). The contact rate of somatizers with a low GHQ is considerably higher (Cohen's $d = .25$). Among the patients with a high GHQ there is no significant difference anymore between somatizers and psycholizers on *T2*.

Reasons for visiting the GP have been categorized as "psychosocial" (ICPC P or Z), "psychosomatic" (ICPC other than P or Z and judged by the GP not to be entirely somatic) and "somatic" (ICPC other than P or Z and judged by the GP to be entirely somatic). Figures 2 and 3 show the total number of complaints presented to the GP in the first and last quarters of the study period. Patients who scored below the GHQ threshold value (Figure 2) and patients who scored above it (Figure 3) have been considered separately.

It would be meaningless to compare the complaints presented by somatizers and psychologizers at *T1*, because they have been designated as such according to the presence or absence of explicitly stated psychosocial complaints. As might be expected given the larger number of contacts, psychologizers also had a larger number of complaints and symptoms at *T1* than did either low-GHQ somatizers or high-GHQ somatizers. The effect sizes are .26 respectively .52.

At *T2*, however, psychologizers with low GHQs (like average patients) presented fewer symptoms and complaints than did somatizers with low GHQs ($d = .25$). The difference can be entirely accounted for by the larger number of entirely physical complaints and symptoms presented by the latter. Patients with high GHQs at *T1* present more complaints at *T2* than the average patient, with psychologizers presenting more complaints than somatizers. Patients with high GHQs presented both more entirely physical complaints and more explicitly stated psychosocial complaints. (This was especially true of psychologizers.)

In brief, although the psychosomatic complaints of the majority of the "somatizers" had vanished by the fourth quarter, explicitly mental complaints and above all purely somatic complaints had taken their place; as a result, the average number of complaints remains above the level of an average population.

At both *T1* and *T2* (Figure 4) GHQ scores show more remitting cases and more continuing cases among psychologizers than among somatizers. The net result is a larger proportion of psychologizers being a GHQ-case in both periods.

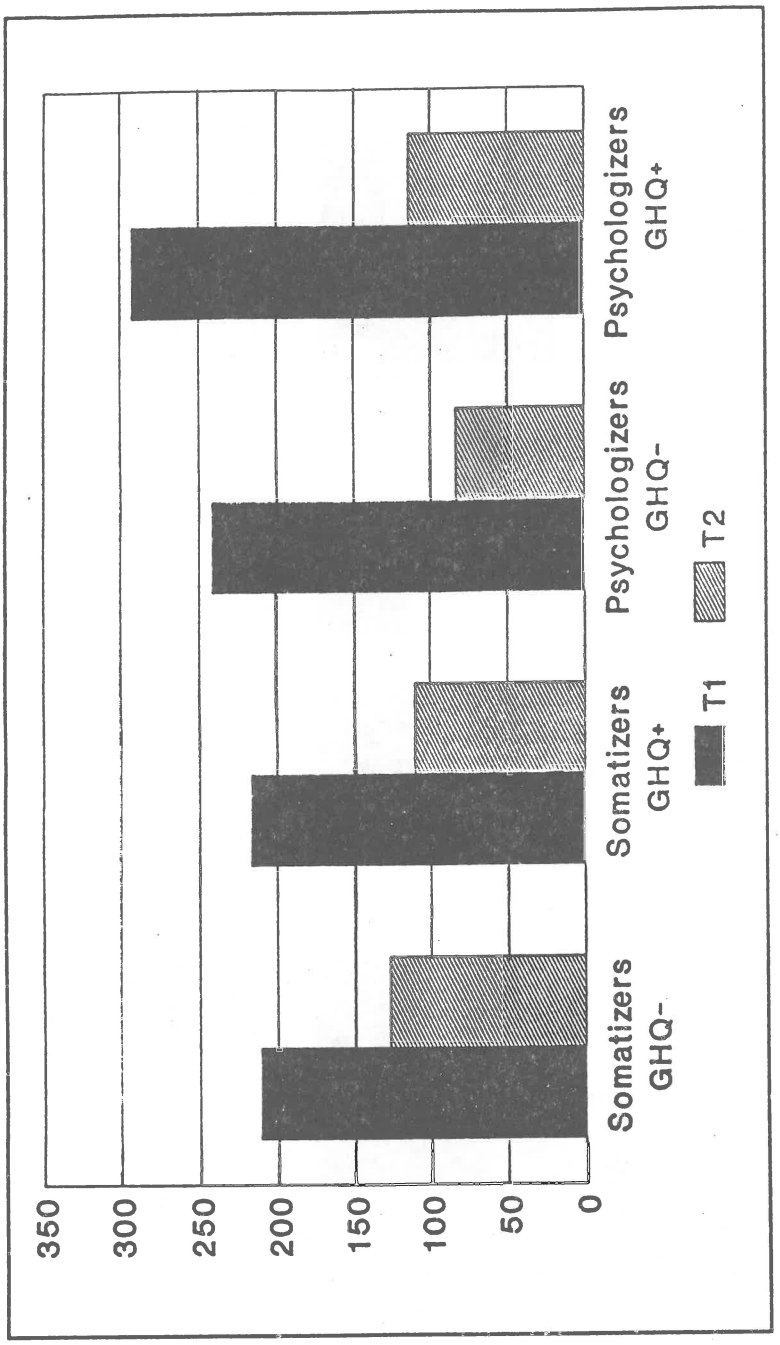


Figure 1. Number of contacts for somatizers and psychologizers on T1 and T2.

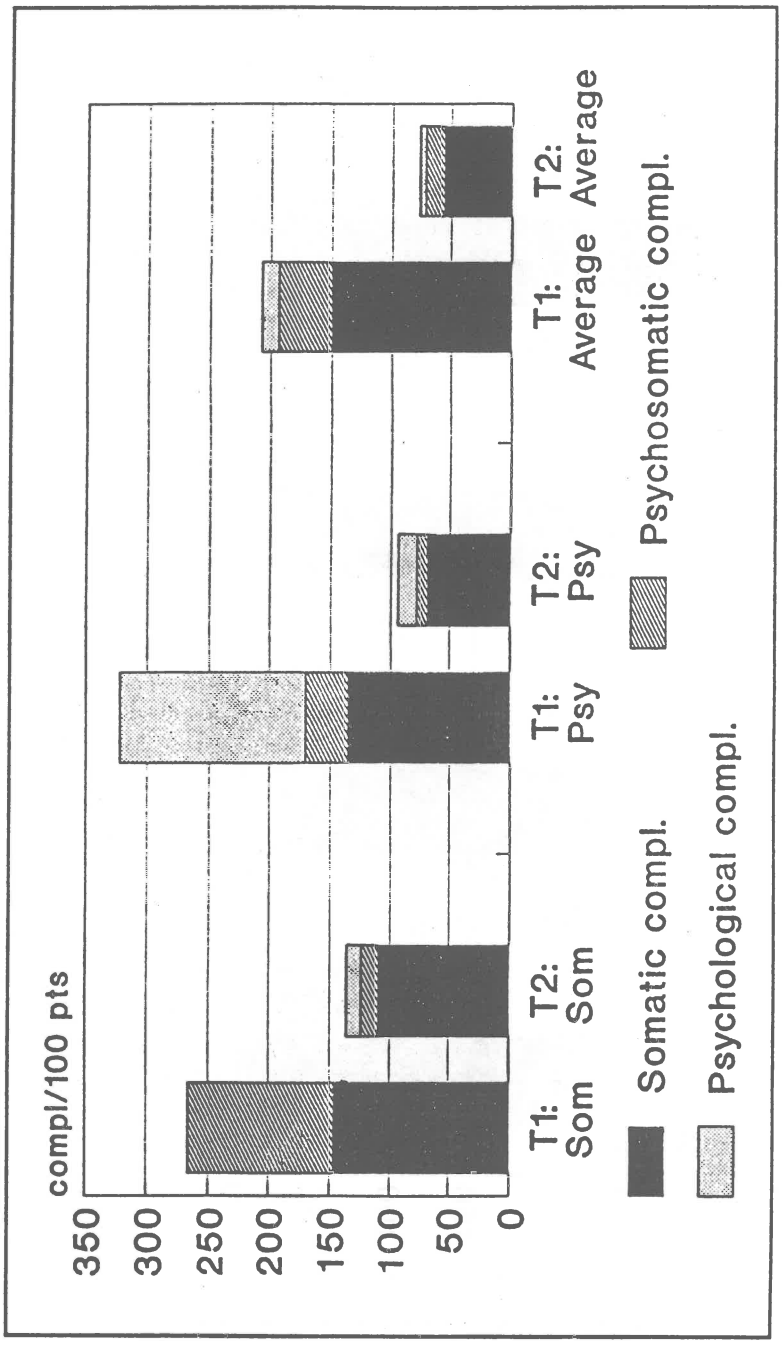


Figure 2. Somatizers and psychologizers: T1 and T2 GHQ–.

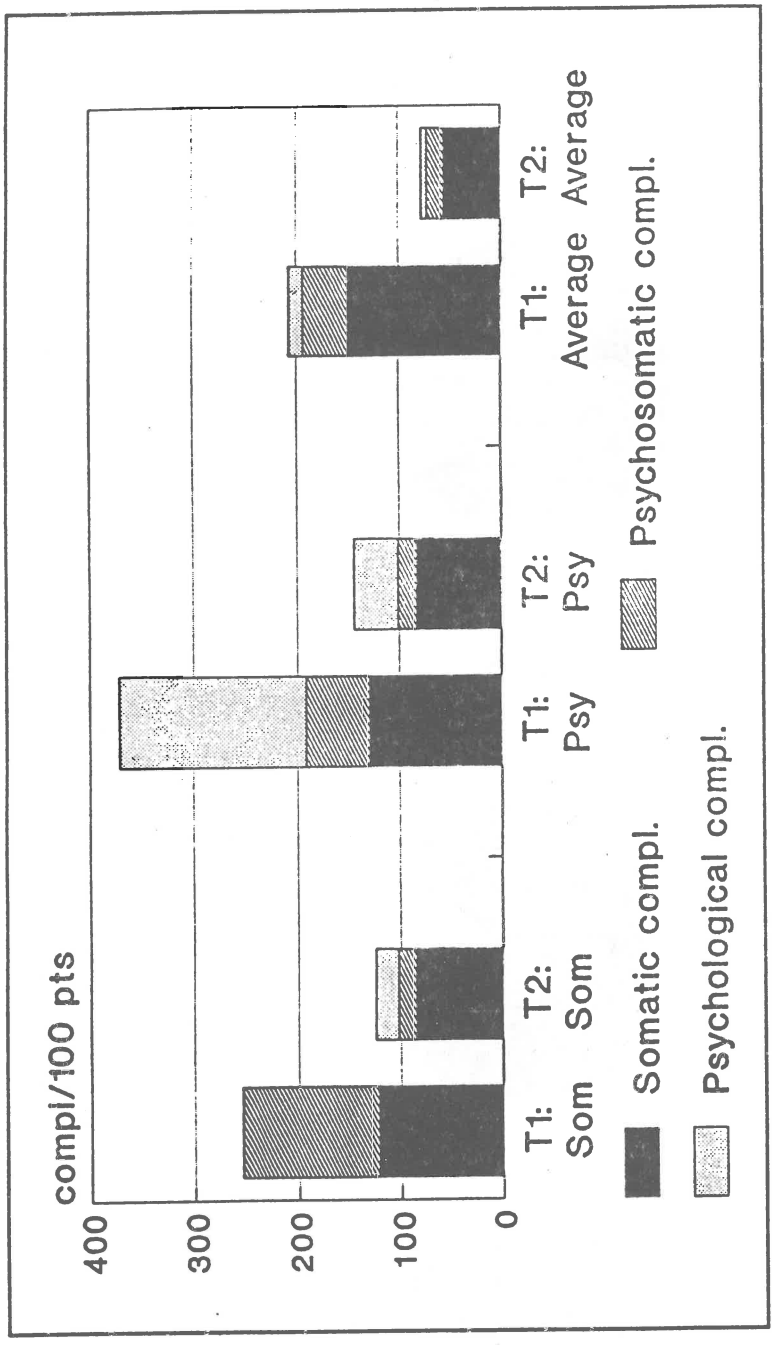


Figure 3. Somatizers and psychologizers: T1 and T2 GHQ+.

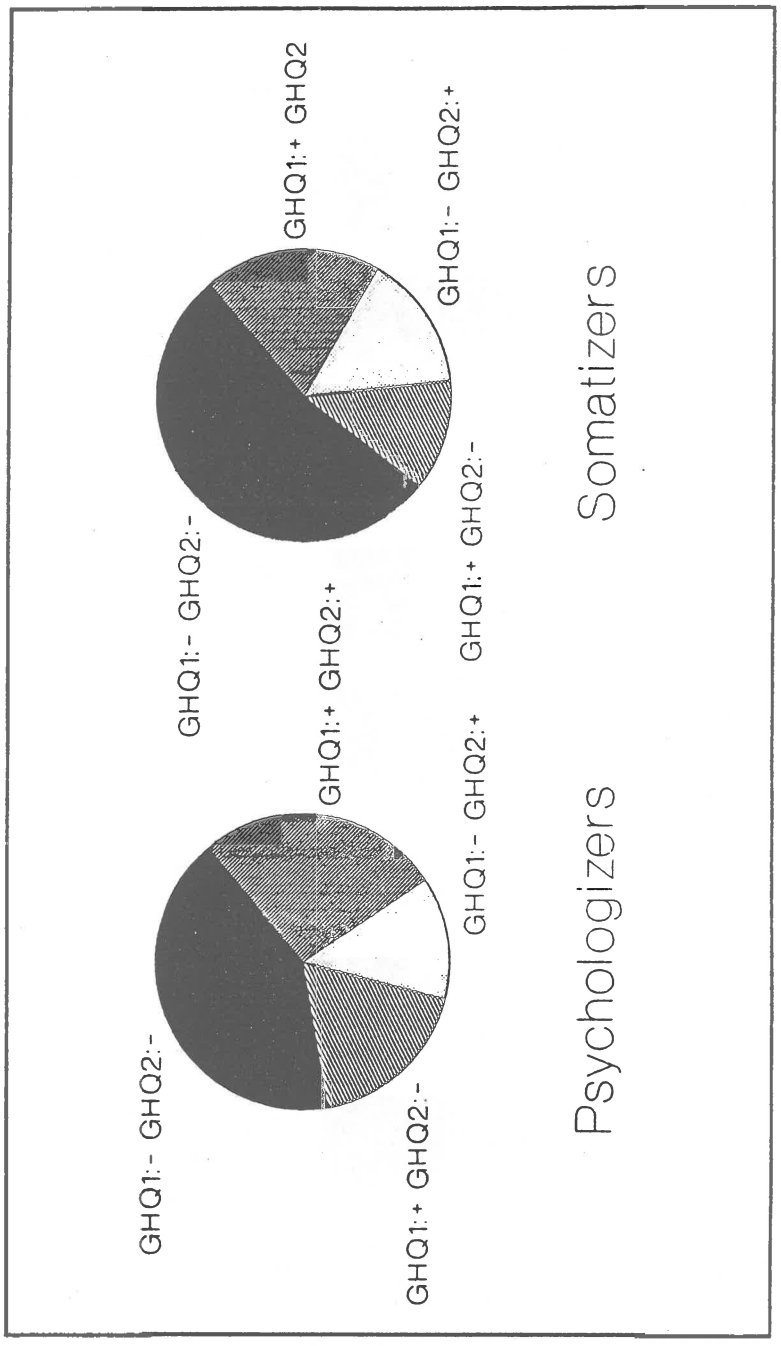


Figure 4. GHQ-caseness on T1 and T2: somatizers and psychologizers.

Treatment of the Psychosocial and Psychosomatic Complaints of Somatizers and Psychologizers

Over the course of the year, individuals in both the somatizing and the psychologizing groups presented explicitly mental complaints and psychosomatic complaints. Table 4 identifies the types of treatment the GPs provided: by group, by GHQ category and by complaint category over the course of the registration year.

Treatment of explicitly mental complaints is notably different from that provided for psychosomatic complaints, regardless of the original classification of the patients as somatizing or psychologizing. Consistent with the evidence in the literature noted in the introduction to this article, patients with psychosomatic complaints received less discussion therapy and were less frequently referred to specialists. Fewer mental health consultations are carried on about them, and with the GPs consultations are shorter. A *t*-test of these differences comparing all psychosomatic complaints to all mental complaints is statistically significant at a 0.05 level.

On the other hand, neither the likelihood of having psychiatric problems (according to the GHQ) nor the initial categorization of patients as somatizers or psychologizers correlated with the type of treatment received. In other words, "somatizers" who later presented explicitly mental complaints received the same treatment as "psychologizing" patients who presented explicitly mental complaints in the first quarter.

Table 4. Treatment of Mental and Psychosomatic Complaints of Somatizers and Psychologizers, Divided according to GHQ-Score

	Somatizers				Psychologizers			
Type of Problem	GHQ+ Ment.	Psom.	GHQ– Ment.	Psom.	GHQ+ Ment.	Psom.	GHQ– Ment.	Psom.
N	41	130	46	279	175	93	216	84
Treatment: Percent of patients with at least one:								
Counseling	76%	61%[a]	76%	53%[a]	85%	68%[a]	76%	55%[a]
Referral	5%	5%	13%	3%[a]	16%	3%[a]	10%	2%[a]
Mental health consultation	0%	0%	2%	0%	10%	0%[a]	5%	1%
Average length of consultation (min.)	12	10[a]	12	10[a]	13	11[a]	11	10[a]

[a]Statistically significant difference between treatment of mental and psychosomatic complaints.

DISCUSSION

One group of patients was selected over a three-month period according to whether they visited their GP in order to present, at least once, physical complaints which the GP diagnosed as being primarily psychosocial. During the selection quarter, the individuals selected for this group did not explicitly mention mental or psychosocial problems. This group was compared to a second group which, during the selection period, explicitly raised mental problems as such to the GP; the groups were then compared to national survey data (the average patient).

The finding that "somatizers" are younger, less often divorced or widowed, and less dependent on the GP than "psychologizers" is inconsistent with the usual characterization of somatizing patients presented in the literature. Patients who express their emotional problems by means of physical complaints are commonly described as older, less educated [13] and more dependent on their physicians than the present study suggests [14]. In fact, the psychologizing group was older, less well educated and more dependent than the somatizing group, although the effect sizes in these respects were only small.

According to the GHQ, somatizers are less likely to have a serious mental disturbance, and according to the BIOPRO they experience fewer psychosocial problems than do patients who present their problems as explicitly mental. They also visited their GPs less frequently during the selection period. These effect sizes are somewhat larger than the former. However, the severity of their complaints is approximately the same as that of the patients who described explicitly mental problems. Wright reports the same results: patients with "clinically significant psychiatric disturbance, who restricted themselves to somatic symptoms without physical explanation" were better off than those who complained explicitly about psychiatric disturbance [31]. Bridges et al. report higher consultation rates for "somatizers" than for "psycholigizers" and lower scores on depression scales and on the GHQ [33].

Psychosomatic complaints are less often treated by discussion, and more specialized mental health care workers are almost never involved irrespective of the group within which the patient was originally classified. Nevertheless, the frequency with which patients from the two groups bring these psychosomatic complaints forward declines strongly. Again, Wright reported the same results [31]. The frequency with which patients in the psychologizing group visit the doctor remains higher than expected, although most of their complaints are assessed by their GPs as purely somatic.

Therefore, somatizing in primary care is in many cases not a matter of somatization disturbance in a psychiatric sense (many symptoms without physical explanation over an extended period) but of "(sub)acute somatization," which can be expected in primary care settings. As others have noted, while chronic somatization disorders (including DSM-III diagnoses of somatization disorders) are rare in primary care settings, one-time incidents of somatization are not [31, 34].

However, a patient who presents a single psychosomatic complaint is unlikely to continue to do so. On the contrary, the patients in our cohort of "somatizers," while frequent visitors to their doctors' offices, presented complaints that were diagnosed as entirely physical. In our opinion it is necessary to distinguish between patients with DSM-III disorders and patients who are (perhaps too) eager to visit their GPs.

Thus somatization need not be an expression of underlying emotional distress. Firstly, the initial GHQ described only one third of the group of "somatizers" as likely to have "serious mental disturbance." Secondly, by far the majority of "somatizing" patients voiced no mental complaints to their GPs in the course of the follow-up year. Thirdly, over the course of the study period the frequency with which psychosomatic complaints were presented fell to the average rate for the general population. And finally, the minority of patients who, according to the GHQ, were more likely to have a serious mental disturbance did not differ in these respects from the majority. While they tended to visit their GPs more frequently at the end of the study than did the psychologizing group, they generally presented purely somatic complaints.

The same patterns observed among the physicians as a group were observed in the case of most of the physicians individually; thus although the behavior studied is partly determined by the treating physicians, the results cannot be ascribed to anomalies of a few individuals.

The gap between the initial presentation of symptoms and the administration of the GHQ ranged from one week to three months, and it is possible that delaying the initial GHQ test until the first month of the second quarter of the study may (partially) account for the fact that most somatizers could not be classified as "psychologically disturbed." However, the same delay occurred in the administering of GHQ tests to the cohort of psychologizers, of whom a significantly larger number were identified as likely to suffer from psychological disturbance. In addition, the GHQ proved to be a rather stable measurement. About 60 percent of the original high-GHQ cases in each group remained high-GHQ cases at *T2*, approximately a year after the original assessment. Moreover, the apparently more "severe" condition of the psychologizers might derive from their higher consultation frequency in the first three months (Table 2) and the larger number of problems they experienced. Theoretically it is possible that, at the time they presented symptoms, all of the somatizers had elevated GHQ scores. However, as other studies have pointed out in general their psychosocial distress is less severe than that of people who explicitly mentioned such distress [31, 32].

It is also possible that the GPs have underreported or misclassified psychosomatic complaints. If the GPs fail to recognize and classify cases of somatization disorder and instead tend to classify the majority of complaints as purely physical, this would affect the accuracy of the regression analysis of psychosomatic complaints. The discrepancy between symptom presentation as assessed by the GP (mostly "entirely physical") and the persistence of elevated GHQ scores would

then indicate a need to reassess the findings. Recent research in The Netherlands indicates that many instances of mental illness that are unrecognized by GPs are borderline cases; such individuals are not severely handicapped and do not demand help [35].

Given our results, we would go one step further and take the position that somatizing need not always refer to underlying unsolved psychosocial problems, but may in itself be an undesirable behavioral pattern. Such a pattern could, in a psychiatric sense, be characterized as "hypochondria," "hysterically cognitive style" or "obsessive behavior" [36]. However, like Portegijs et al. [15] we suggest viewing somatizing less as a disturbance than as a personality trait. Somatizing is then interpreted as a problem-solving mechanism at a cognitive and behavioral level, which should be handled as such.

In other words, it is not the symptom in itself which is relevant but the help-seeking behavior of the patient. As argued above, such patients are unlikely to be more mentally distressed than average patients. They distinguish themselves by help-seeking behavior which might be seen in psychosocial terms but also can be understood as a form of consumerism.

Moreover, this behavioral pattern is not explained by the literature's characterization of the older, submissive patient who, vis-à-vis the physician, resorts to medical language to express discomfort or distress. The patients in our study whose physical complaints were judged to be psychosocial in no way adopted dependent attitudes toward their physicians. It seems more appropriate to characterize them as self-aware consumers. The majority of the "somatizers" are characterized neither by an oversupply of mental problems nor by a dependent attitude, but by frequent visits to the physician and an independent attitude.

Further, although physicians rarely discussed their assessment that particular physical complaints had psychosocial bases with their patients, and although these patients were rarely referred to mental health care workers, such complaints largely disappeared over the course of the year. Somatizing patients continued to be frequent visitors, but possibly the patient, rather than the complaint, calls for attention. If this is the case there seems to be little point in raising the potentially psychosocial basis of a complaint with the patient. Ford emphasizes the importance of a physician-patient relationship in which the physician, by mutual participation (as opposed to a physician-oriented attitude), must make the patient aware of the patient's own responsibilities (cf. the concept "room for the patient") [37-39]. This implies that the general way in which the patient deals with sickness and health should be treated.

Our interpretation of somatizing offers on the other hand a hopeful prospect: the "somatizer" is well-equipped as an equal negotiating partner to make the desired contribution to personal health. On the other hand, he or she may be typified as a modern consumer who might not accept a doctor's authority but who will argue that one has paid for and deserves the maximum amount of care. In this respect,

emphasis should be laid on the doctor-patient relationship, while serious attention should be paid to the complaints as well.

REFERENCES

1. R. A. Schurman P. D. Kramer, and J. B. Mitchell, The Hidden Mental Health Network, *Archives of General Psychiatry, 42*:89-94, 1985.
2. L. G. Kessler, B. J. Burns, S. Shapiro, G. L. Tischler, L. K. George, R. L. Hough, D. Bodison, and R. H. Miller, Psychiatric Diagnoses of Medical Service Users: Evidence from the Epidemiological Catchment Area Program, *American Journal of Public Health, 77*:18-24, 1987.
3. L. G. Kessler, P. D. Cleary, and J. D. Burke, Psychiatric Disorders in Primary Care, *Archives of General Psychiatry, 42*:583-587, 1985.
4. D. Goldberg and K. Bridges, Screening for Psychiatric Illness in General Practice: The General Practitioner versus the Screening Questionnaire, *Journal of the Royal College of General Practitioners, 37*:15-18, 1987.
5. P. Williams, A. Tarnapolsky, D. Hand, and M. Shepher, Minor Psychiatric Morbidity and General Practice Consultations: The West London Survey, *Psychological Medicine Monograph,* Supplement 9, 1986.
6. D. Goldberg, C. Kay, and L. Thompson, Psychiatric Morbidity in General Practice and the Community, *Psychological Medicine, 6*:565-569, 1976.
7. P. Hodiamont, N. Peer, and N. Syben, Epidemiological Aspects of Psychiatric Disorder in a Dutch Health Area, *Psychological Medicine, 17*:495-505, 1987.
8. J. K. Myers, M. M. Wasserman, G. L. Tischler, C. E. Holzer, P. J. Leaf, H. Orvaschel, J. C. Anthony, J. H. Boyd, J. D. Burke, M. Kramer, and R. Stoltzman, Six-month Prevalence of Psychiatric Disorders in Three Communities, *Archives of General Psychiatry, 41*:959-967, 1984.
9. J. L. Vazquez-Barquero, C. Pena, J. F. Diez-Manrique, A. Arenal, R. G. Quinatanal, and C. Samaniego, The Influence of Sociocultural Factors on the Interaction Between Physical and Mental Disturbances in a Rural Community, *Social Psychiatry and Psychiatric Epidemiology, 23*:195-201, 1988.
10. P. F. M. Verhaak, H. J. Wennink, and M. A. R. Tijhuis, The Importance of the GHQ in General Practice, *Family Practice, 7*:319-324, 1990.
11. H. C. Schulberg, M. McClelland, and B. J. Burns, Depression and Physical Illness: The Prevalence, Causation, and Diagnosis of Comorbidity, *Clinical Psychology Review, 7*:145-167, 1987.
12. Z. J. Lipowski, Somatization: The Concept and its Clinical Application, *American Journal of Psychiatry, 145*:1358-1368, 1988.
13. W. Katon, A. Kleinman, and G. Rosen, Depression and Somatization: A Review, Part I, *The American Journal of Medicine, 72*:127-135, 1982.
14. G. Rosen, A. Kleinman, and W. Katon, Somatization in Family Practice: A Biopsychosocial Approach, *The Journal of Family Practice, 14*:493-502, 1982.
15. P. J. M. Portegijs, H. F. Kraan, J. A. Knottnerus, and H. E. J. Stoffers, Wat heet somatisatie? 1. Een zoektocht in de internationale literatuur', *Huisarts en Wetenschap, 35*:18-25, 1992.

16. D. Goldberg, Reasons for Misdiagnosis, in *Psychological Disorders in General Medical Settings,* N. Sartorius, N. Goldberg, and G. de Girolamo, et al. (eds.), Hogrefe & Huber, Toronto, 1990.
17. P. F. M. Verhaak and H. J. Wennink, What Does a Doctor do With Psychosocial Problems in Primary Care?, *International Journal of Psychiatry in Medicine, 20*: 151-162, 1990.
18. D. Goldberg, L. Gask, and T. O'Dowd, Treatment of Somatization: Teaching Techniques of Reattribution, *Journal of Psychosomatic Research, 33*:689-695, 1989.
19. A. Johnstone and D. Goldberg, Psychiatric Screening in General Practice, *Lancet, 20*:605-608, 1976.
20. J. Ormel and R. Giel, Medical Effects of Nonrecognition of Affective Disorders in Primary Care, in *Psychological Disorders in General Medical Settings,* N. Sartorius, D. Goldberg, and G. de Girolamo, et al. (eds.), Hogrefe & Huber, Toronto, 1990.
21. J. Ormel, W. van de Brink, K. van de Meer, M. W. J. Koeter, F. W. Wilmink, and B. Krol, Herkenning, behandeling en beloop van angst en depressie in de huisartspraktijk, *Tijdschrift Sociale Gezondheidszorg, 69*:31-37, 1991.
22. A. Tylee and P. Freeling, The Recognition, Diagnosis and Acknowledgment of Depressive Disorder by General Practitioners in Depression, in *Depression, An Intergrated Approach,* K. Herbert, and E. Paykel, (eds.), Heinemann, London, 1989.
23. F. J. A. Huygen, A Family Physician's Concept of Mental Health and Mental Health Services, in *Mental Health and Primary Care. Dutch and Israeli Experience,* G. J. Visser, J. M. Bensing, B. P. R. Gersons, et al. (eds.), Nivel, Utrecht, 1986.
24. R. Grol (ed.), *Huisarts en somatische fixatie. Theorie en praktijk van de preventie van somatische fixatie.* Nijmegen: Nijmeegs Universitair Huisartsen Instituut, 1981.
25. M. Foets and J. van der Velden, *Nationale Studie naar Ziekten en Verrichtingen in de Huisartsprakjijk. Basisrapport: Meetinstrumenten en procedures,* Nivel, Utrecht, 1991.
26. R. J. G. Halfens, *Locus of Control: de beheersingsoriëntatie in reltie tot ziekte- en gezondheidsgedrag.* Rijksuniversiteit Limburg, Maastricht, 1985.
27. D. P. Goldberg, *The Detection of Psychiatric Illness by Questionnaire,* Oxford University Press, Oxford and London, 1972.
28. C. M. H. Hosman, *Psychosociale problematiek en hulpzoeken,* Swets en Zeitlinger, Lisse, 1983.
29. R. Rosenthal, D. B. Rubin, D. S. P. Schubert, Notes from the Editorial Board: How Big was the Effect? *International Journal of Psychiatry in Medicine, 23*:95-98, 1993.
30. J. Cohen, *Statistical Power Analysis for the Behavioral Sciences,* Lawrence Erlbaum Associates, Hillsdale, 1988.
31. A. F. Wright, A Study of the Presentation of Somatic Symptoms in General Practice by Patients with Psychiatric Disturbance, *British Journal of General Practice, 40*: 459-463, 1990.
32. M. A. R. Tijhuis and P. F. M. Verhaak, *Psychosociale problemen in de huisartspraktijk,* Nivel, Utrecht, 1991.
33. K. Bridges, D. Goldberg, B. Evans, and T. Sharpe, Determinants of Somatization in Primary Care, *Psychological Medicine, 21*:473-483, 1991.
34. C. M. Bass (ed.), *Somatization. Physical Symptoms and Psychological Illness,* Basil Blackwell, Oxford, 1990.

35. E. I. Brilman, W. van del Brink, R. Giel, and J. Ormel, Voorkomen en langetermijnbeloop van psychische stoornissen in de huisartspraktijk, *Tijdschrift voor Psychiatrie, 34*:385-399, 1992.
36. C. V. Ford, *The Somatizing Disorders, Illness as a Way of Life,* Elsevier, New York, 1983.
37. P. S. Byrne and B. E. L. Long, *Doctors Talking to Patients,* HMSO (Her Majesty's Stationary Office), London, 1976.
38. J. M. Bensing, *Doctor-Patient Communication and the Quality of Care,* Dissertation, Rotterdam, Erasmus University, 1991.
39. P. F. M. Verhaak, Detection of Psychologic Complaints by General Practitioners, *Medical Care, 26*:1009-1020, 1988.

Direct reprint requests to:

P. F. M. Verhaak
Netherlands Institute for Primary Health Care (NIVEL)
P.O. Box 1568
3500 BN Utrecht
The Netherlands

SOCIAL SUPPORT AND STRESSFUL EVENTS IN TWO DIMENSIONS: LIFE EVENTS AND ILLNESS AS AN EVENT

M. A. R. TIJHUIS,[1]* H. D. FLAP,[2] M. FOETS[1] and P. P. GROENEWEGEN[1,2,3]

[1]NIVEL, Utrecht, The Netherlands, [2]Department of Sociology, University of Utrecht, Utrecht, The Netherlands and [3]Department of Human Geography, University of Utrecht, Utrecht, The Netherlands

Abstract—Previous research on the buffering effects of social support focused mainly on life events as stressors, and mental illness as outcome. Furthermore, the question as to why support influences illness has not been subjected to theoretical or empirical study much. In this article we develop a hypothesis on the basis of the theory of social capital. We hypothesize that specific types of social resources are more relevant to the consequences of some events than of others. We test this hypothesis in two ways: (1) by taking life events as stressor and occurrence of illness as outcome, and, which is somewhat unusual, (2) by taking illness as stressor and duration and disabilities of illness as the outcome. Analyses of a representative sample of the Dutch population ($N = 10{,}110$) reveal that receiving specific types of support does not lead to better health or less illness in cases of stress. On the contrary, people who are under stress and receive more support, also appear to report more illness, more disabilities and a longer duration. We suggest that in an open sample like ours, the disease level measured is not severe enough to assess buffer effects of social support.

Key words—health, support, life events, illness

INTRODUCTION

Support is implicated in the aetiology of and recovery from both physical and mental illness [1–7]. Two distinct ideas to explain the relationship between support and health have been put forward. First, the direct effect hypothesis argues that support enhances health and well-being irrespective of level of stress. Secondly, according to the buffering hypothesis, support exerts its effects in the presence of stress, by protecting people from the negative consequences of stress. Direct or buffering processes are established in empirical research when different concepts and types of measurement of social support are used. Direct effects tend to be found when support is measured by the degree to which a person is integrated within a social network, while buffering effects tend to be shown when support is indicated by the availability of resources that help one respond to stressful events [8–11]. In this paper we want to contribute to the study of social support as a buffer of the stress of events.

At present, there is an extensive literature on life events and their effects on mental health, and also some literature on the effects on physical health, although the results with regard to the latter are not straightforward [12, 13]. The foregoing argument can also be applied to illness itself. Becoming or being ill is also an event that requires adaptation: the length of the illness or the extent of complaints can be mediated by support. The extent to which people succeed in adapting to this type of stress can be learned from the duration of their illness and the degree of disabilities.

The central questions we will try to answer here are:

(1) When we consider stressful life events, does social support buffer their effects on becoming ill?
(2) When we consider illness as stressor, does social support buffer the level of disabilities and duration of that illness?

We expect that, in the case of life events, illness will occur less often when support is provided. In case of illness as the stressor, we expect that the illness will be of shorter duration or will be accompanied by fewer limitations in respect of daily activities when support is provided.

In investigating these hypotheses we take a somewhat different angle than what seems standard in the literature. We argue that buffering effects can only take place when events to be buffered are present. Therefore, we do not analyse the often used interaction of events and support, but instead select respondents who have met specific events. Finding a buffer effect through interaction actually means that people experiencing more events and more support suffer fewer consequences than people experiencing fewer or no events at all or less support or none at all. But we are not interested in the case of fewer or no events for our analyses here, because then there is nothing to buffer. We believe the central issue of the buffer effect is: does support act as a buffer in case of events?

A shortcoming in existing research thus far has been the lack of theoretically developed hypotheses. Why is

*Author for correspondence at: RIVM, CCM, HDE, P.O. Box 1,3720 BA Bilthoven, The Netherlands.

it that support influences illness? Possible answers to this question are sometimes suggested in a rather *ad hoc* manner [5, 14]. In this paper we want to make a contribution to the debate on buffering effects by providing some theoretical notions that might explain why some researchers find buffer effects and others do not. These notions are presented in the next section.

A THEORY OF BUFFER EFFECTS

Social support is seen as a function of personal relations [13]. Different sides to social support like types of social support, experienced or 'objective' support, positively or negatively experienced support are popular topics in the research literature. Today, the idea of social support has been thoroughly examined, but no one version of this concept has been generally accepted. Here, we propose the theory of social capital, which, we believe, makes a contribution to the empirical and theoretical discussion as it has developed.

The theory of social capital assumes that people have access to resources of the people they know [15–17]. These resources are called social or 'second order' resources, meaning that network-members control the social resources ego has access to [18]. In order to achieve certain goals (like health) people can use their personal (e.g. economic, cultural, physical) and their social resources. People with more resources, including social resources, are better able to achieve their goals [19]. People obtain social capital by investing in their social network by way of entering in new relations, or expanding and preserving old relations [16]. The assumption made is that people invest in social relations to guarantee access in the future, or to repay investments made by others in the past. Therefore, social capital is more than a simple count of all people in the network and their personal resources. Social capital is the result of:

(a) the number of people willing to provide support;
(b) the resources that can be mobilized in this indirect manner; and
(c) the extent in which these people are willing or committed to providing support.

From these assumptions it follows that access to social resources does not imply use. The network-members who control the resources may already have made great investments in ego in the past: ego does not want to become more indebted [20], or these network-members may no longer feel indebted to ego because they see no future repayment from ego [21]. Further, social resources may be goal-specific: "It takes a strong man to carry an invalid" [22]. This may be an explanation for the fact that measures of integration do not, and measures of available resources in times of stress do buffer according to the literature.

In terms of the theory of social capital, the buffer effect concerns those cases in which support is actually mobilized. According to our theory, people may use the social resources they have built up in the past for achieving momentary goals. If their well-being is threatened by stressful events (and we see the occurrence of illness also as a stressor), they can reduce the consequences by resorting to their social resources. When people have more social resources and when they are able to mobilize more of these resources, the consequences of stressful events will be less severe or of shorter duration. This buffer effect will not occur or will do so to a lesser degree if the persons to whose resources one has access to are less socially indebted. Because social resources are often goal-specific it is also possible that the support available is of no use in the specific situation arising. One might suppose that the buffer effect would also be influenced by the availability of professional health care: i.e. when professionals are available, no support from lay network-members is necessary. Since professional providers of support are more or less equally available to all in the Dutch society, we do not imagine that the buffer effect will be influenced to any degree by the presence of professional health care. Studies on the elderly and the chronically ill [22–24], as well as a study on a representative sample of the Dutch [25] give empirical support for this assumption. Moreover, for problems of everyday life there are no standard solutions, which usually are offered by professionals. Social resources are particularly adequate in case of such unpredictable events, in tasks with many contingencies which can not be easily subdivided [22] (p. 10).

Like Litwak [22], Cohen and McKay [26] and Cutrona and Russell [27], who categorized sources and types of social support according to the specific needs of individuals, we assume that specific types of social support are more relevant to the consequences of some events than of others. The characteristics of the specific events and of the types of social support have to match [5, 27]. A general distinction in types of support can be found in the work of House and Kahn [28] and Schonfeld [29]. The most frequently used categories are emotional support, practical support, informational support and social companionship. We hypothesize that emotional support is more relevant when an event causes anxiety. An event will cause more anxiety if people do not know much about it, if the event does not occur frequently, or if a specific event can not be related directly to a specific cause. There is a need for practical support when the event restricts everyday activities. And information is supposed to benefit people who face events that do require professional help, and when the causes of the events are relatively unknown. Social companionship might influence health or illness directly, but will not be useful in case of buffering the stressors because specific types of support and specific events have to match. We consider social companionship non-specific, and not meant to exchange any particular kind of support (compare the notion of instrumental

versus expressive dimensions of social support [30]). Regarding illness as an event, similar hypotheses are proposed, but now with regard to health-related support. We assume that, because events and support have to match, support related to illness is more able to buffer illness as an event than general types of support. Vice versa, general types of support are more able to buffer general types of stress (i.e. other life events than illness) than support related to health.

This theory has much in common with the social support resource theory as developed by Hobfoll [31] and with the theory of optimal matching as developed by Cutrona and Russell [27]. In contrast to these authors [27, 31], we focus on 'behavioral social support'. Schwarzer and Leppin [7] label the activation of support, although usually measured by self-reports of receivers, as behavioral support, as it reflects more concrete experience in specific situations. Perceived support, in contrast, reflects general expectations. Schwarzer and Leppin [7] refer to this last kind of support as 'cognitive'.

> Perceived support may be most important under normal, everyday circumstances where people can usually cope on their own or have to rely only to a limited degree on others' help. The general sense that one is loved and cared for by others and that these others would help once they are really needed should contribute to psychological and physical well-being. [7] (p. 102)

With the help of the theory of social capital we hypothesized that, for a buffer effect to take place, available resources must be used and match with the need raised by the stressor. Since behavioral support refers to concrete situations, we considered this measurement of support to be optimal to test our hypotheses on buffer effects.

These theoretical notions provide us with a tool to examine the relationships among stressful events, support and illness. The goal-specific nature of social resources, like support, may explain why social support has a buffering effect in some studies and not in others. We hypothesize that, for a buffering effect to take place, the characteristics of specific events and of the mobilized types of social support have to match. To test this hypothesis, we conducted multiple regression analyses on several combinations of types of events and support. The next section gives more details about our data and methods.

DATA AND METHODS

Data

Data is gathered as part of the national survey 'Morbidity and Interventions in General Practice' [32] by means of a health interview*. The total sample includes 17,047 respondents, the response rate is 77%. Respondents of all ages are selected, but we have only presented findings for respondents aged 18 and over ($N = 10{,}110$) here, because this group can be considered to have more freedom in choosing their network-members than younger people. We will first describe the variables used in analyses on life events, and then provide information on details of variables used in the analyses of illness as the stressor. Descriptive statistics of the sample are given in Table 1.

Life events in our study were measured with the help of two separate lists: life events experienced in the year before the interview [34] and social problems experienced at the moment of the interview [35]. For each list, a sum-score of the number of events or problems was constructed. We decided to use the number of events or problems instead of specific single events on the basis of a literature study by Cohen and Wills [9] in which they conclude that any single event is less likely to be health-threatening. "It is when multiple problems accumulate [. . .] that the potential for serious disorders occurs" (p. 312). At average, respondents report 1.9 life events and 1.4 social problems (see Table 1). The items on illness in the list of events were not included here, to avoid

Table 1. Descriptive statistics of the sample: percentage of respondents according to gender, age,and to education, prevalence of chronic diseases, minor complaints, psychiatric complaints,and of perceived health status, mean number of social problems, mean number of life events,mean number of days respondents could not carry out their daily routine in the last twomonths (neglecting), mean number of days respondents were confined to bed during the lasttwo weeks (confined to bed), mean number of days respondents have been slowing down inthe last two weeks (quieting) and mean number of disabilities ($N = 10{,}110$)

Percentage of respondents according to:		
Gender		
Male	48.8	
Female	51.2	
Age		
18–24 years	14.3	
25–44years	44.8	
45–64 years	27.0	
65 and above	13.9	
Education		
Primary	19.6	
Lower secondary	28.8	
Middle secondary	15.3	
Upper secondary	22.2	
Tertiary	14.1	
Chronic diseases		
0	44.9	
1 or more	55.1	
Minor complaints		
0	19.9	
1 or more	80.1	
Psychiatric complaints		
1–4 positive items	87.2	
5 or more positive items	12.8	
Perceived health status		
(Very) good	82.9	
Not good	17.1	
Mean number of (and standard deviation)		
Social problems	1.4	(2.1)
Life events	1.9	(1.8)
Days neglecting	2.8	(10.3)
Days confined to bed	0.1	(0.9)
Days quieting	1.1	(3.3)
Disabilities	1.1	(1.9)

*Over 17,000 respondents were selected from the records of 161 general practitioners. Since virtually the entire population of the Netherlands is registered with a general practitioner and since the sample is quite similar to the Dutch population [33], we may consider our sample representative.

contamination with the measurement of the dependent variable, occurrence of illness. Life events in our list range from death of a spouse to starting a new career. Social problems range from problems in the relation with a partner to problems with housing. To indicate the occurrence of illness we used three indicators: minor complaints, psychiatric complaints and experienced state of health. Minor complaints experienced during the last two weeks prior to the interview were measured by a checklist developed by Foets and Van der Velden [33]. Occurrence of illness is computed by summing all reported complaints and coding this in 0 = no complaints, 1 = 1 or more complaints, per respondent. Nearly 80% of the respondents report 1 or more complaints*. The measure for psychiatric complaints used here is the General Health Questionnaire (GHQ-30, [37]). The GHQ usually is used as a first stage screening instrument of non-psychotic psychiatric illness. Anyone reporting five or more positive items is a possible psychiatric 'case' [37, 38]. In our study 13% of all respondents reported five or more positive items. The GHQ must be viewed as measuring present mental health in relation to a 'normal' status. The state of health experienced is measured on a five-point scale [1 = very good through 5 = very bad, recoded to 0 (scores 1 and 2) and 1 (scores 3, 4 and 5)]. 17% of all respondents experience their current health status as not good.

The central variables of this study are indicators of social support. To assess the amount and kinds of social support we used the exchange method [39]. Six name-generating questions had to be answered, each referring to the provision of a specific type of social support†. Respondents were requested to list a maximum of three network-members per question, in reference to people outside the household. Several types of social support can be distinguished: health-related support (see footnote: items A and C) and support not related to health (items B, D, E and F), and within these categories emotional support (items A or B) and practical support (items C or D and E). Item F is supposed to indicate social companionship, but as mentioned earlier, we have excluded this concept for the present problem at hand. We computed sum-scores for each respondent by counting the number of times the respondent reported receiving a specific type of support. Figures for support-variables can be found in our earlier work [40].

A last indicator of support is the amount of medical information people have access to through network-members working (or educated to work) in health care. Nearly half of the respondents had such network-members available to them. We acknowledge the fact that the existence of these social relations in the network does not mean that information is received by the respondent. Although we stated in the section on theory that availability of resources does not imply use of them, and that only use of resources might buffer, we can not analyse this to the full extent, because we have no direct data available on receipt of information.

For the analyses on illness as stressor, we measured illness in terms of four indicators of which we already mentioned three (minor- and psychiatric complaints and experienced health status) above. The fourth indicator concerns chronic diseases. The presence of a chronic disease was measured by a checklist developed by Van den Bos [41]. For each of 25 diseases‡, the respondent filled in whether he or she suffered from this disease. 55.1% of all respondents report one or more chronic diseases. This figure is comparable to the results of other studies carried out by the Central Bureau of Statistics in the Netherlands, using the same list of diseases [42]. We did not relate

*Of the 42 complaints in the list, 2 were excluded for this study because they are not direct indicators of (physical or mental) health: problems at work, and family problems. The list contains the following other complaints: fever; general tiredness; general weakness; sweating problems; headache; vertigo/dizziness; throat complaints; ear pain/earache; buzzing; hearing complaints; sneezing/nasal congestion; nose bleed/epistaxis; cough; pressure/tightness attributed to the heart; palpitations/aware of heartbeat; pain attributed to the heart; nausea; vomiting; diarrhoea; heartburn; stomachache or pain; cramps; constipation; teeth/gum complaints; painful urination; incontinence; neck or shoulder complaints; hip complaints; back complaints; arm-, hand- or finger complaints; leg-, thigh, foot- or toe complaints; menstrual pain. These complaints were distinguished in eight groups with the help of the ICPC [36]: general and unspecified complaints; neurological complaints; respiratory and ear complaints; circulatory complaints; digestive complaints; urinary complaints; musculoskeletal complaints and female genital complaints. Prevalence figures range from 1 to 29%.

†Name-generating items were: 'A. Do you talk with people when you have problems with your health? B. When you have a personal problem, do you talk about this with other people? C. Did you receive any help or assistance from people, with regard to a disease in the last few months? D. Did anyone help you with jobs around the house in the last three months? E. When you are away for a longer period of time, do you ask someone to watch the house, water the plants etc.? F. Did you visit anyone last month, or did anyone visit you, or have you been going out with people?' For each item the next question was: 'Can you tell me the names of at maximum three persons?'.

‡These 25 chronic diseases are (in order of appearance on the check-list): chronic bronchitis/emphysema; asthma; hay fever; hypertension; heart complaints; cardiac failure; haemorrhoids; varices; ulcus cruris; arterio-sclerosis; backache (slipped disc, sciatica); rheumatism/arthritis/arthrosis; neoplasm/cancer; diseases of the nervous system (Parkinson's disease, multiple sclerosis, epilepsy); migraine/chronic headache; chronic gastro-intestinal disorders; diabetes; gall-bladder and liver diseases; kidney diseases; thyroid gland diseases; prostatism/prostatis; menstruation/menopause complaints; chronic skin diseases/eczema; serious consequences of an accident; hereditary handicaps. Respondents with asthma and chronic bronchitis are included together. Respondents suffering from heart complaints or cardiac failure are also included together, because these diseases belong to the same groups in terms of symptoms. The original list consisted also of an 'other' category. This category was excluded because of suspected minor reliability.

life events to the prevalence of chronic diseases, because these diseases generally precede the events and therefore such analyses would be causally incorrect. Several measures were used to indicate the duration of illness. The first is the number of days (in the last two months before the interview) respondents were not able to carry out their daily routine. This variable is labelled 'neglecting'. The second variable is the number of days (in the last two weeks) respondents reported having been 'confined to bed'. The last variable of this kind is the number of days (in the last two weeks before the interview) respondents have been slowing down: 'quieting'. In contrast to the actual duration of an illness these measures can be considered as 'social' duration, or 'social' limitations, because they indicate the consequences illness has for social life. The majority of the respondents (over 80%) reported that they did not neglect their daily routine, had been confined to bed or were 'quieting'. An adjusted version of a checklist developed by the OECD*. (OECD-16, [43]) was used to indicate the physical disabilities people meet with. Respondents could report to what extent they were able to perform certain activities themselves. Over 40% of the respondents experienced one or more disabilities. These may range from not being able to climb the stairs, or not being able to run a 100 m, to not being able to get dressed or cut food. We did not relate minor and psychiatric complaints to the disabilities people experience. These complaints are acute in contrast to the disabilities which may be lifelong. Such analyses would then be causally incorrect. Finally, the support variables employed in the analyses on illness as stressor are the same as the ones we have described above.

Methods

Like Litwak [22] (and with regard to mental illness: Cohen and McKay [26]) we assume that specific types of social support are more relevant to the consequences of some events than of others. The characteristics of the specific events and of the types of social support have to match. Therefore, we categorized our measures of stress (life events, social problems) according to the supposed buffer effects of the types of support (see Table 2). We categorized these measures of stress ourselves. This may be not the best method. We expect that the life events or social problems in our list require support not related to health. Accordingly, we discuss this categorization below.

We also categorized our measures of illness as event (chronic diseases, minor-, psychiatric complaints and experienced health status) according to the supposed buffer effects of the different types of support. Table 3 shows our expectations regarding the needs for support of people experiencing those illnesses. We expect that these types of illness generally only require support that is related to health. We further expect that all chronic diseases require information, because, since these diseases are chronical, professional help might ease the consequences for daily life.

We state clearly that, as we are only considering buffer effects here, and since events or illness are the stressors that are hypothesized to be buffered by social support, we only selected respondents who reported events or illness. We used linear as well as logistic regression analyses. For analyses regarding life events and social problems we are mainly interested in occurrence of illness: does illness not occur when life events or social problems are buffered by social support? Because our data on illness were unable to follow the normal distribution of the

Table 2. Expected need for support per life event or social problem

	Support, not related to health	
	practical	emotional
Life events:		
You moved, within town	+	0
You moved, out of town	+	+
You moved out parents house	+	+
Your child moved out	0	+
Your wife is pregnant, gave birth to a child*	+	+
Your wife had an abortion or miscarriage*	0	+
You stopped working	0	+
Other in family stopped working	0	+
You started working	+	0
Other in family started working	+	0
You started school, college, etc.	+	0
You quitted school, college, etc.	0	+
You got a degree	0	+
You/other in family failed to get a degree	0	+
Other in family married	0	+
Other in family divorced	0	+
You/other in family run away	+	+
You had a financial success	+	0
You had a financial disappointment	+	+
You suffered loss due to theft, fire, violence	+	+
Your partner died	+	+
Your child died	+	+
Your father/mother died	+	+
An important other died	0	+
You married	+	+
You divorced	+	+
You/partner had or has an affair	0	+
You were promoted	0	+
Your pet died	0	+
You/other in family was involved in a lawsuit	+	+
Social problems:		
Financial	+	+
Housing	+	+
Parents	0	+
Education	+	+
Work	+	+
Getting older	0	+
Partner	0	+
Children	0	+
Important others	0	+
Contacts in general	0	+
Sexual	0	+
Religion	0	+
Self-realization	0	+
Yourself	0	+
Future	0	+
Addiction	0	+
Loneliness	0	+
Changes in society	0	+
Neighbourhood	+	+
Spending leisure time	+	+
Life in general	0	+

*For men only.

*OECD stands for Organization for Economic Cooperation and Development.

Table 3. Expected need for support per chronic disease, group of minor complaints, psychiatric complaints, or feeling in bad health

	Health-related support practical	emotional	information
Chronic diseases:			
Arterio-sclerosis	+	+	+
Backache (slipped disc, sciatica)	+	+	+
Chronic bronchitis/emphysema, or asthma	+	+	+
Chronic gastro-intestinal disorders	+	+	+
Chronic skin diseases/eczema	0	+	+
Diabetes	0	+	+
Diseases of the nervous system (Parkinson disease, multiple sclerosis, epilepsy)	+	+	+
Gall-bladder and liver diseases (incl. bilestones)	+	+	
Haemorrhoids	0	0	+
Hay fever	0	0	+
Heart complaints, or cardiac failure	+	+	+
Hereditary handicaps	+	+	+
Hypertension	0	+	+
Kidney diseases (incl. kidney stones)	+	+	+
Menstruation/menopause complaints[a]	+	+	+
Migraine/chronic headache	+	+	+
Neoplasm/cancer (incl. leukaemia)	+	+	+
Prostatism/prostatitis[b]	+	+	+
Rheumatism/arthritis/arthrosis	+	+	+
Serious consequences of accident	+	+	+
Thyroid gland diseases	+	+	+
Ulcus cruris	+	+	+
Varices	+	+	+
Minor complaints:			
General and unspecified	+	+	+
Neurological	+	+	+
Respiratory	0	0	0
Digestive	+	+	0
Circulatory	+	+	+
Musculoskeletal	+	0	+
Urinary	0	0	+
Female genital[a]	+	+	+
Psychiatric complaints	+	+	+
Feeling in bad health	+	+	+

[a]For women only.
[b]For men only.

errors as assumed for linear regression and because we are interested in occurrence (and not number) of complaints, we conducted logistic regression analyses [44]. For the analyses on illness as event we are interested in buffer effects of support on *level* of disabilities and 'social' duration. Since the dependent variable for this type of analysis is continuous (the number of disabilities, or the number of days respondents neglected daily routine, or have been confined to bed, or have been quieting)*, we performed linear regression analyses. Because a majority of the respondents did not report disabilities or neglecting their daily tasks etc., these dependent variables had to be transformed to be able to apply linear regression analyses. Best fitting results with regard to residuals and outliers [45] were obtained by transforming the dependent variable with a square root: $\sqrt{y}+\sqrt{(y+1)}$, where y = dependent variable. These data were now more able to follow the normal distribution of the errors. All coefficients are tested for significance at a one-tail $P<0.05$ level.

Because types of support seem to overlap, we admitted only one type of support in each equation. When practical support is provided, for instance, people most often also experience emotional support. In order to distinguish these different types of support, we elicited them separately and also employed them in separate analyses.

In an attempt to take possible confounders into account, age, gender, education and the number of events or complaints were included in the analyses as control-variables. The number of events or complaints was included to indicate severity. We included the number of events or complaints only for the same category of events or complaints. All variables in regression analyses were used as continuous variables, except gender. Gender is a dummy-variable: 1 = male. Education is coded as 1 = low, 5 = high.

Of course, we also conducted bi-variate analyses. To be able to determine whether events do relate to illness

*Another option for dealing with this problem is to recode the dependent variable in a dichotomous way. Then the question is not whether social support influences the number of disabilities, days neglecting routine etc., but whether social support influences the occurrence of disabilities or neglecting routine etc. Neglecting routine, being confined to bed and 'quieting' must then also be considered as a kind of disability. Such a dichotomized dependent variable demands for logistic regression analyses. Anyway, whether we performed linear or logistic regression analyses, results were always quite the same.

in general, and whether support relates to illness in general, we computed product moment correlations for our total sample. These are shown in Appendices A and B, and will be discussed in the next sections when appropriate.

RESULTS

First, we describe results for our categorization of stress-measures according to the supposed buffer effects of the types of support. Correlations (Pearson product-moment, [46]) between events requiring specific kinds of support (not related to health) and perceiving this kind of support are given in Appendix C. Although the statistical significance of low correlations might be due to the large sample size, these correlations nevertheless indicate that our categorization could be useful. These correlations are calculated for each number of events or problems in combination with the receipt of one particular type of support. All correlations indicate that the experience of more life events or social problems is related to the experience of more support. Relationships of life events and social problems requiring emotional support with receiving emotional support are statistically significant. The relationships of life events or social problems requiring practical support and receiving practical support are not statistically significant. Because the lists of events or problems requiring emotional or practical support partially overlap, the finding that correlations with the receipt of both types of support are statistically significant does not surprise us, although we would have expected stronger correlations with the matching type of support. Further, we see that these stressors are also related to the receipt of health-related types of support. We assume that, although the list of events and problems did not incorparate items on health problems, these correlations are due to the fact that people very often also experience health problems.

Correlations between types of illness and all types of health-related support are also given in Appendix C. Experiencing more chronic diseases, more minor complaints, psychiatric complaints and feeling in bad health is generally related to receiving health-related emotional and practical support. For the availability of sources of information, the correlations with measurements of illness as stressor are mostly negative. Experiencing chronic diseases is generally related to perceiving less sources of information. Experiencing minor and psychiatric complaints, or feeling in bad health is not significantly related to receiving sources of information.

Correlations between one type of support and categories of, for instance, chronic diseases are always quite the same. This is due, of course, to the fact that the chronic diseases in each category overlap: one chronic disease often requires several kinds of support. The same is true for the categories of minor complaints.

Half of the correlations between types of support not related to health (emotional support in particular) and types of illness as stressor are also statistically significant. Suffering from a particular illness, of course, does not mean that only health-related matters are experienced. People who suffer from illness, probably also experience certain other life events or social problems. In further analyses, buffer effects of all kinds of support are estimated with regard to all kinds of stressors.

The next paragraphs in this results section concern the hypothesized buffer effects.

4.1. Social support in relation to life events

That events coincide with more health complaints can be concluded from Appendix A: the number of events or problems and the number of specific complaints are all positively related. The results of testing the hypothesis that support buffers the effects of events or problems on reported complaints are shown in Table 4. The main conclusion to be drawn is that there is just one statistically significant (negative) coefficient in accord with our hypothesis: people who reported social problems requiring emotional support (all problems in our list) and who also have resources available who might give them health-related information, appear to report feeling in bad health less often (coefficient $= -0.38$). What appears more clearly is that people experiencing complaints also receive more support, specifically health-related emotional and practical support.

4.2. Social support in relation to illness as stressor

All types of self-reported illness are positively and statistically significant related to the 'social' duration of illness and disabilities (Appendix B). Findings on the hypothesis about buffer effects of social support on the 'social' duration and disabilities of chronic diseases are in line with the conclusions above (see Table 5)*.

Health-related support has no buffer effects: people who report a longer duration and more disabilities, receive more health-related support than people who report a shorter duration or fewer disabilities. An exception here is information: people with chronic diseases report fewer disabilities in combination with the availability of sources of information (coefficients are: for chronic diseases requiring emotional support -0.15, for chronic diseases requiring practical support -0.17, for chronic diseases requiring information -0.15, all are statistically significant at a level of $P<0.001$).

In all regression equations regarding people with minor complaints, the signs of support-variables are

*To take the possible confounding effect of severity into account, we also conducted analyses for people with and without chronic comorbidity (not shown here). We found no important differences between these analyses.

Table 4. Non-standardized logistic regression coefficients for each type of support per dependent variable, for people with social problems or life events requiring emotional or practical support, analyses are controlled for gender (0 = female, 1 = male), age (continuous), education (1 = low, through 5 = high), and number of problems or events in each group of problems of events

	Occurrence of minor complaints	Psychiatric complaints	Feeling in bad health
(1) For people experiencing social problems requiring emotional support ($N = 3226$)			
Emotional support, not related to health	0.21	0.16	−0.17
Practical support, not related to health	0.15	0.02	−0.01
Emotional support, related to health	0.63***	0.32***	0.50***
Practical support, related to health	0.55	0.68***	0.92***
Information, related to health	0.01	−0.09	−0.38***
(2) For people experiencing social problems requiring practical support ($N = 651$)			
Emotional support, not related to health	0.20	0.28	0.03
Practical support, not related to health	0.15	0.08	0.15
Emotional support, related to health	0.93**	0.32	0.36
Practical support, related to health	0.39	0.36	1.42***
Information, related to health	−0.04	−0.09	−0.28
(3) For people experiencing life events requiring emotional support ($N = 2259$)			
Emotional support, not related to health	0.23*	0.57***	0.07
practical support, not related to health	0.32***	0.06	0.07
emotional support, related to health	0.93***	0.65***	0.57***
practical support, related to health	1.32***	1.01***	1.47***
information, related to health	−0.07	−0.03	−0.12
(4) *For people experiencing life events requiring practical support* ($N = 1733$)			
Emotional support, not related to health	0.40**	0.58***	−0.03
Practical support, not related to health	0.23*	−0.01	0.04
Emotional support, related to health	1.08***	0.83***	0.59***
Practical support, related to health	1.45***	0.91***	1.44***
Information, related to health	0.09	−0.16	−0.08

Statistical significance: *$P < 0.05$; **$P < 0.01$; ***$P < 0.001$, t-test.

positive, indicating that respondents reporting more days neglecting the routine, being confined to bed and quieting experience more support, health-related* or not. Results for people with psychiatric complaints are that health-related support has no buffer effect, and that the possible buffer effects of emotional support not related to health are not statistically significant (coefficients are: for number of days neglecting daily tasks −0.27, for number of days being confined to bed −0.09, for number of days quieting −0.29). That emotional support not related to health might buffer the effects of psychiatric complaints (vs the non-buffering effects of emotional support related to health) may not be strange since our measures of support clearly distinguish between health problems and personal problems. Respondents may view psychiatric complaints as personal problems and not as health problems. Results for people feeling in bad health are all in line with the results for people experiencing minor complaints: people experiencing longer duration perceive more support.

Because most of our respondents report no days of neglecting daily tasks, being confined to bed or quieting at all, and nearly 60% reports having no disabilities, we decided to perform the analyses shown in Table 5 for respondents reporting at least one day of being constrained through illness or at least one disability. This may be also another way of controlling for severity of illness to a certain extent. People who have suffered at least some consequences of their illness, as indicated by the report of a disability or a day of quieting etc., may be all more severely ill than people who have not suffered these consequences recently. In Appendix D we show the results of these analyses. They indicate that, to some extent, buffer effects might exist for this group of respondents. More coefficients of support variables show the expected buffer effect, namely 13 as compared to 3 in Table 5. Nearly half of all coefficients in Appendix D appear to have a negative sign (the 'sign' of buffering), whereas in Table 5 only 11 out of all 135 coefficients do. More in detail, the availability of sources of information appears to be related negatively to the number of days being confined to bed and the number of disabilities. Coefficients of emotional support not related to health reach statistical significance in the number of days quieting in the case of people experiencing chronic diseases requiring emotional support (−0.25), and requiring information (−0.30), and also in the case of people experiencing minor complaints requiring information (−0.22). We will discuss these results in the next section.

*We also carried out separate regression analyses for respondents experiencing minor complaints divided in those with and without paid employment. Perhaps people misinterpreted the question about neglecting their daily routine as only referring to work. There are no important differences in signs and statistical significance of coefficients for support variables between employed and unemployed people. Another possible reason why we did not find any buffer effects could be that the support measures only refer to people outside of the household. For this reason we performed all regression analyses both for people living alone and for those not living alone. Again, no differences were found with regard to buffering minor complaints.

Table 5. Non-standardized linear regression coefficients for each type of support per dependent variable for people with illness as stressor, analyses are controlled for gender (0 = female, 1 = male), age (continuous), education (1 = low, through 5 = high), and number of diseases or complaints in each group of diseases or complaints

	Number of days			
	Neglecting	Confined to bed	Quieting	Number of disabilities
(1) For people experiencing chronic diseases requiring emotional support ($N = 5042$)				
Emotional support, not related to health	0.22*	0.01	0.08	0.01
Practical support, not related to health	0.37***	0.05***	0.20***	−0.03
Emotional support, related to health	0.60***	0.06**	0.41***	0.11**
Practical support, related to health	2.83***	0.41***	1.50***	0.83***
Information, related to health	0.11	0.02	0.09	−0.15***
(2) For people experiencing chronic diseases requiring practical support ($N = 4362$)				
Emotional support, not related to health	0.21	0.01	0.10	−0.01
Practical support, not related to health	0.36***	0.06**	0.22***	−0.03
Emotional support, related to health	0.62***	0.06*	0.41***	0.13**
Practical support, related to health	2.81***	0.44***	1.50***	0.86***
Information, related to health	0.09	0.02	0.07	−0.17***
(3) For people experiencing chronic diseases requiring information ($N = 5509$)				
Emotional support, not related to health	0.17	0.01	0.08	−0.01
Practical support, not related to health	0.37***	0.05***	0.18***	−0.04
Emotional support, related to health	0.61***	0.06**	0.42***	0.11**
Practical support, related to health	2.81***	0.39***	1.43***	0.81***
Information, related to health	0.07	0.02	0.06	−0.15***
(4) For people experiencing minor complaints requiring emotional support ($N = 6264$)				
Emotional support, not related to health	0.07	0.01	0.05	—
Practical support, not related to health	0.32***	0.05***	0.18***	—
Emotional support, related to health	0.48***	0.03	0.36***	—
Practical support, related to health	2.63***	0.41***	1.29***	—
Information, related to health	0.08	0.02	0.10*	—
(5) For people experiencing minor complaints requiring practical support ($N = 7142$)				
Emotional support, not related to health	0.04	0.01	0.01	—
Practical support, not related to health	0.31***	0.04***	0.17***	—
Emotional support, related to health	0.44***	0.03	0.33***	—
Practical support, related to health	2.69***	0.40***	1.31***	—
Information, related to health	0.05	0.02	0.07	—
(6) For people experiencing minor complaints requiring information ($N = 6874$)				
Emotional support, not related to health	0.06	0.01	0.02	—
Practical support, not related to health	0.30***	0.04***	0.17***	—
Emotional support, related to health	0.44***	0.03*	0.33***	—
Practical support, related to health	2.68***	0.41***	1.32***	—
Information, related to health	0.05	0.01	0.06	—
(7) For people experiencing psychiatric complaints ($N = 1221$)				
Emotional support, not related to health	−0.27	−0.09	−0.29	—
Practical support, not related to health	0.59**	0.11*	0.38***	—
Emotional support, related to health	0.66*	0.10	0.52***	—
Practical support, related to health	3.52***	0.65***	1.68***	—
Information, related to health	0.52	0.01	0.28	—
(8) For people feeling in bad health ($N = 1693$)				
Emotional support, not related to health	0.41	0.07	0.25	—
Practical support, not related to health	0.61***	0.12***	0.39***	—
Emotional support, related to health	0.64**	0.06	0.41**	—
Practical support, related to health	2.65***	0.47***	1.42***	—
Information, related to health	0.36	0.03	0.41**	—

Statistical significance: $*P < 0.05$; $**P < 0.01$: $***P < 0.001$, t-test.

Regression analyses for people experiencing minor-, psychiatric complaints and feeling in bad health with number of disabilities are not included: these complaints are acute while disabilities may have been present for a longer period of time. Such analyses then, would be causally incorrect.

CONCLUSION AND DISCUSSION

The main result of this study is that receiving support does not coincide with better health or less illness in cases of stress. On the contrary, especially perceiving practical support related to health is associated with occurrence of illness, a greater number of disabilities and a longer duration. An explanation for these results has been hinted at earlier by Dunkel-Schetter and Wortman [47], Avis *et al.* [48] and Schwarzer and Leppin [7]. With cancer as the stressful event and symptoms and physical functioning as dependent variables, Dunkel-Schetter and Wortman [47] conclude that stress can bring about increased support that could mistakenly be viewed as a negative buffering effect even in longitudinal studies. Schwarzer and Leppin [7] call this phenomenon a 'mobilization' effect. In our case this probably means that we did not take the severity of the complaints sufficiently into account, because 'mobilization' could be caused by the fact that people who are more severely ill also need more support than those less severely ill. Studies by

Ros [49] and Glass and Maddox [50] show that buffereffects are found for severely ill people. Schonfeld [29] argues that when severely threatening life events are rare in a sample, a study does not provide an optimum context for examining stress buffering effects. In this study we have tried to take the severity of complaints into account by controlling for the number of events or complaints. We also conducted separate analyses for people who can be considered more severely ill on grounds of other measurements and those who probably are less severely ill. Finally, for illness as stressor, we also performed analyses for people who reported at least one day of being constrained because of their illness or at least one disability, assuming that these people would be more severly ill than people who did not report these consequences. Although we found an increased number of indications of buffer effects for this group, this number was still not very convincing. However, to examine the effect of severity into more depth, one needs a case-control study in which more specific data on physical and mental condition are available.

Systematic buffer effects are not found in our analyses. Buffer effects are found sometimes in the case of information available through persons working in health care: more health-related information available relates to fewer complaints, shorter duration and fewer disabilities for people suffering from chronic diseases. Nevertheless, the number of relationships indicating buffer effects is too small to be conclusive. Support not related to health does not show buffer effects systematically and this type of support does not often show statistically significant relationships at all.

Methodological objections could be made. For instance, we employ cross-sectional data to test hypotheses on a longitudinal process. Still, we hardly ever find buffer effects for support not related to health, while other researchers in this area, also using cross-sectional data, did find buffer effects, for physical illness too. One reason may be for minor complaints, that they usually are present for such a short period that support simply does not get a chance to buffer its course. One could also criticise our measurement of the length of illness behaviour. This length was limited to two months or two weeks before the interview, but this period may have continued after the interview or may have started before the two month or two weeks period. We do not have a remedy for this problem. Further, all measurements were reported by the respondents themselves. For the support measure for instance this means that only people who experienced the receipt of support report this, for people who did not experience support we have no data on support available. Then, there is the problem of matching types of events and support. We categorized events in a rather *ad hoc* manner. Future studies have to test this categorization. A final remark involves the time-frames of all variables used. Imposing a causal order on the measures used is not that easy, even more so since the support-variables each refer to different time-frames.

But, notwithstanding these remarks, the design of our study has some clear advantages over those of others. For instance, we studied life events as well as illness as stressor. Furthermore, the buffer effects of social support on the occurrence of physical illness have not been studied very often (we can only recall one earlier study: [51]). Even more advanced in this field is our investigation into the duration and accompanying disabilities of the illness with regard to illness as a stressor. This clearly fills a gap in the research literature on buffer effects. Not only is physical illness rarely studied, if it is, researchers only examine utilization of health services or recovery [2]. There are still other methodological advantages in our study worth recalling. The data are gathered for a representative sample. And most important, the idea of necessity of a match between types of support and illness for the occurrence of buffer effects, has been empirically tested in this study. As, in contradistinction to most earlier studies, we have data at our disposal on several types of illness and support, we were able to combine types of support and types of events in our analyses. Although other authors also report on matching events and support [48–49, 52–56], this line of research is rare in the international literature. Unfortunately, we have very few results worth mentioning on the validity of matching types of support and illness.

In considering the properties of our investigations here, we think that the lack of empirical results confirming the buffer effect can not be completely attributed to faults in design and measurement. The only measure of support that shows buffer effects sometimes is 'available information'. These results are open to several possible interpretations. For one, it may be that social support does not buffer the duration or disabilities of illness in an open population. The events or diseases studied might also be not severe enough for the buffer effect to function. The suggestion for future research to be concluded from this is that social support should be asked for more specifically, in relation to specific events, and that also the measure of illness, or duration and disabilities must be much more specific, and of course, longitudinal.

Another option is to withdraw the mechanism of buffer effects as we described it in terms of actual use of resources. Since health-related information is measured indirectly, as access to health professionals, we probably have measured some kind of integration instead of the receipt of support. The mechanism proposed alternatively is of a more psychological nature. This other possible path consists of direct psycho-physiological changes (like elevated blood pressure) instigated by members of the social network, that influence disease susceptibility. Less exposure to social contacts then, is expected to produce a worse physiological or psychological status that could produce occurrence of health complaints [14]. This

possible mechanism is expected to work irrespective of stress level. Previous analyses [57] indicated that other measures of integration, especially household composition, are positively related to health. Some first provisional analyses on a selection of people experiencing chronic diseases (not shown) indicate that not living alone, being member of a church, being member of a voluntary organization, reporting more friends and a higher number of support-givers overall, is related to reporting less disabilities. The causal direction of these results has to be determined by longitudinal analyses, because it is possible that these measures of integration protect health, but it is also possible that it is health status that determines the degree of social integration.

Acknowledgement—The work reported here is supported by a grant (NWO, Grant Number 500-279-202) as part of a larger study on social networks and health.

REFERENCES

1. Leavy R. L. Social support and psychological disorder: a review. *J. Commun. Psychol.* **11**, 3, 1983.
2. Wallston B. S., Alagna S. W., DeVellis B. M. and DeVellis R. F. Reviews. Social support and physical health. *Hlth Psychol.* **2**, 367, 1983.
3. Wortman C. B. and Conway T. L. The role of social support in adaptation and recovery from physical illness. In *Social Support and Health* (Edited by Cohen S. and Syme S.L.), p. 281. Academic Press, New York, 1985.
4. Ganster D. C. and Victor B. The impact of social support on mental and physical health. *Br. J. Med. Psychol.* **61**, 17, 1988.
5. Cohen S. Psychosocial models of the role of social support in the etiology of physical disease. *Hlth Psychol.* **7**, 269, 1988.
6. Cohen, S. Social supports and physical health: symptoms, health behaviours, and infectious disease. In *Life-span Developmental Psychology: Perspective on Stress and Coping* (Edited by Cummings M., Greene A. L. and Karraker K. H.), p. 213. Lawrence Erlbaum, Hillsdale/New York, 1989.
7. Schwarzer R. and Leppin A. Social support and health: a theoretical and empirical overview. *J. Soc. Personal Relationships* **8**, 99, 1991.
8. Cohen S. and Syme S. L. Issues in the study and application of social support. In *Social Support and Health* (Edited by Cohen S. and Syme S. L.), p. 3. Academic Press, New York, 1985.
9. Cohen S. and Wills T. A. Stress, social support and the buffering hypothesis. *Psychol. Bull.* **98**, 310, 1985.
10. Israel B. A. and Rounds K. A. Social networks and social support: a synthesis for health educators. *Adv. Hlth Educ. Promotion* **2**, 311, 1987.
11. Van Sonderen F. L. P. *Het meten van sociale steun* (Measuring social support). Universiteitsdrukkerij, Groningen, 1991.
12. Tausig M. Measuring life events. In *Social Support, Life Events, and Depression* (Edited by Lin N., Dean A. and Ensel W. M.), p. 71. Academic Press, Orlando, 1986.
13. House J. S., Umberson D. and Landis K. R. Structures and processes of social support. *A. Rev. Sociol.* **14**, 293, 1988.
14. Berkman L.F. The relationships of social networks and social support to morbidity and mortality. In *Social Support and Health* (Edited by Cohen S. and Syme S. L.), p. 241. Academic Press, New York, 1985.
15. Bourdieu P. Le capital social. Notes provisoires. [The social capital. Preliminary notes] *Actes Recherche Sci. Soc.* **3**, 2, 1980.
16. Flap H. *Conflict, Loyalty, and Violence*. Peter Lang, Frankfurt, 1988.
17. Coleman J. S. Social capital in the creation of human capital. *Am. J. Sociol.* **94**, supplement s95, 1988.
18. Boxman E. A. W. *Contacten en carrière. Een empirisch-theoretisch onderzoek naar de relatie tussen sociale netwerken en arbeidsmarktposities*. (Contacts and careers. An empirical-theoretical study on the relationship between social networks and labor market positions). Thesis publishers, Amsterdam, 1988.
19. Campbell K. E., Marsden P. V. and Hurlbert J. S. Social resources and socioeconomic status. *Soc. Networks* **8**, 97, 1986.
20. Joosten J., Van der Horst F. and De Witte L. *Chronische patienten en hun sociale contacten* (Chronic patients and their social contacts). Paper for the Dutch-Flemish days for Sociologists, 1986.
21. Rounds K. A. and Israel B. A. Review. Social networks and social support: living with chronic renal disease. *Patient Educ. Counseling* **7**, 227, 1985.
22. Litwak E. *Helping the Elderly. The Complementary Roles of Informal Networks and Formal Systems*. Guilford Press, New York/London, 1985.
23. Van den Brink-Muinen A. and Sixma H. *De zorg voor ouderen in Almere* (Care for the elderly in Almere). NIVEL, Utrecht, 1990.
24. Janssen M. *Personal Networks of Chronic Patients*. Datawyse, Maastricht, 1992.
25. Felling A. J. A., Fiselier A. A. M. and Van der Poel M. G. M. *Primaire relaties en sociale steun. Achtergronden van de behoefte aan steun, de aard en de omvang van informele steunverlening en daarbij opgedane ervaringen* (Primary relationships and social support. Backgrounds of the need for support, the type and amount of informal care-giving, and experiences with informal care). Instituut voor Toegepaste Sociale Wetenschappen, Nijmegen, 1992.
26. Cohen S. and McKay G. Social support, stress and the buffering hypothesis: a theoretical analysis. In *Handbook of Psychology and Health*, Vol. 4 (Edited by Baum A., Singer J. E. and Taylor S. E.), p. 253. Erlbaum, Hillsdale/New York, 1984.
27. Cutrona C. E. and Russell D. W. Type of social support and specific stress: toward a theory of optimal matching. In *Social Support: An Interactional View* (Edited by Sarason B. R., Sarason I. G. and Pierce G. R.), p. 319. Academic Press, New York, 1990.
28. House J. S. and Kahn R. L. Measures and concepts of social support. In *Social Support and Health* (Edited by Cohen S. and Syme S. L.), p. 83. Academic Press, New York, 1985.
29. Schonfeld I. S. Dimensions of functional social support and psychological symptoms. *Psychol. Med.* **21**, 1051, 1991.
30. Lin N. Conceptualizing social support. In *Social Support, Life Events, and Depression* (Edited by Lin N., Dean A. and Ensel W. M.), p. 17. Academic Press, Orlando, 1986.
31. Hobfoll S. E., Freedy J., Lane C. and Geller P. Conservation of social support resources: social support resource theory. *J. Soc. Personal Relationships* **7**, 465, 1990.
32. Foets M., Van der Velden J. and de Bakker D. De *Survey Design. National Study of Morbidity and Interventions in General Practice*. NIVEL, Utrecht, 1992.
33. Foets M. and Van der Velden J. *Een Nationale Studie van ziekten en verrichtingen in de huisartspraktijk. Basisrapport: Meetinstrumenten en procedures* (National Study of

morbidity and interventions in general practice. Basis-report: Measurements and procedures). NIVEL, Utrecht, 1990.
34. Ormel J. and Koeter M. W. J. *De Groningse gebeurtenissen lijst* (The Groningen events list). Academic Hospital, Department of Social Psychiatry, Internal publication, Groningen, 1985.
35. Furer J. W. and Tax B. *Somatische klachten, psychiatrische symptomen en psychosociale problemen. Prevalentie, overlappingen en verbanden met demografische variabelen* (Somatic complaints, psychiatric symptoms and psychosocial problems. Prevalences, overlap en relations with demografic variables). Eindrapport van het Regioprojekt Nijmegen, deel 1. Instituut voor Sociale Geneeskunde, Katholieke Universiteit Nijmegen. ISG-publikatie nr. 17, 1987.
36. Lamberts H. and Woods M. *International Classification of Primary Care*. Oxford University Press, Oxford, 1987.
37. Goldberg D. *The Detection of Psychiatric Illness by Questionnaire*. Maudsley Monographs No. 21. Oxford University Press, London, 1972.
38. Tarnopolsky A., Hand D. J., McLean E. K., Roberts H. and Wiggins R. D. Validity and uses of a screening questionnaire (GHQ) in the community. *Br. J. Psychiat.* **134,** 508, 1979.
39. Fischer C. S. *To Dwell among Friends: Personal Networks in Town and City*. Chicago University Press, Chicago, 1982.
40. Tijhuis M. A. R., Flap H. D., Foets M. and Groenewegen P. P. Netwerken in Nederland. Een onderzoek naar persoonlijke netwerken van Nederlanders (Networks in the Netherlands. A study of personal networks of the Dutch). *Mens en Maatschappij* **67,** 5, 1992.
41. Van den Bos G. A. M. *Zorgen van en voor chronisch zieken* (Care for the chronically ill). Bohn, Scheltema and Holkema, Utrecht, 1989.
42. Van den Berg J. and Van den Bos G. A. M. Het (meten van het) voorkomen van chronische aandoeningen, 1974–1987 [The (measurement of the) prevalence of chronic conditions, 1974–1987]. *Maandbericht Gezondheid* **3,** 4, 1989.
43. Van Sonsbeek J. L. A. Methodische en inhoudelijke aspecten van de OESO-indicator betreffende langdurige beperkingen in het lichamelijk functioneren (Methodological and substantial aspects of the OECD-indicator concerning long-lasting limitations in bodily functioning). *Maandbericht Gezondheid* **6,** 4, 1988.
44. Hosmer D. W. and Lemeshow S. *Applied Logistic Regression*. John Wiley and Son, New York, 1989.
45. Weisberg S. *Applied Linear Regression*. John Wiley and Son, New York, 1985.
46. Hays W. L. *Statistics*. Holt, Rinehar and Winston, New York, 1981.
47. Dunkel-Schetter C. and Wortman C. B. The interpersonal dynamics of cancer: problems in social relationships and their impact on the patient. In *Interpersonal Issues in Health Care* (Edited by Friedman H.S. and DiMatteo M.R.), p. 69. Academic Press, New York, 1982.
48. Avis N. E., Brambilla D. J., Vass K. and McKinlay J. B. The effect of widowhood on health: a prospective analysis from the Massachusetts women's health study. *Soc. Sci. Med.* **33,** 1063, 1991.
49. Ros W. J. G. *Sociale steun bij kankerpatienten* (Social support in cancer patients). Thesis Publishers, Amsterdam, 1990.
50. Glass T. A. and Maddox G. L. The quality and quantity of social support: stroke recovery as psycho-social transition. *Soc. Sci. Med.* **34,** 1249, 1992.
51. Lin N. and Ensel W. M. Life stress and health: stressors and resources. *Am. Sociol. Rev.* **54,** 382, 1989.
52. Schaefer C., Coyne J. C. and Lazarus R. S. The health-related functions of social support. *J. Behav. Med.* **4,** 381, 1982.
53. Cohen S., Mermelstein R., Kamarck T. and Hoberman H. M. Measuring the functional components of social support. In *Social Support: Theory, Research and Applications* (Edited by Sarason I. G. and Sarason B. R.), p. 73. Martinus Nijhoff Publishers, Dordrecht, 1985.
54. Dooley D., Rook K. and Catalano R. Job and non-job stressors and their moderators. *J. Occup. Psychol.* **60,** 115, 1987.
55. Ensminger M. E. and Celentano D. D. Unemployment and psychiatric distress: social resources and coping. *Soc. Sci. Med.* **27,** 239, 1988.
56. Buehler C. and Legg B. H. Mothers' receipt of social support and their psychological well-being following marital separation. *J. Soc. Personal Relationships* **10,** 21, 1993.
57. Tijhuis M. A. R., Flap H. D., Foets M. and Groenewegen P. P. Kenmerken van sociale relaties en gezondheid (Characteristics of social relations and health). *Tijdschrift voor Sociale Gezondheidszorg* **7,** 341, 1994.

APPENDIX A

Pearson's product moment correlations for stressful events, types of support, minor complaints, psychiatric complaints and feeling in bad health ($N = 9662$–$10{,}110$)

	Minor complaints	Psychiatric complaints	Feeling in bad health
Emotional support, not related to health	0.08**	0.10**	−0.09**
Practical support, not related to health	0.05**	0.04**	−0.02*
Emotional support, related to health	0.19***	0.13***	0.09**
Practical support, related to health	0.19***	0.17***	0.16***
Information, related to health	0.01	0.01	−0.11**
Social problems requiring emotional support	0.42***	0.53***	0.21***
Social problems requiring practical support	0.26***	0.33***	0.12***
Life events requiring emotional support	0.11**	0.13***	0.01
Life events requiring practical support	0.11**	0.12**	−0.01

Statistical significance: *$P<0.05$; **$P<0.01$; ***$P<0.001$.
Correlations between life events and chronic diseases are not included, because these diseases generally precede the events, and therefore these analyses would be causally incorrect.

APPENDIX B

Pearson's product moment correlations for illness as stressor, types of support, and consequences of illness ($N = 9106$–$10{,}110$)

	Number of days			
	Neglecting	Confined to bed	Quieting	Number disabilities
Emotional support, not related to health	−0.01	0.01	0.01	−0.12**
Practical support, not related to health	0.04**	0.03**	0.05**	−0.03**
Emotional support, related to health	0.08**	0.04**	0.10***	−0.03**
Practical support, related to health	0.20***	0.16***	0.21***	0.22***
Information, related to health	−0.01	−0.01	−0.01	−0.15***
Chronic diseases requiring emotional support	0.17***	0.11***	0.18***	0.41***
Chronic diseases requiring practical support	0.17***	0.12***	0.19***	0.39***
Chronic diseases requiring information	0.16***	0.10***	0.17***	0.38***
Minor complaints requiring emotional support	0.12***	0.14***	0.21***	—
Minor complaints requiring practical support	0.15***	0.14***	0.26***	—
Minor complaints requiring information	0.15***	0.13***	0.24***	—
Psychiatric complaints	0.22***	0.15***	0.26***	—
Feeling in bad health	0.23***	0.13***	0.25***	—

Statistical significance: $*P<0.05$; $**P<0.01$; $***P<0.001$.

Correlations for minor-, psychiatric complaints and feeling in bad health with number of disabilities are not included because these complaints are acute in contrast to the disabilities which may have been present for a longer period of time. These analyses would be causally incorrect.

APPENDIX C

Correlations between receipt of support and expected need for support

	Support, not related to health		Health-related support			
	Practical	Emotional	Practical	Emotional	Information	N
Number of:						
Social problems requiring emotional support	0.01	0.08*	0.07**	0.16***	0.04***	3257
Social problems requiring practical support	−0.02	0.01	−0.01	0.09**	0.02*	661
Life events requiring emotional support	0.03	0.06**	0.02	0.07**	0.08**	2283
Life events requiring practical support	0.03	0.05	0.01	0.10**	0.08**	1748
Chronic diseases requiring emotional support	−0.02	−0.05**	0.13***	0.01	−0.07**	5099
Chronic diseases requiring practical support	−0.01	−0.04*	0.13***	0.02	−0.05**	4412
Chronic diseases requiring information	−0.01	−0.05**	0.12***	0.03*	−0.06**	5570
Minor complaints requiring emotional support	0.01	0.06***	0.13***	0.13***	0.01	6335
Minor complaints requiring practical support	0.02	0.06**	0.14***	0.14***	0.01	7220
Minor complaints requiring information	0.03*	0.05**	0.15***	0.13***	−0.01	6948
Psychiatric complaints	−0.02	0.02	0.08**	−0.01	−0.04	1240
Experienced health status (1 = very good, 5 = very bad)	−0.01	−0.05*	0.09**	−0.01	−0.05	1720

Statistical significance: $*P<0.05$; $**P<0.01$; $***P<0.001$.

APPENDIX D

Non-standardized linear regression coefficients for each type of support per dependent variable for people with illness as stressor and at least one day of neglecting, being confined to bed, or quieting, or at least one disability, analyses are controlled for gender (0 = female, 1 = male), age (continuous), education (1 = low, through 5 = high), and number of diseases or complaints in each group of diseases or complaints

	Number of days			
	Neglecting	Confined to bed	Quieting	Number of disabilities
(1) For people experiencing chronic diseases requiring emotional support				
N =	977	256	984	2544
Emotional support, not related to health	−0.11	−0.11	−0.25*	−0.01
Practical support, not related to health	0.30	−0.04	0.01	0.05
Emotional support, related to health	0.55	0.13	−0.02	0.09*
Practical support, related to health	1.30***	0.68***	0.78***	0.77***
Information, related to health	−0.38	−0.48*	−0.21	−0.11*
(2) For people experiencing chronic diseases requiring practical support				
N =	882	239	904	2241
Emotional support, not related to health	−0.05	−0.17	−0.25	−0.02
Practical support, not related to health	0.31	−0.04	0.06	0.06
Emotional support, related to health	0.52	0.11	−0.07	0.09*
Practical support, related to health	1.41***	0.63**	0.74***	0.76***
Information, related to health	−0.46	−0.50*	−0.22	−0.13**
(3) For people experiencing chronic diseases requiring information				
N =	1062	271	1042	2638
Emotional support, not related to health	−0.16	−0.15	−0.30*	−0.02
Practical support, not related to health	0.28	−0.01	−0.02	0.04
Emotional support, related to health	0.62*	0.11	−0.03	0.07
Practical support, related to health	1.31*	0.64***	0.68***	0.76***
Information, related to health	−0.43	−0.45*	−0.20	−0.12*
(4) For people experiencing minor complaints requiring emotional support				
N =	1318	343	1228	—
Emotional support, not related to health	−0.26	−0.07	−0.16	—
Practical support, not related to health	0.24	0.07	0.07	—
Emotional support, related to health	0.68**	0.08	0.15	—
Practical support, related to health	1.49***	0.77***	0.77***	—
Information, related to health	−0.45	−0.46**	−0.10	—
(5) For people experiencing minor complaints requiring practical support				
N =	1471	360	1349	—
Emotional support, not related to health	−0.22	−0.09	−0.20	—
Practical support, not related to health	0.20	0.04	0.05	—
Emotional support, related to health	0.64**	0.09	0.11	—
Practical support, related to health	1.56***	0.73***	0.80***	—
Information, related to health	−0.43	−0.37*	−0.14	—
(6) For people experiencing minor complaints requiring information				
N =	1430	350	1323	—
Emotional support, not related to health	−0.21	−0.08	−0.22*	—
Practical support, not related to health	0.21	0.06	0.06	—
Emotional support, related to health	0.62**	0.10	0.09	—
Practical support, related to health	1.56***	0.74***	0.79***	—
Information, related to health	−0.46*	−0.40*	−0.17	—
(7) For people experiencing psychiatric complaints				
N =	390	121	415	—
Emotional support, not related to health	−0.81	0.05	−0.19	—
Practical support, not related to health	−0.07	−0.07	0.07	—
Emotional support, related to health	−0.31	−0.18	−0.32	—
Practical support, related to health	1.19*	0.76*	0.62**	-
Information, related to health	−0.28	−0.10	0.09	—
(8) For people feeling in bad health				
N =	488	139	550	—
Emotional support, not related to health	−0.08	0.02	−0.08	—
Practical support, not related to health	0.11	0.05	−0.08	—
Emotional support, related to health	0.23	0.30	−0.08	—
Practical support, related to health	0.39	0.40	0.37*	—
Information, related to health	0.12	−0.52	−0.16	—

Statistical significance: *$P<0.05$; **$P<0.01$; ***$P<0.001$, t-test.

Regression analyses for people experiencing minor-, psychiatric complaints and feeling in bad health with number of disabilities are not included because these complaints are acute in contrast to the disabilities which may have been present for a longer period of time. These analyses would be causally incorrect.

Determinants of the help-seeking process: Goldberg and Huxley's first level and first filter

PETER F. M. VERHAAK[1]

From the Netherlands Institute of Primary Health Care, Utrecht, The Netherlands

SYNOPSIS This paper discusses minor psychiatric morbidity in the community and its relation to help-seeking. The research is aimed at identifying the demographic and social characteristics that enhance the likelihood of minor psychiatric morbidity, as measured by the General Health Questionnaire, and to reveal how these determinants relate to health status. In the second stage of the study, the same characteristics are related to their effect on the decision to consult a general practitioner and, more specifically, on presenting psycho-social problems to the GP. At the population level, the likelihood of a high GHQ score was greater for women, divorced persons and the unemployed, other variables remaining equal. However, GHQ score was contaminated by poor subjective health perception, especially for unemployed persons. Chronic physical illness did not have an independent effect on the GHQ score. Chronic physical illness did influence help-seeking. In addition, several socio-demographic characteristics showed an independent effect on consulting behaviour for both GHQ-positive and GHQ-negative patients. Women, unemployed persons and publicly insured patients with a minor psychiatric disturbance were more apt to visit their GP, regardless of their physical health status. However, no socio-demographic characteristics increased the likelihood that a GHQ-positive patient would present unambiguous psycho-social complaints as a reason to see their doctor.

INTRODUCTION

The filter-model of Goldberg & Huxley was originally designed to describe 'the pathway to psychiatric care'. It depicts psychiatric morbidity encountered among patients on several levels (population, primary care, psychiatric settings). Moreover, it identifies points of decision, the so-called filters between the different levels: which psychologically disordered individuals will seek care; which ones will have their disturbances detected; who will be treated in a primary care setting and who will be referred for psychiatric care? (Goldberg & Huxley, 1980).

The model, however, might also be understood as a research programme. In this respect, much epidemiological research has been done already at the several levels and in different countries, e.g. the UK, USA, Australia, Spain and the Netherlands (the ECA-studies; Goldberg *et al.* 1976; Finlay-Jones & Burvill, 1977; Marks *et al.* 1979; Goldberg & Bridges, 1987; Hodiamont *et al.* 1987; Vázquez-Barquero *et al.* 1987; Regier *et al.* 1988, 1993; Wilmink, 1989; Bensing & Beerendonk, 1990; Ormel *et al.* 1990, 1991; Verhaak *et al.* 1990; Robins & Regier, 1991; Verhaak & Tijhuis, 1992; Verhaak, 1993).

In regard to the points of decision (the filters in the model) much attention has been directed to the second filter: recognition of mental illness by the general practitioner (Marks *et al.* 1979; Ormel *et al.* 1990; Sartorius *et al.*, 1990; Cooper & Eastwood, 1992). However, concerning the decision of the disturbed patient to seek care, a paucity of research literature on psychiatric morbidity in the community and its relation to help-seeking behaviour was reported as late as 1990 (Giel *et al.* 1990). In another review, Williams *et al.* (1990) discussed the weak theoretical foundation of research in this area; the authors recommended taking recourse to general models of health service utilization, as developed by Andersen & Newman (1973).

One of the few studies carried out according to this recommendation is the secondary analysis

[1] Address for correspondence: Dr Peter F. M. Verhaak, Netherlands Institute of Primary Health Care, PO Box 1568, 3500 BN Utrecht, The Netherlands.

by Williams *et al.* (1986) of the West London survey. They determined the joint effects of minor psychiatric morbidity (estimated by the GHQ-30), socio-demographic variables (gender, age, marital status, socio-economic group (SEG) and employment status), and health status (self-assessment, chronic illness and social impairment) on general practitioner consultation. Each of the health related variables (including GHQ score) had an independent effect; socio-demographic variables had no independent effects. The effect of minor psychiatric morbidity did not vary according to sex, marital status or another socio-demographic variable.

Vásquez-Barquero and his colleagues (Vázquez-Barquero *et al.* 1990) described minor psychiatric morbidity (measured by the GHQ-60) in the population. Psychiatric morbidity was higher for females than for males and increased with age. Previously married men had a higher rate of minor psychiatric morbidity than currently married or single men. Unemployed men had higher rates than employed men, as did poorly educated women compared to highly educated women. These investigators also examined the effect of psychiatric morbidity on the probability of being in contact with primary care physicians and the socio-demographic variables which influenced that effect. They found independent effects of a positive GHQ score and a number of socio-demographic variables (age, sex, marital status for females and employment status for men), but no interaction between GHQ status and socio-demographic variable, could be demonstrated. A more recent publication (Vázquez-Barquero *et al.* 1992), calls attention to the strong association between physical illness and medical consultation for people with probable minor psychiatric morbidity.

These investigations suggest a relatively high probability that patients with a minor psychiatric disorder (as expressed by the GHQ) will pass Goldberg & Huxley's first filter. Furthermore, a number of socio-demographic variables are related to the likelihood of a high GHQ score as well as that of consulting a general practitioner. In the latter case, however, those variables disappear as independent factors when minor psychiatric disorder and health status are considered together.

A number of issues remain unsolved, one of these is the possible interrelationship of GHQ score and health status. As we recently found, patients with an elevated GHQ score felt less healthy, experienced more acute symptoms during the previous 2 weeks, reported more chronic diseases and had a higher level of absenteeism than patients with a low GHQ score. Also, they reported a higher level of medical attention over the previous year, mentioning not only an increased use of mental health services, but also of medical specialists, physiotherapists, and alternative medicine (Bensing & Verhaak, 1994). This is in line with the results of Vázquez-Barquero *et al.* (1992). One conclusion is that ill health is a consequence of mental distress; alternatively the GHQ may be sensitive to mental distress as well as to ill health. In the latter case, it is imperative to investigate whether ill health acts as a confounder in the relation between GHQ and other determinants.

The second issue to be discussed in this paper is the interrelation between health status, socio-demographic factors and the likelihood of consulting a GP. Vázquez-Barquero *et al.* (1990) do not consider health status in their study. The independent effects of socio-demographic variables that they found for GHQ-positive (GHQ+) as well as GHQ-negative (GHQ−) cases, might be attributed for example to a poorer healthy status of women or older patients. They did incorporate health status in an exploratory analysis in their recent study Vázquez-Barquero *et al.* (1992) and they found a sizeable effect of serious physical illness. Williams *et al.* (1986) also considered this interrelationship, but looked at the undivided sample, and treated the GHQ score as variable in their sample. In contrast to this approach, we followed GHQ+ as well as GHQ− cases, and have analysed each group for the influence of health status and socio-demographic characteristics on the decision to consult a GP.

A final objective of this paper is to specify the concept 'help-seeking behaviour'. In the research mentioned above, help-seeking was defined as visiting a doctor, or more specifically, a GP. Help-seeking behaviour was measured by asking the respondents if they had seen a doctor/GP during the last... weeks. The data provided here allow us to use a more specific measure. We have a 3 month registration of

doctor–patient contacts at our disposal. From these records we can determine with certainty whether there have been one or more contacts with the GP and whether the patient has presented any psycho-social symptoms during one of those contacts. Of course, this measure does not cover the whole concept of 'help-seeking behaviour'. A patient might express his distress in physical symptoms. Nevertheless, it might be a better approximation of help-seeking than the patient's response to direct questioning.

To elucidate the interrelationship mentioned above, we first considered the socio-demographic determinants of minor psychiatric morbidity, as expressed by the GHQ score. By adding indicators of health status subsequently, we analysed whether the GHQ measures 'health' as well as 'minor psychiatric morbidity'.

Secondly we considered the use of health services by two separate groups of patients: those with either high or low GHQ scores. Following the model of Andersen & Newman (1973), we used three groups of explaining variables: predisposing variables; enabling variables; and illness level. Thirdly we analysed the presentation of psychological and social symptoms by those two groups of patients. This third analysis was only carried out on the subsample of patients who actually visited their GP.

In this study we have asked the following questions.

1 Which demographic and social characteristics are related to minor psychiatric morbidity in the population, as measured by the GHQ?

2 In what way are these relationships modified by perceived health status?

3 Which demographic, social and enabling characteristics are related to help-seeking by patients with either an elevated or a low GHQ score?

4 How does perceived health status interfere in this respect?

5 Which demographic, social and enabling characteristics are related to the presentation of psycho-social symptoms by patients with an elevated GHQ score who visit their GP?

METHOD

Data collection

The data were derived from the Dutch National Study of Morbidity and Interventions in General Practice. During this nationwide survey, 161 GPs registered all their doctor–patient contacts over a period of 3 months. Participating practices were randomly selected according to a stratification procedure that guaranteed a sufficient number of practices from each region and each level of urbanization. For a detailed description of this morbidity survey, see Foets *et al.* (1982).

General practitioners in The Netherlands have fixed lists, as is the case in the UK. For each participating GP, approximately 100 patients were selected from these lists at random. In total 17342 patients were approached; with 13340 of them a 2 h health interview was held, resulting in a positive response rate of nearly 77%. The main reasons for non-participation were patients having moved, patients not found at home on several occasions, and refusal. There is a slight over-representation among the respondents of the 25–44 age group, compared with the Dutch population as a whole. However, the sample can be considered representative of the Dutch population with respect to age and sex. Interviews took place during the second month of the morbidity registration.

Variables used in the analysis

This National Survey was designed to provide data for an extensive array of health services research questions. For the aim of this present study, only a small number of variables gathered in the survey have been used. The health interviews provide excellent data to answer the first two research questions. The first dependent variable considered is 'minor psychiatric morbidity'. The score on the General Health Questionnaire (30-items-version) was used as an indicator of mental illness. The GHQ was in fact developed as a first-stage screening instrument. Taking a psychiatric clinical instrument as a reference standard, specificities between 0·74 and 0·86 and sensitivities between 0·72 and 1·00 have been established (Goldberg, 1985). The usual cut-off point between 4 and 5 'positive' items was used to discriminate between 'cases' and 'non-cases'.

Table 1. *Proportion of respondents scoring high on the GHQ, according to socio-demographic characteristics*

	N (10305)*	GHQ+ (*N* = 1293) (%)	χ^2	*P*
Age (years)				
15–24	2005	10·8	9·67	< 0·05
25–44	4342	12·8		
45–64	2605	13·6		
65–75	912	11·6		
⩾ 75	62	14·1		
Sex				
Male	5092	9·3	98·56	< 0·0001
Female	5213	15·8		
Marital status				
Married	6469	11·5	82·63	< 0·0001
Divorced	370	25·6		
Widowed	606	18·2		
Never married	2833	12·0		
Employment status				
Homemaker	1404	16·1	73·16	< 0·0001
Unemployed	227	18·5		
Incapacitated	317	24·0		
Retired	1296	12·3		
Working/student/ military service	7061	11·1		
Education				
Primary	5510	13·0	2·80	NS
Secondary	3265	11·8		
Higher	1053	12·2		
Insurance				
Public	6924	13·4	13·64	< 0·001
Private	3310	10·8		

* For most variables *N* will not add up to 10305, due to missing values.

The following predisposing, enabling and illness variables (Andersen & Newman, 1973) were assessed in the course of the interview. Predisposing/demographic variables were: age, sex, marital status (married/co-habiting, divorced, widowed, never married). Predisposing/social variables were: employment status (house-wife/husband, unemployed, incapacitated, retired, working (including students and the military); and education level, primary (primary school, technical and vocational training for 12–16 years), secondary (technical and vocational training for 16–18 years, secondary schools), and higher (technical and vocational training for ⩾ 18 years, university). Enabling variables were: health insurance, private/public. Illness variables were: general health feeling (a 5-point scale, stating 'my health status in general is: very good/good/not good/not bad/bad/very bad); number of acute symptoms and complaints during the previous 2 weeks (the respondent had to mark each of the symptoms from a list of 45 which had given trouble during the previous 2 weeks); number of chronic conditions (the respondent had to mark each chronic condition from which he/she suffered from a list of 28).

To answer the third, fourth and fifth research questions, a link has to be made between minor psychiatric morbidity and perceived health status assessed during the health interview and consulting behaviour as registered during the morbidity registration respectively. Consulting behaviour was defined operationally as a patient having had at least one consultation or home visit with a general practitioner during the 3 month registration period.

All reasons for a health visit initiated by the patient in the contacts registered during this registration period, were recorded in the words of the patient and coded according to the International Classification of Primary Care. If any of these reasons for visit of a patient was classified within the chapter Psyche or Social, the patient was considered as presenting psychosocial symptoms.

As the health interview took place during the second month of registration, only registration data from this second month have been used. In this way, assessment of GHQ status and perceived health took place within a few weeks of the index consultation.

Analysis

To answer the first two questions descriptive statistics were first computed for each of the questions. The influence of the socio-demographic variables and health status on GHQ score was then evaluated by logistic regression. First, the influence of all socio-demographic variables was estimated by step-forward selection. Subsequently health status variables were added to the significant socio-demographic determinants. SPSS routine logistic regression was used.

To control any interaction effect between the significant socio-demographic determinants and health status, the following model was fitted:

$$\ln(p/[1-p]) = \text{constant} + \text{socdem} + \text{hlth} + \text{socdem*hlth},$$

where p is the proportion of GHQ cases re-

spectively of people who visit their general practitioner, socdem is the socio-demographic variable under consideration and hlth is health status of the patient, as assessed by the patient.

To answer questions three and four, the same procedure was followed, for patients with a high and a low GHQ score separately. For patients with a high GHQ who actually visited their GP, logistic regression was carried out with socio-demographic factors as independent variables. This should supply us with the answer to question five.

RESULTS

Approximately 12·5% of the population at large are likely to be mentally ill, according to the GHQ. This probability increases gradually with age, with a decline for the 65–74 year age group. About twice as many women as men have high GHQ scores. People suffering a loss, either through divorce or widowhood, are clearly apt to have a high GHQ score, as are the unemployed and the incapacitated. GHQ cases are fairly evenly distributed among the different levels of education. More publicly than privately insured people are found among the GHQ 'cases'.

Logistic regression with GHQ 'caseness' as the dependent variable and the socio-demographic measures as independent variables resulted in the effects depicted in Table 2 (first three columns). The relationships that seem to be the strongest in the bivariate cross-tabulations also remain the most important ones in the multivariate analysis. Other variables being equal, it is women, divorced and widowed people, the unemployed and the incapacitated who are more at risk of being mentally ill, as indicated by the GHQ.

The logistic regression analysis was repeated, this time including the health indicators mentioned above, in order to assess the degree of confounding caused by this relationship. The results are shown in the last three columns of Table 2. Two out of three health indicators have a significant effect on GHQ 'caseness': patients with a subjective feeling of ill health and patients who had experienced acute complaints during the last 2 weeks have a higher chance of a positive GHQ score. Reported chronic (physical) illness does not have a significant effect.

Furthermore, this analysis indicates weaker effects of the determinants that were previously strong: being a woman or being divorced. These effects are still significant, however. The effects of being unemployed or incapacitated disappear when general health status is held constant. However, negative effects of being married or having retired appear. In other words, the chances that these groups would have a high GHQ score decrease when general health status is held constant.

There were no significant interactions between health status on the one hand and sex, employment status or marital status on the other.

With regard to research questions 1 and 2, we can conclude that sex and marital status exert effects on GHQ status, independent of the effect of health status. Persons who have lost their job,

Table 2. *Determinants of GHQ 'caseness', before and after adding health indicators*

	Before adding health			After adding health			
Determinant	*B*	*P*	Odds	*B*	*P*	Odds	In respect of
Marital status							Never married
Married	−0·10	NS	0·91	−0·14	NS	0·87	
Divorced	0·76	0·000	2·15	0·50	0·001	1·64	
Widowed	0·35	0·009	1·42	0·12	NS	1·13	
Sex							Female
Male	−0·59	0·000	0·56	−0·31	0·000	0·73	
Employment status							Employed
Unemployed	0·56	0·002	1·74	0·34	NS	1·40	
Incapacitated	1·03	0·000	2·79	0·06	NS	1·06	
Retired	−0·07	NS	0·93	−0·42	0·000	0·66	
General health feeling	—	—	—	0·55	0·000	1·73	
N of acute complaints	—	—	—	0·55	0·000	1·73	

Table 3. *Socio-demographic characteristics of people scoring high and low on the GHQ, visiting or not visiting a doctor*

	GHQ+				GHQ−			
	Visit+ (N = 352) (%)	Visit− (N = 941) (%)	χ^2	P	Visit+ (N = 1400) (%)	Visit− (N = 7612) (%)	χ^2	P
Age (years)								
15–24	14	18	14·18	> 0·01	16	20	60·15	< 0·0001
25–44	40	44			36	43		
45–64	28	27			29	24		
65–75	10	7			12	8		
⩾ 75	8	4			6	4		
Sex								
Male	29	39	12·73	< 0·001	45	52	25·45	< 0·0001
Female	71	61			55	48		
Marital status								
Married	59	57	13·92	< 0·01	66	63	25·39	< 0·0001
Divorced	8	7			3	3		
Widowed	12	7			8	5		
Never married	21	29			24	28		
Employment status								
Homemaker	23	15	50·59	< 0·0001	16	13	129·97	< 0·0001
Unemployed	3	3			2	2		
Incapacitated	8	5			4	3		
Retired	19	10			20	11		
Working/student/ military service	46	67			58	72		
Education								
Primary	66	56	10·07	< 0·01	61	55	19·91	< 0·0001
Secondary	26	33			31	34		
Higher	8	11			8	11		
Insurance								
Public	80	69	13·39	< 0·001	70	66	8·33	< 0·01
Private	20	31			30	34		

be it because of ill health or other reasons, are more likely to have a high GHQ score. But this risk is plausible, as these persons are also less healthy in general.

Table 3 shows the relationship between visits to the doctor (by GHQ 'cases' and 'non-cases') and socio-demographic variables and health status. People who had not visited their general practitioner in a 3-month period are more likely to be young, male, never married, privately insured, working, and better educated than those who did visit their GP in the same time period. These factors are the same for GHQ+ as for GHQ− cases, although GHQ+ cases in general have a higher consultation rate than GHQ−.

In the case of a probable mental illness (GHQ+), logistic regression demonstrated that when the effects of the socio-demographic variables are taken in conjunction, gender, employment status and insurance have the strongest independent effects (see Table 4). Variables indicating health status were added to the model, as they were for the prediction of GHQ 'caseness' in Table 2. Only one of the indicators – the number of chronic illnesses – had a small additive effect, but the βs or odds ratios were not substantially changed. There were no interactions between sex, employment status and insurance sector. We therefore conclude that for GHQ 'cases', the chances of visiting a doctor are higher for the publicly insured, for females and for people without regular work, regardless of their health status.

The same analysis was carried out for people who were not GHQ cases. The same socio-demographic variables were selected by step-forward selection, but all health status indicators were accepted in the second stage (see Table 5). For persons who, on account of their GHQ scores, are not considered to be possible 'cases' of mental illness, health variables, especially the general feeling of health are of greater weight

Table 4. *Determinants of visiting a doctor given GHQ 'caseness', before and after adding health indicators*

Determinant	Before adding health			After adding health			In respect of
	B	*P*	Odds	*B*	*P*	Odds	
Insurance							Public
Private	−0·44	0·005	0·64	−0·40	0·011	0·67	
Sex							Female
Male	−0·33	0·026	0·72	−0·32	0·031	0·72	
Employment status							Employed
Homemaker	0·61	0·000	1·86	0·51	0·004	1·66	
Unemployed	0·16	NS	1·17	0·14	NS	1·16	
Incapacitated	0·89	0·000	2·43	0·71	0·007	2·04	
Retired	1·00	0·000	2·70	0·79	0·000	2·20	
N of chronic conditions	—	—	—	0·16	0·000	1·17	

Table 5. *Socio-demographic determinants of visiting a doctor, given non-GHQ 'caseness' (health status added)*

Determinant	*B*	*P*	Odds	In respect of
Sex				Female
Male	−0·16	0·012	0·85	
Employment status				Employed
Homemaker	0·33	0·000	1·39	
Unemployed	0·19	NS	1·10	
Incapacitated	0·25	0·000	1·29	
Retired	0·55	0·000	1·79	
N of acute complaints	0·05	0·009	1·05	
N of chronic conditions	0·09	0·000	1·10	
General health feeling	0·33	0·000	1·38	

and socio-demographic differences are less important than they are for patients who are considered to be 'cases'.

Logistic regression on the presentation of psycho-social symptoms or complaints, among patients with minor mental disturbance (GHQ+) who actually visited their GP in the 3-month period, with all socio-demographic variables as possible determinants, did not reveal any significant effect.

DISCUSSION

Approximately 13·0% of the population may be called 'mentally disturbed' according to their scores on the GHQ. Women, divorced persons, widows and widowers, and those who are unemployed, either as a result of illness or for other reasons, are more likely to have a high GHQ score. The initial high risk of patients with compulsory insurance or a lower level of education disappears after other socio-demographic variables have been taken into account. The high GHQ score of these groups seems partly to be a result of perceived inferior health, since the effects of the socio-demographic variables decrease after general health indicators have been added. It must be stressed that only health indicators with a strong subjective component (health feeling, number of acute symptoms experienced) appeared to be related to a high GHQ score. The presence of chronic physical illness was not related to the GHQ score. The GHQ score seems to be more closely related to the feeling of illness than to the presence of disease.

The likelihood of visiting a doctor is higher for women, unemployed people and those with compulsory insurance. The effect of these socio-demographic variables is stronger for persons who are considered to be mentally ill according to their GHQ scores. Furthermore, health status does less to explain why a doctor was consulted in the case of a high GHQ. The only health status variable with a significant effect is the 'hard' operationalization, that is to say the presence of chronic disease. Incidentally, this is the health status variable which is least contaminated with the GHQ. In other words, in the event of mental illness, a contact with the doctor depends more upon socio-demographic variables than upon subjective health status. It may be possible that, given the relationship between subjective health status and GHQ, the GHQ+ subgroup is too homogeneous with respect to

subjective health status to permit any further effect of this variable on help-seeking behaviour. In the GHQ− subgroup, the subjective perception of general health status has a greater impact. Although some socio-demographic variables seem to affect the decision to visit a general practitioner, they have no further impact on the presentation of the psycho-social problems in the event of a consultation.

The prevalence of mental illness as measured by the GHQ is less than that reported by most other population studies using similar instruments. Finlay-Jones & Burvill (1977) reported 16·3% of their respondents scoring above the threshold level; Vázquez-Barquero *et al.* (1990) reported 18·5%; and Hodiamont *et al.* (1987) found 22·7%. Only Goldberg *et al.* (1976) arrived at a comparable 11·3%. One explanation for this may be that the GHQ was administered unobtrusively among many other questionnaires in the course of a comprehensive interview. Apparently the threshold at which our respondents could get a positive GHQ score was relatively high. This may have had consequences for the severity of selected cases.

Another methodological problem discussed here is the probable time span of a few weeks between the health interview and the index consultation. In order to reduce this time span to a minimum we have restricted ourselves to consultations during the second month of registration. This approach yields the same results as an analysis in which the whole registration period was taken into account. Limiting the possible time span between health interview and index consultation to 2 weeks makes no difference for the GHQ group (Table 5), but it does leave us with too few cases in the GHQ+ group. As a result, in that analysis the effect of 'sex' as a determinant for a doctor's visit cannot be replicated. For the other factors, the results remain exactly as they were.

Most of the relationships found, however, are compatible with those described by Vázquez-Barquero *et al.* (1990) and Williams *et al.* (1986): women and divorced or unemployed people are more at risk of mental illness. Results from former studies with respect to age are inconclusive: Finlay-Jones & Burvill (1977) reported a peak for men between 30 and 39 years; Williams *et al.* (1986) noted a peak for men younger than 30 years; and Myers *et al.* (1984) saw a sharp decline in DIS cases from 45 years onward; Vázquez-Barquero *et al.* (1992) detected an increase for older people. Our data appear to support the latter finding, although age effects disappeared after considering other socio-demographic variables jointly.

As we expected, however, a high score on the GHQ was contaminated by poor health, especially by the patient's subjective self-assessment. As subjective health measures had a higher impact than chronic physical disease, we might conclude that GHQ is contaminated less with disease but more with subjective feeling of illness. The result that divorced people and women remain at a relatively high GHQ level after taking health perception into account suggests that these categories at least are in any event susceptible to mental illness. In addition, GHQ 'cases' had a relatively high medical consumption, regardless of health status. This result tends to support the use of the GHQ as an instrument that does not merely indicate health status. Likewise, Vázquez-Barquero *et al.* (1992), reported a strong effect of physical illness on consulting behaviour of GHQ-positive patients and they defined physical illness as 'non-transient disorder of a serious nature: the majority of those affected had a *chronic*... disorder' (p. 501). In other words, their measurement was comparable to our 'chronic disease', which indeed exerted an effect on consulting behaviour.

In the terminology of Andersen & Newman (1973), the type of insurance acts as an enabling variable and determines part of a person's medical consumption regardless of health status. This effect is particularly so for individuals with a minor psychiatric disturbance. Gender and employment status are clearly predisposing variables, that operate in a way that is comparable to health insurance. Of course this mechanism, especially the role of insurance, is restricted to the Dutch health-care system. Its effects cannot be extrapolated indiscriminately to other health-care systems. Nevertheless, the discrepancy with the results of Williams *et al.* (1986) (they did not find any independent effect of socio-demographic variables after holding physical and mental health status constant) cannot be explained by different health care systems alone. The only real difference between our approach and theirs is that we fitted the

model for GHQ− and GHQ+ groups separately. However, when we tried to fit their model to our data (using insurance as a replacement for their variable 'socio-economic group', and without their health-related variable 'social impairment'), we found comparable effects of GHQ status, subjective health feeling and chronic illness, and also independent effects for gender and employment status. The interaction effects reported by Williams *et al.* (1986) of gender with marital status, employment status and health feeling, and a three-way interaction of gender, insurance and GHQ status could not be replicated.

On the evidence presented, women and unemployed or disabled people combine a high risk of mental illness with a high tendency (in the case of mental illness) to consult a doctor. Divorced and widowed people, on the other hand, are also at higher risk. But if they do become mentally ill, the chance of their seeing a doctor is no better than average. In this respect, they form that segment of the population which most deserves extra attention.

A general practitioner should generally be aware of the higher risks of problems among people who are deprived in one way or another: the unemployed, widowed and divorced people, and people from lower socio-economic classes. He/she should also be aware of the possible attribution of the general feeling of misery to physical dysfunction. Perhaps their affliction is not strictly a mental health problem but a difficult life. A possible interpretation of the contamination of GHQ and ill health might be that the complaints of GHQ+ people do not indicate their objective physical status but rather their urge to complain in general. And given most of the risks associated with high GHQ – unemployment, divorce, ill health, low socio-economic status – they may indeed have good reason to complain.

REFERENCES

Anderson, R. & Newman, J. (1973). Societal and individual determinants of medical care utilization in the United States. *Millbank Memorial Fund Quarterly* **51**, 95–124.

Bensing, J. M. & Beerendonk, P. (1990). Psychosociale problemen in de huisartspraktijk. *Maandblad Geestelijke Volksgezondheid* **45**, 595–618.

Bensing, J. M. & Verhaak, P. F. M. (1994). Goldbergs pyramide vanuit een generalistisch perspectief (Goldberg's pyramid from a generalistic point of view). *Nederlands Tijdschrift voor Geneeskunde* **138**, 130–135.

Cooper, B. & Eastwood, R. (eds.) (1992). *Primary Health Care and Psychiatric Epidemiology*. Tavistock/Routledge: London.

Dohrenwend, B. P. (1990). The problem of validity in field studies of psychological disorders revisited. *Psychological Medicine* **20**, 195–208.

Finlay-Jones, R. A. & Burvill, P. W. (1977). The prevalence of minor psychiatric morbidity in the community. *Psychological Medicine* **7**, 475–489.

Foets, M., Velden, J. v.d. & Bakker, D. de (1992). *Dutch National Survey of General Practice, Survey Design*. NIVEL: Utrecht.

Giel, R., Koeter, M. W. J. & Ormel, J. (1990). Detection and referral of primary care patients with mental health problems: the second and third filter. In *The Public Health Impact of Mental Disorder* (ed. D. Goldberg and D. Tantam), pp. 25–34. Hogrefe & Huber: Toronto.

Goldberg, D. P. (1985). Identifying psychiatric illness among general medical patients. *British Medical Journal* **291**, 161–162.

Goldberg, D. & Bridges, K. (1987). Screening for psychiatric illness in general practice: the general practitioner *versus* the screening questionnaire. *Journal of the Royal College of General Practitioners* **37**, 15–18.

Goldberg, D. & Huxley, P. (1980). *Mental Illness in the Community. The Pathway to Psychiatric Care*. Tavistock Publications: London.

Goldberg, D., Kay, C. & Thompson, L. (1976). Psychiatric morbidity in general practice and the community. *Psychological Medicine* **6**, 565–569.

Hodiamont, P., Peer, N. & Syben, N. (1987). Epidemiological aspects of psychiatric disorder in a Dutch health area. *Psychological Medicine* **17**, 495–505.

Marks, J. N., Goldberg, D. P. & Hillier, V. F. (1979). Determinants of the ability of general practitioners to detect psychiatric illness. *Psychological Medicine* **9**, 337–353.

Myers, J. K., Wasserman, M. M., Tischler, G. L., Holzer, C. E., Leaf, P. J., Orvaschel, H., Anthony, J. C., Boyd, J. H., Burke, J. D., Kramer, M., & Stoltzman, R. (1984). Six-month prevalence of psychiatric disorders in three communities. *Archives of General Psychiatry* **41**, 959–967.

Ormel, J., Brink, W. vd., Koeter, M. W. J., Giel, R., Meer, K. vd., Willige, G. vd. & Wilmink, F. W. (1990). Recognition, management and outcome of psychological disorders in primary care: a naturalistic follow-up study. *Psychological Medicine* **20**, 900–923.

Ormel, J., Koeter, M. W. J., Brink, W. vd. & Willige, G. vd. (1991). Recognition, management and course of anxiety and depression in general practice. *Archives of General Psychiatry* **48**, 700–706.

Regier, D. A., Boyd, J. H., Burke, J. D., Rae, D. S., Meyers, J. K., Kramer, M., Robins, L. N., Goerge, L. K., Karno, M. & Locke, B. Z. (1988). One-month prevalence of mental disorders in the United States: based on five Epidemiologic Catchment Areas sites. *Archives of General Psychiatry* **45**, 977–986.

Regier, D. A., Narrow, W. E., Rae, D. S., Mandeerscheid, R. W., Locke, B. Z. & Goodwin, F. K. (1983). The de factor US mental and addictive disorders service system. *Archives of General Psychiatry* **50**, 85–94.

Robins, L. N. & Regier, D. A. (eds.) (1991). *Psychiatric Disorders in America*. Free Press: New York.

Sartorius, N., Goldberg, D., Girolamo, G., Costa e Silva, J., Lecrubier, Y., & Wittchen, U. (eds.) (1990). *Psychological Disorders in General Medical Settings*. Hogrefe & Huber: Toronto.

Vázquez-Barquero, J. L., Diez Manrique, J. F., Pena, C., Aldama, J., Samaniego, C., Menendez Arango, J. & Mirapeix, C. (1987). A community mental health survey in Cantabria. *Psychological Medicine* **17**, 227–241.

Vázquez-Barquero, J. L., Wilkinson, G., Williams, P., Diez-Manrique, J. F. & Pena, C. (1990). Mental health and medical consultation in primary care settings. *Psychological Medicine* **20**, 681–694.

Vázquez-Barquero, J. L., Diez Manrique, J. F., Gaite, L., Iglesias Garcia, C., Artal, J., Roberts, S. E. & Wilkinson, G. (1992). Why

people with probable minor psychiatric morbidity consult a doctor. *Psychological Medicine* **22**, 495–502.

Verhaak, P. F. M. (1993). Analysis of referrals of mental health problems by general practitioners. *British Journal of General Practice* **43**, 203–208.

Verhaak, P. F. M., Wennink, H. J. & Tijhuis, M. A. R. (1990). The importance of the GHQ in general practice. *Family Practice* **7**, 319–324.

Verhaak, P. F. M. & Tijhuis, M. A. R. (1992). Psychosocial problems in primary care: some results from the Dutch National Study of Morbidity and Interventions in General Practice. *Social Science and Medicine* **35**, 105–110.

Williams, P., Tarnapolsky, A., Hand, D. & Shepherd, M. (1986). *Minor Psychiatric Morbidity and General Practice Consultations: the West London Survey. Psychological Medicine Monograph Supplement 9*, Cambridge University Press: Cambridge.

Williams, P., Wilkinson, G. & Arreghini, E. (1990). The determinants of help-seeking for psychological disorders in primary care settings. In *Psychological Disorders in General Medical Settings* (ed. N. Sartorius, D. Goldberg, G. de Girolamo, J. A. Costa e Silva, Y. Lecrubier and H.-U. Wittchen), pp. 1–37. Hogrefe & Huber: Toronto.

Wilmink, F. W. (1989). *Patient, Physician, Psychiatrist. Assessment of Mental Health Problems in Primary Care*. Thesis: Groningen.

Musculoskeletal disorders in children: a study in Dutch general practice

Alexander S. Vijlbrief[1], Marc A. Bruijnzeels[1], Johannes C. van der Wouden[1] and Lisette W.A. van Suijlekom-Smit[2]

[1] Department of General Practice, Erasmus University, Rotterdam, [2] Department of Pediatrics, Erasmus University and University Hospital/Sophia Children's Hospital, Rotterdam, The Netherlands.

Vijlbrief AS, Bruijnzeels MA, van der Wouden JC, van Suijlekom-Smit LWA. Musculoskeletal disorders in children: a study in Dutch general practice. Scand J Prim Health Care 1995;13:105–11.

Objective **– This study focuses on the occurrence of musculoskeletal disorders in children presented in general practice. Known epidemiological studies addressing musculoskeletal diseases in childhood are scarce and based on a low number of episodes.**

Design and setting **– Prospective study of all patient contacts in general practice. A total number of 161 general practitioners participated, divided into four groups, registering during four consecutive three-month periods.**

Patients **– All children younger than 15 years of age who visited their GP during the registration period. All diagnoses and working hypotheses concerning musculoskeletal disorders were selected.**

Results **– The total number of children in the study was 64 198. Disorders of the musculoskeletal system accounted for 3 699 (7.5%) of all 49 309 contacts and for 3 046 (7.5%) of all 40 340 episodes. Of the 3 046 episodes registered for ICPC-chapter L (musculoskeletal), 2 562 (84%) were new episodes, i.e. not presented to the GP before.**

Fifty-four percent of all new episodes were acute injuries. In 22% of the new episodes the general practitioner made a symptom diagnosis. Differences by age and sex were found for a limited number of diagnosis categories.

Conclusion **– Children present disorders of the musculoskeletal system less often than adults; they also present different disorders to their general practitioners. The majority of disorders presented by children are acute injuries, mostly sprains and strains.**

Key words: **children, musculoskeletal disorders, general practice, The Netherlands.**

JC van der Wouden, MA, PhD, Department of General Practice, Erasmus University, Room Fe319, PO Box 1738, 3000 DR Rotterdam, The Netherlands.

Musculoskeletal disorders in childhood are either congenital in origin or subsequently acquired (1). The acquired disorders can be subdivided in traumatic and non-traumatic disorders. Within the group of traumatic disorders we distinguish acute and overuse injuries. Infections, tumours, and acquired deformities are examples of acquired non-traumatic disorders.

The aim of this study was to gain insight into the occurrence of musculoskeletal disorders in children presented in Dutch general practice.

Several studies show that both the occurrence of and medical consumption for musculoskeletal disorders are high. The Danish Health and Morbidity Survey of 1986-87 demonstrates a high prevalence of musculoskeletal disorders in the population over 15 years of age (2). The Dutch general practitioner (GP) is confronted with musculoskeletal disorders frequently (3). Lamberts (4) reports that musculoskeletal disorders account for 15% of all episodes (all ages). A Finnish study shows a higher increase in hospital admissions and outpatient consultations for musculoskeletal disorders (including trauma) than for other conditions during the period from 1970 to 1985 (5).

Studies of musculoskeletal morbidity in children are usually restricted to specific disorders (1,6). Data relating to the total morbidity of musculoskeletal diseases in childhood are scarce and based on a relatively low number of episodes (4,7).

In this study we address three questions:

(1) which musculoskeletal disorders do children present in general practice?
(2) what is their part in the total morbidity in children?
(3) are there age and/or sex specific differences?

Patients and methods

From 1 April 1987 to 31 March 1988, the Dutch National Survey of Morbidity and Interventions in General Practice was performed by the NIVEL (Netherlands Institute of Primary Health Care) (8). For this survey 103 general practices (161 GPs) recorded all contacts with patients during one of the four consecutive periods of three months (contact registration). Data were registered for each contact concerning reason(s) for the consultation, working hypothesis, and differential diagnosis. The reason(s) given for the consultation and the differential diagnosis were written down literally by the GP and subsequently coded by a medically trained clerk of the NIVEL using the International Classification of Primary Care (ICPC). The version used by NIVEL (9) differs slightly from the original classification (10). With respect to chapter L (musculoskeletal disorders) the NIVEL version of the ICPC contains more diagnosis codes. In the original ICPC, diagnoses are often clustered into one diagnosis code whereas the NIVEL version allows extra scores for separate diagnoses. For example the original version has only one code for congenital disorders (L 82), where the modified version has separate codes for congenital dislocation of the hip (L 82.1) and other congenital disorders (L 82.9). Another important difference is that in the NIVEL version a diagnosis code can be assigned to a particular joint/part of the musculoskeletal system. This enabled us to discriminate a distorsion/contusion at the shoulder (L 79.1) from one at the wrist (L 79.2).

This National Survey includes an episode-oriented registration of morbidity. Census data, such as birth date and sex, were recorded for the whole population under study (patient registration).

For our study we have used the data of all children younger than 15 years of age in the National Survey. All diagnosis codes within the ICPC-chapter L (musculoskeletal disorders) were grouped into eight diagnostic categories (Table I). The diagnostic category Other Diagnoses contains a heterogeneous group of diagnoses that cannot be classified elsewhere. We distinguished a separate diagnostic category Symptom Diagnoses for the symptom codes. In making such a diagnosis the GP confines himself to describing the symptom, not further defining the underlying disorder.

First, all contacts were classified into the various diagnostic categories. Next we studied new episodes of illness. A new episode was defined as an episode of which the first contact took place within the registration period. In order to assign the new episodes into the various diagnostic categories we used the diagnosis or working hypothesis of the last registered contact of the episode (episode diagnosis). We calculated incidence rates dividing the number of new episodes by the time the children were followed. Because of the heterogeneity we did not calculate incidence rates for the diagnostic categories Other Diagnoses and Symptom Diagnoses.

For the Congenital Disorders and Congenital Hip Dislocation we determined the prevalence for infants under the age of one year and the sex distribution.

Table I. Diagnostic categories ICPC-chapter L

DIAGNOSTIC CATEGORY	DIAGNOSES
Acute Injuries	sprains/strains, whiplash of calf muscles, fractures, dislocations and (sub)luxations
Overuse Injuries	shoulder syndromes, bursitis, tendinitis, synovitis, chronic knee problems and osteochondroses
Infections	Bornholm disease, osteomyelitis, osteïtis, abscess and infections NEC*
Tumours	benign, malign, not further specified
Acquired Deformities	spine deformities, flat feet, hallux valgus, genua vara or vara and limbs NEC*
Other Diagnoses	arthrosis, osteoporosis, low back pain with radiating symptoms, ganglion, rheumatoïd arthritis and allied conditions, musculoskeletal syndromes and diseases NEC*
Symptom Diagnoses	symptoms/complaints of musculoskeletal system not further specified, including muscle pain, fear of cancer and disability.

*NEC not elsewhere classified

Results

The number of children under the age of 15 in the Dutch National Survey amounted to 64 198. Of these children 27 462 had a total of 49 309 contacts with their GPs during the registration period. Disorders of the musculoskeletal system accounted for 3 699 of these contacts (7.5%).

Table II shows the number of episodes for the various ICPC-chapters. Disorders of the musculoskeletal system accounted for 3 046 (7.5%) of all 40 340 registered episodes.

Of the 3 046 episodes registered for ICPC-chapter L (musculoskeletal) 2 562 were new episodes (84%). On average a new episode consisted of 1.2 contacts. Table III lists the number of new episodes and the calculated incidence rates for the diagnostic categories of the musculoskeletal disorders. For acute injuries 1 392 new episodes (54.3% of all new episodes) were registered. In 558 new episodes (22%) the GP made a symptom diagnosis.

Table II. Top 10 ICPC-chapters (episodes)

ICPC-chapter	n	%
R – respiratory	10930	27.1
A – general	7286	18.0
S – skin	6689	16.6
D – digestive	3465	8.6
H – hearing	3441	8.5
L – musculoskeletal	3046	7.5
F – eye	1477	3.7
P – psychological	645	1.9
U – urology	771	1.6
N – neurological	573	1.4

Table III. Incidence rates per diagnostic category of musculoskeletal disorders in children (per 1000 person-years)

DIAGNOSTIC CATEGORY	Number	%	Incidence rate
Acute Injuries	1392	54.3	82.0
Symptom Diagnoses	558	21.7	–
Overuse Injuries	205	8.0	12.7
Acquired Deformities	202	7.9	12.7
Other Diagnoses	133	5.3	–
Congenital Disorders	59	2.3	0.1*
Infections	10	0.4	0.6
Tumours	3	0.1	0.1
TOTAL	2562	100.0	

*prevalence (%)

The 15 diagnosis codes with the highest incidence rate are listed in Table IV. These codes accounted for 1 681 of all 2 562 new episodes (66%). Of these 15 diagnosis codes, nine concerned acute injuries, three acquired deformities, two symptom diagnoses, and one overuse injury.

The age and sex specific incidence rates for Acute Injuries are shown in Figure 1. We calculated an incidence rate of 77.6 per 1 000 person years for boys, 87.1 per 1 000 person years for girls, and 82.0 per 1 000 person years for the total group. The age and sex specific incidence rates for Overuse Injuries and Acquired Deformities are shown in Figures 2 and 3, respectively. The incidence rates of Overuse Injuries for boys and

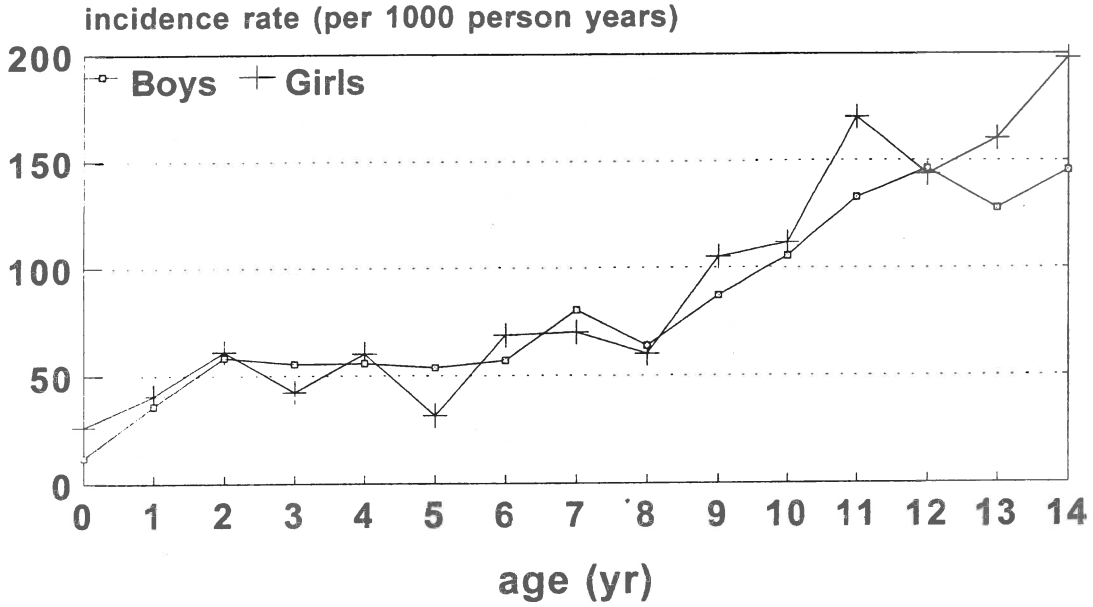


Figure 1. Age-specific incidence for acute injuries (n = 1392)

Table IV. Top 15 of incidence rates for separate diagnosis codes (per 1 000 person-years)

DIAGNOSIS (ICPC-CODE)	Episodes	Incidence rate
Sprain/strain wrist/hand/finger (L29.2)	276	17.5
Sprain/strain ankle (L77)	203	12.9
Sprain/strain foot/toe (L79.3)	193	12.2
Myalgia/muscle cramps (L18)	187	11.7
Sprain/strain musculoskeletal system NEC (L79.9)	157	9.9
Sprain/strain knee (L78)	124	7.9
Sprain/strain shoulder/arm/elbow (L79.1)	108	6.8
Tendinitis/synovitis (L93.9)	84	5.3
Disability/impairment (L28)	67	4.2
Flat feet (L98.1)	64	4.1
Fracture radius/ulna (L72)	62	3.9
Fracture (meta)carpal/(meta)tarsal bone (L74)	59	3.7
Acquired deformities limbs (L98.9)	56	3.5
Fracture NEC (L76.9)	51	3.2
Acquired deformities spine (L85)	45	2.9

girls were both 12.8 per 1 000 person years. The incidence rates of Acquired Deformities were 12.3 per 1 000 person years for boys, 13.3 per 1 000 person years for girls, and 12.7 per 1 000 person years for the total group.

There were 59 (new) episodes registered for Congenital Disorders, of which 33 (56%) referred to Congenital Hip Dislocation. In 29 cases the child was under one year of age. We calculated a prevalence of 0.7% for children under one year of age (4 451 children were under one year of age). The sex distribution within this group was 12 boys and 17 girls.

Discussion

The incidence and prevalence rates we calculated are based on data collected in general practice and not on the whole population. Hence, our data give an insight into the musculoskeletal disorders that children present to their GP. Considering the iceberg of illness (11), and due to the fact that after a trauma Dutch patients go to hospital directly, we expect the calculated incidence rates to be underestimations of the occurrence in the population.

This study indicates that, compared with adults, musculoskeletal disorders account for a smaller proportion of total morbidity in children.

In about one in every five consultations and

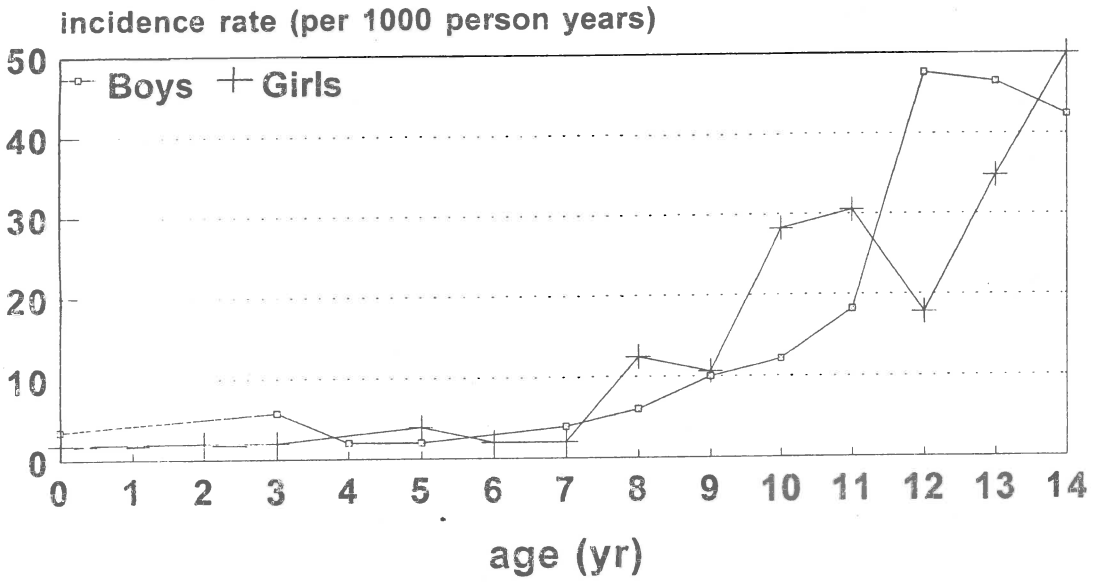


Figure 2. Age-specific incidence of overuse injuries (n = 205)

episodes concerning a musculoskeletal problem the GP confines himself to a symptom diagnosis. We think, as proposed by Lamberts (4), that this relatively high percentage is the result of the low specificity of musculoskeletal symptoms.

Since the numbers are small our findings concerning tumours and infections of the musculoskeletal system must be interpreted carefully. It is also important to realize that a GP can only suspect such a diagnosis. Nevertheless our data show that a GP rarely sees a child with a tumour or an infection of the musculoskeletal system.

Of all musculoskeletal disorders, the GP is confronted mostly with sprains and strains in children. Sprains/strains of the wrist/hand/finger have the highest incidence rate, followed by sprains/strains of the ankle. The latter is the form most frequently seen in adults (3,4). As in the study of Landin (12) we found that fracture of the forearm is the fracture most often seen in children. Since acute injuries usually result from accidents it is not surprising that the age distribution resembles the age distribution for accidents in children (13,14). The higher incidence rate in our study for acute injuries in girls is remarkable since accidents occur more often in boys (13). The fact that we limited ourselves to acute injuries of the musculoskeletal system and did not take all accidents into account may explain the sexdifference.

In this study overuse injuries were rarely seen in children under the age of eight years, but above this age their occurrence increased rapidly. Like Kannus et al. (15) we did not find a sexdifference for this age group. Other studies state that boys are more active at sports than girls (16,17). The equal sex incidence rates for overuse injuries in boys and girls may result from a greater vulnerability in girls for these kinds of injury (15,17).

Of all congenital disorders, the GP most often sees congenital dislocation of the hip. The prevalence in children under the age of one year in our study was lower than that found in other Dutch studies (18,19). This probably results from the fact that the GP does not see the children in whom the congenital hip dislocation is detected directly after birth in the hospital. These studies also demonstrate a female preponderance for dislocation of the hip.

Not only do children present disorders of the musculoskeletal system less often than adults, they also present different disorders to their GP. The majority of musculoskeletal disorders presented by children are acute injuries, mostly sprains and strains.

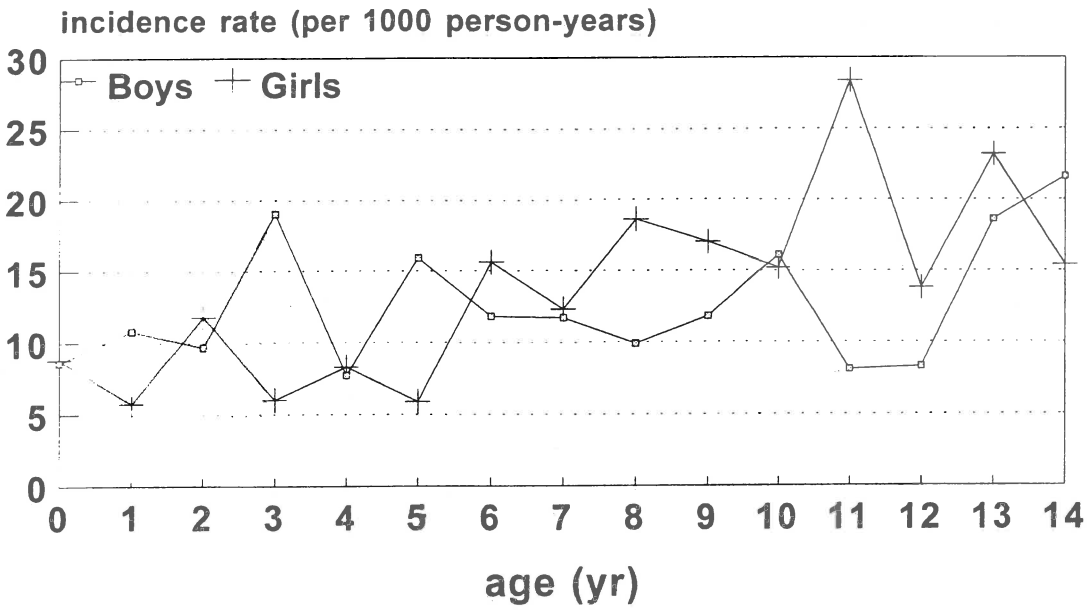


Figure 3. Age-specific incidence for acquired deformities (n = 202)

Acknowledgement

The authors thank J. van der Velden of the Netherlands' Institute of Primary Health Care.

References

1. Silman AJ. Musculoskeletal disorders in childhood. Br Med Bull 1986;42:196–9.
2. Bredkjær SR. Musculoskeletal disease in Denmark. Acta Orthop Scand 1991;62 (Suppl 241):10–12.
3. Van de Lisdonk EH, Van den Bosch WJHM, Huygen FJA, Lagro-Jansen ALM eds. Ziekten in de huisartspraktijk (Diseases in general practice). Utrecht: Bunge, 1990.
4. Lamberts H. In het huis van de huisarts. (In the house of the general practitioner). Lelystad: Meditekst, 1991.
5. Slätis P, Ruusinen A. Orthopedic diseases and trauma in Finland: trends in consumption of health services 1970–1985. Acta Orthop Scand 1991;62 (Suppl 241): 13–16.
6. Fleuren MAH. Aandoeningen van het bewegingsapparaat bij kinderen van 6–15 jaar. (Musculoskeletal disorders in 6–15 yr old children). Amsterdam: Vrije Universiteit, 1990.
7. Van den Hoogen HJM, Huygen FJA, Schellekens JWG, Straat JM, Van der Velden HGM eds. Morbidity figures from General Practice. Nijmegen: Nijmegen University Department of General Practice, 1985.
8. Foets M, Van der Velden J, Bakker DH. Dutch National Survey in General Practice. A summary of the survey design. Utrecht: NIVEL, 1992.
9. Van der Velden J, Schellevis F, Van der Steen J. International Classification of Primary Care. Utrecht: NIVEL, 1989.
10. Lamberts H, Wood M (eds). International classification of primary care. Oxford: Oxford University Press, 1987.
11. Last JM. The iceberg: "Completing the clinical picture" in general practice. The Lancet 1963;ii:28–31.
12. Landin LA. Fracture patterns in children. Acta Orthop Scand 1983;202 (suppl):4–89.
13. Fife D, Barancik JI, Chatterjee BF. Northeastern Ohio Trauma Study: II. Injury rates by age, sex and cause. AJHP 1984;74:473–8.
14. Gallagher SS, Finison JI, Guyer B, Goodenough S. The incidence of injuries among 87 000 Massachusetts children and adolescents. AJHP 1984;74:1340–7.
15. Kannus P, Niittymäki S, Jarvinen M. Athletic overuse injuries in children. Clin Ped 1988;27:333–7.
16. Zaricznyj B, Shattuck LJ, Mast TA, Robertson RV, D'elia G. Sports-related injuries in school-aged children. Am J Sports Med 1980;8:318–24.
17. Kvist M, Kujala UM, Heinonen OJ, Vuort IV, Pajulo

O, Hintsa A, et al. Sports-related injuries in children. Int J Sports Med 1989;10:81–6.

18. Burger BJ, Burger JD, Bos CFA, Obermann WR, Rozing PM, Vandenbroucke JP. Neonatal screening and staggered early treatment for congenital dislocation or dysplasia of the hip. Lancet 1990;336:1549–53.

19. Van Hees-van der Laan ZJ, Huttinga-Edens MM. Congenitale dysplasie van het heupgewricht bij zuigelingen (Congenital dysplasia of the hip in infants. English summary). Ned Tijdsch Geneesk 1981;125:1913–7.

Received July 1994
Accepted November 1994

EXPLORING THE ICEBERG OF MORBIDITY: A COMPARISON OF DIFFERENT SURVEY METHODS FOR ASSESSING THE OCCURRENCE OF EVERYDAY ILLNESS*

S. E. KOOIKER

The Netherlands Institute of Primary Health Care (NIVEL), P.O. Box 1568, 3500 BN Utrecht, The Netherlands

Abstract—This paper examines problems in measuring the occurrence of acute symptoms of ill health. Health interview surveys and health diaries often lead to different results. Two key hypotheses assume that: 1, interviews using checklists are more sensitive to the respondent's psychological distress than are the open-ended questions of health diaries; and 2, health diaries demand high levels of compliance leading to underreporting of symptoms. An additional 3rd hypothesis assumes that the effect of psychological distress on response patterns is strong for reporting psychological symptoms but insignificant for musculoskeletal symptoms.

The hypotheses were tested and explored with data from the Dutch Survey of General Practice, a nationwide study among 161 GPs. A random sample of 100 patients per GP was approached for a health interview and asked to keep a structured health diary during three weeks. Symptoms were recorded during the interview with a checklist and queried in the health diary with open-ended questions. The occurrence of symptoms was modelled with logistic regression.

High levels of psychological distress increase the likelihood of recording symptoms for both instruments, but the increase is greater for the interviews. Respondents who have only received limited education, heavy smokers and those who suffer from chronic conditions have a significantly lower likelihood of recording symptoms in the diary as compared to the questionnaire. There was no significant effect of taking an interest in health matters, gender, and work and domestic role obligations. Taking the nature of symptoms into account, it was found that psychological distress had indeed a great effect on the response pattern for psychological symptoms, but not for musculoskeletal symptoms.

The criticism that symptom checklists are sensitive to psychological distress rather than to physical illness alone, is confirmed in this study. Open-ended questions prevent biased responses, but result in fewer symptoms being recorded. Health diaries with open-ended questions 'produce' more symptoms but take more effort to complete, requiring sufficiently motivated respondents. It is recommended that a less biased specific list for the assessment of acute symptoms be developed.

Key words—health interview survey, health diary, symptom measurement

INTRODUCTION

Estimating the health status of the general population has never been an easy task in either epidemiological or sociological research. One of the main problems is that most symptoms of ill health are not brought to the attention of medical professionals. In the 1960s, Wadsworth and his colleagues discovered that 95% of their sample experienced health complaints in the two weeks preceding the interview, but only 20% of the respondents saw a physician for these complaints [1]. In the three decades since Wadsworth's study, other studies reported similar or even lower consultation rates [2–4]. Unreported health problems consist mainly of minor everyday illnesses like headaches, fatigue and the common cold, although some serious conditions also remain unreported [4, 5]. Using a well known metaphor, medical practitioners speak of unreported health problems as the "clinical iceberg" [6]. Epidemiologist and social scientists draw on the same metaphor when they speak of the "iceberg of morbidity" [7].

Medical practitioners are interested in the early detection of diseases and to them the clinical iceberg consists of undiagnosed but serious conditions like diabetes, glaucoma or tuberculosis [6]. These are conditions, that remain undetected because their development is asymptomatic for the patient. Epidemiologists are less concerned with clinical practice, but share the clinician's focus on disease. Epidemiologists aim at obtaining accurate estimates of the amount and distribution of disease in the population, which requires an assessment of unreported morbidity [8].

Social scientists, but also researchers in General Practice or Community Medicine, focus on 'illness' or

*Revised version of a paper presented on the *Fifth European Health Services Research Conference*, Maastricht, 17–18 December 1993.

perceived morbidity [5, 9, 10]. To them the iceberg metaphor relates to reported (above water) and unreported (below water) symptoms.

The clinical iceberg and the iceberg of morbidity overlap, but are not identical. This paper is primarily concerned with the methodology of measuring perceived morbidity.

Social scientists and researchers in General Practice, have tried to estimate the total amount of perceived morbidity in the general population [5, 7, 10]. Their research typically consists of an oral or postal survey, asking a random sample of the population to report the experience of symptoms of ill health during a certain period. Only rarely are these surveys accompanied by physical examinations carried out by trained physicians [11].

Comparing the results of these surveys leads to a troubling discovery: estimates of the occurrence of health complaints differ substantially, not only *between* countries but *within* countries. Most of these differences can be attributed to the method of inquiry, as demonstrated in Table 1.

This table summarizes the results of several studies, done in the Netherlands, that used either open-ended questions about symptom experiences or symptom checklists. Most studies used a retrospective questionnaire, with a reference period of 14 days, although on occasion prospective methods like health diaries were used. It should be noted that when the same instrument is used in different studies (e.g. questionnaire with checklist and 14 day reference period) a high level of agreement is achieved. A slight difference in the wording of the question however, can have large effects on the outcome. The study from Nijmegen put specific emphasis on reporting minor complaints in the wording of its open-ended question, leading to the higher rates of reporting. It is also clear that the prospective health diary leads to a higher probability of symptom reporting, as compared to the retrospective questionnaire. Compared to the instrumental effects are the time trends that may have occurred in the period 1983–1986 of minor importance.

With the conflicting results of these studies in mind, one must ask how the truest picture of population symptom experiences can be obtained. Verbrugge and Ascione addressed this question in their influential paper on the iceberg of morbidity [7]. They mention a number of requirements that should be fulfilled in health surveys in order to obtain the most valid assessment of population symptoms. Their recommendations encompass data collection, coding and publication. In their view, the instrument must allow for a full scope of symptoms, and reports of symptoms should not depend on whether or not any action was taken for them. They prefer open-ended queries over a symptom checklist and favour prospective queries (health diary) over retrospective interviews. This list of requirements is helpful, but as all researchers know, to list requirements is one thing, to meet them is quite another. Financial and time constraints limit the choice of methods and often a retrospective interview with a symptom checklist is the best a researcher can get for his or her money. What harm does it do to settle for this option? In the next section, I explore the advantages and disadvantages of different methods of inquiry.

A COMPARISON OF METHODS

Researchers designing a health survey to assess the occurrence of symptoms of ill health are confronted

Table 1. The occurrence of health complaints in a period of 7 or 14 days in different studies from the Netherlands

Study	Checklist 14 days (%)	Checklist 7 days (%)	Open-ended question (%)
CBS 1983[1]	78		42
Uniken Venema 1986[2]			41
Survey General Practice 1987/1988[4]	81		63[3]
Nederland oke 1983[5]		65	
Regio Nijmegen 1983[6]			51
Huygen *et al.* (1983)[7]	81		

[1] Central Bureau of Statistics. Respondents ⩾16 years, private households ($N = 1054$). Weighted sample. Checklist with symptoms similar to studies 4 and 5 [12].

[2] Regional sample of Dutch and Turkish respondents, ages 16–69 years ($N = 489$). The percentage with health complaints in 14 days was the same for both groups [13].

[3] First two weeks in the health diary (data mentioned below).

[4] Dutch National Survey of General Practice. Sample of respondents listed with General Practitioner ($N = 13{,}014$). For comparison with CBS respondents ⩾16 years in private households are presented. Weighted sample. Checklist similar to studies 1 and 5 [14].

[5] Respondents from a panel-study, ages 20–67 years, living in private households ($N = 884$). Checklist of symptoms similar to studies 1 and 4 [15].

[6] Regional sample of Nijmegen and environs. Respondents in private households, ages 18–64 years ($N = 3245$) [16].

[7] One General Practice in the Nijmegen area. Sample of families with children ⩾12 years. Families with a chronically ill patient were excluded ($N = 857$). Checklist differs from studies 1, 4, 5 [4].

with at least two fundamental choices:

(1) the use of either a retrospective assessment of symptoms during an interview or a prospective method (i.e. a health diary); and
(2) the use of either a symptom checklist or an open-ended question about symptom experiences.

The decisions made most often depend not on theoretical premises but on a number of practical considerations like availability of resources, time constraints and sample size.

Several authors have considered the strengths and weaknesses of various methods. Table 2 summarizes their work [7, 12, 17–21].

A comparison of the characteristics of different methods shows that a retrospective interview with a checklist is both inexpensive to carry out and easy to process for data-analysis. There are, however, some doubts about the validity of the results. It is doubtful that respondents will remember the occurrence of symptoms in a reference period correctly [22]. They may also completely forget some symptoms. This is especially true for symptoms at the bottom of the list or those not explicitly mentioned but solicited in an auxiliary, open-ended category. With checklists some bias seems inevitable: symptoms at the beginning of the list are overrated, while those at the bottom are underreported. Another problem is the choice and the wording of the symptoms: should only symptoms be mentioned (fever, running nose) or should the list contain common names for frequent diseases (flu or cold)? Most symptom lists are derived from questionnaires intended for clinical settings, like the Cornell Medical Index (CMI) [23]. As such they are designed to establish the medical history of individual patients and are ill-suited for epidemiological purposes.

Table 2. A comparison of data collection methods in health surveys

1. Type of questions: checklist vs open-ended	
1(a). Checklist with symptoms	
Advantages:	—uniform definition and coding of symptoms and easy analysis
	—aided recall for minor symptoms
Disadvantages	—selection of symptoms
	—unknown symptom terminology
	—tedious for respondent and tail effect
	—triviality of symptoms
	—complaining on request
1(b). Open-ended questions	
Advantages	—all types of complaints possible
	—easy to answer
	—recording of non-trivial symptoms
Disadvantages	—no uniform definition of symptoms, respondents mix symptoms and diagnoses
	—coding problems
2. Type of questionnaire: retrospective (interview) vs prospective diary	
2(a). Retrospective interview	
Advantages:	—easy to administer (e.g. telephone) with minimum trouble for respondent
	—interviewer can check date quality
	—inexpensive
Disadvantages	—recall error: telescoping and memory lapse
	—difficult to combine complaints with illness action
2(b). Diary	
Advantages	—minimum recall error
	—monitoring complaints and illness action from day to day
Disadvantages	—conditioning: sensitization and fatigue
	—cooperation is time consuming for respondent
	—data quality is like postal survey
	—high costs
	—difficult to exploit data adequately in analysis

Checklists are often criticized for measuring psychological distress or neuroticism rather than actual illness or disease [24–28]. Mechanic argues that symptom reporting reflects a pattern of illness behaviour which is to a great extent influenced by the affective state of the individual. These reports are not necessarily a sign of an underlying disease. Support for Mechanic's position is found in the strong correlation between scales that measure 'neuroticism' or 'psychological distress' and the number of symptoms reported on a symptom list [26, 29]. Furthermore, in a number of experimental studies, Skelton and Pennebaker found that: "there is no one-to-one correspondence between responses that occur at the physiological level and our experiences of bodily sensations and symptoms" [30]. Pennebaker and his colleagues also observed that patients who were informed about their high blood pressure or high blood glucose levels, indicating diabetes, began to experience symptoms that they had not experienced before [31].

What alternatives are available to a researcher who wishes to avoid the problems associated with retrospective interviews with symptom lists? Open-ended questions about symptom experiences seem less sensitive to psychological distress since they are less of an open invitation to complain, but empirical evidence about this assumption is lacking. When open-ended questions are used on a retrospective questionnaire, it is likely that many minor complaints are forgotten and hence not reported. Although many of these symptoms are not serious, recording these symptoms provides valuable information on the extent to which illness behaviour, like self-medication, is driven by symptoms. The problem of memory lapse can be addressed with a prospective instrument, like a health diary, but the use of a health diary demands more time and effort on the part of the respondent [32]. People with very little time to spare or those who lack motivation are likely to refuse participation. Moreover, participation does not mean compliance with the task of completing the diary every day. Health diary studies therefore frequently use compliance enhancement strategies, like weekly phone-calls or lotteries, all of which have an effect on the outcome of the study [18, 19, 33, 34]. A further problem with the open-ended question is coding. Respondents are invited to define their complaint in their own words: symptoms, lay terminology for common diseases and medical diagnoses, are all included in health diaries.

Coding becomes very time consuming and may not be reliable.

In sum, there is no ideal method to assess the occurrence of common daily health problems. The best one can hope for is to develop a set of rules to guide researchers in their choice of a particular method for a particular situation. In order to develop these rules, one must have a thorough understanding of the actual operation of different data gathering techniques. This paper contributes to this understanding by comparing reports of symptom experiences obtained with a health diary using open-ended questions with the results of a health survey using a retrospective checklist, both completed by the same respondents. The respondents were approached for a face-to-face interview, after which they were asked to keep the health diary for the following three weeks. This study was part of the Dutch Survey of General Practice, and the main results on the occurrence of symptoms were mentioned in Table 1 [14]. The interested reader will find details on both diary and questionnaire in the data section and in Appendix 1 and 2. Previous research allowed me to formulate some hypotheses about the differences that are likely to occur between health diary and questionnaire. Since the questionnaire only contains a checklist and the diary only open-ended questions, it is not possible to compare all four combinations that come with these two instruments (Ref. Table 2).

HYPOTHESES

I conduct this comparison by exploring two key hypotheses. The first hypothesis relates to the assumption that in particular checklists measure psychological distress or neuroticism.

Hypothesis 1

The increase in the likelihood of reporting health complaints with increasing levels of psychological distress will be disproportionally high for the interview data with the checklist method as compared to the health diary data with open-ended questions.

If it is true that a symptom checklist is biased towards psychological distress, especially when compared to open-ended questions about symptoms in a health diary, then the following situation should apply. Respondents who are in a state of psychological distress will have a high probability of reporting complaints when completing the checklist. They also have a more than average probability of reporting complaints in the health diary. However, the odds of reporting complaints will be higher for the checklist than for the diary, since the checklist is more sensitive to psychological distress. Respondents who are emotionally stable will have a low probability of reporting complaints, both on the checklist and in the diary. Figure 1 illustrates this hypothesis.

The rates obtained with the checklist are placed on the left hand side because the questionnaire precedes the diary in time. Table 1 has shown that the symptom rate of the questionnaire is higher than the rates in the health diary. This finding is unusual in this type of research and a comparison of diary studies has shown that the symptom rate in this diary study is relatively low among the elderly and towards the end of the diary-keeping period (see Appendix 2).

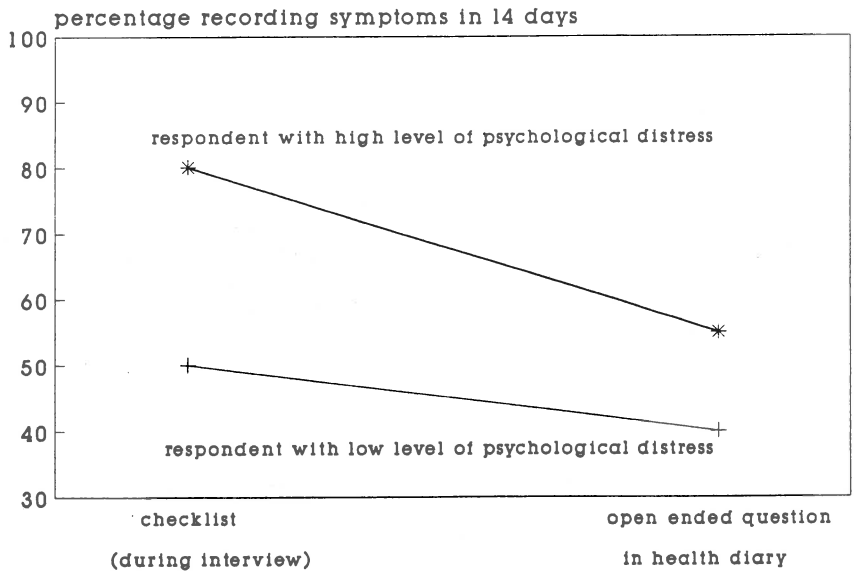


Fig. 1. Expected probability of recording symptoms when either checklist or open-ended questions are used.

The second hypothesis deals with the personal characteristics of respondents that are associated with compliance with the regimen of diary keeping. The comparison of data collection methods showed that compliance is a crucial issue in diary studies. The chances that the health diaries are completed properly at the end of each day is to a certain extent determined by the motivation and the time constraints of the respondents. Although the data do not allow compliance to be assessed directly, we can infer compliance by comparing results obtained with the questionnaire to those from the diary. The null hypothesis assumes equal probabilities of symptom reporting for both questionnaire and diary. The alternative hypothesis assumes that lack of compliance leads to relatively low symptom rates in the diary.

Hypothesis 2

Respondents who are not interested in health matters and/or have many time constraints or role obligations are less likely to complete the health diary accurately. The lack of compliance leads to relatively low symptom rates in the diaries as compared to the interviews.

The effect of compliance on the answering pattern for both the health diary and the interview differs from the previously mentioned effect of psychological distress. Responding to the symptom list during the interview is not very demanding for the respondent. Consequently, individual differences in motivation will not lead to large differences in the probability of mentioning symptoms during the interview. In the diary study on the other hand will motivation have larger effects on the probability of mentioning symptoms. This effect is illustrated in Fig. 2.

The data that are used in this paper contain an assessment of the interest that the respondent takes in health matters. They also contain information on the daily activities of the respondents with which time pressures (like family or work) can be inferred. Individual characteristics like gender, education and lifestyle seem of importance too. They determine the individual's concern with health and illness and are expected to be related to compliance.

Women are more attentive to their health than men, a fact confirmed by leafing through some popular women's (or men's) magazines, but reported in the literature as well [35]. Some researchers who used health diaries have gone as far as to approach only women as informants about the health of their families [20, 21, 34, 36]. Evidently, this procedure may lead to a serious underestimation of men's health problems [37]. In this study with both male and female respondents we expect a higher level of compliance among women.

Studies of response behaviour in social surveys have shown that people with only lower education give less reliable answers to socio-medical questions [38]. Extending this finding to this health diary study we expect higher compliance among those who have attained a higher educational level.

The concern with the relationship of lifestyle and symptom reporting is based on the assumption that those who have a relatively 'unhealthy' lifestyle (e.g. smoking, excessive alcohol consumption) are less interested in health matters and less attentive to bodily experiences that may signal ill health. An 'unhealthy' lifestyle leads to *cognitive dissonance*, caused by the disparity between knowledge about health risks and the awareness of actually engaging in unhealthy behaviour [39]. Dissonance may be reduced by dis-valuing health, which in turn leads to a lower level of compliance. The individual may also seek to reduce or prevent dissonance by ignoring

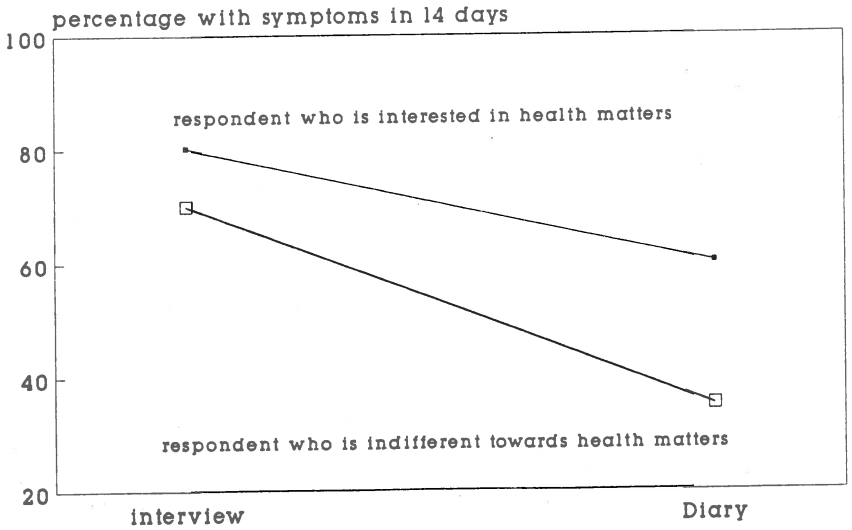


Fig. 2. Expected probability of recording symptoms during interview or with health diary.

bodily experiences that provide evidence that one is actually damaging one's own health. As a result, these bodily experiences are not framed as illness. It is therefore less likely that they are mentioned in the diary with its open-ended question about illness experiences. The task of completing the diary is likely to provide a threshold for symptoms that are considered part of everyday life. On the other hand, when symptoms as coughing, nausea, headache or nervousness are mentioned one by one on a checklist they are more likely to be acknowledged, since the checklist is more of an open invitation to complain.

It is possible to develop some extension of these two key hypotheses. For instance, so far the nature of the symptoms has not been taken into account. By grouping symptoms, a refinement can be made of the previous assumptions. Psychological symptoms and musculoskeletal symptoms provide interesting examples for this comparative study. Symptoms like dizziness, fatigue and nervousness, which are readily identified as related to psychological distress should, in accordance with the first hypothesis, result in larger differences between instruments. This phenomenon was in fact observed in a methodological study of Van Sonsbeek, who compared the results obtained with a questionnaire using a symptom checklist with those that were obtained with a similar questionnaire employing open-ended questions about acute symptoms [12]. In his study the checklist and the open-ended question showed the highest similarity for the occurrence of musculo-skeletal symptoms. With this result in mind, a third hypothesis may be stated:

Hypothesis 3

Psychological distress plays a minor role in explaining the difference in symptom rates between health diaries with open-ended questions and questionnaires with a symptom checklist for musculo-skeletal symptoms, but not for psychological symptoms.

An extension of this comparative study that is of an explorative nature concerns the effect of long-standing health problems on the answering pattern of questionnaire and diary. The presence of chronic conditions and their effects on daily functioning are receiving increasing interest from the socio-medical research community [40]. But this is an area where it is difficult to formulate hypotheses. Opposing effects seem plausible. On the one hand, it may be assumed that individuals with one or more chronic conditions have become sensitized to the effects of their health problem on day to day functioning, causing them to monitor their health rather closely. When it comes to completing a health diary, the answering pattern of these individuals would resemble the answering pattern of highly motivated respondents, increasing the probability of recording symptoms. Equally plausible, however, is that the presence of chronic conditions leads to a decrease in the probability of recording symptoms. The symptoms that accompany these conditions have become familiar and are not sufficiently unusual to be reported in the diary. In this case, the answering pattern would resemble the pattern of those with an unhealthy lifestyle, since these individuals have also learned to treat their symptoms as not unusual.

DATA

My study consists of a secondary analysis of data from the Dutch National Survey of General Practice; a nationwide study conducted in 1987/1988 among 161 general practitioners in 103 surgeries using a non-proportional stratified sampling scheme with urbanization, region and distance to the hospital as stratifying variables [14]. A random sample of 100 listed patients of each general practitioner was approached for a health interview. In the Netherlands nearly the entire population is listed with a GP. The patients were visited at their homes. A letter from their GP helped to gain entry for an interview. At the end of the interview, the respondents were asked to keep a structured health diary about the daily occurrence of health complaints and illness behaviour for the following three weeks. The short interval between these two measurements allows the assumption that the health status of the respondents did not change dramatically. The interviewer made an appointment to collect the diary from the respondents who agreed to take part in the diary study. During the diary keeping period the interviewer phoned twice to check if there were any problems with the completion of the health diary. The interviewers checked the entries of the respondents when collecting the diaries. The response rate is 77% for the questionnaire and of those respondents 85% completed the health diary for the entire 3 weeks (Total $N = 11{,}038$). For children of ages up to 14-years-old proxy interviews were held with one of the parents or guardian. These data are not used in this paper.

The health diary consists of a 21 paged booklet with a simple one-page questionnaire to be completed each day. The respondents were asked to provide a daily rating of their health, mood and activities and whether or not any health complaints occurred that day. Those who had complaints were asked to describe the nature of their complaints (with a maximum of two) in their own words, followed by a series of precoded questions on the assessment of the complaint and the illness action prompted by the complaint on that day (Appendix 1 lists the questions of the diary). In other words: on days with complaints there was more work to do than on days without complaints, an artifact which is one of the explanations for the relatively low symptom rate in this study.

The health interview provided the information on the respondent, including health status and background variables like sex, age, education etc. (see Appendix 1). During the interview the respondents

were asked to complete several checklists on their health status, with one list of acute symptoms. This list contained 42 precoded items and 3 open-ended items for additional complaints. The respondents were asked to indicate for each item if they experienced that symptom/condition during the previous fortnight. This checklist was originally developed for a study designed to have physicians estimate the severity of daily symptoms and was used in several other studies, of which the results are shown in Table 1 [12, 15]. The present version of the checklist contains some additional symptoms and some minor alterations. The respondents were also asked to complete a checklist on the presence of chronic conditions. In the present analysis, the total number of chronic conditions was used, with weights for the severity of these conditions (see Appendix 1).

Data related to the effect of psychological distress or neuroticism on symptom reporting were gathered with several instruments assessing the mental health status of the respondents: Goldberg's General Health Questionnaire (GHQ) and the BIOgraphic PROblem list (BIOPRO), a scale aimed at measuring the presence of psycho-social problems [41, 42]. The survey did not contain an instrument specifically designed for measuring neuroticism like the Eysenck Neuroticism scale or its Dutch equivalent. Goldberg's questionnaire is a screening instrument for use in general practice settings. It performs well in detecting anxiety and depression but is not intended to be used for the detection of psychoses [43]. While several authors have described the GHQ as an indicator of psychological distress or neuroticism, it should be noted that the GHQ is intended to register acute psychiatric symptoms based on the respondents experiences in the previous four weeks and not meant to measure neuroticism as a personality trait [44–46]. Researchers using the GHQ and a Eysenk's Neuroticism scale simultaneously, found similar associations with socio-demographic variables but report that there are some marked differences between the instruments among the elderly (75 years and over) [47]. These results have led to the decision to exclude the elderly of 75 years and over from the analyses. Because some of the questions on the 30 item version of the GHQ resembled questions on the symptom checklist (e.g. both lists contain questions about sleeping problems and nervousness), it was necessary to validate the results obtained with the GHQ. The Biographic Problem list contains 22 questions covering a wide variety of common problems ranging from 'difficulties in establishing relationships with others' to 'worries about the future' (see Appendix 1 for a list of all items). The BIOPRO has satisfying scaling properties. In a panel-study on psycho-social problems, Cronbach's α was 0.76 and 0.78 in the first and second wave respectively [48]. Previous users of the problem list argue that in several studies a simple counting of problems was strongly correlated with indicators of distress [16, 42, 49–51].

The analysis of the relationship between available time and diary keeping requires some measure of time use. In this study time constraints were defined differently for men and women; for men heavy time constraints were primarily related to the demands of work whereas for women combining work and the rearing of children was considered as such. For both sexes combining a job with being a head of a one-parent family was considered as a living condition with severe time limitations. It was assumed that (disability) pensioners, the unemployed and women who are housekeeping in a family without children, face very little time constraints and consequently have less trouble keeping the diary on a daily basis. Accurate completion of the health diary also depends on the motivation of the respondent. Respondents who show more interest in health matters were assumed to be keener on completing the diary than others. The respondent's interest in health matters was rated by the interviewer at the end of the interview.

In order to get a measure of the effect of health habits on reporting daily symptoms, respondents were asked about their current and previous alcohol and tobacco consumption. Three levels of drinking were distinguished: abstaining, light or moderate drinking and excessive drinking [52]. There were only very few abstainers who reported heavy drinking in the past. Previous drinking habits were therefore not considered. For smoking habits, a distinction was made between former and current smokers (light, moderate and heavy smokers) and respondents who have never smoked.

Both the checklist and the health diary could be used to assess the occurrence of health complaints in a period of 14 days: the checklist uses 14 days as a reference period and from the diary study we use the first two weeks. Both studies then provide us with a comparable estimate about whether or not any health complaints occurred. When it comes to the nature of complaints, checklist and health diary differ: the checklist records the *number* of different complaints in two weeks while the health diary records the *number of days* with complaints, restricting the number of different complaints on each day to two. It follows that only the dichotomous variable (yes or no complaints; specified or unspecified) is comparable for both instruments.

The diary contained two open-ended questions on the nature of the symptoms of ill health requiring a coding system to be developed. The health complaints of the diary were coded with the aid of the coding system that is widely used in General Practice research [53].

Three comparisons between checklist and diary were made. For hypothesis 1 and 2, the reporting of complaints was compared, irrespective of the nature of the complaints. For the third hypothesis, I made two additional comparisons on the occurrence of specific complaints: (1) complaints related to psychological

distress; and (2) musculo-skeletal complaints. The following procedure was used to group these symptoms. The checklist contains a number of complaints like dizziness, fatigue, nervousness that could easily be labelled as related to psychological distress or 'psychological symptoms' for short. With other complaints, like headache or extreme perspiration, there was less certainty. A Principal Components analysis on all items of the checklist helped in distinguishing a set of complaints that have psychological distress as a common denominator (these complaints are listed in Appendix 1). A similar procedure was carried out to distinguish musculo-skeletal problems. After defining these groups for the checklists the closest matching titles from the 97 categories in the diaries were grouped under the same headings.

METHODS

This section deals with the statistical analysis of the hypotheses. Under H_0 the questionnaire with symptom checklist and the health diary with open-ended questions about health complaints are equivalent instruments and both should result in the same percentage of respondents that record symptoms over a period of 14 days. The alternative hypotheses assume that these percentages differ. How should these differences be assessed? For both instruments the dichotomous variable, occurrence or non-occurrence of symptoms, could be modelled with logistic regression, using the variables pertaining to the hypotheses as predictors. The effect of each of these predictors on the occurrence of symptoms in either health diary or questionnaire is of limited interest, however. The primary concern of this paper lies with the (magnitude and sources of) the difference between instruments. These differences can be modelled in a single equation when a dummy variable for instrumental effects is created and interactions of this dummy variable with each of the predictors are incorporated in the model [54].

The respondents in this study completed both questionnaire and health diary and these response variables could therefore not be treated independently in one equation. Independent observations were obtained by assigning respondents randomly to two groups, which contributed to the analyses with either the diary or the questionnaire data on the occurrence of symptoms. Three logistic regressions were carried out in this way: the first analysis used the probability of the occurrence of symptoms, irrespective of its nature as response variable, whereas the two other analyses looked at the occurrence of specific symptoms. The following independent variables were entered in the analyses simultaneously:

(1) psychological distress (GHQ or BIOPRO);
(2) interest in health matters;
(3) educational level;
(4) time- and role obligations;
(5) smoking and drinking habits;
(6) the number of chronic conditions;
(7) sex;
(8) the dummy variable for the instrument that is used; and finally
(9) age is entered as control variable.

The design of this analysis required that interactions with the dummy variable for method are included as well. The inclusion of all independent variables simultaneously ensures that the variables that are included for the first hypothesis on psychological distress act as 'controls' for the analysis of the second hypothesis on compliance and vice versa.

RESULTS

The following tables display the results from these logistic regressions. They show the main effects of the predictors pertaining to the hypotheses, and interactions of these variables with the dummy variable for the method of enquiry. Table 3 shows the results for analysis of the occurrence of symptoms in general.

In this table the diary data are used as a reference category. The significant dummy variable 'method' indicates that the likelihood of reporting complaints is significantly higher for the questionnaire. What does the analysis show with respect to the hypotheses? First of all, the analysis clearly points out that psychological distress (GHQ) is an important predictor of reporting health complaints for both health diary and questionnaire. The strong interaction of 'method' with GHQ shows that the effect of distress is much stronger for the questionnaire than for the diary, which confirms the first hypothesis. An analysis that uses the questionnaire as a reference category shows much larger main effects (and interactions of the same magnitude, but opposite sign) for distress than the diary data do, which is in accordance with the expectations shown in Fig. 1.

Regarding the second hypothesis, only educational level leads to clearly significant differences between instruments, with a much lower probability of reporting symptoms in the diary for those with only limited education. A comparison of main effects shows that education has much larger effects for the diary than for the questionnaire data (for which they are not significant), which is in accordance with Fig. 2. Of further interest is that heavy smokers have a significantly higher probability of reporting symptoms in the questionnaire than in the diary, which partly confirms the predicted effect of lifestyle. Other categories of smoking and also drinking behaviour do not have the expected effects. The predictions about the effects of time pressures, gender and the interest that individuals have in health matters were not confirmed in this analysis. It is noteworthy that having one or more chronic

Table 3. Logistic regression of the probability of having health complaints in 14 days, recorded in either health diaries or during an interview with a retrospective questionnaire.[1] Data from Dutch National Survey of General Practice ($n = 7993$; ages: 15–74 years)

Independent variables main effects	Regr. coeff.	Standard error	*P*-value	Independent variables interaction with method (*m*)	Regr. coeff.	Standard error	*P*-value
GHQ				*GHQ*m*			
1–2	0.3729	0.0894	0.000	1–2	0.6742	0.1508	0.000
3–4	0.6537	0.1565	0.000	3–4	0.8906	0.2983	0.003
⩾5	1.2574	0.1434	0.000	⩾5	1.0247	0.3327	0.002
Not interested in health matters	−1.1254	0.3349	0.001	Not interested**m*	−0.0228	0.4953	0.963
Interested in health matters	0.0329	0.0717	0.646	Interested**m*	−0.0855	0.1138	0.452
Lower education	−0.3903	0.0834	0.000	Lower education**m*	0.3432	0.1316	0.009
Higher education	0.2499	0.1016	0.014	Higher education**m*	−0.1136	0.1572	0.470
Low time pressure	0.0696	0.1161	0.549	Former smoker**m*	−0.0865	0.1535	0.573
High time pressure	0.0108	0.0996	0.914	Light smoker**m*	0.2372	0.1785	0.184
Former smoker	0.3761	0.0977	0.000	Moderate smoker**m*	0.0177	0.1723	0.918
Light smoker (⩽10 c/d)	0.1449	0.1098	0.187	Heavy smoker**m*	0.5219	0.2140	0.015
Moderate smoker (11–20 c/d)	0.3030	0.1119	0.007	Light/mod alc.**m*	0.2115	0.1471	0.150
Heavy smoker (⩾20 c/d)	0.1839	0.1289	0.154	Heavy alc.**m*	0.2257	0.2957	0.445
Light/mod alc. cons.	0.1917	0.0914	0.036	Chronic disease**m*	0.1866	0.0371	0.000
Heavy alc. cons.[2]	0.0515	0.1836	0.779	*Age*m*			
Chronic disease (weighted)	0.2134	0.0189	0.000	25–34	−0.4841	0.1783	0.007
Age				35–44	−0.3393	0.1844	0.066
25–34	0.1378	0.1134	0.224	45–54	−0.2351	0.2111	0.265
35–44	−0.1321	0.1170	0.259	55–64	0.0944	0.2443	0.699
45–54	−0.2980	0.1333	0.025	65–74	0.1026	0.2806	0.715
55–64	−0.6018	0.1516	0.000	Male**m*	0.0069	0.1193	0.954
65–74	−0.9552	0.1789	0.000	Low time press.**m*	−0.2405	0.1829	0.189
Male	−0.6616	0.0744	0.000	High time press.**m*	−0.0385	0.1564	0.806
Method (questionnaire)	0.5341	0.2006	0.008				

[1] Logistic regression with indicator coding, reference categories: GHQ: 0; interest in health matters: neutral; education: intermediate; time pressure: intermediate; smoking: never smoked; alcohol consumption: abstainer; age: 15–24 yr; sex: female. The number of chronic diseases is a numerical variable; severity is weighted in 3 categories ref. Appendix 1.

[2] Heavy alc. cons., ♂ 4 glasses alcohol or more every day or 5 glasses alcohol or more almost every day; ♀ 3 glasses alcohol or more every day or 4 glasses alcohol or more almost every day. Abstainer, no alcohol consumption during past 6 months. Light/moderate alc., consumption of alcohol during past 6 months but less than heavy drinker.

diseases comes with a significant lower probability or reporting complaints in the diary as compared to the interview.

The regression analyses for the occurrence of specific symptoms may be interpreted in a similar vein as the previous analysis, with some important differences for the response variable. The dichotomous response variable treats respondents with one or more days with specific symptoms (or entries on the checklist), as the occurrence category, whereas the non-occurrence category consists of respondents who either did not experience symptoms at all or only symptoms that do not belong to the specific category at which the analysis aims.

During the interview, 59% of the respondents (age 15–74 years) reported to have experienced one or more psychological symptoms in the previous fortnight. In contrast, only 32% of the respondents reported psychological symptoms on one or more days in the first two weeks of the diary. The results of the logistic regression for the occurrence of symptoms related to psychological distress is shown in Table 4.

Evidently, higher scores on the GHQ increase the likelihood of reporting psychological symptoms with a strong excess of reporting complaints during the interview. Again the educational level is a variable that has strong effects on the answering pattern in the diaries but much less so in the questionnaire. The interaction is statistically significant for all categories of education combined (Wald statistic: $P = 0.013$) but not for separate dummy variables. Respondents who have attained a higher educational level are apparently more willing and/or able to report psychological symptoms in a health diary than others. During the interview smokers reported more psychological symptoms than non-smokers do, with heavy smokers having the highest probability of reporting symptoms. This relationship is virtually absent in the health diaries. Respondents with one or more chronic conditions are more likely to report psychological symptoms during the interview than in the diary. Respondents who were indifferent towards health matters have a very low probability of entering psychological complaints in the diary, as compared to the interview. Again, the interaction is not significant, although the differences between instruments are much larger than in the previous analysis. Women report more psychological symptoms than men do but without the expected difference between instruments.

Musculo-skeletal problems were also more often reported during the interview than in the diary. About 38% of the respondents ticked one or more of these symptoms on the checklist, while only 20% wrote about musculo-skeletal symptoms in the diary.

Table 4. Logistic regression of the probability of having psychological symptoms in 14 days, recorded in either health diaries or during an interview with a retrospective questionnaire.[1] Data from Dutch National Survey of General Practice (n = 7993; ages: 15–74 years)

Independent variables main effects	Regr. coeff.	Standard error	P-value	Independent variables interaction with method (m)	Regr. coeff.	Standard error	P-value
GHQ				GHQ*m			
1–2	0.2625	0.0905	0.004	1–2	0.7473	0.1271	0.000
3–4	0.4025	0.1420	0.005	3–4	1.1882	0.2202	0.000
⩾5	0.9638	0.1117	0.000	⩾5	1.3534	0.2102	0.000
Not interested	−1.4641	0.5376	0.007	Not interested*m	1.0829	0.6439	0.093
Interested	0.0839	0.0738	0.256	Interested*m	−0.1286	0.1045	0.218
Lower education	−0.2770	0.0863	0.001	Lower education*m	0.2208	0.1216	0.069
Higher education	0.3497	0.0974	0.000	Higher education*m	−0.1837	0.1394	0.188
Low time pressure	−0.1455	0.1173	0.215	Former smoker*m	0.0089	0.1395	0.949
High time pressure	0.0914	0.1004	0.362	Light smoker*m	0.3303	0.1624	0.042
Former smoker	0.0837	0.0977	0.392	Moderate smoker*m	0.2115	0.1587	0.183
Light smoker (⩽10 c/d)	−0.0976	0.1139	0.392	Heavy smoker*m	0.3629	0.1878	0.053
Moderate smoker (11–20 c/d)	−0.0226	0.1144	0.843	Light/mod alc.*m	0.2270	0.1334	0.089
Heavy smoker (⩾20 c/d)	0.0032	0.1308	0.980	Heavy alc.*m	0.1829	0.2739	0.504
Light/mod alc. cons.	−0.0283	0.0921	0.759	Chronic disease*m	0.1162	0.0259	0.000
Heavy alc. cons.[2]	0.0168	0.1936	0.931	*Age*m			
Chronic disease (weighted)	0.1171	0.0161	0.000	25–34	−0.2310	0.1592	0.147
Age				35–44	−0.1523	0.1677	0.364
25–34	0.1658	0.1119	0.139	45–54	−0.0603	0.1928	0.755
35–44	0.0065	0.1191	0.957	55–64	0.1266	0.2245	0.573
45–54	−0.1344	0.1367	0.325	65–74	0.0753	0.2708	0.781
55–64	−0.2913	0.1601	0.069	Male*m	−0.0030	0.1073	0.978
65–74	−0.7920	0.1975	0.000	Low time press.*m	0.0415	0.1652	0.802
Male	−0.8503	0.0756	0.000	High time press.*m	−0.1937	0.1431	0.176
Method (questionnaire)	0.5530	0.1849	0.003				

[1] Logistic regression with indicator coding, reference categories: GHQ: 0; interest in health matters: neutral; education: intermediate; time pressure: intermediate; smoking: never smoked; alcohol consumption: abstainer; age: 15–24 yr; sex: female. The number of chronic diseases is a numerical variable; severity is weighted in 3 categories ref. Appendix 1.

[2] Heavy alc. cons., ♂ 4 glasses alcohol or more every day or 5 glasses alcohol or more almost every day; ♀ 3 glasses alcohol or more every day or 4 glasses alcohol or more almost every day. Abstainer, no alcohol consumption during past 6 months. Light/moderate alc., consumption of alcohol during past 6 months but less than heavy drinker.

The results from the logistic regression are shown in Table 5.

These results differ in a number of ways from the previous analyses. Firstly, there is no significant interaction between psychological distress and the instrument of inquiry, although it remains true that the probability of reporting symptoms increases with higher levels of distress. This finding is in accordance with the third hypothesis. Respondents with severe time limitations or role constraints are less likely to report musculo-skeletal symptoms in the diary than during the interview, which was not found in the previous analyses. Smokers are again more likely to report symptoms during the interview than in the diary. The differences between the sexes are smaller than in the previous analyses.

All of these logistic regressions have been repeated with scores on the Biographic Problem list (BIOPRO) as indicator of psychological distress. The results of these analyses are very similar to those presented above with the GHQ and therefore not shown in separate tables, but may be obtained from the author. The expected interaction of distress with method of inquiry holds equally for the occurrence of symptoms in general and for the occurrence of psychological symptoms. For musculo-skeletal symptoms the interaction is larger than in the analysis with Goldberg's GHQ, but not clearly significant (Wald statistic: $P = 0.08$ for all categories combined).

CONCLUSION AND DISCUSSION

With the goal of improving our understanding of the 'iceberg of morbidity', this paper addressed the issue of the differences that occur in survey research that aims at measuring the amount of self reported morbidity in the population. Two key hypotheses were tested:

(1) interviews using checklists are more sensitive to the respondent's psychological distress than are the open-ended questions of health diaries; and
(2) health diaries demand high levels of compliance leading to underreporting of symptoms among those who are less likely to comply.

Compliance was assumed to be related to taking an interest in health matters, time restrictions, gender, education and lifestyle. These two hypotheses were tested with a comparison of the reported occurrence of symptoms during a face-to-face interview and in a health diary. The nature of the reported symptoms was not taken into account. A third hypothesis assumed that psychological distress is of particular importance when assessing the occurrence of psycho-

Table 5. Logistic regression of the probability of having musculo-skeletal symptoms in 14 days, recorded in either health diaries or during an interview with a retrospective questionnaire.[1] Data from Dutch National Survey of General Practice (n = 7993; ages: 15–74 years)

Independent variables main effects	Regr. coeff.	Standard error	P-value	Independent variables interaction with method (m)	Regr. coeff.	Standard error	P-value
GHQ				*GHQ*m*			
1–2	0.3259	0.1008	0.001	1–2	0.1023	0.1334	0.443
3–4	0.2792	0.1616	0.084	3–4	0.1904	0.2127	0.371
⩾5	0.5562	0.1214	0.000	⩾5	0.2444	0.1659	0.141
Not interested	−1.1369	0.5955	0.056	Not interested*m	−0.0753	0.7405	0.919
Interested	0.0019	0.0831	0.982	Interested*m	−0.0199	0.1104	0.857
Lower education	−0.0625	0.0968	0.519	Lower education*m	0.1568	0.1280	0.221
Higher education	0.0795	0.1147	0.489	Higher education*m	−0.1831	0.1517	0.228
Low time pressure	−0.0254	0.1256	0.840	Former smoker*m	−0.0235	0.1478	0.874
High time pressure	−0.3107	0.1198	0.010	Light smoker*m	0.3593	0.1726	0.037
Former smoker	0.1666	0.1099	0.130	Moderate smoker*m	0.0304	0.1663	0.855
Light smoker (⩽10 c/d)	−0.0284	0.1318	0.829	Heavy smoker*m	0.2893	0.1962	0.140
Moderate smoker (11–20 c/d)	0.1503	0.1269	0.236	Light/mod alc.*m	0.0882	0.1406	0.530
Heavy smoker (⩾20 c/d)	0.0129	0.1484	0.931	Heavy alc.*m	0.4175	0.2860	0.144
Light/mod alc. cons.	0.0522	0.1044	0.617	Chronic disease*m	0.1399	0.0248	0.000
Heavy alc. cons.[2]	−0.0570	0.2172	0.793	*Age*m*			
Chronic disease (weighted)	0.1461	0.0166	0.000	25–34	−0.1629	0.1758	0.354
Age				35–44	−0.2434	0.1830	0.184
25–34	0.1956	0.1354	0.149	45–54	−0.2491	0.2020	0.218
35–44	0.2834	0.1403	0.043	55–64	0.1503	0.2353	0.523
45–54	0.5318	0.1522	0.001	65–74	0.0275	0.2766	0.921
55–64	0.1302	0.1793	0.468	Male*m	−0.0940	0.1131	0.406
65–74	−0.0489	0.2093	0.815	Low time press.*m	−0.0380	0.1675	0.820
Male	−0.1625	0.0852	0.056	High time press.*m	0.3282	0.1562	0.036
Method (questionnaire)	0.5128	0.2062	0.013				

[1] Logistic regression with indicator coding, reference categories: GHQ: 0; interest in health matters: neutral; education: intermediate; time pressure: intermediate; smoking: never smoked; alcohol consumption: abstainer; age: 15–24 yr; sex: female. The number of chronic diseases is a numerical variable; severity is weighted in 3 categories ref. Appendix 1.

[2] Heavy alc. cons., ♂ 4 glasses alcohol or more every day or 5 glasses alcohol or more almost every day; ♀ 3 glasses alcohol or more every day or 4 glasses alcohol or more almost every day. Abstainer, no alcohol consumption during past 6 months. Light/moderate alc., consumption of alcohol during past 6 months but less than heavy drinker.

logical symptoms, but much less so, for musculo-skeletal symptoms. The effect that the presence of chronic conditions may have on symptom reporting was also assessed, but in an explorative manner. The results of the analyses are summarized in Table 6.

This table clearly shows that most empirical support is found for the first hypothesis. We may conclude that symptom checklists, like the one that was used in this paper, are likely to lead to biased results, due to psychological distress. Why? An examination of the contents of the symptom list may help in the explanation of this result. The Dutch checklists mentioned in this paper (Table 1) have origins of measuring both psychological and physical symptoms. They are modifications of the 'VOEG' (Questionnaire for assessment of subjective health) checklist that was developed in the 1960s to measure distress of workers in an industrial environment [55]. In turn, the VOEG took 32 of its items from the Cornell Medical Index (CMI) [12]. The CMI was designed as a checklist of a patient's medical history which the patient could complete independently in the waiting room [23]. Subsequent research with the CMI showed that it was particularly well suited to measure psychological problems that physicians do not detect during a consultation [56]. The checklists that are currently used in the Netherlands contain

Table 6. Summary of results of comparison of diaries and interviews (sig. $P < 0.05$ two tailed)

	Nature of symptoms		
Effect	General	Psychological	Musculo-skeletal
Psychological distress	High distress High prob. interv.	High distress High prob. interv.	Not significant
Time pressure	Not significant	Not significant	High time press. Low prob. diary
Interest in health	Not significant	Not significant	Not significant
Gender	Not significant	Not significant	Not significant
Education	Low education Low prob. diary	Not significant	Not significant
Lifestyle	Heavy smoker Low prob. diary	Light smoker Low prob. diary	Light smoker Low prob. diary
Chronic conditions	More chron. cond. Low prob. diary	More chron. cond. Low prob. diary	More chron. cond. Low prob. diary

up to ten 'psychological' symptoms which are prominently placed at the beginning of the list (see Appendix 1). There is reason to believe that a different checklist with a list of well-defined physical symptoms would be less prone to the psychological distress bias. Our comparison of the reporting of musculoskeletal complaints in the diaries and in the questionnaire is a case in point: levels of distress are still significant predictors in both diary and questionnaire but the difference in sensitivity to distress is slight. Future research of the occurrence of health complaints would be well advised to develop a new symptom list, rather than using one of the prevailing lists.

The second hypothesis on compliance received much less support. It only became clear that a health diary study is less suited for examining the symptom experience of respondents with only a few years of education. Using diaries for the assessment of the health status of the entire population means running the risks of an underrepresentation of disadvantaged groups like the elderly with only limited education.

The analyses also showed that respondents with a chronic condition and smokers resemble each other in their response to the two instruments. Both groups report far more symptoms during the interview than in the diary. Although this may be interpreted as lack of compliance, this response pattern may also provide a clue about the way that respondents have used the health diary. It seems that the health diary is primarily used to report symptoms that are outside the range of normal experience. Smokers and respondents with a chronic condition may have become insensitive to minor disturbances in their daily health and therefore do not enter these symptoms in the diary. They have raised the threshold for labelling a physical symptom as illness. The respondent's assessment of the symptoms reported in the diary supports this interpretation. The structured questionnaire of the health diary not only required mentioning the symptoms of that day but also asked to provide an assessment of each symptom (i.e. is the symptom: new or unknown, bothering or irritating, reason for worry? see Appendix 1). During new episodes of illness, symptoms were rated as bothering/irritating on $>70\%$ of the days with illness. It is plausible that the respondents did not take the trouble to mention symptoms in the diary that interfere only very little with daily life, which may also add to the explanation for the lower symptom rate in the diary as compared to the interview.

This paper began by asserting the need to fully explore the iceberg of morbidity. This study shows that the amount of self-reported morbidity is not a quantity that lends itself to precise measurements. Respondents may report (and think) of their symptoms differently, depending on the method with which these symptoms are assessed. In metaphoric language this would mean that the boundaries of that part of the iceberg that lies below water cannot be clearly delineated. What is ice to some is water to others. Not surprisingly, estimates of the size of the iceberg do not exist independently of the instruments used to make the estimates. For clinicians the iceberg metaphor remains apt: for them reported morbidity remains some small percentage of the total amount of diseases in a population. But for social scientists whose interest lies below the surface, the metaphor of the iceberg melts, precisely because what they find there is much too unsolid to be referred to as ice. Reports about the occurrence of symptoms of ill health should therefore always contain detailed information on how these symptoms were assessed.

REFERENCES

1. Wadsworth M. E. J., Butterfield W. J. H. and Blaney R. *Health and Sickness, the Choice of Treatment*. Tavistock, London, 1971.
2. Banks M. H., Beresford S. A. A., Morell D. C., Waller J. J. and Watkins C. J. Factors influencing demand for primary medical care in women aged 20–44 years: a preliminary report. *Int. J. Epidemiol.* **4,** 189, 1975.
3. Folmer H. R. Huisarts en ijsberg. (dissertation) Utrecht, 1986.
4. Huygen F. J. A. van den Hoogen H. and Neefs W. J. Gezondheid en ziekte; een onderzoek van gezinnen. *Ned. Tijdschr. Geneeskunde* **127,** 1612, 1983.
5. Hannay D. R. *The Symptom Iceberg. A Study of Community Health*. Routledge & Kegan Paul, London, 1979.
6. Last J. M. The iceberg: completing the picture in general practice. *The Lancet* **ii,** 28, 1963.
7. Verbrugge L. M. and Ascione F. J. Exploring the iceberg. Common symptoms and how people care for them. *Med. Care* **25,** 539, 1987.
8. MacMahon B. and Pugh T. F. *Epidemiology. Principles and Methods*. Little, Brown & Company, Boston, 1970.
9. Eisenberg L. Disease and illness. Distinctions between professional and popular ideas of sickness. *Culture, Med. Psychiat.* **1,** 9, 1977.
10. Van de Lisdonk E. H. Perceived and presented morbidity in general practice. *Scand. J. Prim. Hlth Care* **7,** 73, 1989.
11. Mootz M. Health indicators. *Soc. Sci. Med.* **22,** 255, 1986.
12. Van Sonsbeek J. L. A. Het meten van veel voorkomende gezondheidsklachten in de Nederlandse bevolking. *Maandbericht gezondheidsstatistiek CBS* **9,** 4, 1990.
13. Uniken Venema H. P. Toen ik hier kwam was ik kerngezond. (dissertation) EUR, Rotterdam, 1989.
14. Foets M., van der Velden J. and de Bakker D. *Dutch National Survey of General Practice. A Summary of the Survey Design*. Netherlands Institute of Primary Care, Utrecht, 1992.
15. Halfens R., Drop M. J. and Philipsen H. *Leefwijzen en subjectieve gezondheid van een panel uit de Nederlandse bevolking*. RU Limburg, Maastricht, 1984.
16. Furer J. W. and Tax B. (Eds) *Somatische klachten, psychiatrische symptomen en psychosociale problemen*. ISG/Nijmegen, 1987.
17. Freer C. B. Health diaries: a method of collecting health information. *J. R. Coll. Gen. Prac.* **30,** 279, 1980.
18. Van de Lisdonk E. H. Ervaren en aangeboden morbiditeit in de huisartspraktijk. (dissertation) Nijmegen, 1985.
19. Norman G. R., McFarlane A. H., Streiner D. L. and Neale K. Health diaries: strategies for compliance and relation to other measures. *Med. Care* **20,** 623, 1982.

20. Roghmann K. J. and Haggerty R. The diary as a research instrument in the study of health and illness behavior. *Med. Care* **10,** 143, 1972.
21. Verbrugge L. M. Health diaries. *Med. Care* **18,** 73, 1980.
22. Sudman S. and Bradburn N. M. *Response Effects in Surveys: A Review and Synthesis.* Aldine, Chicago, 1974.
23. Brodman K., Erdman A. J., Lorge I., Wolff H. G., Broadbent T. H. *et al.* The Cornell Medical Index; an adjunct to medical interview. *JAMA* **140,** 530, 1949.
24. Cassee E. T. Naar de dokter. *Enkele achtergronden van ziektegedrag en gezondheidszorg.* Boom, Meppel, 1973.
25. Mechanic D. Correlates of physician utilization: why do major multivariate studies of physician utilization find trivial psychosocial and organizational effects. *J. Hlth Soc. Behav.* **20,** 387, 1979.
26. Mechanic D. The experience and reporting of common physical complaints. *J. Hlth Soc. Behav.* **21,** 146, 1980.
27. Mechanic D. Social psychological factors affecting the presentation of bodily complaints. Repr. in Mechanic D. *Painful Choices. Research and Essays on Health Care.* Transaction Publishers, New Brunswick, 1989.
28. van der Zee J. De vraag naar diensten van de huisarts. (dissertation) Utrecht, 1982.
29. van Sonsbeek J. L. A. *De VOEG: klaaglijst of lijst met gezondheidsklachten?* Centraal Bureau voor de Statistiek. Den Haag, 1990.
30. Skelton J. A. and Pennebaker J. W. The psychology of physical symptoms and sensations. In *Social Psychology of Health and Illness* (Edited by Sanders G. S. and Suls J.). Lawrence Erlbaum Assoc., Hillsdale, 1982.
31. Pennebaker J. W., Gonder-Frederick L., Cox D. J. and Hoover C. W. The perception of general vs. specific visceral activity and the regulation of health-related behavior. In *Advances in Behavioral Medicine* (Edited by Katkin E. S. and Manuck S. B.). Jai Press, Greenwich & London, 1985.
32. Sudman S. and Bradburn N. M. *Asking Questions. A Practical Guide to Questionnaire Design.* Jossey-Bass Inc., San Francisco, 1982.
33. Dahlquist G., Wall S., Ivarsson J. I., Sterky G. and Tengvald K. Health problems and care in young families—an evaluation of survey procedures. *Int. J. Epidemiol.* **13,** 221, 1984.
34. Bentzen N., Christiansen T. and Pedersen K. M. *The Danish Health Study.* Odense University, Odense, 1988.
35. Waldron I. Gender and Health-Related Behavior. In *Health Behavior: Emerging Research Perspectives* (Edited by Gochman D. S.). Plenum Press, New York, 1988.
36. Moll van Charante A. W. Ziektegevoel: Ziektegedrag. (dissertation) Utrecht, 1980.
37. Kirscht J. P. Social and Psychological Problems of Surveys on Health and Illness. *Soc. Sci. Med.* **5,** 519, 1971.
38. Dijkstra W. and van der Zouwen J. *Response Behaviour in the Survey Interview.* Academic Press, New York, 1982.
39. Pervin L. A. and Yatko R. J. Cigarette smoking and alternative methods of reducing dissonance. *J. Personality Soc. Psychol.* **2,** 30, 1965.
40. Van den Bos G. A. M., Mohrs J., Habbema J. D. F., Van der Maas P. J. *et al. Chronische aandoeningen, hulpbehoevendheid en zorggebruik.* ISG/Amsterdam & IMG/Rotterdam, 1986.
41. Goldberg D. P. *The Detection of Psychiatric Illness by Questionnaire.* Oxford University Press, London, 1972.
42. Hosman C. M. H. *Psychosociale problematiek en hulpzoeken.* Swets & Zeitlinger, Lisse, 1983.
43. Bowling A. *Measuring Health. A Review of Quality of Life Measurement Scales.* Open University Press, Milton Keynes, 1991.
44. Duncan-Jones P., Grayson D. E. and Moran P. A. P. The utility of latent trait models in psychiatric epidemiology. *Arch. Gen. Psychiat.* **39,** 391, 1985.
45. Jorm A. F. and Duncan-Jones P. Neurotic symptoms and subjective well-being in a community sample: different sides of the same coin? *Psychol. Med.* **20,** 647, 1990.
46. Marks J. N., Goldberg D. P. and Hillier V. F. Determinants of the ability of general practitioners to detect psychiatric illness. *Psychol. Med.* **9,** 337, 1979.
47. Cox B. D., Blaxter M., Buckle A. L. J., Fenner N. P. *et al. The Health and Lifestyle Survey.* Health Promotion Research Trust, London, 1987.
48. Tijhuis M. A. R., Verhaak P. F. M. and Wennink H. J. *Psychosociale problemen in de huisartspraktijk.* NIVEL, Utrecht, 1991.
49. Andrews F. M. and Withey S. B. *Social Indicators of Well-being; Americans' Perceptions of Life Quality.* Plenum Press, New York, 1976.
50. Greenley J. R. and Mechanic D. Patterns of seeking care for psychological problems. In Mechanic D. *The Growth of Bureaucratic Medicine: An Inquiry into the Dynamics of Patient Behaviour and the Organization of Medical Care.* Wiley-Interscience, New York, 1976.
51. Mechanic D. Sociocultural and social-psychological factors affecting personal responses to psychological disorder. *J. Hlth Soc. Behav.* **16,** 393, 1976.
52. Royal College of General Practitioners. *Alcohol—A Balanced View.* Report from General Practice No. 24. Royal College of General Practitioners, London, 1986.
53. Lamberts H. and Wood M. (Eds) *International Classification of Primary Care.* Oxford University Press, Oxford, 1987.
54. Kleinbaum D. G., Kupper L. L. and Muller K. E. *Applied Regression Analysis and Other Multivariable Methods.* PWS-KENT, Boston, 1988.
55. Dirken J. M. *Het meten van 'stress' in industriële situaties.* J. B. Wolters, Groningen, 1967.
56. Brodman K., Erdman A. J., Lorge I., Gershenson C. P. and Wolff H. C. The Cornell Medical Index-Health Questionnaire: III. The evaluation of emotional disturbances. *J. Clin. Psych.* **8,** 119, 1952.

APPENDIX 1

Variables in the Dutch Survey of General Practice used in this Paper

*1. Checklist of complaints experienced in the 14 days prior to the interview:**

Q: This form mentions various complaints from which everybody may suffer once in a while. Please, indicate for each complaint if you suffered from that complaint in the past 14 days.

Checklist contains the following symptoms:
Dizziness, headache, fever, fatigue, nervousness, insomnia, general weakness, aggressiveness (frustration), unbalanced nerves, extreme perspiration, sore throat, ear pain, buzzing in the ear, hearing problems, nasal congestion, nose bleeding, cough, palpitations, swollen ankles, nausea, chest ache, vomiting, diarrhoea, heartburn, stomach ache, cramps, constipation, gaining weight, eating disorders, toothache, painful urination, incontinence, menstrual problems, complaints about neck shoulder or hip, back pain, complaints about limps, problems at work, family problems, problems with contraception.

*The current health problems in this checklist are described in the terms that a lay person would use. The translation into English is intended to remain as close to the Dutch wording as possible. It is inevitable that divergences exists between the Dutch and English lay persons terminology of health complaints.

In this paper the following subdivisions are made for separate analysis:

1(a) Complaints related to psychological distress: dizziness, headache, fatigue, nervousness, insomnia, general weakness, aggressiveness (frustration), unbalanced nerves, extreme perspiration.

1(b) Musculo-skeletal complaints: complaints about neck shoulder or hip, back pain, complaints about limps.

*2. Checklist with chronic conditions:**

2(a) Not very limiting or severe (weight = 1): hayfever, haemorrhoids, varicose veins, eczema nervousness, allergy, migraine.

2(b) Moderately limiting or severe (weight = 2): hypertension, ulcis cruris, chronic back pain, rheumatism, migraine, stomach and bowel problems, bile and liver diseases, diseases of thyroid gland, disorders of eye or ear, disorders of the joints, prostate problems (♂ only), menstrual problems (♀ only).

2(c) Limiting or more severe (weight = 3): bronchitis or asthma, heart problems, arteriosclerosis, cancer, diseases of nervous system (e.g. epilepsy), diabetes, diseases of the kidney, lasting consequence of an accident, hereditary handicap.

3. Health diary

Respondents were asked to rate their self-perceived health, their mood and their daily activities on a 5 point scale. After that they were asked about their symptoms.

Q: Did you have any complaints about your health today?

Respondents who answered 'yes' were asked to mention their complaints in their own words, with a maximum of two complaints on each day in separate boxes. When experiencing more than two complaints they were asked to mention the two most important complaints. Related complaints (e.g. coughing, sneezing, fever) should be entered together in one box. The respondents were allowed to use their own words. They used terminology that could be related to both symptoms and diagnosis.

Complaints were coded in 97 categories. The coding scheme was based on ICPC chapters. This scheme was also used in the questionnaire to assess reasons for encounter with the GP, as reported by the patient. In this paper the following subdivisions are made for separate analysis (category in parenthesis):

3(a) Complaints related to psychological distress: (57) fatigue, insomnia, nervousness; (58) dizziness, vertigo; (59) headache, migraine; (89) depression, other psychiatric complaints.

3(b) Musculo-skeletal complaints: (49) back pain; (50) joints; (51) muscles; (52) sprain, lacerations; (54) fractures; (55) bruises; (61) neck; (62) shoulder; (63) arm, hand, finger; (66) leg, foot, ankle; (85) bone; (87) surgery of the back; (88) surgery of the limbs.

Following the nature of their complaints, respondents were asked to assess their symptoms as either: (1) new/unknown, (2) lasting >1 year, (3) irritating, troubling, (4) reason for worry, (5) self-limiting, (6) known why existing.

Finally they were asked to choose on one or more of 16 different actions to alleviate their symptoms.

*Most of the chronic conditions on the checklist are also formulated in the lay persons terminology. Occasionally a medical term has been used (which is translated into English for this paper) when it may be assumed that this term is very common. The rating of severity of chronic conditions is based on a paper from Van den Berg & Van den Bos (1989) with some minor alterations.

4. Goldberg's General Health Questionnaire

The 30 item version was used which was translated in Dutch. (0) No items marked, (1) 1–2 items marked, (2) 3–4 items marked, (3) 5 or more items marked.

5. BIOgraphic PROblem list

This list contains 22 problems. Respondents could answer with yes, no or does not apply. Problem list: financial, housing, parents, education, job, ageing, partner or marital problems, children, other relationships, establishing relationships, sexual, religious, self-development, self-image, worry about the future, abuse of alcohol, medicine, illegal drugs, loneliness, societal change, neighbourhood, leisure time, worry (in general), other problems. (0) No items marked, (1) 1–2 items marked, (2) 3–4 items marked, (3) 5 or more items marked.

6. Lifestyle

Smoking: (0) never-smoked, (1) former smoker, (2) light smoker: ⩽10 cigarettes daily, (3) moderate smoker: 11–20 cigarettes daily, (4) heavy smoker: ⩾21 cigarettes daily.

Alcohol consumption: (0) abstainers: no alcohol during the past 6 months, (1) moderate or light alcohol consumption: drinking alcohol during the past 6 months but less than heavy, (2) heavy drinking: ♂ ⩾4 glasses daily or ⩾5 glasses almost every day, ♀ ⩾3 glasses every day or ⩾4 glasses almost every day.

7. Education

Highest attained educational level (not necessary completed). (0) Intermediate general education or vocational training (O-level), (1) primary school only or lower vocational level, (2) secondary education (A-level) or higher vocational training or university.

8. Demographic variables

Sex: (0) female, (1) male.

Age: (0) 15–24 yr, (1) 25–34 yr, (2) 35–44 yr, (3) 45–54 yr, (4) 55–64 yr, (5) 65–74 yr.

9. Role and time constraints

♂ (0) intermediate = default; (1) low time pressure: unemployed, (disability) pensioner, housekeeping; (2) heavy time pressure: self-employed/working at managerial level/head of a one-parent family combined with paid employment or enrolment in full-time education.

♀ (0) intermediate = default; (1) low time pressure: unemployed, (disability) pensioner, housekeeping; (2) heavy time pressure: self-employed/working at managerial level/head of a one-parent family combined with paid employment or enrolment in full-time education/combining work and housekeeping in a family with children.

10. Interest in health matters

Interest of the respondent in topics discussed in the health interview (as rated by the interviewer): (0) neutral, (1) not interested, (2) interested.

APPENDIX 2

A Comparison of Diary Studies

After her extensive review of available health diary studies, Verbrugge asserted that diary studies always lead to higher rates of symptom prevalence than retrospective interviews [1]. Table 1 of my article showed that this is not the case for the diaries and interviews from the Dutch Survey of General Practice: the interview data show the highest rates. Table 1 also shows that the rates that were obtained with the questionnaire do not differ much from similar studies. Consequently, we have to

Table A1. A comparison of the observed percentages of respondents with symptoms in health diary studies with the expected percentage on the basis of the diary study of the Dutch Survey of General Practice

Study	Percentage observed	Period (weeks)	Percentage expected extrapolating day 1–21[1]	R^2	Percentage expected extrapolating day 8–21	R^2	Percentage expected extrapolating day 15–21	R^2
Verbrugge and Ascione								
⩾18 years	89	6	1. 88%	0.91	1. 89%	0.97	1. 70%	0.99
			2. 90%	0.95				
			3. 92%	0.62				
⩾18 years	95	6	1. —	0.86	1. 90%	0.95	1. 83%	0.97
			2. 96%	0.94				
			3. 98%	0.64				
Van de Lisdonk	86	4	1. 82%	0.89	1. 74%	0.95	1. 76%	0.98
15–60 years families			2. 84%	0.94				
			3. 89%	0.61				
Stoller *et al.* independent elderly (⩾65 years)	83	3	51% (observed)					

[1]Three different models are used: 1. $S(t) = \exp(-bt)\text{-}a$; 2. Ln $\{S(t)\} = -bt$ (OLS); 3. $S(t) = \exp(-bt)$.

ask ourselves how typical our data are in comparison with other studies that use health diaries. In the Netherlands, there exists only one diary study with which comparison is possible [2]. We therefore included two studies from the U.S. [3, 4]. One of these studies and the Dutch study use longer periods of diary keeping, which rules out a direct comparison. A comparison can however be made, by extrapolating the occurrence rate for three weeks to an estimated occurrence rate for four or six weeks. For this purpose we fitted several versions of the (piecewise) exponential model to our data [5]. This model assumes that the duration or 'survival' without complaints during the diary keeping period is governed by a process in which the probability of obtaining complaints (the hazard rate) is constant over (certain parts of) the diary keeping period. The following equation, shows the survival for either the entire period or pieces of that period.

$$S(t) = \exp(-b*t) - a;$$

$S(t)$ = survival without having had symptoms in the diary keeping period; a = intercept indicating proportion surviving until certain days in different versions of the model; b = hazard rate; indicating the daily decline in the proportion surviving.

Three versions of the piecewise model were estimated with unconstrained non-linear regression, using different time intervals to estimate the regression coefficients (day 1–21; day 8–21; day 15–21). A model without intercept: $S(t) = \exp(-b*t)$ was estimated with OLS regression of Ln $\{S(t)\}$ or non-linear regression of $S(t)$. This model uses the entire period of day 0–21, assuming the absence of symptoms before the beginning of the study.

Several models were estimated because the hazard rate (b) may not be the same during the diary keeping period: at the beginning of the study respondents may become sensitized to monitoring symptoms whereas towards the end, respondents might grow weary of completing the diary every day. The complement of the estimated survival curve, the cumulative distribution function: $F(t) = 1 - S(t)$, allowed us to estimate the percentage of the respondents with complaints in the periods of either four or six weeks.

The next table shows the results of several health diary studies (percentages of the sample with symptoms) along with estimates of these percentages for similar sub-samples of data from the Dutch Survey of General Practice.

This last comparison of this table demonstrates that in particular the percentage of elderly respondents with health complaints in a period of three weeks is relatively low in the Dutch data. It should be noted that the study of Stoller *et al.* uses health diaries with a checklist whereas the Dutch data use open-ended questions. The comparison with other studies has to be made with extrapolations of the Dutch data. An extrapolation of the survival function of the entire 3 week period leads to results that are similar to the diary studies that asked their respondents to keep the diary for 4 or 6 weeks. When on the other hand only the last week of the diary keeping period is used, which is more realistic for extrapolation than the entire period, the results do differ. Extending the hazard rate of the last week leads to lower estimates of the percentage of respondents who will have had complaints in 4 or 6 weeks. This means that towards the end of the diary keeping period in the Dutch study only few respondents (of those who did not have recorded complaints during the previous days) have mentioned complaints. An extension of our diary study to 4 or 6 weeks would most likely lead to a lower percentage of respondents with health complaints than is observed in other studies.

There are a number of plausible explanations for the lower rates. One of them is the lay-out of the diary: on days with complaints there was more work for the respondent than on days without symptoms. Respondents were in this way rewarded for not entering their complaints. They probably entered symptoms that were unequivocally experienced as illness, while skipping minor symptoms. Verbrugge's diary study also meant more work on days with symptoms, but these differences were relatively small in comparison to our study [6]. Another difference is that in our study only complete records are used, whereas Verbrugge used data from respondents who completed the diary for one week or more. The analysis of cooperation with our study showed that respondents who completed only part of the study are in poorer health than respondents who completed the diary for the entire three weeks. The other Dutch study of Van der Lisdonk used a health diary that is very similar to ours. In fact, Van de Lisdonk was consulted for the design of our study. His study however, differed in two ways. First, his sample of respondents was not representative, because he aimed at having equally large groups of high, low and average users of medical care. Second, his study was carried out in January and February, whereas our study took place in different periods throughout the year. It is very likely that in wintertime more people have health complaints.

Despite the lower rates in our study, it should be noted that our diary study with open-ended questions did indeed show a higher probability of recording symptoms than health interviews with similar open-ended questions (ref. Table 1). When using the same method of query there is some 'gain' in using a diary. Unfortunately, our data do not allow us to compare diaries with questionnaires that both use checklists.

References

1. Verbrugge L. M. Health diaries. *Med. Care* **18,** 73, 1980.
2. Van de Lisdonk E. H. Ervaren en aangeboden morbiditeit in de huisartspraktijk. (dissertation). Nijmegen, 1985.
3. Stoller E. P., Forster L. E. and Portugal, S. Self-care responses to symptoms by older people. *Med. Care* **31,** 24, 1993.
4. Verbrugge L. M. and Ascione F. J. Exploring the iceberg. Common symptoms and how people care for them. *Med. Care* **25,** 539, 1987.
5. Namboodiri K. and Suchindran C. M. *Life Table Techniques and their Applications*. Academic Press, New York, 1987.
6. Verbrugge L. M. Triggers of symptoms and health care. *Soc. Sci. Med.* **20,** 85, 1986.

2

UTILIZATION OF CARE

WHAT DOES A DOCTOR DO WITH PSYCHOSOCIAL PROBLEMS IN PRIMARY CARE?

P. F. M. VERHAAK, PH.D.

H. J. WENNINK, PH.D.

Netherlands Institute for Primary Health Care (NIVEL), Utrecht

ABSTRACT

Though a lot has been published on the prevalence of psychosocial disorders in primary care, less is known about the actual treatment, given by primary care providers. This article describes treatment given to complaints which are considered by the physician as being psychosocial by nature. Treatment has been assessed by means of observation. A database of approximately 1500 videotaped consultations, sampled from thirty general practitioners has been used for this purpose. The possibilities and limitations of (generalist) psychosocial treatment in primary care are discussed.

(*Int'l. J. Psychiatry in Medicine* 20:151-162, 1990)

Key Words: general practitioner, primary health care, psychosocial problems, treatment.

What mental health care is being delivered in primary care? To answer this question we need an understanding of the way primary care providers actually treat the psychosocial problems which are presented to them.[1] This understanding is lacking at the moment. Though a lot of research has been done

[1] For the sake of convenience we use the term "psychosocial problems" to indicate all not purely physical problems. In this way the term designates psychiatric as well as milder psychoneurotic and social problems; they might be presented as mental complaints as well as co-existent with somatic complaints.

about the psychosocial morbidity presented and recognized in primary care, not so much is known about the way a family physician or general practitioner handles those problems [1]. Knowledge about primary health care treatment of psychosocial problems is based upon small numbers of cases or data from self-registration (cf. for e.g., the data from the National Ambulatory Medical Care surveys, [2]). In order to get a more reliable and detailed picture of mental health treatment in primary care, we studied doctor's behavior by observation of videotaped consultations, in which psychosocial complaints were presented.

Most of the mental illness episodes identified are initially presented in the primary care sector. A majority of these problems is not treated in any other health care sector [3, 4]. Although many episodes are presented explicitly as mental problems (anxiety, depressive feelings), there is also a substantial proportion of cases with somatic comorbidity [2, 5]. Assessing and treating these cases does seem to pose a substantial problem to the average physician. Some authors consider this comorbidity as one of the major reasons for primary care physicians' failure to detect mental illness, as the somatic presentation may mask the psychological background [5–7]. Primary health care plays a substantial role in the treatment of both psychosocial problems and comorbidity cases. In Dutch research both groups have been estimated as comprising 10 percent to 14 percent of all complaints presented [8, 9]. In Sweden, Swartling et al. report 21 percent psychosocial problems, of which 6 percent were diagnosed mainly psychiatric [10]. This being the case the question arises as to which treatment a primary care physician should offer, should he treat or should he refer? In particular in countries like the United Kingdom and the Netherlands, where specialized mental health care is only accessible after referral by a primary care physician, this seems to be the classical dilemma.

On the one hand, there are indications that psychosocial problems, measured in primary care populations, tend to show a high rate of spontaneous recovery. It is assumed that in using standardized mental illness screening instruments, which are highly sensitive but not very specific, a great many false positives are also detected [1, 5]. This finding seems to argue for a reserved referral policy. On the other hand, it has been argued that patients with depressive disorders for instance, which are very common, should be referred quickly to prevent the problem from becoming chronic [11]. Since in the primary care situation somatic problems are the most common reasons for visits, there may not be sufficient opportunity for psychosocial problems to come to the fore. When both doctor and patient share the same somatic definition of the situation psychosocial problems are likely to be ignored and hence not treated. Is it also probable that, in the case of comorbidity referral is less likely to occur than in clear-cut psychosocial cases. Data from a Dutch study conducted by the Netherlands Institute of primary health care (NIVEL) have shed some light on this issue.

METHOD

To study doctors' treatment of psychosocial problems a research was conducted among thirty Dutch general practitioners. The data consisted of 1569 videotaped doctor-patient consultations, and a questionnaire about each contact filled in by the GP. The reason for visit, as put forward by the patient, was derived by observing and coding the encounter by means of the Reason for Encounter classification, RFE [12], a predecessor of the ICPC, the International Classification of Primary Care [13]. The physicians' assessment of the psychosocial character of the RFE was based on a four-point scale, containing the following categories:

1. strictly somatic,
2. somatic with psychosocial side-effects,
3. somatic presentation, with a suspected psychosocial background,
4. mainly psychosocial

This categorization is derived from McWhinney in a slightly adapted form [14]. Goldberg and Bridges mention a categorization which similarly tries to do justice to the fact that many complaints are presented somatically (and accordingly have a somatic RFE) while there is a suspicion of a psychosocial background [15].

The consultations were sampled by videotaping all doctor-patient encounters of thirty Dutch general practitioners until sixty recordings had been made. About 15 percent of the patients refused to participate, 16 percent of these patients had psychosocial problems. The sample of thirty GP's appeared to be slightly more oriented towards general-medicine, as opposed to specialist medicine, and had attended more post-graduate training courses on applied medical-psychology than the average Dutch GP [16]. An implication of this sample characteristic could be a psychosocial bias in the sense that the selected GP's show a greater tendency to focus on the psychosocial aspects than more clinically oriented colleagues.

The treatment of psychosocial problems was coded by the observers[2] with the following categories:

- prescribing psychotropic drugs,
- passive counseling (encouraging, comforting, listening),
- active counseling (exploring, giving insight),
- referral to somatic specialists,
- referral to mental health agencies,
- advice.

[2] For a more detailed description of the complete observation procedure, consult [16].

To assess the inter-observer reliability, a number of consultations was observed by all five of the observers. The correspondence between pairs of them (*N* of pairs varied between 20 and 100) varied between 81 percent and 94 percent. Cohen's Kappa, which accounts for correspondence, attributable to chance, varied between .59 and .76, with one exception: prescribing of psychotropic drugs, which yielded a Kappa of only .26. This last feature has been observed again in all those consultations with non-somatic assessments.

RESULTS

The GP Caseload

Table 1 shows the way the complaints were assessed. We will start the analysis with two so-called background variables, sex and age, to determine whether the assessment might be influenced by patient characteristics.

When broken down by sex there appeared to be no significant effect (Table 2). Although about two-thirds of the complaints are presented by women there is not much sex-difference in the distribution of assessments. Looking at the assessments we see a slight overrepresentation of strictly somatic assessments for men and mixed assessments for women. Having established no substantial

Table 1. Assessment of 2001 Complaints Presented to Thirty General Practitioners in a Dutch NIVEL Study, 1986

Assessment by GP	*N*	*Percent*
1. Strictly somatic	1057	53
2. Somatic with psychosocial side effects	353	18
3. Somatic with psychosocial background	326	16
4. Mainly psychosocial	265	13

Table 2. Assessment of 2001 Complaints Presented to Thirty GPs Broken Down by Sex

	Male Patients		*Female Patients*		
Assessment	*N*	*(Percent)*	*N*	*(Percent)*	*Total*
1. Strictly somatic	409	56	648	51	1057
2. Somatic with psychosocial side effects	115	16	238	19	353
3. Somatic with psychosocial background	115	16	211	17	326
4. Mainly psychosocial	94	13	171	13	265
Total	733		1268		2001

$\chi^2 = 4.72; p = 0{,}2$

Table 3. Assessment of 1999 Complaints Presented to Thirty GPs Broken Down by Age

	Patients < 40		*Patients > 40*		
Assessment	*N*	*(Percent)*	*N*	*(Percent)*	*Total*
1. Strictly somatic	594	57	461	48	1055
2. Somatic with psychosocial side effects	156	15	197	21	353
3. Somatic with psychosocial background	164	16	162	17	326
4. Mainly psychosocial	128	12	137	14	265
Total	1042		957		1999

$\chi^2 = 30.1; p = 0{,}001$

relationship between the sex of the patient and the assessment of the complaint we turn to the influence of the "age" variable.

There is a greater difference in age groups. Most detected psychosocial problems are presented by middle-aged patients (40 to 59 years). The somatic-psychosocial ratio decreases with an increase of age, i.e., the average doctor encounters more psychosocial problems in the older age-groups. The data seem to suggest a tendency to perceive more somatic problems when dealing with younger patients and more psychosocial problems when dealing with older patients. This difference in assessment might also be due to a greater number of psychosocial complaints being presented by older patients. We will enlarge on this issue later on when discussing the actual treatment given. Now we will turn our attention to what could be called the crucial issue in the doctor-patient relationship, that is the degree of consonance between the-complaint-presented and the-complaint-perceived. In other words: do patient and doctor agree upon the nature of the problem (Table 4).

Table 4 shows that of all 2001 complaints a mere 8.7 percent were presented as psychosocial problems. A much larger percentage, however, was perceived as one. This could mean that the average general practitioner is well aware of psychosocial comorbidity, it does not mean that he actually treats the psychosocial aspect; this rather vital difference will be discussed later on. Although one might conclude that on the whole doctors and patients seem to agree, their agreement is restricted to the purely somatic cases. Doctor and patient do seem to agree on the true nature of these complaints. There appears to be however an interesting difference of opinion concerning the comorbidity cases (values 2, 3 and 4 in the first column). Especially noteworthy are the 6 percent somatic RFEs which were assessed as mainly psychosocial. Whereas the table shows 8.7 percent psychosocial RFEs (the *patients'* presentation), from the point of view of the *physician* a substantial 47 percent of the problems had a psychosocial aspect to them.

Table 4. Relationship between 2001 Reasons for Encounter (As Stated by the Patient) and Assessment on the Somatic-Psychosocial Axis by Thirty GPs

	Reason for Encounter (patient)					
	Somatic		*Psychosocial*		*Total*	
Assessment (GP)	*N*	*(Percent)*	*N*	*(Percent)*	*N*	*(Percent)*
Strictly somatic	1057	53	0	0	1057	53
Somatic/psychosocial side effects	350	17	3	0	353	18
Somatic presentation/suspicion of psychosocial background	305	15	21	1	326	16
Mainly psychosocial	115	6	150	7	265	13
TOTAL	1827	91	174	9	2001	

GP TREATMENT

We now turn to the main issue of this article: what is the treatment given, especially in comorbidity cases? A clear distinction can be seen in Figure 1 between complaints considered to be mainly psychosocial and those considered to be somatic with psychosocial side effects or background. The former are seldom neglected, the latter in about 40 percent of the cases.

Figures 2 through 4 present a more detailed picture of treatment for each of the three non-somatic assessments apart.

Complaints which are assessed "somatic, with psychosocial aspects" were relatively often referred to somatic specialists. Referral seems to decrease as organic relevance in the assessment decreases. Prescribing and advising on the other hand increase when psychosocial aspects come into play. A somewhat startling pheonomenon is the relatively high rate of psychotropic prescriptions in the case of somatic RFEs with a possible psychosocial origin. Referral to mental health agencies seems to be restricted to mainly psychosocial complaints.

When the physician actually treats psychosocial complaints, this mostly (80% of the cases) comes down to some or other form of counseling, or a good heart to heart talk. Interestingly enough, this counseling had the most active, exploratory character when the complaint was perceived as being somatic with a possible psychosocial background. Analysis of variance did not show significant differences between men and women. The only difference concerns the active-passive counseling rate, i.e., doctors were more passive towards male patients, given a somatic complaint with psychosocial side-effects, and more active towards female patients with somatically presented complaints with a suspected psychosocial background.

Differences between age groups were more marked. The response to psychosocial problems appears to be proportionate to the age of the patient,

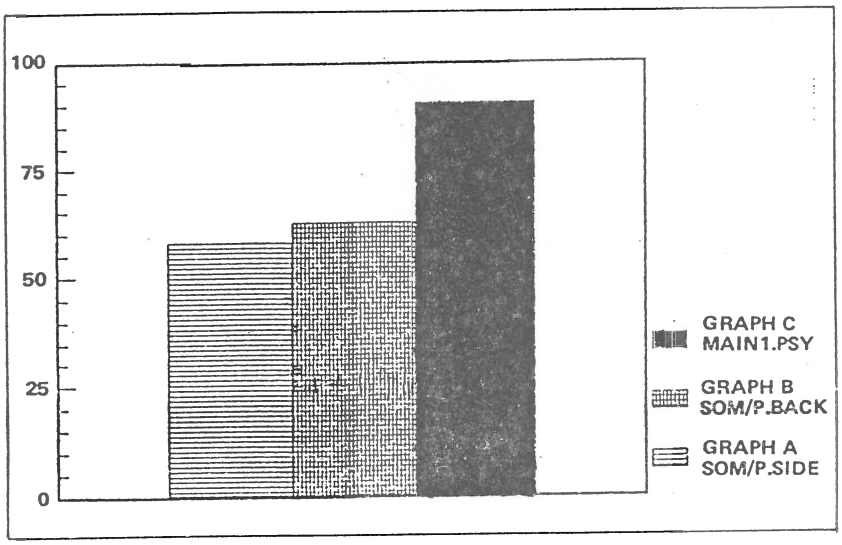


Figure 1. Proportion reactions to complaints assessed as A: somatic/ps.socs side-effects; B: somatic/possible ps.soc background; C: mainly psychosocial.

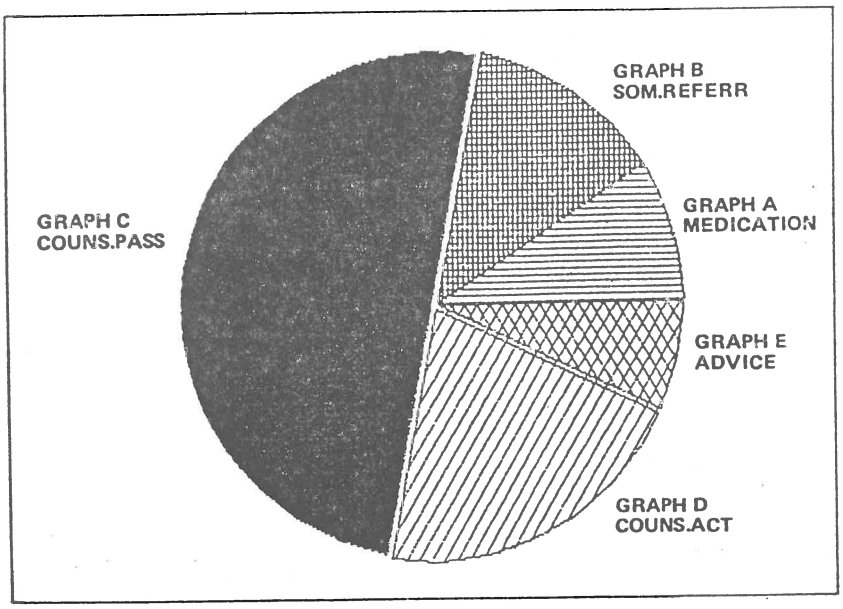


Figure 2. Treatment when complaint is "somatic with psychosocial side-effects."

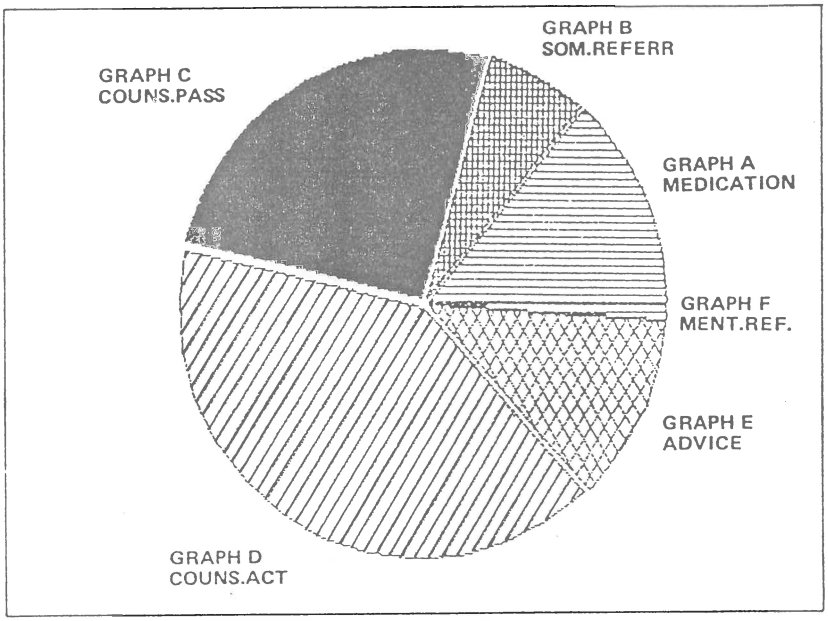


Figure 3. Treatment when complaint is "somatic with possible psychosocial background."

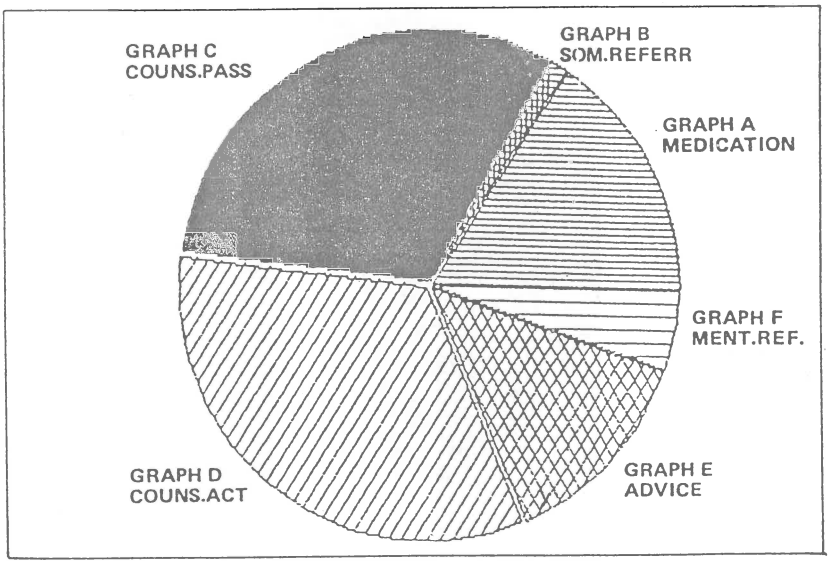


Figure 4. Treatment when complaint is "mainly psychosocial."

i.e., more response was observed in the younger age groups. Treatment consists in the younger categories less often in medication, more often in counseling. The counseling itself has a more active nature when the patient is younger.

DISCUSSION

We can summarize the findings of our research as follows. Psychosocial complaints in primary care cover a greater area then one would suspect from the number of psychosocial reasons for encounter, put forward by the patient. The data direct our attention to two related phenomena typical of the average primary care situation. In the first place the doctor perceives more problems as being psychosocial by nature than are actually presented as such. In the second place much of his time is devoted to treating comorbidity. These findings are in accord with the reported prevalence rates for mental illness, measured by screening-instruments; these rates are higher than the number of psychosocial reason for visit a physician reports [1]. They are also in accord with Jencks' findings which showed that a lot of therapeutic listening and psychotopic drugs prescribing occurs without a psychiatric diagnosis. Schurman et al. [2] have pointed to a sharp contrast between the caseload of non-psychiatric agencies and psychiatric agencies: 27.8 percent psychosocial complaints as compared to 72.2 percent. Goldberg and Bridges [15] find 26 percent not entirely physical illnesses among 590 consecutive new patients in general practice, of which only 5 percent are entirely psychiatric illnesses, while 21 percent are labeled "physical illness with secondary psychiatric illness," "unrelated physical and psychiatric illness," and "psychiatric illness with somatic symptoms."

Our findings underline the problematic character of this average-GP-caseload and highlight the unclear situation in which he has to practice his art.

Contrary to the psychiatrist who may devote himself to clear cut psychiatric problems, the primary care physician has to deal with comorbidity. Furthermore it seems to be the case that doctor and patient often start their encounter with rather different views of the nature of the problem.

Purely psychosocial complaints seldom appear to be totally neglected; at least 90 percent got attention. On the other hand, the limitations of primary care became quite clear also: the main focus of treatment is directed to consolation, comfort or reassurance. Results which are in accord with the differences Schurman et al. found between non-psychiatrist and psychiatrist treatment [2]. Non-psychiatrists restrict themselves much more than psychiatrists do, to counseling and prescribing drugs. Psychotherapy in the strict sense of the word was hardly observed. One remarkable finding was the rare occasion that a patient was referred to a psychiatrist or psychotherapist. Detailed analysis showed that many patients with complaints like depression and anxiety were very reluctant to get more specialized help [18]; they considered mental health

institutions anonymous and stigmatizing and preferred their well known family doctor, unless their suffering was unbearable. The same preference for a family doctor was found by Kiraly a.o [19].

In the second place we considered the cases where a somatic presentation by the patient is accompanied by a suspicion of a psychosocial background. This category occurs rather frequently; in view of the often mentioned poor performance of general physicians with respect to detection of hidden mental illness [20] we may assume that this category is in fact much larger. It is reassuring that the GP often manifests an exploratory tendency in respect of this group of patients. It is alarming though that in 40 percent of the cases no action at all was taken. Goldberg distinguishes between patients with psychosocial distress who do not require specific intervention and those who do [21]. The group which does need intervention is described as having affective disorders like anxiety and depression, often accompanied by somatic complaints which require intervention. Our figures indicate that a number of these patients get some attention, but it may be too little. The fact that only one patient in this group was referred to a mental health agency indicates the lack of confidence a GP has in specialized mental health care, in cases that are masked by a somatic presentation. It remains to be seen however whether this is due to general practitioners' failure to recognize the necessity of referral or to the lack of appropriate facilities to refer to. More research is needed to establish the value of different therapeutic approaches, depending on the presentation of the complaint and with an emphasis on recognition and treatment of psychological complaints, masked by comorbidity.

The suggestion seems to be that the average GP does indeed treat many psychosocial problems, but mostly when the patient has explicitly stated a psychosocial reason for encounter. That is to say only when the situation is clearly defined in psychosocial terms. In the more ambiguous case of comorbidity the GP seems to have a tendency to neglect the psychosocial side of the problem in a considerable amount of cases (up to 50%). When an appeal is made to his time-honored skills as a physician he pays more attention to the somatic aspects. But then again, our findings also indicate not so much a failure to detect, but a tendency to neglect those aspects in favor of somatic problems: a doctor's prediliction.

The association between age and treatment seems to be congruent with the overall picture in the psychiatric field where as a whole more attention is paid to younger patients (the ideal Young, Attractive, Verbal, Intelligent and Sociable client). If we combine this finding with the smaller prevalence of psychosocial assessments of younger patients, we may conclude that younger patients are less at risk for psychosocial problems, but get a more active treatment in case they are assessed as such. The opposite holds true for older patients: the GP has a tendency to focus on the somatic side of their problem, though the chances on psychosocial (co-)morbidity are higher.

To determine the actual effects of the treatment given, a longitudinal study is being conducted at this moment. In this research project the doctor-patient relationship is considered to be a negotiation in which both partners have to reach common objectives: diagnosis, treatment, prescription etc. Videotaped consultations will be analyzed to determine the influence of doctors' interactional style, patients' assertiveness and the nature of the complaint on the success of the treatment. To analyze these consultations additional data are being collected:

- registration by sixteen GPs of all consultations with a sample of 100 patients with psychosocial problems during one year,
- a follow-up study to determine the mental health, attitudes and beliefs of the patient sample.

From this study we hope to deduce possible strategies for the general practitioner to cope with the treatment of psychosocial problems, be they accompanied by somatic comorbidity or not.

REFERENCES

1. H. C. Schulberg and B. J. Burns, Mental Disorders in Primary Care: Epidemiologic, Diagnostic and Treatment Research Directions, *General Hospital Psychiatry*, *10*, pp. 79–87, 1988.
2. R. A. Schurman, P. D. Kramer, and J. B. Mitchell, The Hidden Mental Health Network. Treatment of Mental Illness by Non-Psychiatrist Physicians, *Archives of General Psychiatry*, *42*, pp. 89–94, 1985.
3. D. I. Regier, D. Goldberg, and C. Taube, The Defacto U.S. Mental Health Services System, *Archives of General Psychiatry*, *35*, pp. 685–693, 1978.
4. D. B. Kamerow, Research on Mental Disorders in Primary Care Settings: Rationale, Topics and Support, *Family Practice Research Journal*, *6*, pp. 5–11, 1986.
5. H. C. Schulberg, N. Saul, M. McClelland, M. Ganguli, W. Christy, and R. Frank, Assessing Depression in Primary Medical and Psychiatric Practices, *Archives of General Psychiatry*, *42*, pp. 1164–1170, 1985.
6. D. Goldberg, Detection and Assessment of Emotional Disorders in a Primary Care Setting, *International Journal of Mental Health*, *8*, pp. 30–38, 1979.
7. M. R. Eastwood and M. H. Trevelyan, Relation between Physical and Psychiatric Disorder, *Psychological Medicine*, *2*, pp. 363–372, 1972.
8. F. J. A. Huygen, H. J. M. v. d. Hoogen, A. Th. v. d. Logt, and A. J. A. Smits, Nerveus-functionele klachten in de huisartspraktijk, *Nederlands Tijdschrift voor Geneeskunde*, *128*, pp. 1321–1327; 1372–1376, 1984.
9. H. Lamberts and B. Hartman, Psychische en sociale problemen in de huisartspraktijk, *Huisarts en Wetenschap*, *25*, pp. 333–342; 376–393, 1982.
10. P. G. Swartling, L. Kebbon, and B. Smedby, Mental Health Problems in Primary Health Care as Seen by Doctors, *Scandinavian Journal of Primary Care*, *5*, pp. 201–204, 1987.

11. H. C. Schulberg, M. McClelland, and B. J. Burns, Depression and Physical Illness: The Prevalence, Causation and Diagnosis of Comorbidity, *Clinical Psychology Review*, *7*, pp. 145–167, 1987.
12. J. M. Bensing, The Use of the RFE-classification System in Observation Studies—Some Preliminary Results, *Proceedings of the 10th WONCA Conference*, Singapore, 1983.
13. H. Lamberts and M. Wood, *ICPC*, Oxford Medical Publications, Oxford, 1986.
14. I. R. McWhinney, Special Article beyond Diagnosis. An Approach to the Integration of Behavioural Science and Clinical Medicine, *The New England Journal of Medicine*, *283*, pp. 384–387, 1972.
15. D. Goldberg, and K. Bridges, Screening for Psychiatric Illness in General Practice: The Genreal Practitioner Versus the Screening Questionnaire, *Journal of the Royal College of General Practitioners, 37,* pp. 15-18, 1987.
16. P. F. M. Verhaak, Variations in the Diagnosis of Psychosocial Disorders: A General Practice Observation Study, *Social Science and Medicine*, *23*, pp. 595–604, 1986.
17. S. F. Jencks, Recognition of Mental Distress and Diagnosis of Mental Disorder in Primary Care, *Journal of the American Medical Association*, *253*, pp. 1903–1907, 1985.
18. J. T. van Busschbach and P. F. M. Verhaak, Hoe hanteren huisartsen psychische klachten, *Maandblad Geestelijke Volksgezondheid*, *41*, pp. 475–492, 1986.
19. D. A. Kiraly, C. J. Coulton, and A. Graham, How Family Practice Patients View Their Utilization of Mental Health Services, *The Journal of Family Practice*, *15*:2, pp. 317–323, 1982.
20. D. Goldberg and P. Huxley, *Mental Illness in the Community*, Tavistock Publications, London, 1980.
21. D. Goldberg, The Concept of a Psychiatric "Case" in General Practice, *Social Psychiatry*, *17*, pp. 61–65, 1982.

Direct reprint requests to:

Peter F. M. Verhaak
Netherlands Institute for Primary Health Care
P.O. Box 1568
3500 BN, Utrecht
The Netherlands

Dutch General Practice Care for Patients with Epilepsy: Results from the Dutch National Survey of Morbidity and Interventions

LOEK STOKX, DIEDERIK KERSTEN AND KOOS VAN DER VELDEN

Stokx L, Kersten D and van der Velden K. Dutch general practice care for patients with epilepsy. Results from the Dutch National Survey of Morbidity and Interventions. *Family Practice* 1991: **8**: 125–128.

This study aimed to assess the share which Dutch general practice has in the care of patients with epilepsy. During a 3-month period 400 000 patient contacts in 103 general practices with a total list of 335 000 patients were registered, 1536 of these, concerning 1059 patients, concerned epilepsy.

Contacts with patients with known epilepsy in Dutch General Practice were handled mainly by the practice nurse and most involved repeat prescriptions. Patients suspected of having epilepsy had more attention from the General Practitioner. The involvement of General Practitioners in the care of epilepsy was found to be small, but not unimportant. GPs are in a crucial position as regards the detection of epilepsy. They can enlarge their role in respect of patients with known epilepsy and improve continuity of care.

INTRODUCTION

In this study we were concerned with the way in which patients with epilepsy are seen by general practices in the Netherlands. Not much research has been done on this subject in the Netherlands; however the impression conveyed is that specialist facilities are far more important and the role of the general practitioner minimal. Our primary concern was to assess this assumption about the role of the Dutch GP.

First we will give some background information about specialist facilities in the Netherlands for epilepsy. There are about 30 hospital-based neurologists per million inhabitants (compared to the UK, with about 3 per million[1]). Secondly, there are three specialized epilepsy hospitals in the Netherlands (total population about 15 million) with a total capacity of 1350 beds. There are also 15 outpatient consultation services connected to these hospitals spread throughout the country: these give care to the more serious and difficult cases. At the moment such care is supplied to an estimated 13% of the patients with epilepsy in the Netherlands.[2]

A questionnaire showed that Dutch GPs do not know who is in charge of follow-up for 78% of the epileptic patients in their practices.[3] Two other questionnaire surveys conclude that the role of the Dutch GP seems to be small, and that patients want it extended.[4,5] Thus far, the care given to patients with epilepsy by Dutch general practice has not been recorded, and our principal research question was 'what share does Dutch general practice have in the care of patients with epilepsy?'

METHODS

We have used data from the Dutch National Survey of Morbidity and Intervention in General Practice,[6] a research project which is intended to assess the role of general practice within the Dutch health care system.

An important instrument of the National Survey is recording contacts, an exercise in which 163 GPs and their practice nurses (in the Netherlands, this is mainly an administrative position) have cooperated. The National Survey covers 103 practices all over the country with a total list of 335 000 patients. All contacts with patients during a 3-month period, between April 1, 1987 and March 31, 1988 were recorded in terms of the main complaints, the (differential) diagnoses made by GPs, underlying illness(es), diagnostic procedures, medication prescribed, referrals, other treatment measures, contact with other care providers and follow-up appointments. The complaints, the (differential) diagnoses and the underlying illness(es) in these contacts were coded with the help of the International Classification of Primary Care (ICPC).[7]

Netherlands Institute of Primary Health Care, Utrecht, The Netherlands. Correspondence to: Dr L J Stokx, Netherlands Institute of Primary Health Care, PO Box 1568, 3500 BN Utrecht, The Netherlands.

From the 400 000 contacts gathered in this way, we have analysed those for which the (differential) diagnoses or the underlying illness were recorded as epilepsy (all types)—neonatal seizures were excluded.

RESULTS

Activities in General Practice

A total of 1536 contacts concerning 1059 patients were identified as above. After correction for differences between sample and population in terms of urbanization grade, health region and distance to the nearest hospital,[6] and assuming there are no major seasonal influences on epilepsy, we find that one out of every 250 contacts in Dutch general practice concerns epilepsy. The average Dutch practice has about 40 epilepsy contacts per year.

Table 1 shows the diagnostic, therapeutic and follow-up activities undertaken in general practice for the patients in our survey. Diagnostic work-up refers to physical examination and simple diagnostic procedures done in general practice; external diagnostic refers to tests in laboratories outside general practice. Whenever the conversation between patient and health care provider was used as therapy (therapeutic talks, reassurance, patient education), it was called therapeutic counselling. Drug prescription is divided into new and repeat prescriptions: a repeat prescription can have been prescribed initially by either the GP or the specialist. As time passes the distinction is often unclear. 'Specified' follow-up appointments were distinguished from 'routine' follow-up appointments. A follow-up appointment was defined as 'specified', when diagnostic measures were taken in the doctor–patient contact, in the case of therapeutic counselling or a decision to use a new prescription. All other follow-ups were defined as 'routine'. Referral to specialist care was divided into new referrals and repeat referrals. Contact with specialist refers to direct (mostly telephone) contact between GP and a hospital specialist about a certain patient. The first column gives figures for all the patients in our survey, while the second to fourth columns divide patients into suspected epilepsy, known epilepsy and underlying epilepsy. The group of patients with known epilepsy is by far the largest. A diagnostic work-up was performed in 57% and 60% of patients with suspected epilepsy or underlying epilepsy, respectively. Patients with known epilepsy had a diagnostic work-up much less often (5%). The same pattern is noticeable for external diagnostics, therapeutic counselling, new prescription, specified follow-up and new referrals. All of these actions, involving a clear decision in general practice about the kind of treatment, are far less often seen in patients with known epilepsy. On the other hand, repeat prescriptions and repeat referrals are given to more patients with known epilepsy than those in whom epilepsy is either suspected or underlying. Not only specified follow-up but also routine follow-up appointments are made much more often with patients suspected epilepsy than with patients with known epilepsy. The same goes for contacts with specialists.

Main Complaints Related to Epilepsy

The main complaints of patients suspected of epilepsy and with known epilepsy are presented in Table 2.

The most important complaint in the group with 'suspected epilepsy' was fainting/loss of consciousness (27%). There was also a large group of other/miscellaneous complaints (16%). The most important reason by far for patients with known epilepsy attending general practice was a request for antiepileptic drugs followed by a request for a specialist referral card.

Patients with known epilepsy who in fact visit the general practice for complaints other than their epilepsy (underlying epilepsy group) present a wide range of complaints not shown in the table. Their complaints, as well as the diagnoses, are different from the total sample of the population in the Dutch National Survey. Neurological, psychiatric and pregnancy, childbearing and family planning are related problems encountered significantly more often, as are general and social problems (z-scores between 2.0 and 7.6, $P<0.05$).

Referral to Specialists (Table 3)

Of all the patients in our survey 19% were referred to a specialist. Of the three groups, this percentage is high-

TABLE 1 *Activities related to epilepsy performed in general practice. Percentage of patients*

Activities:	All survey patients (n=1059)	Suspected epilepsy (n=171)	Known epilepsy (n=809)	Underlying epilepsy (n=197)[a]
Diagnostic work-up	23	57	5	60
External diagnostics	3	6	1	6
Therapeutic counselling	29	59	11	68
New prescription	9	13	3	28
Repeat prescription	64	9	80	12
Specified follow-up	19	39	5	51
Routine follow-up	9	16	5	5
Contact with specialist	4	9	2	5
New referral	11	37	3	15
Repeated referral	7	6	9	1

[a] 118 of these patients are also seen for their known epilepsy as such in the same period.

TABLE 2 *Main complaints of patients with suspected and known epilepsy in general practice. Percentage of patients*

Complaints/reason for encounter:	Suspected epilepsy (*n*=171)	Known epilepsy (*n*=171)
Request for anti-epileptic drugs	6	79
Request for a referral card	2	9
Follow-up	4	4
Fainting/loss of consciousness	27	1
'Do I have epilepsy?'	8	0
Convulsions	7	1
Vertigo	6	0
Abnormal movements	5	0
Neurological counselling	3	1
Other neurological complaints	8	2
Psychiatric complaints	8	1
Other/miscellaneous complaints	16	2
Total	100	100

est for those with suspected epilepsy (43%). Most patients with suspected and known epilepsy were referred to a neurologist: the pediatrician comes in a long way behind in second place. Patients with underlying epilepsy were mostly referred to other specialists. As far as the intention of the referral is concerned, we can say that in more than half of the cases it was to assess diagnosis as well as treatment.

GP and Practice Nurse Activities for Patients With Known Epilepsy

The largest group of patients was made up of those known to have epilepsy (N = 809). These patients are not always seen by the GP himself: they can also visit or call the practice nurse directly, and 77% were taken care of by the practice nurse. The GP sees 19% himself, and 4% are seen by both of them. Table 4 shows the details of each of their activities.

TABLE 3 *Referral to hospital patients concerning epilepsy. Percentage of patients*

Specialist:	All patients (*n*=1059)	Suspected epilepsy (*n*=171)	Known epilepsy (*n*=809)	Underlying epilepsy (*n*=197)
Neurologist	13	33	10	4
Pediatrician	2	5	1	1
Internist	0	1	0	0
Psychiatrist	1	1	0	2
Other specialist	3	3	1	9
Subtotal: total referrals	19	43	12	16
no referrals	81	57	88	84
Total:	100	100	100	100

The most frequent GP intervention is writing repeat prescriptions (48%); therapeutic counselling comes second (43%) and diagnostic work-ups, mostly neurological examination combined with taking blood pressure, comes third (23%). External diagnostics (blood level monitoring) are performed rarely (3%). One in every six prescriptions was new. Follow-ups were given to 40% of the patients. In 23% of the cases a specified follow-up was given and in 17%, routine follow-up: 13% of the patients were given a referral. Contact with a hospital specialist about the patient was made in 9% of the cases. These contacts were in the nature of a consultation or were to arrange a hospital admission.

The interventions of the practice nurse are limited to giving repeat prescriptions for antiepileptic drugs and handing out repeat referrals (87% and 9% respectively).

DISCUSSION

The average Dutch general practice will have about 40 contacts for epilepsy annually; about three-quarters of these will concern patients with known epilepsy. Most of these patients are handled by the practice nurse and are mainly seen for repeat prescriptions of anti-epileptic drugs. Only a small number of these patients are also seen by the GP in the same period. The patients with known epilepsy, seen by the GP himself, generally come for repeat prescriptions, follow-up or for different neurological complaints. As well as a repeat pre-

TABLE 4 *Activities performed by the GP and the practice nurse for patients with known epilepsy. Percentage of patients*

Activities:	Known epilepsy patients (all) (*n*=809)	Known epilepsy patients seen by the GP (*n*=187)	Known epilepsy patients seen by the practice nurse (*n*=657)
Diagnostic work-up	5	23	0
External diagnostics	1	3	0
Therapeutic counselling	11	43	1
New prescription	3	11	1
Repeat prescription	81	48	87
Specified follow-up	5	23	0
Routine follow-up	5	17	1
Contact with specialist	2	9	0
New referral	3	9	1
Repeated referral	8	4	9

scription, they receive therapeutic counselling, diagnostic work-ups and other forms of care. The GP is seldom involved in the start or change of anti-epileptic drug regimens. Blood level monitoring was only rarely used by GPs in our survey, as was also found earlier in British and Dutch studies.[4,8]

In patients with suspected epilepsy, the GP performs physical examinations most of the time. A large number of these patients are referred to a specialist, generally to the neurologist, for further diagnosis and treatment. The GP sees more than half these patients again.

Our survey gives support to the thesis that the Dutch GP is little involved in the diagnosis and treatment of epilepsy. On the other hand however, we see that the GP (discounting the merely administrative interventions of the practice nurse) has a role as a gatekeeper. In other words, it is the GP who has to deal with patients with fairly general complaints like fainting and vertigo and the question, 'do I have epilepsy?' first. The GP has to consider next whether this patient could have epilepsy or something else and whether the patient should be referred to a specialist. Furthermore, patients with known epilepsy visit the GP not only to seek help for their epilepsy as such but also for other complaints or illnesses. In these cases the GP has an excellent opportunity of monitoring the epilepsy as well as associated problems.

Compared to the Netherlands much more research on the care for epileptic patients has been done in the UK.[1,8–19] It has been suggested that the British GP has an important role in the care of epileptic patients and that the specialist is seldom involved in day-to-day care or in routine follow-up. The main themes of these surveys address the need for better communication between GP and specialist, the GP's role in providing continuity of care, medication evaluation, patient information cards and the planning of specialized local epilepsy clinics. A large cohort study started a few years ago in the UK to acquire more information about somatic and psychosocial aspects of epilepsy as well as the provision of services for epileptic patients.[20]

Although there is a difference in the number of specialized epilepsy facilities favouring the Netherlands and the number of surveys about the care for epileptic patient favouring the UK, the need for a follow-up by the GP is considered important and stressed in both countries.[3–5,8–10,13–15] The Dutch GP can probably expand the follow-up option for known epilepsy patients and secure continuity of care by instructing the practice nurse to guide patients asking for a repeat prescription into the consulting room. Frequently the Dutch GP will not find this guidance necessary because, ample specialist care is at hand and because many patients coming for prescriptions do not have seizures anymore. Precisely these patients however many need a regular follow-up by the general practitioner to establish whether they are (still) being seen by a specialist at all; to possibly reduce or change their medication in consultation with a specialist or to detect and take care of epilepsy related psychological and social problems. Of course this brings extra work, but it may also improve the quality of care.[4,16]

REFERENCES

[1] Sander J W A S, Shorvon S D. Epilepsy in the community. *J R Coll Gen Pract* 1988; **38:** 51–52.

[2] Becker H A, van de Klashorst J P, van Schaik-Colijn H G, Hermkens P L J. *Epilepsiezorg in Nederland 1970/2010.* 1st ed. Heemstede: De Christelijke Vereniging voor de Verpleging van Lijders aan Epilepsie, 1987.

[3] Bongers E, Coppoolse J, Meinardi H, Posthuma E P S, van Zijl C H M. *A survey of epilepsy in Zeeland, the Netherlands.* Nederland 1976. Heemstede: Instituut voor epilepsiebestrijding 'Meer en Bosch'–'De Cruquiushoeve'.

[4] Crobach M J J S, Niezink G M, van der Leden J, Springer M P. Epilepsie en huisarts: toeval of hoofdzaak? *Ned Tijdschr Geneeskd* 1988; **132:** 1888–1892.

[5] Rutgers M J. *Geneeskundige en maatschappelijke aspecten van de zorg voor epilepsiepatiënten in Nederland* (Dissertation). Rotterdam, Nederland: Erasmus Universiteit Rotterdam, 1984.

[6] Foets M, van de Velde J. *Basisrapporten bij de Nationale Studie.* Nederland 1989. Utrecht: NIVEL.

[7] Lamberts H, Wood M. *International Classification of Primary Care.* Oxford: Oxford Medical Publications, 1987.

[8] Goodridge D M G, Shorvon S D. Epileptic seizures in a population of 6000. *Br Med J* 1983; **287:** 641–646.

[9] Burton M H, Williams D D R. Treatment and follow up of patients with epilepsy in two general practices. *Fam Pract* 1986 **3:** 235–239.

[10] Cooper G L, Huitson A. An audit of the management of patients with epilepsy in thirty general practices. *J R Coll Gen Pract* 1986; **36:** 204–208.

[11] Department of Health and Social Security Welsh Office. *People with Epilepsy.* Reid Report, London HMSO, 1969.

[12] DHSS Working Group. *A report to the Department of Health and Social Security, Department of Education and Science and the Welsh Office.* London: HMSO, 1986.

[13] Goodman I. Auditing care of epilepsy in a group practice. *Practitioner* 1983; **227:** 435–436.

[14] Hartropp P. Audit of epilepsy in a group practice. *Update* 1987; 299–303.

[15] Hopkins A, Scambler G. How doctors deal with epilepsy. *Lancet* 1977; 183–6.

[16] Jones A L. Medical audit of the care of patients with epilepsy in a group practice. *J R Coll Gen Pract* 1980; **30:** 396–400.

[17] Taylor M P. Epilepsy in a Doncaster practice: audit and change over eight years. *J R Coll Gen Pract* 1987; **37:** 116–9.

[18] White P T, Buckley E G. The management of epilepsy—an audit of two practices. *Health Bull* 1981; **39:** 82–88.

[19] Zander L I, Graham H, Morell D C, Fenwick P. Audit of care for epileptics in a general practice. *Br Med J* 1979; **ii:** 1035.

[20] Hart Y M, Sander J W A S, Shorvon S D. National general practice study of epilepsy and epileptic seizures: objectives and study methodology of the largest reported prospective cohort study of epilepsy. *Neuroepidemiology* 1989; **8:** 221–227.

Incidence and management of transient synovitis of the hip: a study in Dutch general practice

A S VIJLBRIEF

M A BRUIJNZEELS

J C VAN DER WOUDEN

L W A VAN SUIJLEKOM-SMIT

SUMMARY. Little is known about transient synovitis of the hip in the community. Using data from the Dutch national survey of morbidity and interventions in general practice, a study was undertaken to look at the incidence and management of transient synovitis of the hip in children under 15 years of age. Transient synovitis of the hip was diagnosed in 19 children, 17 of whom were new cases. The mean age of the children was six years six months with a sex ratio of 2.8:1.0 boys to girls. An incidence rate of 1.1 per 1000 person years was calculated. General practitioners prescribed drug treatment for six children and bed rest was advised for six children. Two children were referred for an x-ray examination. Clear follow-up arrangements were made for 16 of the 19 children. It appears that general practitioners preferred to adopt a wait-and-see approach to transient synovitis of the hip rather than referring children for diagnostic ultrasound or x-ray examination.

Keywords: synovitis; hips; morbidity; children and infants.

Introduction

A child presenting with a limp and complaining of pain and restricted movement in the hip represents a diagnostic problem for the doctor. The differential diagnosis may include transient synovitis of the hip, Legg-Calvé-Perthes disease, septic arthritis, osteomyelitis, juvenile chronic arthritis, a tumour or, depending on the child's age, congenital hip dislocation or epiphysiolysis of the femur head.[1,2] Diagnosis is also complicated by the fact that a problem in the hip may manifest itself as a painful knee, while pain felt in the hip may arise from conditions in the lumbar spine. Of all these disorders, transient synovitis of the hip has the highest incidence.

Research concerning the incidence of transient synovitis of the hip is scarce; a literature search revealed only two studies. Landin and colleagues reported an incidence of 2.0 per 1000 person years for children under 15 years of age in Malmö, Sweden.[3] A study of children aged 16 years and under in Helsinki, Finland, reported an incidence of 0.52 per 1000 person years.[4]

Transient synovitis of the hip is a short lived inflammatory condition where there is effusion in the hip joint. In the study by Haueisen and colleagues the symptoms disappeared within one week in 67% of patients, while in 12% the symptoms persisted for more than one month.[5] Transient synovitis of the hip occurs more often in boys than in girls,[5] and is usually found in children between the ages of three and 10 years.

The pathogenesis of the condition remains obscure. As there are no specific diagnostic criteria only a provisional diagnosis of transient synovitis of the hip may be made initially and then confirmed after the symptoms have disappeared.[6,7] Therefore the diagnosis is usually one of exclusion.

In hospital, ultrasound examination is used increasingly as a diagnostic tool to detect hip disorders because of its high sensitivity for demonstrating effusion in the hip joint.[8-10] Meradji and Diepstraten considered pelvic x-ray to be unnecessary in uncomplicated cases with clear sonographic and clinical findings.[9] Bickerstaff and colleagues devised a protocol for the management of hip complaints in children,[10] proposing that ultrasound examination be performed on presentation and, if the effusion persisted for at least 10 days, an x-ray examination should be performed in order to exclude Legg-Calvé-Perthes disease. In this way, the radiation load of children suffering from transient synovitis of the hip could be reduced while other, possibly more serious, hip disorders could still be identified at an early stage. For x-ray as well as for ultrasound examination Dutch general practitioners have to refer the patient to a radiologist. When requested, the radiologist will perform the examination the same day. As not every radiologist has the experience to perform a diagnostic ultrasound examination of the hip in children, general practitioners may have to consult radiologists at major hospitals for this.

Treatment of transient synovitis of the hip consists of a few days of bed rest, preferably with the affected hip in flexion and slight abduction.[1] The *Dutch textbook of orthopaedics* states that if after a fortnight of bed rest the symptoms have not disappeared or if the symptoms recur, further clinical examination is necessary.[11] However, the orthopaedic surgeon, Visser, suggested performing x-ray examination of every child that limped and for whom internal rotation was restricted and/or painful.[1]

The aims of this study were first, to investigate how frequently transient synovitis of the hip was diagnosed by Dutch general practitioners, and secondly to find out how Dutch general practitioners managed these cases.

Method

Between 1 April 1987 and 31 March 1988, the Dutch national survey of morbidity and interventions in general practice was carried out by the Netherlands institute of primary health care.[12] For this survey 103 general practices (161 general practitioners), divided into four groups, recorded all contacts with patients during one of the four successive three month fieldwork periods (contact registration). The practices were a stratified random sample across the whole of the Netherlands.

As well as collecting demographic data on all patients, data were recorded for each consultation (including home visits, follow-up visits and so on) concerning reasons for the consultation, provisional diagnosis, diagnostic tests, treatment, referral to a specialist, paramedic or hospital and any follow-up appointment. Both the reasons given for the visit and the working hypothesis, together with the differential diagnosis were written down by the general practitioner on the contact registration

A S Vijlbrief, MD, general practitioner trainee; M A Bruijnzeels, MA, data analyst; and J C van der Wouden, MA, research coordinator, Department of General Practice, Erasmus University, Rotterdam. L W A van Suijlekom-Smit, MD, paediatrician, Department of Paediatrics, Erasmus University, and University Hospital, Sophia Children's Hospital, Rotterdam, The Netherlands.

Submitted: 4 October 1991; accepted: 18 February 1992.

form and subsequently coded by a researcher using the *International classification of primary care (ICPC).* The version used by the Netherlands institute of primary health care differs slightly from the original classification by Lamberts and Wood.[13,14]

The national survey was based on an episode oriented registration of morbidity. An episode was considered to be the period during which a patient was ill, from the appearance up to the disappearance of the symptoms. If, after a period without symptoms, the patient consulted the general practitioner with the same symptoms, this was considered to be a relapse and the beginning of a new episode. The preceding episode could have taken place prior to the registration period.

Since transient synovitis of the hip is hardly ever found in patients over the age of 14 years, only data for children under 15 years were used in this study. Only those patients for whom the general practitioner had made a provisional diagnosis of transient synovitis of the hip were included. The diagnosis of transient synovitis of the hip was provisionally classified under 'other disorders of the soft tissues and joints' (code L91.9[14]). On closer study of the data, a number of patients with the provisional diagnosis of transient synovitis of the hip were found to have been given the code 'infections of musculoskeletal system not classified elsewhere' (code L70.3[14]). By examining the original forms it could be established whether the general practitioner had actually used the diagnosis of transient synovitis of the hip as a working hypothesis. Patients were included in the study if the first contact of the episode had taken place within the registration period and at least a fortnight prior to the end of the registration period in order to allow study of the follow up of the disease. General practitioners' management of children with transient synovitis of the hip was studied on the basis of data relating to diagnosis, treatment, referral and follow up. The incidence rate was calculated by dividing the number of new cases by the product of the number of children under study and the length of the registration period.

Results

There were 64 198 children under the age of 15 years in the national study. Of these children, 27 462 had a total of 49 309 contacts with their general practitioners during the registration period. Nineteen patients had a provisional diagnosis of transient synovitis of the hip. Two cases were relapses, the patient having had a previous episode which occurred before the registration period. An incidence rate for new cases was calculated to be 1.1 per 1000 person years. There were 14 boys and five girls (sex ratio 2.8:1.0) aged between one year four months and 14 years five months, mean age six years six months. Fourteen of the 19 children (74%) were more than three years and less than 10 years old.

Twelve patients had one consultation with the general practitioner, four patients had two and three patients were seen three times. The total number of contacts was therefore 29. The symptoms presented by the patients at the initial contact, as written down by the general practitioner, are listed in Table 1. Five patients complained of two symptoms and the other 14 reported one symptom. At the initial contact every patient had a history taken and a physical examination of the hip performed. The general practitioner also performed a physical examination of the hip at subsequent consultations. One patient had an x-ray following the initial consultation with the general practitioner and another patient had an x-ray after a subsequent consultation.

The general practitioners prescribed a drug for six patients; two patients were prescribed paracetamol, three patients were prescribed calcium carbaspirin (a form of aspirin) and one patient was given diclofenac sodium. The general practitioners advised bed rest for six patients (one of these had also received medication). Except for the two diagnostic referrals to the radiologist none of the patients was referred to a specialist and none of the general practitioners consulted a specialist for therapeutic advice.

Of the 29 contacts, 25 were concluded with clear follow-up arrangements (Table 2). In 16 of the 19 initial contacts (84%) a definite follow-up arrangement was made. Follow-up arrangements were not known for one patient after the initial consultation. Data were missing for four of the five follow-up arrangements where the patient was asked at the initial consultation to telephone the doctor, and for one of the five follow-up arrangements where the patient had been asked at the initial consultation to return to the practice, even though the follow-up period fell within the registration period in all five cases. Six patients were asked to return after the initial contact if their symptoms did not improve; one patient returned.

Table 1. Symptoms presented by the 19 patients at their initial consultation with the general practitioner.

ICPC description	% of patients[a] ($n = 19$)
Symptoms/complaints leg/thigh	42
Symptoms/complaints hip	37
Restriction/handicap	32
Symptoms/complaints knee	5
Other localized abdominal pain[b]	11

n = number of patients. [a] Five patients presented with two complaints. [b] Pain in groin.

Table 2. Follow-up arrangements made by the general practitioners at the 19 initial consultations and the 10 subsequent consultations.

Follow-up arrangement for patient	No. of initial consultations	No. of subsequent consultations
To return if no improvement	6	2
To return to the practice	5	3
To telephone the doctor	5	1
Need not come back	0	3
No follow-up appointment	2	1
Not known	1	0

Discussion

As the national survey was not designed to look at transient synovitis of the hip specifically and as there is no separate code for transient synovitis of the hip within the *International classification of primary care,* the original contact registration forms were examined. In view of the diagnostic problems of transient synovitis of the hip, the possibility cannot be ruled out that some general practitioners only described the symptomatology of the child and did not make a diagnosis. This probably led to an underestimation of the actual number of patients with transient synovitis of the hip, so the incidence rate found in this study (1.1 per 1000 person years) should be viewed as a minimum.

The ratio of boys to girls (2.8:1.0) found in this study corresponds with those of other studies. Landin and colleagues reported a ratio of 2.6:1.0.[3] In their article, Haueisen and colleagues gave a summary of literature in the field of transient synovitis of the hip and found a ratio of 1.86:1.0.[5] The average age of the children in this study (six years six months) was higher than that reported by Haueisen and colleagues (5.9 years).[5]

Relapses in children suffering from transient synovitis of the

hip have been recorded previously.[4,15] Two patients in this study were suffering a relapse, a preceding episode having occurred before the registration period. Since the study only took place over a three month period it was not possible to monitor patients long enough to establish the rate of relapse reliably.

Rather than consulting a paediatrician or orthopaedic surgeon for an early opinion, it appeared that Dutch general practitioners chose to await developments in patients with a provisional diagnosis of transient synovitis of the hip.

The diagnosis of transient synovitis of the hip made by general practitioners was based on history taking and physical examination. Two patients also had an x-ray examination. General practitioners did not use ultrasound scanning as a diagnostic tool for patients suffering from hip complaints. This may have been the result of any of the following factors: the general practitioners may have been unfamiliar with ultrasound scanning as a diagnostic tool in the case of hip complaints, there may have been limited availability of ultrasound scanning for general practitioner patients at the time of the national survey, and general practitioners may have lacked confidence in ultrasound as a means of diagnosis in hip disorders. In 32% of the cases, general practitioners advised bed rest for patients. When prescribing medication, general practitioners seem to prefer non-steroidal anti-inflammatory drugs, which was logical in view of the inflammatory nature of the complaint.

The wait-and-see approach of the general practitioners necessitated definite follow-up arrangements. Clear follow-up arrangements were made for 16 children after the initial consultations. Where data were missing for five follow-up appointments, it is assumed that the patients did not keep these appointments. The self limiting nature of the condition underlies this assumption. The chance of patients returning for the follow-up appointment may be decreased as their symptoms may not affect their ability to perform daily activities.[16] However, it could be that follow-up consultations took place without having been recorded.

In conclusion, the estimate of the incidence rate of transient synovitis of the hip found in this study is similar to that found in Nordic countries.[3,4] Dutch general practitioners practise a wait-and-see approach when suspecting the diagnosis. The question of whether the wait-and-see policy was justified can not be answered on the basis of this study.

References

1. Visser JD. The limping child. *Ned Tijdschr Geneeskd* 1988; **132:** 1644-1647.
2. Zemel L. Letter. *J Pediatr Orthop* 1986; **6:** 755.
3. Landin LA, Danielsson LG, Wattsgard C. Transient synovitis of the hip. *J Bone Joint Surg (Br)* 1987; **69:** 238-242.
4. Kunnamo I, Kallio P, Pelkonen P. Incidence of arthritis in urban Finnish children. *Arthritis Rheum* 1986; **29:** 1232-1238.
5. Haueisen DC, Weiner DS, Weiner SD. The characterization of transient synovitis of the hip in children. *J Pediatr Orthop* 1986; **6:** 11-17.
6. Illingworth CM. 128 limping children with no fracture, sprain or obvious cause. *Clin Pediatr (Phila)* 1978; **17:** 139-142.
7. Ambrosia RDD. *Muscoskeletal disorders.* Philadelphia, PA: Lippincott, 1986.
8. Marchal GJ, Holsbeek MT, Raes M, *et al.* Transient synovitis of the hip in children: role of US. *Radiology* 1987; **162:** 825-828.
9. Meradji M, Diepstraten AFM. Coxitis fugax: sonographisches und radiologisches Bild in 65 Fällen. *Radiologie* 1988; **28:** 473-478.
10. Bickerstaff DR, Neal LM, Booth AJ, *et al.* Ultrasound examination of the irritable hip. *J Bone Joint Surg (Br)* 1990; **72:** 549-553.
11. Kingma MJ. *Dutch textbook of orthopaedics.* Utrecht, The Netherlands: Bohn, Scheltema and Holkema, 1977.
12. Foets M, van der Velden J. *A national survey of morbidity and interventions in general practice. Basic report: research methods.* Utrecht, The Netherlands: NIVEL, 1990.
13. Lamberts H, Wood M (eds). *ICPC. International classification of primary care.* Oxford University Press, 1987.
14. Van der Velden J, Schellevis F, van der Steen J. *International classification of primary care.* Utrecht, The Netherlands: NIVEL, 1989.
15. Illingworth CM. Recurrences of transient synovitis of the hip. *Arch Dis Child* 1983; **58:** 620-623.
16. Van der Steen JJ. *Compliance in follow-up appointments in general practice [dissertation].* Leiden, The Netherlands: Rijksuniversiteit Leiden, 1987.

Acknowledgements

The authors thank J van der Velden of the Netherlands institute of primary health care.

Address for correspondence

J C van der Wouden, Department of General Practice, Erasmus University, Mathenesserlaan 264a, 3021 HR Rotterdam, The Netherlands.

Prescription of Antibiotics and Prescribers' Characteristics. A Study into Prescription of Antibiotics in Upper Respiratory Tract Infections in General Practice

MARIJKE KUYVENHOVEN,*,** RUUT DE MELKER* AND KOOS VAN DER VELDEN***

Kuyvenhoven M, de Melker R and van der Velden K. Prescription of antibiotics and prescribers' characteristics. A study into prescription of antibiotics in upper respiratory tract infections in general practice. *Family Practice* 1993; **10**: 366–370.

There is a growing concern about rational prescribing of antibiotics. That is why a secondary analysis of prescribing antibiotics in upper respiratory tract infections has been conducted by means of a nationwide study of morbidity and interventions in The Netherlands. The mean percentage of antibiotic prescriptions varied from about 20% for acute otitis media and acute upper respiratory tract infections to about 70% for sinusitis and tonsillitis. Only attitude—toward prescribing antibiotics in sore throat—and years of settlement were important predictor variables. The other characteristics studied—type of practice, list size, frequency of use of Het Farmacotherapeutisch Kompas, containing national pharmacotherapeutical guidelines, and urbanization level were not. The importance of attitude, however, was less for general practitioners who went into practice after 1975. This means that the influence of a personal characteristic as attitude might have become less influential since the introduction of vocational training for general practice.

INTRODUCTION

In The Netherlands about 60% of all consultations in general practice result in the prescription of drugs.[1] Both here and abroad the feeling grows that the prescription of drugs in general and that of antibiotics in particular should be dictated by rational, i.e. medical motives. Two-thirds of all antimicrobial drugs prescribed by Dutch general practitioners (GPs) are prescribed for upper respiratory tract infections.[2]

In general practice 25% of new cases are upper respiratory tract infections.[1-3] These infections are mainly self-limiting diseases. Most upper respiratory tract infection labels, such as acute otitis media, acute tonsillitis and sinusitis are used to denote a wide range of virological and bacteriological infections in the middle ear, nose and throat. The effectiveness of antibiotic prescriptions in many cases of upper respiratory tract infections can be questioned.[5-11] GPs vary a great deal in their readiness to prescribe antibiotics.[5,12,13] There is evidence that individual characteristics of physicians and their practice correlate with prescription behaviour. In some studies young doctors of more recent training proved to be better[14] but more expensive prescribers.[15] Extra training through postgraduate courses was found to be correlated with better prescribing.[16]

Practice characteristics may affect prescription behaviour too. GPs in single-handed practices proved to be more inclined to prescribe antibiotics in cases of upper respiratory tract infections.[10] Other studies showed a limited correlation between physician and practice characteristics on the one and prescription behaviour on the other hand,[17] or no correlation at all.[18] Lately more and more attention is paid to rational prescribing policies for antibiotics.[19-21] Het Nederlands Huisartsen Genootschap, the Dutch

*Department of General Practice, University of Utrecht, ***The Netherlands Institute of Primary Health Care. **Correspondence to: Ms Marijke Kuyvenhoven, Department of General Practice, University of Utrecht, Bijlhouwerstraat 6, 3511 ZC Utrecht, The Netherlands.

College of General Practitioners, has developed a standard for acute otitis media and sore throat within the framework of a general standard programme to set standards for diagnosis and management in general practice.[22] Strict application of the standards set may limit the cases indicated for antibiotics. Besides, 'Het Farmacotherapeutisch Kompas', containing the therapeutic guidelines of the 'Ziekenfonds' (The Netherlands Sick Fund) Board, aims at improving and securing the quality of prescription behaviour.[23]

Insight into factors influencing the prescription of antibiotics may contribute to a more rational prescription behaviour through education programmes like the vocational training of GPs and implementation of national standards and guidelines. Considerations of this kind underlie the commission of this study into the correlation between both physician and practice characteristics and the prescription of antibiotics for upper respiratory tract infections.

METHODS

Data

A secondary analysis of the data of the Dutch National Survey of General Practice has been performed; 161 GPs serving 335 000 people participated in this study. They formed a random disproportionate stratified sample of the total population of GPs in The Netherlands.[24] The GPs were selected from three independent stratified random samples from all Dutch GPs. Stratification variables were: region, urbanization level and distance from practice to hospital. Non-response analysis showed over-representation of younger GPs, female GPs and partnerships. The GPs were divided into four groups of 40 physicians who were consecutively involved in the research for 3 months from 01/04/87 to 31/03/88. They recorded data on all their contacts with patients. Morbidity data were centrally coded in ICPC and ICPC-compatible systems.[25]

Upper Respiratory Tract Infection Cases

The study was limited to first encounters with upper respiratory tract infection cases: acute otitis media/myringitis (H71), acute upper respiratory tract infections (R74), acute/chronic sinusitis (R75) and acute tonsillitis (R76).

Prescriptions of Antibiotics

The medication was classified according to the Anatomical Therapeutical Chemical Classification System.[26] The inclination to prescribe antibiotics for each GP was measured by means of the percentage of contacts with new cases in which he or she prescribed antibiotics; the possible range varied from 0 to 100%, an interval scale.

Physician Characteristics

The 161 GPs completed a questionnaire on their physician and practice characteristics:

- years of settlement (1: $\leqslant$1968; 2: 1969–1975; 3: 1976–1980; 4: 1981–1988);
- type of practice (1: solo practice; 2: duo-group practices or health centres);
- list size (1: $\leqslant$1500 patients; 2: 1501–2000; 3: 2001–2300; 4: 2301–2700; 5: $\geqslant$2701);
- urbanization level (1: rural; 2: urbanized/rural area; 3: small town; 4: city);
- frequency of use of 'Het Farmacotherapeutisch Kompas' as a source of drug information (1: never; 2: $\leqslant$once a week; 3: a few times a week; 4: nearly daily);
- attitude toward prescribing antibiotics for sore throat (1: (very) positive to 4: very negative).

Analysis

All data were analysed by computer; the Statistical Package for the Social Sciences (SPSS-X) was used. First, the mutual correlations between physician and practice characteristics and the correlations with the inclination to prescribe antibiotics were described. The value of physician and practice characteristics relating to the inclination to prescribe antibiotics was measured by means of multiple regression analyses, which are suitable for the interval level of the dependent variable. The standardized B-coefficient (ß) and explained variance (R^2) were calculated in relation to each predictor variable with regard to the mean percentage of antibiotic prescriptions for all upper respiratory tract infection cases combined, and for each group of upper respiratory tract infection cases. The standardized regression coefficient ß is the slope of the least squares line when both X and Y are expressed as Z-scores. The ß-coefficient does not reflect in absolute sense the importance of the various independent variables, but it represents the relative importance of these variables in the five regression equations. Interaction effects were checked. Those variables were included in the multiple regression analyses that had at least a significant ($P \leqslant 0.05$) correlation with the inclination to prescribe antibiotics.

RESULTS

During the 3 months of registration the 161 GPs were confronted with about 1800 new cases of acute otitis media, 8300 new cases of acute upper respiratory tract infections, 2500 cases of sinusitis and 1700 cases of acute tonsillitis. The physicians had, on average, 11–16 encounters for acute otitis media, sinusitis and tonsillitis and about four times more for acute upper respiratory tract infections. The inclination to prescribe antibiotics depended upon diagnosis. The mean percentage of antibiotic prescriptions in the four diagnostic groups varied from about 20% for acute otitis media and acute upper respiratory tract infections to about 70% for sinusitis and tonsillitis.

Physician Characteristics and Practice Characteristics

There were several mutual correlations between physician and practice characteristics. List size cor-

related with each of the other characteristics (Table 1). The greater the general GPs' list size, the longer they had set up, the more they practised single-handed, the more their region was rural, the less frequently Het Farmacotherapeutisch Kompas was used, the more frequently they were positive about antibiotics. Years of settlement correlated with most other characteristics; the less the years of settlement, the less the practitioners practised single-handed, the smaller was the list size, the more frequently Het Farmacotherapeutisch Kompas was used, and the more negative their attitude was toward prescribing antibiotics in sore throat cases. This attitude correlated with the frequency of use of the Farmacotherapeutisch Kompas too; the more negative the physicians' attitude was, the more frequently the Farmacotherapeutisch Kompas was used.

Prescription of Antibiotic Drugs

The more recently the GPs had set up in practice, the more frequently they practised in group practices or health centres, the smaller their list size was, the more frequently they used Het Farmacotherapeutisch Kompas and the more frequently they were negative about prescribing antibiotics in sore throat cases on the one hand, and, on the other hand the less frequently GPs prescribed antibiotic drugs in all upper respiratory tract infection cases combined (Table 2). This also applied for acute otitis media cases and to a lesser degree for acute upper respiratory tract infection cases, sinusitis and tonsillitis cases.

Physician and practice characteristics accounted for more than a quarter of the variance of prescribed antibiotics in all upper respiratory tract infections (Table 3). When the four diagnostic groups were separately analysed this proved to hold only for acute otitis media; the explained variance (R^2) was low in the other three groups. The attitude toward prescribing antibiotics in sore throat cases was the most influential characteristic with regard to the mean percentage of antibiotics prescribed in first contacts for all upper respiratory tract infection cases and for the four separate diagnostic groups. Years of settlement were of value for all upper respiratory tract infection cases and for the separate diagnostic groups except for

TABLE 1 *The mutual correlations (Pearson's r) between physician and practice characteristics (n = 161)*

	1	2	3	4	5	6
1. Years of settlement (≤1968 ↔ recently)	–					
2. Type of practice (solo ↔ partnership)	0.23*	–				
3. List size (less ↔ more)	−0.34*	−0.31*	–			
4. Urbanization level (rural ↔ city)	0.10	−0.05	−0.19*	–		
5. Farmacother. Kompas (never ↔ nearly daily)	0.29*	0.08	−0.23*	−0.06	–	
6. Attitude toward antibiotics (pos ↔ neg)	0.33*	0.15	−0.29*	0.08	0.26*	–

*$P \leq 0.05$.

TABLE 2 *The correlations (Pearson's r) between physician and practice characteristics and the inclination to prescribe antibiotics according to diagnosis (n = 161)*

	Years of settlement	Type of practice	List size	Urbanization level	Farmac. Kompas	Attitude to antibiotics
Acute otitis media	−0.30*	−0.20*	0.26*	0.02	−0.22*	−0.33*
Acute URTI[a]	−0.24*	−0.18*	0.18*	−0.11	−0.12	−0.29*
Sinusitis	−0.15	0.00	0.16	−0.06	−0.13	−0.27*
Acute tonsillitis	−0.11	−0.02	0.19*	−0.16	−0.07	−0.19*
All URTIs[a]	−0.29*	−0.23*	0.23*	−0.14	−0.18*	−0.47*

[a]URTI, upper respiratory tract infections.
*$P \leq 0.05$.

TABLE 3 *Multiple regression analyses of antibiotic prescriptions in new cases of upper respiratory tract infections (n = 161); standardized B-coefficients (ß) and explained variance (R^2)*

	Acute otitis media		Acute URTI[a]		Sinusitis		Tonsillitis		All URTIs[a]	
	ß	R^2	ß	R^2	ß	R^2	ß	R^2	ß	R^2
1. Years of settlement (⩽1968 ↔ recently)	−0.56	0.10	−0.21	0.06	−0.13	0.02	−0.17	0.02	−0.47	0.09
2. Type of practice (solo ↔ partnership)	−0.09	0.12	0.02	0.06	0.08	0.02	0.08	0.02	−0.07	0.10
3. list size (less ↔ more)	0.09	0.14	0.04	0.07	0.18	0.06	0.20	0.06	0.06	0.12
4. Use of Farm. Kompas (never ↔ nearly daily)	−0.10	0.15	−0.03	0.07	−0.04	0.07	0.03	0.06	−0.03	0.12
5. Attitude to antibiotics (pos ↔ neg)	−0.55	0.22	−0.29	0.12	−0.34	0.12	−0.27	0.09	−0.64	0.26
1 × 5	0.59	0.23	0.09	0.12	0.20	0.12	0.19	0.09	0.49	0.27
Explained variance R^2	23%		12%		12%		9%		27%	

[a]URTI, upper respiratory tract infections.

sinusitis. The other characteristics, type of practice, list size and the use of Het Farmacotherapeutisch Kompas were not valuable predictor variables with regard to the inclination to prescribe antibiotics in these analyses. There was an interaction between years of settlement and attitude toward prescribing antibiotics in sore throat cases for all groups except for acute upper respiratory tract infection cases. Attitude was an important characteristic for physicians who had set up in practice before January 1975, but not for those who had set up after that year. This fact especially applied for acute otitis media. GPs who had started practice before 1975 and who were positive toward prescribing antibiotics in sore throat cases, more frequently prescribed antibiotics in upper respiratory tract infection cases and especially in acute otitis media cases. This correlation with attitude was nearly absent with GPs who had started practice after that date.

DISCUSSION

Prescription of antibiotics mainly depends on diagnosis. The inclination to prescribe antibiotics is low for acute otitis media and acute upper respiratory tract infections and high for sinusitis and tonsillitis. The low prescription rate for acute otitis media is in accordance with the standard set by the Dutch College of General Practitioners and the results of recent trials.[27,28] The high percentage for sinusitis is surprising since the effectiveness of antibiotics has been called into question where this condition is concerned.[29,30] The percentage for acute tonsillitis is relatively low in comparison with other countries.[31]

Our research results are based on a registration of patient–doctor encounters by 161 GPs over a period of 3 months. They must give a fairly reliable picture of the daily routine in Dutch general practices. The 161 GPs might have different age–sex frequencies in their practice populations, but these patient characteristics hardly correlate with the percentages of antibiotic prescriptions.[32] We described comparable percentages of antibiotic prescriptions for upper respiratory tract infections in an earlier study with written vignettes.[10] In this country GPs are restrictive in their antimicrobial policy in comparison with the figures for other countries.[8,10,31]

The inclination to prescribe antibiotics can best be accounted for in the case of acute otitis media. The variance that can be accounted for by physician and practice characteristics is higher than that found by De Maesseneer in Belgium.[17] Especially the attitude towards prescribing antibiotics in sore throat cases is a relatively good predictor variable of the inclination to prescribe antibiotic drugs for acute upper respiratory tract infections. However, this variable interacts with years of settlement. This attitude is a good predictor variable for the inclination to prescribe antibiotics with GPs who had set up in practice before January 1975, the very year when vocational training for general practice was started in The Netherlands. Most GPs who established themselves as doctors before 1975 have not received a vocational training for general practice, unlike those who started to practise after that date. This may imply that the influence of personal characteristics as attitudes has decreased since the introduction of vocational training for general practice;

vocationally trained GPs more frequently use Het Farmacotherapeutisch Kompas as a source of drug information, are less often to be found in solo practices and less frequently prescribe antibiotics.

For sinusitis and tonsillitis the correlation with physician and practice characteristics is low. This may show that doctors feel there is not much choice in the prescription of drugs for these complaints. From another study we have learnt, however, that in the case of acute tonsillitis physicians tend to highly overestimate the probability of a positive culture for group A streptococci for patients with sore throats.[33] Obviously there is a need for further studies on the epidemiological knowledge of doctors on probabilities of bacteriological and virological infections, as well as studies on prospective criteria in order to predict the course of an upper respiratory tract infection. Findings from both types of studies can be used in the process of implementation of national guidelines and standards.

REFERENCES

1 Groenewegen PP, de Bakker DH, van der Velden J. Een nationale studie naar ziekten en verrichtingen in de huisartspraktijk. Basisrapport verrichtingen in de huisartspraktijk. Utrecht, Nivel, 1992.

2 Sampers GHMA, Sturm AW. Antimicobiele middelen in de eerste lijn bij luchtweginfecties. *Huisarts Wet* 1990; **33:** 220–222.

3 *Ziekten in de huisartspraktijk* (red EH vd Lisdonk) Utrecht, Bunge 1990.

4 van der Velden J, de Bakker DH, Claessens AAMC, Schellevis FG. A Dutch National Survey of General Practice: morbidity in general practice. Utrecht, Nivel, 1992.

5 Irvine DA. The general practitioner and upper respiratory tract infection in childhood. *Fam Pract* 1986; **3:** 126–131.

6 Norreland AN. Behandling af sinusitis i almenpraksis. *Ugeskr Laeg* 1978; **140:** 2792–2795.

7 von Sydow C, Axelsson A, Jensen C. Acute maxillary sinusitis—a comparison between 27 different treatment modes. *Rhinol* 1982; **20:** 223–229.

8 Froom J, Culpepper L, Grob P, *et al.* Diagnosis and antibiotic treatment of acute otitis media: report from International Primary Care Network. *Br Med J* 1990; **300:** 582–586.

9 Bain J. Justification for antibiotic use in general practice. *Br Med J* 1990; **300:** 1006–1007.

10 de Melker RA, Kuyvenhoven MM. Management of upper respiratory tract infection in Dutch general practice. *Br J Gen Pract* 1991; **41:** 504–507.

11 Pichichero ME. Culture and antigen detection tests for streptococcal tonsillopharyngitis. *Am Fam Physician* 1992; **45:** 199–205.

12 Howie JGR, Richardson JM, Gill G, Durno D. Respiratory illness and antibiotic use in general practice. *J R Coll Gen Pract* 1971; **21:** 657–663.

13 Howie JGR. Clinical judgement and antibiotic use in general practice. *Br Med J* 1976; **2:** 1061–1067.

14 Becker MH, Stolley PD, Lasagna L, *et al.* Differential education concerning therapeutics and resultant physician prescribing patterns. *J Med Educ* 1972; **47:** 118–127.

15 Dunnell K, Cartwright A. *Medicine takers prescribers and hoarders.* London: Routledge and Kegan, 1970.

16 Becker MH, Stolley PD, Lasagna L, *et al.* Correlates of physicians prescribing behaviour. *Inquiry* 1972; **9:** 30–42.

17 De Maesseneer J. *Huisartsgeneeskunde: een verkenning* (dissertatie). Gent: Centrum voor huisartsgeneeskunde, 1989.

18 Quik-van Milligen MLT, Kuyvenhoven MM, de Melker RA, Touw-Otten FWMM, Koudstaal PJ, van Gijn J. Transient ischaemic attacks and the general practitioner: diagnosis and management. *Cerebrovasc Dis* 1992; **2:** 102–106.

19 Stott NCH. Management and outcome of winter upper respiratory tract infections in children aged 0–9 years. *Br Med J* 1979; **280:** 29–31.

20 Stephenson RJ, Henry N, Norman GR. Factors influencing antibiotic use in acute respiratory tract infections in family practice. *Can Fam Phys* 1988; **34:** 2149–2152.

21 Wyatt TD, Passmore CM, Morrow NC, Reilly PM. Antibiotic prescribing: the need for a policy in general practice. *Br Med J* 1990; **300:** 441–444.

22 Grol R. National standard setting for quality of care in general practice: attitudes of general practitioners and response to a set of standards. *Br J Gen Pract* 1990; **40:** 361–364.

23 van der Kuy A (ed.) *Farmacotherapeutisch kompas 1990–1991. Medisch farmaceutische voorlichting. Uitgave van de Centrale Medisch Pharmaceutische Commissie van de Ziekenfondsraad.* Amstelveen: Ziekenfondsraad, 1990.

24 Foets M, van der Velden J, de Bakker DH. *The Dutch national survey of morbidity and interventions in general practice. Study design.* Utrecht: NIVEL, 1992.

25 Lamberts H, Wood M. *ICPC: international classification of primary care.* Oxford: Medical Publications, 1987.

26 Anonymous. *Nordic statistics on medicines 1975–1977. Part II Statistical reports of the nordic countries no. 36.* Oslo: Nordiska Läkemedelsnämnden, 1979.

27 van Buchem FL. Peeters MF, van 't Hof MA. Acute otitis media: a new treatment strategy. *Br Med J* 1985; **290:** 1033–1037.

28 Appelman CLM, Claessen JQPJ, Touw-Otten FWMM, Hordijk GJ, de Melker RA. Efficacy of amoxicillin–Clavulanate in recurrent acute otits media. A placebo controlled study. *Br Med J* 1991; **303:** 1450–1452.

29 Berg O, Bergstedt H, Lind MG, *et al.* Discrimination of purulent from non-purulent maxillary sinusitis. *Ann Otol Rhinol Laryngol* 1981; **90:** 272–275.

30 Gwaltney J, Sydnor A, Sandi MA. Etiology and antimicrobial treatment of acute sinusitis. *Ann Otol Rhinol Laryngol* 1981; **90:** 68–71.

31 Touw-Otten FWMM, Steahr Johansen K. Diagnosis, antibiotic treatment and outcome of acute tonsillitis: report of a WHO regional office for Europe study in 17 European countries. *Fam Pract* 1992; **9:** 255–262.

32 de Melker RA, Kuyvenhoven MM, van der Velden J. Voorschrijf-en verwijsgdrag van huisartsen bij bovenste luchtweginfecties. *Huisarts Wet* 1993; **36:** in press.

33 Posen RM, Cedul RD, Collins M, Fager S. The accuracy of experienced physicians, probability estimates for patients with sore throats. *J Am Med Assoc* 1985; **254:** 925–929.

Analysis of referrals of mental health problems by general practitioners

PETER F M VERHAAK

SUMMARY. The majority of people in the community who have a psychiatric disorder will consult their general practitioner. Referrals from general practice to specialist services are, however, relatively rare. The filter between primary care and specialist care has been characterized by Goldberg and Huxley as the least permeable of the filters separating psychiatrists and other specialists from the populations they serve. These referrals form the subject of this study in the Netherlands. Using a large database of doctor–patient contacts, the proportion of mental health disorders resulting in a referral and the characteristics of the patient and general practitioner that are involved in such a referral have been determined. In addition, the type of mental health institution or specialist to which referrals were directed and the characteristics influencing this choice were examined. Only 6% of patients presenting with a psychiatric disorder during surgery hours were referred to specialist care. Younger patients, male patients and patients with severe diagnoses had a greater probability of being referred. The percentage of patients referred was higher in urban areas than in rural areas. Doctors with a limited task perception regarding mental treatment tended to refer more often. Although the diagnosis did have some relationship with the institutions to which patients were referred (psychotic conditions to psychiatric services and social/material problems to social workers), the most prevalent diagnoses (neurotic conditions and relationship problems) seemed to be more or less randomly distributed over the various possibilities. Preferences appeared to be related to the existence of regular meetings between general practitioners and specialists and a positive evaluation by general practitioners of the institution concerned.

Keywords: referral to psychiatric services; referral patterns; referral rates; referral reason; psychiatric disorders.

Introduction

MOST mental health problems come to the attention of the general practitioner; the majority are recognized and treated by the general practitioner, but a small minority are referred by the general practitioner to specialized mental health workers, be they social workers, psychotherapists or psychiatrists. These facts were established 25 years ago by Shepherd and colleagues, and have been confirmed by a number of studies since then.[1-4] There are large variations between general practitioners in terms of referral rates to psychiatric services[5] and their preferences concerning the discipline or institution to which they refer patients.[6,7]

The chances of being referred are not equal for all patients. Patients with serious psychiatric complaints,[4,8] or with a diagnosis of psychosis[1,5,9,10] are referred relatively often, in contrast to patients with neurotic complaints. Men are more likely to be referred than women, and younger patients (especially those aged 25–35 years) are more likely to be referred than elderly patients.[2,11-13] The characteristics of the general practitioner also play a part in the chances of referral. Older general practitioners, those practising in urban areas and those working in single handed practices make more referrals to psychiatric services than younger doctors or those working in rural areas and in group practices.[2,3] Robertson reports fewer psychiatric referrals and a preference for psychological and social work referrals among doctors who show an interest in psychotherapy.[5] Creed and colleagues confirm these results: those general practitioners who write more detailed referral letters show a lower referral rate to psychiatric services and a higher referral rate to psychologists than those writing poorer letters.[7]

An important determinant in the mental health referral process could be the doctor–patient relationship. As Morgan pointed out, for only 40% of referred psychiatric patients did clinical indications only become a decisive factor in relation to the decision to refer: the ineffectiveness of previous treatment, often accompanied by a mutual loss of confidence, was often a general practitioner's stated reason for referral.[14] In an analysis of videotaped consultations in which there was a psychiatric referral, it was observed that it was not the type of complaint that determined whether a referral was proposed, but the feeling that all previous efforts had failed.[6,7] Robertson reports that about 35% of referrals are made because the patient is not responding to the general practitioner's treatment.[5]

Although a comprehensive picture of general practitioner referrals to mental health professionals appears to emerge from the literature, the picture is composed of fragmentary evidence. Most of the studies cited above are either outdated or based on relatively small samples. In most cases the studies are restricted to referrals to psychiatrists, while little information about the patterns of referrals from general practitioners to paramedical providers of mental health care exists. Wilkinson concludes in his review that the proportion of patients with mental health problems who are referred by general practitioners to psychiatrists and paramedical mental health workers is unknown.[15]

The aim of this study was, therefore, to provide a description of mental health referrals by Dutch general practitioners in order to answer the following questions: What proportion of mental health disorders result in a referral? What characteristics of the patient and general practitioner determine whether a mental health referral is made? To what type of mental health institution or specialist are referrals directed? What factors influence the type of institution or discipline to which the referral is made?

Method

Data were collected from April 1987 to April 1988 within the framework of the national survey of morbidity and interventions in general practice, conducted by the Netherlands Institute of Primary Health Care (NIVEL).[16] A total of 103 Dutch general practices (161 general practitioners) were selected for this study and details of all contacts with patients over a period of three months were recorded by the general practitioners. The three month recording periods were distributed over the whole year to exclude seasonal effects. Data collected included the reason for the patient's visit, the diagnosis, the treatment and whether or not

P F M Verhaak, PhD, psychologist and sectoral manager of mental health research, Netherlands Institute of Primary Health Care, Utrecht, The Netherlands.

Submitted: 6 January 1992; accepted: 4 May 1992.

the patient was referred. In addition, the 161 participating general practitioners completed an extensive questionnaire, which included questions about perception of tasks regarding mental health care, performance of tasks, opinions about the possible psychosocial nature of illness, and the mental health care services in their region (social work, ambulatory mental health care and psychiatric services).

The participating practices were a randomly selected non-proportionally stratified sample from the population of all Dutch general practices (4755 practices, 6288 general practitioners in 1988). The sample was stratified to include a balance of practices in all regions of the Netherlands, and of rural and urban practices. Group practices in the Netherlands often have an undivided practice list. As a result, colleagues of a selected general practitioner in such a practice were also asked to participate and this procedure resulted in an overrepresentation of general practitioners from group practices and health centres, of general practitioners aged less than 40 years and of women general practitioners. However, the practice population (335 000 patients) can be considered to be representative of the Dutch population.

Dependent variables

The diagnosis made by the general practitioner during normal surgery consultations was coded using the *International classification for primary care (ICPC).*[17] Those diagnoses coded within chapter 'P' (Psychological) or 'Z' (Social) are considered here. As there are about 50 difference symptoms and diagnoses within each chapter, these have been clustered in six larger groups:[16] neurotic disorders (for example, depression, anxiety, phobia, mental strain); psychotic disorders; other symptoms within *ICPC* chapter 'P' (for example, addiction); other diagnoses within *ICPC* chapter 'P' (for example, dementia); *ICPC* chapter 'Z', relationship problems; *ICPC* chapter 'Z', material and social problems (for example, housing). The diagnosis is a characteristic of the episode of illness which might cover only one visit to the general practitioner, but which might cover many visits. It is a characteristic of general practice, and of psychological problems presented in general practice in particular, that many diagnoses remain at symptom level.[18]

Referrals considered relevant to this study were referrals to hospital psychiatrists, psychiatric outpatient clinics, private psychiatrists, mental health hospitals, regional institutions for ambulatory mental health care, institutions for alcohol and drug problems, private psychologists and social workers. In the Dutch health care system, a referral from a general practitioner is mandatory in order for a patient to obtain specialist medical help. Regional institutions for ambulatory mental health care provide several forms of care (social psychiatric treatment, crisis intervention, psychotherapy and counselling). The professional staff includes specialized social workers, psychiatrists, social psychiatric nurses and psychotherapists. In formal terms, access to ambulatory mental health care should also be mediated by a general practitioner. In practice, however, only about 50% of all clients of ambulatory mental health care arrive via the general practitioner. Although social work is freely accessible, and private psychologists are, in most cases, beyond any kind of legislation, the general practitioner is the most important referring agency for these disciplines too. At the time of this study, the costs of private psychologists were not reimbursed by public health insurance companies; the other alternatives were covered in one way or another. As the first four referral possibilities listed above are dominated by psychiatrists, they have been considered together. Alcohol and drug institutions provide ambulatory care, so referrals to these institutions are considered together with those to institutions for ambulatory mental health care. This reduces the total number of categories of possible referrals to four: psychiatric referrals, referrals to ambulatory mental health care, referrals to private psychologists and referrals to social workers.

A referral ratio was calculated for each general practitioner. The referral ratio is the number of referrals made by the general practitioner to any mental health specialist, divided by the number of episodes of illness that the general practitioner has given a diagnosis from chapters 'P' or 'Z' of the *ICPC.* A referral ratio was calculated only for those general practitioners who had made at least 100 such diagnoses over the three month period (127 of the 161 participating doctors). The referral ratio is expressed as the number of referrals per 100 diagnoses.

The preference for the four referral categories was calculated for each of the general practitioners who had made at least five referrals to a mental health specialist (83 doctors). For example, preference for psychiatry was taken as the number of psychiatric referrals divided by the total number of referrals to a mental health specialist, expressed as a percentage.

Independent variables

The type of practice (single handed, two partner practice, group practice, or health centre) and location (degree of urbanization) were assessed.

In the questionnaire completed by participating general practitioners the questions on perception of tasks consisted of a number of items expressing psychosocial activities, such as treatment of agoraphobia, counselling on sexual problems and discussing a work related problem. For each item the general practitioner rated the activity on a five point scale from 'Definitely a general practitioner's task' (five) to 'Definitely not a general practitioner's task' (one). In order to determine the performance of tasks the same items were rated again on a five point scale from 'I always carry out this activity' (five) to 'I never carry out this activity' (one). The questions for the perception of the possible psychosocial nature of illness listed a number of complaints and diagnoses, to be rated on a five point scale from 'Not influenced by psychosocial factors' (one) to 'Very much influenced by psychosocial factors'(five). As a second indication of general practitioners' bias regarding the psychosocial nature of illness they were simply asked to estimate the proportion of all problems presented to them that were not entirely physical in nature.

Affiliation with certain institutions or contact with specialists, and the evaluation of them, might influence the choice general practitioners make once they have decided to make a referral. Therefore, the general practitioners were asked about their regular appointments (regular meetings at fixed times) with social workers, ambulatory mental health care workers, psychiatrists and private psychologists. The four categories of specialist referral were also evaluated in respect of a number of aspects (adequacy of help, waiting lists, negative experiences in the past, geographical accessibility, appropriate only for certain patients, for example only those who are sufficiently articulate). The general practitioners were asked to rate each aspect of each specialist category on a 10 point scale from very negative (one) to very positive (10).

Analysis

The relationship between general practitioners' characteristics and the referral ratio or the four preference scores has been analysed using analysis of variance, as the predictors are discrete variables and the criterion variables are continuous. The distribution of referrals over age, sex and diagnostic categories has been analysed by means of hierarchical log linear analysis. Chi square has been used to test the goodness of fit.

Results

The mean age of the 127 participating general practitioners for whom referral ratios could be calculated was 42 years (range 33 to 63 years). Most of the general practitioners worked in rural (36.2%) and suburban (39.4%) areas, more or less according to the distribution of the population.[16] Only 18.9% worked in urban areas and 5.5% in large cities. Of the 127 general practitioners 36.2% were single handed, 30.7% worked in a two partner practice, 20.5% in a group practice and 12.6% in a multidisciplinary health centre.

The mean scores for perception and performance of tasks show that the general practitioners had a slightly positively biased task perception and task performance (Table 1). Both scales were normally distributed. The perception and performance of tasks appeared to be strongly intercorrelated (product-moment correlation coefficient $r = 0.73$; $P<0.001$). The scale for the perception of the possible psychosocial nature of the listed complaints was again normally distributed around the mean score of 3.1. The estimate of the proportion of presented symptoms that were not entirely physical was normally distributed with a mean of 45%. The latter is related to perception of tasks ($r = 0.32$; $P<0.001$).

Of the 127 general practitioners 59.8% reported regular contact with social workers, whereas only 20.5% had regular appointments with ambulatory mental health care workers, 4.7% with psychiatrists and 5.5% with psychologists. The mean evaluation of ambulatory health care is clearly lower than for the other three disciplines (Table 1), a picture that also emerges on the several subscales that comprise the overall score.

Table 1. Attitudes of the participating general practitioners.

	Mean	(SD)	Range
Score on five point scale			
Task perception ($n = 126$)	2.7	(0.6)	1.3–4.1
Task performance ($n = 125$)	2.8	(0.6)	1.4–4.2
Perception of possible psychosocial nature ($n = 127$)	3.1	(0.6)	1.7–4.9
% of symptoms rated as not entirely physical (n = 127)	*45*	*(21)*	*1–99*
Score on 10 point scale[a] *evaluating:*			
Social work ($n = 102$)	6.9	(1.4)	1.0–9.4
Ambulatory mental health care ($n = 104$)	5.5	(1.3)	1.8–8.2
Psychiatric services ($n = 35$)	6.6	(1.2)	3.6–8.8
Psychologist ($n = 54$)	6.9	(1.2)	4.0–9.4

n = number of respondents. SD = standard deviation. [a]Mean of mean score for each GP.

Proportion of episodes of illness referred

A total of 19 286 episodes of illness with a psychological or social diagnosis were recorded. A total of 1106 referrals were recorded during surgery hours (surgery visits and home visits taken together) and included in the analysis. Overall, there were 1310 referrals to mental health care and social work. Considering only referrals made during surgery hours, 5.7% of episodes of illness were referred.

Referrals, by patient characteristics

Table 2 shows the proportion of referrals in the six diagnostic groups, by the age and sex of the patient. Statistical analysis revealed that the referral rates were not independently distributed over diagnostic category, age and sex ($\chi^2 = 628$, 39 degrees of freedom, $P<0.001$) — age and diagnosis, age and sex, and sex and diagnosis interact. Psychotic disorders and other psychological diagnoses were the most likely diagnoses to be referred. No differences could be found in this respect between the various type of psychosis — schizophrenia, affective psychosis, puerperal psychosis and organic psychosis. However, the numbers in these categories were small. Neurotic disorders, such as depression, anxiety and stress disorders, were much more common and only 5.5% were referred.

The distribution of referrals by age and sex (Table 2) reveals that men were more likely to be referred than women and that younger people (less than 40 years of age) were more likely to be referred than older patients. It should be noted that a diagnosis of mental disorder was most common in women aged 40 years or more (40.7% of all mental illness diagnoses) and least common in men under 40 years of age (14.2% of all diagnoses).

Referral ratio, by general practitioner characteristics

Table 3 summarizes a number of analyses of variance, with the referral ratio as the dependent variable. The referral ratio increased with the degree of urbanization. The figures for cities, in particular, are considerably higher than those for the countryside. The referral ratio of doctors in health centres was higher than for doctors who work in single handed, two partner or group practices. Doctors who did not consider psychosocial treatment as their task referred slightly more patients than doctors who did consider this to be their task.

Table 2. Prevalence of mental health disorders and rates of referral age and sex.

	% of episodes referred (total no. of episodes in group)				
		Female patients aged:		Male patients aged:	
Symptoms/diagnoses	Total	40+ years	<40 years	40+ years	<40 years
Psychological problems					
Neurotic disorders	*5.5* (9256)	*3.7* (3675)	*7.5* (2464)	*5.1* (1800)	*7.4* (1317)
Psychotic disorders	*16.4* (365)	*11.4* (149)	*15.3* (72)	*16.3* (92)	*32.7* (52)
Other symptoms	*3.9* (4711)	*1.4* (1895)	*5.5* (1116)	*2.7* (971)	*9.6* (729)
Other diagnoses	*12.0* (676)	*9.7* (309)	*17.1* (123)	*8.4* (167)	*22.1* (77)
Social disorders					
Relationship problems	*7.3* (2653)	*3.5* (1272)	*13.2* (645)	*6.4* (488)	*12.9* (248)
Social/material problems	*4.7* (1607)	*3.9* (535)	*7.3* (395)	*2.2* (359)	*5.7* (318)

Table 3. Referral ratio, by general practitioner characteristics.

Independent variable	Referral ratio	*F*
Age (years)		1.98
41+ (*n* = 67)	5.30	
⩽ 40 (*n* = 60)	6.31	
Practice area		4.46**
Rural (*n* = 46)	4.73	
Suburban (*n* = 50)	5.75	
Urban (*n* = 24)	6.35	
Large city (*n* = 7)	10.21	
Type of practice		4.34**
Single handed (*n* = 46)	4.79	
Two partner (*n* = 39)	5.99	
Group practice (*n* = 26)	5.23	
Health centre (*n* = 16)	8.70	
Task perception[a]		2.78
Low score (*n* = 61)	6.34	
High score (*n* = 65)	5.17	
Task performance[a]		2.31
Low score (*n* = 67)	6.22	
High score (*n* = 58)	5.15	
Perception of psychosocial nature[a]		0.53
Low score (*n* = 67)	5.99	
High score (*n* = 60)	5.46	
Estimate of % of symptoms not entirely physical		0.01
Low (0–40%) (*n* = 63)	5.77	
High (41–99%) (*n* = 64)	5.71	

n = number of general practitioners in group. ** *P*<0.01. [a] High and low scores are divided by the median score.

Preference for kind of mental health institution/specialist

Table 4 shows the destination of those patients who were referred according to the diagnosis. Referrals with different diagnoses were not equally distributed over the four referral possibilities ($\chi^2 = 334$; 20 df, $P<0.001$). The predominance of neurotic disorders and psychological symptomatology in general practice, which is also reflected in the absolute referral figures, results in a majority of this kind of disorder in the caseload of each of the referral options — more than half of all the referrals to each discipline have these diagnoses. When general practitioners referred a patient with a psychotic disorder, in most cases psychiatric services were preferred (Table 4). The majority of social and material problems were referred to a social worker. In the case of relationship problems, general practitioners seemed to use two major options: ambulatory mental health care or social work; in the case of neurotic disorders and psychological symptomatology three options were chosen: a psychiatrist, ambulatory mental health care and, somewhat less often, a social worker.

Table 5 shows the age–sex distribution for the four referral options. Again, the distribution contradicts the assumption of independence of age, sex and option for referral ($\chi^2 = 246$, 13 df, $P<0.001$). Controlling for diagnosis does not alter this situation. Older patients were referred to psychiatric services more often than younger patients. Younger men were overrepresented within ambulatory mental health care and a relatively large proportion of the younger women were referred to social workers.

Table 6 shows the preferences of the general practitioners for referral, by their characteristics. Regular appointments with a specialty resulted in an increased share in referrals in the case of social workers, psychologists and ambulatory mental health care. A positive evaluation had a critical effect on referrals to social workers and ambulatory mental health care. In the case of psychiatric services and psychologists, however, data were available from a minority of respondents only. The practice area did not have an effect on any of the preferences, and type of practice shows only one clear effect: doctors in health centres preferred to refer to social workers (who are part of the health centre).

Table 4. Destination of patients who were referred, by diagnosis.

	% of referrals					
	Psychological problems				Social problems	
	Neurotic disorders (*n* = 510)	Psychotic disorders (*n* = 60)	Other symptoms (*n* = 186)	Other diagnoses (*n* = 81)	Relationship problems (*n* = 194)	Social/material problems (*n* = 75)
Psychiatric services	*37.6*	*68.3*	*38.7*	*42.0*	*9.3*	*16.0*
Ambulatory mental health care	*29.4*	*30.0*	*40.9*	*30.9*	*39.2*	*20.0*
Psychologist	*10.6*	*1.7*	*7.0*	*9.9*	*7.2*	*8.0*
Social work	*22.4*	*0.0*	*13.4*	*17.3*	*44.3*	*56.0*

n = total number of referrals.

Table 5. Destination of patients who were referred, by their age and sex.

	% of referrals			
	Female patients aged (years)		Male patients aged (years)	
	40+ (*n* = 274)	<40 (*n* = 392)	40+ (*n* = 186)	<40 (*n* = 252)
Psychiatric services	*38.3*	*28.6*	*45.2*	*26.6*
Ambulatory mental health care	*28.1*	*30.6*	*30.1*	*42.1*
Psychologist	*7.7*	*9.7*	*6.5*	*9.9*
Social work	*25.9*	*31.1*	*18.3*	*21.4*

n = total number of referrals.

Discussion

Of all the episodes of mental illness presented during surgery hours in this study, 5.7% resulted in a referral. Whitehouse similarly reported that of all consultations with psychosocial problems 4.8% of patients are referred to a consultant, men 1.9 times as frequently as women.[19] The figure of 5.7% determined here is also not dissimilar to the 6.6% reported by Wilkin and Smith for all types of referral.[20] The mean referral rate in the Dutch national morbidity survey (mean number of referrals per 100 episodes of illness) is 10.9 (Verhaak P, unpublished results).

In this study 33 per 10 000 of the population at risk were referred to a mental health specialist or social worker in a three month period. When all referrals, including those outside normal surgery hours, are considered this increases to 39 per 10 000 of the population at risk. These figures are compatible with those obtained by Shepherd and colleagues[1] and Kessel.[21] They are, however, much lower than Italian figures. Tansella[22] reported that 22% of patients identified as having conspicuous psychiatric morbidity by general practitioners were referred to a specialist, while Arreghini and colleagues presented a one day prevalence figure for general practitioner referral to specialist psychiatric services of 7.3%, which is equivalent to 17.6 per 10 000 of the population at risk being referred in one day.[23]

The age and sex distribution of the referred patients in this study showed the same characteristics as reported in other studies.[11-13] Although psychiatric morbidity was less frequently identified among younger men, this group was most frequently referred. It might be that this group of patients is overrepresented in the 'hidden psychiatric category', and that as a result the illness of the identified sample is on average more severe. Younger men were more commonly referred to ambulatory mental health care, while elderly patients were more commonly referred to psychiatric services. As the former favours a multidisciplinary approach while the latter constrict themselves to a medial framework, this difference might reflect a difference in general practitioners' perception of the 'treatability' of older and younger patients.

Psychotic disorders and other 'classical' psychiatric diagnoses were referred most frequently. However, although the likelihood of being referred is higher for these serious mental disorders, the majority of patients with these conditions remain under the care of the general practitioner. In a longitudinal study, 391 patients with psychological complaints were monitored during one year.[24] Of the patients 13% were referred to a mental health specialist during that year. The likelihood of referral was higher if the patient experienced more problems, had a higher score on the general health questionnaire, or was aged between 25 and 44 years. Patients referred presented with more psychosocial complaints over the study year than non-referred patients. These results indicate that referral is related to the burden a patient feels and the severity of his or her situation.

The referral ratio in this study is also clearly influenced by the geographical area: large cities induce more referrals than rural areas. A common finding has again been replicated.[2] General practitioners working in health centres also tended to refer more of their patients. This seems to be a result of their preference for their social worker colleagues in the health centre. The results suggest that general practitioners with an interest in psychological treatment (expressed by their task perception) do more treatment themselves and hence refer fewer of their patients. The age of the general practitioner and other personal characteristics did not have an effect on the referral ratio. This finding is similar to that of Wilkin and Smith who also did not find significant relationships between doctor characteristics and referral rates.[18] It is possible that relationships at the level of the general practitioner are obscured by the differences in the case mix of the individual general practitioners. Although general practitioners recording fewer than 100 episodes of mental illness were excluded from the analysis, such differences may have played a part.

It is noteworthy that the area where general practitioners practise had little effect on their preference for where to refer. One

Table 6. Preferences of general practitioners for the four types of specialties, by general practitioner characteristics.

	Social work			Ambulatory mental health care			Psychiatric services			Psychologist		
Independent variable	No. of GPs	Preference ratio (%)	*F* ratio	No. of GPs	Preference ratio (%)	*F* ratio	No. of GPs	Preference ratio (%)	*F* ratio	No. of GPs	Preference ratio (%)	*F* ratio
Regular appointments with specialty			15.60***			4.16*			1.19			8.88**
Yes	48	*31.1*		14	*43.6*		3	*21.5*		5	*23.8*	
No	35	*15.1*		69	*31.6*		80	*34.5*		78	*6.8*	
Evaluation of specialty (mean score)			10.27***			4.61*			0.11			3.45
⩾6	54	*29.4*		28	*29.4*		16	*44.4*		28	*15.6*	
<6	15	*12.4*		43	*39.4*		7	*41.4*		8	*3.6*	
Practice area			0.36			0.14			1.29			2.58
Rural	24	*22.1*		24	*34.6*		24	*39.5*		24	*2.8*	
Suburban	36	*25.1*		36	*33.6*		36	*29.5*		36	*11.8*	
Urban	18	*26.1*		18	*31.6*		18	*36.5*		18	*5.8*	
Large city	5	*18.1*		5	*29.6*		5	*39.5*		5	*12.8*	
Type of practice			5.73***			1.78			2.31			1.96
Single handed	30	*15.1*		30	*39.6*		30	*35.5*		30	*9.8*	
Two partner	22	*24.1*		20	*33.6*		22	*35.5*		22	*5.8*	
Group practice	18	*29.1*		18	*27.6*		18	*40.5*		18	*2.8*	
Health centre	13	*38.1*		13	*28.6*		13	*20.5*		13	*12.8*	

***$P<0.001$; **$P<0.01$; *$P<0.05$.

would expect a preference for social work with its geographically intricate structure in rural areas, and for psychiatric services in urban areas. Private psychologists claim to fill in the geographical gaps of mental health services and would be expected to predominate in rural areas. These effects did not occur.

The discriminatory power of diagnostic categories was limited to the extremes: most of the psychotic disorders and other clearly psychiatric diagnoses were referred to the psychiatric services, whereas social/material problems were referred to social workers. As the likelihood of referral lessened, the distinction between referral possibilities disappeared. The most common psychological disorders, that is neurotic disorders and other psychological symptoms, were more or less equally distributed between psychiatric services, ambulatory mental health care and social workers. An important result was the overall positive evaluation of social work and the relatively negative evaluation of ambulatory mental health care, which seems to influence referral preferences. As a consequence, the less specialized social workers appear to constitute a reasonable alternative in the case of less pronounced mental health problems that quantitatively play a major role in the epidemiology of mental health problems in primary care.

Earlier it was suggested that the doctor–patient relationship might influence the referral decision; referral might be induced by the burden a general practitioner feels. The data presented here throw no further light on this. Much still needs to be explained and the doctor–patient relationship might shed some light on these matters. Further study is directed at a comparison between consultations with patients suffering from neurotic depression who were referred to a mental health specialist or social worker and consultations with patients having the same diagnosis, who were not referred.

This study has shown that most psychological and social problems are treated by general practitioners and diagnostic labels and clinical features are of only secondary importance in the decision to refer. Thus, one should not be too 'prescriptive' when defining the types of problems general practitioners should not deal with. This is contrary to somatic medicine where, for example, is it quite clear that a patient with a suspected heart attack should be referred to a cardiologist. The relationship is less straightforward with psychological diagnoses which may be complicated by social phenomena such as the doctor–patient relationship, the support offered by the community and the characteristics and interests of the general practitioner.

References

1. Shepherd M, Cooper B, Brown AC, Kalton GW. *Psychiatric illness in general practice.* London: Oxford University Press, 1966.
2. Goldberg D, Huxley P. *Mental illness in the community: the pathway to psychiatric care.* London: Tavistock Publications, 1980.
3. Orlean CS, George LK, Houpt JL, Brodie HKH. How primary care physicians treat psychiatric disorders: a national survey of family practitioners. *Am J Psychiatry* 1985; **142:** 52-57.
4. Giel R, Koeter MWJ, van de Meer K, Ormel J. De filters in het model van Goldberg en Huxley in de praktijk [The filters in the Goldberg and Huxley model in actual practice]. *Huisarts en Wetenschap* 1991; **34:** 169-173.
5. Robertson NC. Variations in referral patterns to the psychiatric services by general practitioners. *Psychol Med* 1979; **9:** 355-364.
6. Busschbach JT, Verhaak PFM. Huisarts en psychische klachten [General practitioners and psychological complaints]. *Maandblad Geestelijke Volksgezondheid* 1986; **41:** 475-492.
7. Creed F, Gowrisunkur J, Russell E, Kincey J. General practitioner referral rates to district psychiatry and psychology services. *Br J Gen Pract* 1990; **40:** 450-454.
8. Munk-Jorgensen P. General practitioner's selection of patients for treatment in community psychiatric services. *Psychol Med* 1986; **16;** 611-619.
9. Hull J. Psychiatric referrals in general practice. *Arch Gen Psychiatry* 1979; **36:** 406-408.
10. Regier DA, Goldberg ID, Burns BJ, *et al.* Specialist/generalist division of responsibility for patients with mental disorders. *Arch Gen Psychiatry* 1982; **39:** 219-224.
11. Shapiro S, Skinner A, Kessler LG, *et al.* Utilization of health and mental health services. *Arch Gen Psychiatry* 1984; **41:** 971-978.
12. Leaf PJ, Bruce ML. Gender differences in the use of mental health-related services: a re-examination. *J Health Soc Behav* 1987; **28:** 171-183.
13. Leaf PJ, Bruce ML, Tischler G, Holzer CE. The relationship between demographic factors and attitudes toward mental health services. *J Community Psychol* 1987; **15;** 275-284.
14. Morgan D. Psychiatric cases: an ethnography of the referral process. *Psychol Med* 1989; **19:** 743-753.
15. Wilkinson G. Referrals from general practitioners to psychiatrists and paramedical mental health professionals. *Br J Psychiatry* 1989; **154:** 72-76.
16. Foets M, van der Velden J. *Nationale studie naar ziekten en verrichtingen in de huisartspraktijk: basisrapport meetinstrumenten en procedures* [*National survey of morbidity and interventions in general practice: instruments and methodology*]. Utrecht. The Netherlands: NIVEL, 1991.
17. Lamberts H, Wood M (eds). *ICPC. International classification for primary care.* Oxford University Press, 1987.
18. Wood M. The systematic assessment of psychological and social factors affecting health. In : Lipkin M, Kupka K (eds). *Psychological factors affecting health.* New York, NY: Praeger, 1982.
19. Whitehouse CR. A survey of the management of psychosocial illness in general practice in Manchester. *J R Coll Gen Pract* 1987; **37:** 112-115.
20. Wilkin D, Smith AG. Variation in general practitioners' referral rates to consultants. *J R Coll Gen Pract* 1987; **37:** 350-353.
21. Kessel WIN. Psychiatric morbidity in a London general practice. *Br J Prev Soc Med* 1962; **14;** 67-22.
22. Tansella M (ed). Community based psychiatry: long term patterns of care in south Verona. *Psychol Med* 1991; **21:** supplement 19.
23. Arreghini E, Agostini C, Wilkinson G. General practitioner referral to specialist psychiatric services: a comparison of practices in north and south Verona. *Psychol Med* 1991; **21:** 485-494.
24. Verhaak PFM, Tijhuis MAR. Psychosocial problems in primary care: some results from the national survey of morbidity and interventions in general practice. *Soc Sci Med* 1992; **35:** 105-110.

Address for correspondence

Dr P F M Verhaak, Netherlands Institute of Primary Health Care, PO Box 1568, 3500BN Utrecht, The Netherlands.

Sex differences among recipients of benzodiazepines in Dutch general practice

Fransje W van der Waals, Jacob Mohrs, Marleen Foets

Abstract

Objective—To analyse sex differences among recipients of benzodiazepines in Dutch general practice.

Design—Study of consultations and associated interventions as recorded in the Dutch national survey of general practice.

Setting—Practices of 45 general practitioners monitored during 1 April to 30 June 1987.

Subjects—61 249 patients (29 035 (47·4%) men in the age groups 19-44, 45-64, and 65 years and over.

Main outcome measures—Symptoms among recipients of repeat as well as new benzodiazepine prescriptions stratified by sex and age.

Results—Prescriptions for benzodiazepines were found to be significantly more common among women than among men, (*a*) after correcting for the sex distribution of the total patient population, and (*b*) in the two oldest age groups after correcting for the number of consultations. Of all prescriptions for benzodiazepines, 89% (6055/6777) were repeats and 70% (4759/6777) requests. Only 9% (439/4759) of these were authorised by the general practitioner, the rest being issued by the general practitioner's assistant after he or she had referred to the diagnosis in the patient's record. In contrast, only three (1%) of the 492 first time recipients of benzodiazepines had requested a prescription and were not seen by the general practitioner. Women (43/96; 45%) aged 45-64 years received their first prescription for benzodiazepines almost twice as often as men (15/63; 24%) without symptoms or a diagnosis being an indication (female to male relative risk 1·88 (95% confidence interval 1·15 to 3·08)).

Conclusions—The sex difference among first time recipients of benzodiazepines seems to be due to general practitioners being less stringent when prescribing this drug for women. The difference continues in repeat prescriptions, physicians failing to check adequately the need for these.

Women's Health Studies, Medical Faculty, University of Amsterdam, 1105 AZ Amsterdam, The Netherlands
Fransje W van derWaals, *general practitioner*

General Practice Department, Medical Faculty, University of Amsterdam
Jacob Mohrs, *database manager*

Netherlands Institute of Primary Health Care, PO Box 1568, 3500 BN Utrecht, The Netherlands
Marleen Foets, *medical sociologist*

Correspondence to:
Dr van der Waals.

BMJ 1993;307:363-6

Introduction

In the Netherlands, as in the United States and the rest of Europe, some 10% of the population use tranquillisers and hypnotics.[1-2] In 1989 roughly one third of Dutch patients had been taking tranquillisers or hypnotics, mostly benzodiazepines, continually for 180 days or longer and were designated long term users.[3] Most long term users of benzodiazepines in the United States and Europe are elderly women.[4-5] National household surveys on the use of tranquillisers and hypnotics in the Netherlands in 1985 and 1989 disclosed that women in the age group 15-44 used these drugs at about the same low frequency (2%) as did men.[6-7] At ages 45-64 years and 65 and over, however, 10% and 16% of the women respectively used tranquillisers or hypnotics in 1985; in the older age group this figure increased to 18% in 1989. These figures contrasted sharply with those for men; in two older age groups only 6% and 9% of men used these drugs.

Benzodiazepines are mainly prescribed by general practitioners.[4-5] After drugs for heart and circulatory problems benzodiazepines are the drugs most frequently prescribed by Dutch general practitioners.[8] Exact figures on the use of benzodiazepines and the reasons for prescription are not available as there is no central register of medication and consumption. Most assessments come from doctor-patient registrations,[9] health surveys,[5,10-12] pharmacy records,[13] and the Dutch national health service.[14]

Recently a national survey of general practice was conducted in the Netherlands.[15] It included the complete records of all consultations with 161 representative general practitioners and their assistants during 1 April 1987 to 31 March 1988. That study enabled us to analyse the circumstances of the initial and repeat prescriptions for benzodiazepines. Firstly, we investigated whether the figures on the use of benzodiazepines gathered from the national survey of general practice paralleled the figures from the national household surveys. Secondly, we tried to discover why doctors prescribe benzodiazepines more to women than to men.

Subjects and methods

DESIGN OF NATIONAL SURVEY OF GENERAL PRACTICE

General practitioner sampling

From all 5826 general practitioners in the Netherlands on 1 Janaury 1985 a random, stratified, non-proportional sample was selected. Stratification was deemed necessary because of geographic factors contributing to morbidity data and procedures in general practice. The country was divided into three regions and four levels of urbanisation. Finally, the distance from the doctor's practice to a hospital was noted.

Comparison of the total general practitioner population in the Netherlands with the 161 doctors selected showed that 24 (15%) of the sample group but only 466 (8%) of all general practitioners were women. In addition, the sample group was younger and included a larger proportion of doctors in group practices.

We studied the practices of the 45 general practitioners for whom during the first three month period (1 April to 30 June 1987) details of all repeat prescriptions were recorded.

Doctor-patient contacts

Contacts with the general practitioner were recorded during consecutive three month periods of the survey year. A full registration form was used for contacts with the doctor and a shorter version of the form for contacts with the assistant to the general practitioner. No significant differences were found in regional

distribution, urbanisation, or distance to the hospital between the first group of doctors (our study group) and doctors in the other three groups. Also other variables such as characteristics of the practices, numbers of episodes, reasons for consultations, sex distribution of the patients, and types of health insurance did not differ. Seasonal influences on the data were minimal.

The contact registration form included the general practitioner's identification code; the patient's identification code; the description of contact (first consultation, repeat consultation, relapse); morbidity data (reasons for consultation (reported as symptoms and diagnosis)); intervention data; and, finally, prescription data. We used records of first and repeat consultations, symptoms, diagnosis, and subsequent new or repeat prescriptions. These variables were recorded by the general practitioner and his or her assistant. For each general practitioner one fieldworker was appointed to be responsible for recording, classifying, and coding data. This centralised procedure resulted in an inter-observer agreement score of 0·85 for reasons for consultation and 0·80 for rare or complicated diagnoses. The symptoms and diagnoses were classified according to the International Classification of Primary Care,[21] with some minor changes.[22]

Patient registration

The 45 general practitioners cared for a total of 61 249 patients. Their age and sex distribution did not differ significantly from that of the population as a whole. A patient registration form including information on socioeconomic status, was completed by 91% of the study population.

Benzodiazepine prescriptions

We selected for our study the following symptoms and diagnoses recorded in the International Classification of Primary care as indications for benzodiazepine: anxiety or nervousness or feelings of inadequacy, or both; acute stress or panic disorders, or both; irritability or anger, or both; restlessness or agitation, or both; insomnia. As an additional category we studied requests for a prescription or repeat prescription.

Computer analysis and statistics

We evaluated prescriptions in relation to reasons for consultation (reported as symptoms and diagnoses) stratified by sex and age. The SPSS package was used to compute the risks and 95% confidence intervals. In this study the relative risk ratio is the same as the female to male relative risk ratio.

Results

AGE AND SEX DISTRIBUTION OF BENZODIAZEPINE RECIPIENTS

The 61 249 patients in the study population comprised 29 035 men and 32 214 women. Analysis of the frequency of consultations with the general practitioner showed that 36 047 patients (14 400 male, 21 647 female) visited their general practitioner at least once during the three months.

Benzodiazepines were prescribed at least once to 4231 patients (1324 (31%) male, 2907 (69%) female). The numbers of patients prescribed benzodiazepines increased with age among both men and women (table I). In the older age groups (45-64 and 65 and over) the absolute increase was greater for women (from 12% to 20%) than for men (from 6% to 11%). Even after correcting for sex distribution in the total patient population significantly more women than men were recorded as receiving benzodiazepine prescriptions (relative risk 1·68-1·92).

When we corrected for the sex difference in visits to the general practitioner significantly more women (relative risk 1·57 for the age group 45-64 years, 1·67 for age 65 and over) than men were prescribed benzodiazepines (table II).

INDICATIONS FOR CONTINUING BENZODIAZEPINES ASSESSED BY GENERAL PRACTITIONER

A total of 1448 (69%) male patients and 3311 (71%) female patients received a prescription for benzodiazepines on request. Because most patients' contacts with the practice entailed a request (4759 (70%)) for benzodiazepines or a repeat (6055 (89%)) prescription we decided to study to what extent general practitioners reassessed their patients' needs for repeat prescriptions.

In 90% of the men requesting benzodiazepines and 91% of the women the prescriptions were issued by the general practitioner's assistant, who referred to the diagnosis in the patient's record. These proportions were significantly higher than those in contacts with assistants that resulted in non-benzodiazepine prescriptions (data not shown).

We studied visits to the practice by benzodiazepine recipients for reasons unrelated to benzodiazepines. Most of these patients were seen by the general practitioner. No significant differences between men and women were observed.

Given that most benzodiazepine prescriptions were issued by the assistant and not by the general practitioner, the correlation among symptoms, diagnosis, and actual prescriptions was remarkably good in all age groups. No differences between men and women were observed. Between 73% and 85% of all symptoms considered indications for benzodiazepines were diagnosed as such. Also between 65% and 85% of all

TABLE I—*Numbers (percentages) of patients in general practice population prescribed benzodiazepines*

Age group (years)	No (%) of total patient population prescribed benzodiazepines		Female to male relative risk (95% confidence interval)
	Women	Men	
19-44	695/18 065 (4)	393/17 112 (2)	1·68 (1·48 to 1·89)
45-64	997/8172 (12)	501/7879 (6)	1·92 (1·73 to 2·13)
≥65	1215/5977 (20)	430/4044 (11)	1·91 (1·75 to 2·12)

TABLE II—*Numbers (percentages) of patients visiting general practitioners who were prescribed benzodiazepines*

Age group (years)	No (%) of general practitioner consulting patient population prescribed benzodiazepines		Female to male relative risk (95% confidence interval)
	Women	Men	
19-44	695/11 456 (6)	393/7306 (5)	1·13 (1·00 to 1·27)
45-64	997/5399 (18)	501/4267 (12)	1·57 (1·42 to 1·74)
≥65	1215/4792 (25)	430/2827 (15)	1·67 (1·51 to 1·84)

TABLE III—*Numbers (percentages) of patients initially prescribed benzodiazepines for symptoms and diagnoses*

Age group (years)	Patients given benzodiazepines for symptoms			Patients given benzodiazepines for diagnoses		
	No (%) of women	No (%) of men	Female to male relative risk (95% confidence interval)	No (%) of women	No (%) of men	Female to male relative risk (95% confidence interval)
19-44	73/149 (49)	46/89 (52)	0·95 (0·73 to 1·23)	89/249 (60)	61/89 (69)	0·87 (0·72 to 1·06)
45-64	45/96 (47)	41/63 (65)	0·72 (0·54 to 0·95)	56/96 (58)	43/63 (68)	0·85 (0·67 to 1·08)
≥65	31/70 (44)	13/25 (52)	0·85 (0·85 to 1·35)	34/70 (49)	13/25 (52)	0·93 (0·60 to 1·46)

TABLE IV—*Numbers (percentages) of patients with and without symptoms as indication for benzodiazepines as proportion of all new diagnoses*

Age group (years)	Patients with symptoms as proportion of all incident diagnoses			Patients without symptoms as proportions of all incident diagnoses		
	No (%) of women	No (%) of men	Female to male relative risk (95% confidence interval)	No (%) of women	No (%) of men	Female to male relative risk (95% confidence interval)
19-44	42/89 (47)	33/61 (54)	0·87 (0·63 to 1·20)	47/89 (53)	28/61 (46)	1·15 (0·82 to 1·61)
45-64	28/56 (50)	28/43 (65)	0·77 (0·55 to 1·08)	28/56 (50)	15/43 (35)	1·43 (0·88 to 2·33)
≥65	20/34 (59)	7/13 (54)	1·09 (0·61 to 1·94)	14/34 (41)	6/13 (46)	0·89 (0·44 to 1·82)

prescriptions for benzodiazepines were backed by a relevant diagnosis.

INITIAL BENZODIAZEPINE PRESCRIPTIONS AND SEX DIFFERENCES IN SYMPTOMS AND DIAGNOSES

As only 32% of benzodiazepine prescriptions (2166/6777) were written by the general practitioner it was difficult to study the doctor-patient interaction in long term users. We therefore studied 492 incident cases—that is, prescriptions that were annotated as "new" prescriptions separately. Of these patients, 489 (99%) were seen by the general practitioner. In contrast with 79% of repeats of all benzodiazepine prescriptions that were requests (4756/6055), among the 492 incident cases only three (1%) requests for benzodiazepines were recorded. The high frequency of contacts with general practitioners allowed us to study the impact of doctor-patient interactions on new benzodiazepine prescriptions (table III). In all age groups more men than women had symptoms classified as indications for benzodiazepines. In the age group 45-64 years the difference was 65% *v* 47%. Likewise, doctors diagnosed more in men than in women a condition requiring benzodiazepines, particularly at ages 45-64 years.

We also compared diagnoses based on symptoms considered to be indications for benzodiazepines with diagnoses made without symptoms (table IV). In the youngest age group (19-44 years) and in the oldest age group (65 and over) about half of the diagnoses were made in patients without relevant symptoms. No significant differences were observed between men and women. In the age group 45-64 years diagnoses warranting benzodiazepines were made less frequently in women. This difference did not, however, reach statistical significance. Surprisingly, benzodiazepines were prescribed to men for headache, general malaise and fatigue, chest pain, and sensory disturbances and to women for general malaise and fatigue, vertigo, depression and headache, and chest pain.

Table V shows the frequency with which doctors prescribed benzodiazepines when no symptoms were identified or a diagnosis made warranting the treatment. A significant difference was observed in the age group 45-64. Women in this group received benzodiazepines without indication almost twice as often as men—45% of women (43/96) *v* 24% of men (15/63).

In men of all age groups the most frequent symptoms were back pain, accident trauma, headache, sensory disturbances, and non-specified pain. The most frequent diagnoses in men were lumbago, alcohol misuse, myalgia, problems at work, and problems resulting from the death of a family member (other than a child or partner). In women the most frequent symptoms were headache, chest pain, general malaise and fatigue, abdominal pain, palpitations, back pain, and diarrhoea. The most frequent diagnoses in women were headache, hypertension, family problems, back pain, myalgia, cystitis, and problems resulting from the death of a partner.

TABLE V—*Numbers (percentages) of patients initially prescribed benzodiazepines without any indication for their use, based on symptoms or diagnosis*

Age group (years)	No (%) of patients		Female to male relative risk (95% confidence interval)
	Women	Men	
19-44	58/149 (39)	32/89 (36)	1·08 (0·77 to 1·52)
45-64	43/96 (45)	15/63 (24)	1·88 (1·15 to 3·08)
≥65	39/70 (56)	12/25 (48)	1·16 (0·73 to 1·84)

Though lumbago and back pain might be considered legitimate reasons to prescribe benzodiazepines, we did not include these symptoms or diagnoses in our study; muscle relaxation is not mentioned in efficacy studies as an indication for the drug. When women with back pain (9/315; 3%) and men with lumbago (18/177; 10%) were excluded the difference between women and men was even greater—132 (42%) of the women and 25 (14%) of the men received benzodiazepines without having relevant symptoms or diagnoses.

Discussion

Our general practice survey, confirming the results of other studies, showed that women are more often prescribed benzodiazepines than men.[2,22] The results of our analysis agree exactly with the results of the Dutch national household surveys, in which respondents reported their use of benzodiazepines.[2,3] In all age groups, but particularly in the older age groups far more women received benzodiazepines than did men. This was true even after correction for the female to male ratios in the general practice as a whole. The female to male ratios of all ages were consistently higher than among the total patient population under study. This strongly indicates that there are specific causes for the greater numbers of female benzodiazepine users in Dutch general practices. In our study the greater number of women receiving benzodiazepine prescriptions is not explained by overdiagnosis of benzodiazepine indications in women. Analysis of the indications recorded for benzodiazepines showed that most were simply requests for a first or repeat prescription. From the absence of sex differences in indications for repeat prescriptions we tentatively conclude that the diagnostic process leading to the continuation of benzodiazepine use does not explain the greater proportion of women prescribed these drugs.

General practitioners generally do not see patients when they request a prescription for benzodiazepines but see these patients when they have symptoms or disorders for which other treatments would be used. That diagnoses preceding a repeat prescription are mainly registered by the general practitioner's assistant suggests that general practitioners rarely re-evaluate their patients' continuing need for benzodiazepines. Dutch general practitioners often delegate requests for repeat prescriptions to assistants.[24] This is especially worrying because half of all doctors' assistants in the Netherlands do not have the relevant training.[24]

Prescribing further benzodiazepines without re-evaluating the patient may explain in part why this drug is prescribed excessively to men and women but does not explain the sex difference. It simply maintains the difference. We suggest that the sex difference among recipients of benzodiazepines originates from the time of the initial prescription.

In the age group 45-64 years significantly more men

Clinical implications

- Benzodiazepines are prescribed for 10% of the population, of whom one third are long term users; only short term use has been shown to be effective
- Women use twice as much benzodiazepine as men, and its use increases with age
- Most prescriptions are repeats (89% in this study) and are given by the doctor's assistant after requests (70%) from patients
- More women are initially given benzodiazepines for conditions other than anxiety, stress, and insomnia
- Excessive prescribing of benzodiazepines might be reduced if they were no longer given for reasons other than anxiety, stress, and insomnia and requests for repeats were thoroughly evaluated by the general practitioner

than women had legitimate reasons for starting treatment with benzodiazepines. Doctors frequently diagnosed women with symptoms such as headache and general fatigue (not legitimate indications for benzodiazepines) as suffering from anxiety, stress or insomnia (which are legitimate indications) and thus as candidates for benzodiazepines. Even more striking was that the doctors also frequently diagnosed women as having headache for which subsequently benzodiazepines were prescribed. In addition, initial prescriptions in this age group were more often given to women than to men when neither symptoms nor diagnosis warranted the drug.

More research is needed to see if similar prescribing practices occur among general practitioners in other countries. In the Netherlands the sex difference in recipients of benzodiazepines is probably the result of a two step process. More women than men are initially prescribed benzodiazepines for conditions other than anxiety, stress, or insomnia. As benzodiazepines do not relieve other conditions, most patients asking for repeat prescriptions do so because they continue to have complaints. Subsequently general practitioners continue to prescribe benzodiazepines without fully re-evaluating their patients' needs.

Three steps could be taken that might reduce the excessive prescribing of benzodiazepines to women. Strict criteria should be set for prescribing benzodiazepines initially. Conditions other than anxiety, stress, and insomnia should be excluded as indications for the drug. And requests for repeat prescriptions of benzodiazepines should be thoroughly re-evaluated by the general practitioner.

We thank Jaap Goudsmit, Joke Haafkens, Claartje Moerman, Jan van Ree, and Bert Schadé for advice and critical reading of the manuscript.

1 Balter MB, Levine J, Mannheimer DI. Cross-national study of the extent of antianxiety/sedative drug use. *N Engl J Med* 1974;290:769-74.
2 Balter MB, Mannheimer DI, Mellinger GE, Uhlenhuth EH. A cross-national comparison of antianxiety/sedative drug use. *Curr Med Res Opin* 1984;8(suppl 4):5-20.
3 Herings RMC. *Effecten van chronisch en gecombineerd geneesmiddelengloruik.* Utrecht: University of Utrecht, 1989.
4 Brekel EJG van den. Prescribed medicines 1981-1985. *Maandberichten Gezondheidsstatistiek* 1987;3:5-14.
5 Asselt JW van. *Slaapverwekkend. Een onderzoek naar het gebruik van slaap- en kalmeringsmiddelen in Lelystad.* Lelystad: GGD, 1989.
6 Central Bureau of Statistics. *Vademecum gezondheidsstatistiek Nederland.* The Hague: Staatsdrukkerij Uitgeverij, 1989.
7 Central Bureau of Statistics. *Vademecum gezondheidsstatistiek Nederland.* The Hague: Staatsdrukkerij Uitgeverij, 1993.
8 Mellinger GD, Balter MB, Uhelenhuth EH. Anti-anxiety agents: duration of use and characteristics of users in the USA. *Curr Med Res Opin* 1984;8(suppl 4):21-35.
9 Mellinger GD, Balter MB, Uhelenhuth EH. Prevalence and correlates of the long-term regular use of anxiolytics. *JAMA* 1984;251:375-9.
10 Catalan J, Gath DH, Bond A, Edmonds G, Martin P, Ennis J. General practice patients on long-term psychotropic drugs: a controlled investigation. *Br J Psychiatry* 1988;152:399-405.
11 Nolan L, O'Malley K. Patients, prescribing and benzodiazepines. *Eur J Clin Pharmacol* 1988;35:225-9.
12 Rodrigo EK, King MB, Williams P. Health of long-term benzodiazepine users. *BMJ* 1988;296:603-6.
13 Morgan K, Dallasso H, Ebrahim S, Arie T, Fentem PH. Prevalence, frequency and duration of hypnotic drug use among the elderly living at home. *BMJ* 1988;296:601-2.
14 Woods JH, Katz JL, Winger G. Abuse liability of benzodiazepines. *Pharmacol Rev* 1987;4:251-413.
15 Institute of Medical Statistics. *Diagnose informatie en medische statistiek.* Slaapstoornissen, The Hague: Institute of Medical Statistics, 1989.
16 Swinkels H. Huisarts en patient in cijfers. Enkele gegevens uit de Gezondheids enquete van het CBS. *Huisarts Wet* 1990;33:504-10.
17 Haafkens JA, Moerman C. Gender differencesin the use of benzodiazepines in Amsterdam. *Contemporary Drug Problems* 1992;Fall:505-32.
18 Timmermans AJGM, Dolmans JMEV, Bruinink RJ, van Ommeren GJ, van den Brink G. Benzodiazepinegebruik in Utrecht. *Medical Contact* 1983;38:1331-3.
19 Post D. Flunitrazepam op de derde plaats. *Pharm Weekbl* 1988;123:522-5.
20 Bensing JM, Foets M, van der Velden J, van der Zee J. De nationale studie van ziekten en verrichtingen in de huisartsenpraktijk. *Huisarts en Wetenschap* 1991;34:51-61.
21 Lamberts H, Wood M, eds. *International classification of primary care.* Oxford: Oxford Unversity Press, 1987.
22 van der Velden J, Schellevis F, van der Steen J. *International classification of primary care. Tabulaire lijst ten behoeve van de nationale studie van ziekten en verrichtingen in de huisartsenpraktijk.* Utrecht: NIVEL, 1989.
23 Blijenberg-Ruis B, Garretsen HFL, Schuurman JH, Verdonk ALT. Het gebruik van slaap en kalmerende middelen. Een secundaire analyse van een Rotterdams onderzoek. *T Soc Gezondheidszorg* 1986;64:491-7.
24 Nijland A. De praktijkassistente in de huisartspraktijk. Progressie in professionalisering (dissertation). Groningen: Rijks Universiteit Groningen, 1991.

(Accepted 26 May 1993)

A MEMORABLE PATIENT

Just a bullet in the heart

He was in his mid-70s and had been admitted with a ruptured Achilles tendon. He had jumped down from a hedge on his farm, heard a loud crack, and thought he had broken his leg. His injury required a surgical repair. His preoperative chest *x* ray examination showed an elongated opaque foreign body in the left cardiac area. Screening showed that this moved during cardiac contraction.

"Did you know," I said, "that you have a bullet in your heart?"

"Yes," said he, "it is a German 300 and it has been with me for 50 years. In fact," he continued, "I should have died on three occasions."

He had graduated from Aberystwyth University College in June 1914 and in August he joined the army and was posted to an infantry battalion. He was at the Western Front in October of that year and was out on patrol when he was shot in the chest. He thought then that he would die. He lay there for some time until a German patrol came across him. He remembered clearly a German soldier swearing at him and lifting his rifle to shoot him again, when mercifully a German officer appeared and with a pistol in his hand ordered the soldier to stop and to get a stretcher.

Some two weeks later he was in a German military hospital when the surgeon there, whom he thought was Sauerbruch, told him that he had a bullet in his heart and he proposed to remove it by open operation. Sauerbruch was one of the most eminent surgeons in the world, particularly in experimental surgery of the thorax. Nevertheless, our hero adamantly refused an operation. This was the third time, he thinks, he would have died.

He recovered from his wounds and spent the rest of the war as a prisoner. Subsequently he had a distinguished career as a chief inspector of schools. On retirement he took up farming in south Pembrokeshire. When he died many years later he left his farms and his personal fortune to the University College of Wales, Aberystwyth.—R L REES *is a retired orthopaedic surgeon in Dyfed*

Primary care registration systems in Europe: in search of a standardised framework. A case study on three European data systems

C. van Weel*, B. Bolibar**, E. Carrillo**, S. Juncosa**, W. Boerma***, K.van der Velden***

Summary

The objective of the study is to describe the difficulties of developing harmonised data sets in primary care throughout Europe in order to obtain comparable information on morbidity and interventions.

A case study on three European general practice data systems was undertaken analysing their objectives, items collected, classifications used, methods of data collection and analysis, units of data aggregation and quality control.

The systems studied were the Continuous Morbidity Registration (CRM) of the Nijmegen University and the National Study on Morbidity and Interventions in General Practice of the NIVEL both from the Netherlands, and the SIG-7 System of the Catalan Institute of Health including two specific data bases (ANAC-1 and ANAC-2) from Spain.

We found major differences between the data systems, partly explained by distinct objectives and the structure of the health care system. Although the type of data collected is similar, comparison of data is difficult because of the different aggregation levels and strategy chosen.

It is concluded that harmonisation does not necessarily mean developing a single minimum data set at a European level. Rather, our work suggests the use of a common framework with a specific weight on particular data types depending on the objectives and with a different level of development depending on the data collection resources of each system.

Key words: Data collection, information systems, cross-cultural comparison

Introduction

General practice is an ideal source of information for comprehensive epidemiological, as well as managerial and resource use, relevant at local, regional and national level[1,2]. There is a wide number of experiences of general practice delivery of continuous information for country-wide surveillance of specific infectious (influenza, hepatitis, salmonellosis...) or non-infectious (suicide attempt,..) diseases and ad hoc information on interventions performed (prescription, referrals, etc). But experience with the management of large-scale data collection in primary care is mainly restricted to local, or regional initiatives, mostly with a limited objective. Existing data bases are therefore difficult to compare or exchange. On the other hand those developing data bases have shown remarkable ingenuity in solving the many practical problems and in overcoming the difficulties of data standardisation and exchange. The development of morbidity classifications for primary care such as ICHPPC-2-D[3] (International Classification of Health Problems in Primary Care-2nd version-defined) and ICPC[4] (International Classification of Primary Care) is probably the best example of the potential in this respect.

There is an evident need for specific data from ambulatory care to support clinical, managerial and political decisions[5]. Information technology developments can set the stage for an extensive collection of data at regional/national level enabling

* *Department of General Practice, University of Nijmegen, The Netherlands*
** *Unitat d'Investigació de l'Area Centre, Institut Català de la Salut, Terrassa, Spain*
*** *NIVEL, Netherlands Institute of Primary Health Care, Utrecht, The Netherlands*

Address for correspondence and reprint requests: Dr. B. Bolibar, Unitat d'Investigació de l'Area Centre, Institut Català de la Salut, c Torrebonica s/n, 08227 Terrassa (Barcelona), Spain

Int. J. Health Sciences 1993, Vol. 4, no. 2, pp. 51-60

comparison on morbidity and interventions in primary care with other countries and health care facilities: but a preliminary step on the standardisation of the information to be collected is necessary.

Most discussions on data collection and data analysis in ambulatory care have so far been based upon theoretical considerations and concepts. Standardisation of data, the choice of a relevant denominator, quality control of collected data, the development of software for data storage and linkage have been identified as substantial problems.

A case study on three European general practice data sets was undertaken by the McACE project* sponsored by the Commission of the European Communities within the AIM (Advanced Informatics in Medicine) programme exploratory phase[6]. This practical study was undertaken to complement the modelling work of the McACE project[7]. Its general aim is to illuminate our models by using of real world examples, looking at existing data sets: their requirements, content, purposes, limitations and other features.

Three of the McACE partners were involved in routine data base management and although the selection is limited to only two countries a comparison of these data bases seemed useful. The aim of this study is to describe similarities and differences in items collected, populations, providers, services, methods of data collection, methods of analysis, of quality control and collected items.

Methods

The data bases chosen for analysis cover the broad field of ambulatory care, addressing the areas of 'general practice', 'primary care' and 'ambulatory care' at local, regional or national level.

- *The Continuous Morbidity Registration (CMR) of the University of Nijmegen (The Netherlands)* collecting morbidity episodes in a well-defined stable population of four general practices since 1967[8].
- *The Management Information System of the Area 7 (SIG-7) of the Catalan Institute of Health (Spain)*, based on two complementary data bases, collects data on utilisation of ambulatory care services (*ANAC-1 data base*; Analysis of Activity-1, and more in depth data on presented morbidity and the use of general practice facilities (*ANAC-2 data base)*[9].
- *The National Study of Morbidity and Interventions in General Practice of the NIVEL Institute (The Netherlands)* collecting data on presented morbidity in primary care, the use of diagnostic and therapeutic facilities and the health status of the practice population[10,11].

Each information system reported its specific details on a standard review format, including:

1. Objectives of collecting data for a set of specific activities
2. The catchment areas and extent of data collection (local/regional or national; all General Practitioners (GPs) vs samples or volunteers; all patients or a sample, etc)
3. Items collected on:
   - Population (socio-demographic data)
   - Clinical data (symptoms, diagnoses, causes of death). The following details were required on clinical data:
     = Specificity and standardisation used;
     = Standard uses of data and analyses, particularly the denominator (population – contact – episodes – provider);
     = Protection of privacy of patients and providers
   - Services provided (contacts, visits, referrals, hospital admissions and therapeutic interventions)
4. Techniques of data collection (information sources; central vs on the spot coding; direct extraction of data from the provider vs non- intrusive indirect extraction)
5. Units of data aggregation
6. Quality control of data
7. Manpower figures and cost of the information system.

The analysis of the findings was made by the authors against a number of theoretical models developed in McACE[7]. Particular attention was paid to the potentialities of developing transnational information standards.

* McACE Consortium consisted of MARI Group Lts (prime contractor) and Newcastle University from the UK, a team of the Institut Catalan de la Salus (ICS) from Spain, and Nijmegen University, Nivel Primary Care Research Centre, and SIG Services BV from the Netherlands

Results

Table 1 and 2 summarise the most relevant features of the three data systems. There are many differences between them, particularly in objectives, the scale of work and regional dimension.

The objectives of the studies show a range of variation from regional and local planning and daily management (SIG-7), general national planning and research (NIVEL Study) to academic research in general practice (CMR). The profile of institutions which generated the systems match very well these requirements: a regional health authority in the SIG-7; a Primary Care Research Institute (NIVEL) and an Academic Department of Family Medicine (Nijmegen University).

According to the objectives, the systems also show major variability in their extent and design. CMR collects data continuously from a small number of practices in order to enable longitudinal analysis. NIVEL Study is a national cross-sectional study for planning and research. ANAC-1 is based on routine collection of data from all ambulatory practitioners in a region for daily management; and ANAC-2 collects samples of visits from volunteer health centres for planning purposes.

The differences in the practical approaches to data collection are also of interest. Data from the NIVEL Study and ANAC-2 are collected in specific registration forms by the practitioners involved. NIVEL Study is based on a national sample of practitioners that collect data from all their contacts while ANAC-2 is based on a sample of contacts (visits) in a number of voluntary health centres from a region. CMR data are also collected by the participating practitioners but by means of the automatisation of their practices. On the other hand, ANAC-1 collects routine data from different sources external to the practitioner (non-intrusive collection).

Units of aggregation are also different. ANAC-1 minimal aggregate is the practitioner activity monthly level. ANAC-2 data set is contact-based. Further aggregation at episode or patient level is not possible because it is based on a sample of visits. Neither is aggregation at practitioner level because the sample technique is representative of the overall health centre. CMR data-set minimal unit is the episode, and further aggregation can be made at patient or family level. No specific information on each contact is recorded. NIVEL Study shows a wider intermediate picture with data items based on contacts which are aggregated at episode, patient and practitioner level.

For such different systems, the type of data collected (Table 2) shows a picture of surprising simi-

Table 1. Description of the four data bases of the three data systems

	SIG-7		NIVEL STUDY	CMR NIJMEGEN
	ANAC-1	ANAC-2		
Area	Regional	Regional	National	Local
Objectives	Planning Management	Planning Management	Planning Research	Research
Special features	Consultations in ambulatory care	Temporary sample of consultations	Cross sectional (1987-1989)	Longitudinal patient data (1967 ongoing)
Type of health care	General practice and ambulatory specialties	General practice	General practice	General practice
Extent	All practitioners	All practitioners volunteer practices	Sample of practitioners	Volunteer practitioners
Population denominator	Practice list Census data	Practice list Census data	Practice list Census data	Practice list
Unit of aggregation	Practitioner	Contact General practices	Contact Episode Physician	Episode Physician
Method of collection	Indirect (from outside contact)	Registration form	Registration form	Automatisation
Quality control	+	+	++	+++

Table 2. Comparison of the data types and data items of the four data sets

DATA TYPES	Data items	ANAC-1	ANAC-2	NIVEL ST.	CMR
Provider	Identification	+	+	+	+
	Characteristics	+	–	+	–
	Linkage with other specific data sets	+	–	+	–
Patiënt	Identificator	*	+	+	+
	Characteristics	*	+	+	+
	Linkage with other specific data sets	–	–	+	+
Contact	Date	*	+	+	*
	Characteristics	*	+	+	*
Health Problems	Reason for encounter	*	+	+	–
	Diagnosis / Problem	*	Main problem ICHPPC-2-D	All problems ICPC	All episodes E-Book
	Severity of morbidity	*	–	+	+
Interventions	Diagnosis procedures	+	+	+	–
	Referrals	+	+	+	+
	Therapy	+	+	++	–
	Others	+	+	–	–
Disposition	Follow-up	–	–	+	–

+ = Information collected
– = Information non collected
* = Not pertinent

larities even though the level of detail for each type of data and other features make it extremely difficult to compare actual data (Table 3). ANAC-1 collects basic quantitative data on contacts and use of external resources (lab tests, X-ray, work certificates, referrals). ANAC-2 introduces socio-demographic characteristics and morbidity data plus a very rough description of resource use (number of prescriptions, ordering or not X-ray or lab tests, etc). NIVEL Study includes a very similar spectrum of data with a higher level of detail including all health problems of the contacts and accurate details on tests or procedures performed; furthermore, it can be linked to other important data sets. CMR data set includes data on episodes, patient characteristics and health problems, but no major interest is paid to resource use.

The four data bases use population data (practice list and/or census) as denominator of the results.

All three systems use some form of standardised classification for socio-demographic and morbidity data, even though there are differences due to history of the systems and their objectives. The Continuous Morbidity Registration, for example, started recording morbidity in 1967, 12 years before ICHPPC-2 was published, and the main objective of the data base (longitudinal analysis of morbidity) limits the possibility of changing the classification system. The ANAC-2, designed in the middle 80s uses ICHPPC-2-D. The more recent National Study of Morbidity in the Netherlands analyses the broad area of health problems in primary care and consequently has to include symptoms as well as diagnoses using ICPC. So far, ICHPPC-2-D and ICPC appear prominent as relevant standards of recording morbidity and health problems.

The control of data quality has been given great attention in all four data bases. The large scale of the ANAC-1 only allows for somewhat loose types of quality control. The more limited scale of ANAC-2 allows the introduction of better surveillance and data quality and relies on the help of an external coder and computer routines. The NIVEL Study introduces central coding to minimise intercoder variation and participants training. Furthermore, the field workers who visited 'their' practices about every fortnight, played a major role in maintaining quality of data collection. The objectives of the CMR require very precise control of input -monthly meetings to discuss classification problems, annual checks on defaults in coding,

Table 3. Comparison of the four data sets

Type of Information	Data items	ANAC-1 Existence	ANAC-1 Codification	ANAC-2 Existence	ANAC-2 Codification	NIVEL STUDY Existence	NIVEL STUDY Codification	CMR NIJMEGEN Existence	CMR NIJMEGEN Codification
Provider	Identification	+	own	+	own	+	own	+	own
	Centre	+	own	+	own	+	own	+	own
Linkage with other data sets		**+ Provider data set** (name speciality, social security cards, etc)		–		**+ Provider questionnaires** name, age, sex, number of years in practice, univ. of education, etc.			
Patient	Identification	ø		+	own	+	own	+	own
	Age			+	age	α	date of birth	α	date of birth
	Sex			+	m/f	α	m/f	α	m/f
	Profession			+	INE	α	own	–	
	Type of assurance			+	own	–		–	
	Resident			+	own	–		–	
Linkage with other data-sets				–		**+ Census data Health Interview** (family situation, income profession, education, religion, morbidity, life-style, use of service, etc.		**+ Practice list** (date of birth, sex, family composition, social class	
Minimum unit of aggreg.		CONTACT		CONTACT		CONTACT		EPISODE	
	Date	+	month	+	date & day of the week	+	date	Onset and resolution of the episode	
	Type of contact	+	own	+	own	+	own	–	
	Initiator of c.	–		–		+	own	–	
	Length of c.	–		–		+	minutes	–	
	Time from previous	ø		+	own	∞		–	
	Contact in the episode	ø		+	own	∞		–	
Morbidity	Reason for encounter	ø		+	own	+	ICPC	–	
	Diagnosis/problem	ø		+ (main problem)	ICHPPC-2-Def	+ (all)	ICPC	+ (all epis.) E-book	
	Underlying cause disease/problem	ø		–		+	ICPC/own	–	
	Severity of morb.	ø		–		+	own	+	own
	Mortality	ø		–		–		+	

Table 3. Comparison of the four data sets (cont.)

Type of information	Data items	ANAC-1 Existence	ANAC-1 Codification	ANAC-2 Existence	ANAC-2 Codification	NIVEL STUDY Existence	NIVEL STUDY Codification	CMR NIJMEGEN Existence	CMR NIJMEGEN Codification
INTERVENTIONS									
Diagnosis	Physical exam.	–		–		+	own	–	
Procedures	Tests at surgery	–		–		+	own	–	
	Lab tests	α	number	+	number	+	own	–	
	X-ray	α	number	+	number	+	own	–	
	Other procedures	–		+	number	+	own	–	
	Reasons for diag.	–		–		+	own	–	
Referrals	Ref. Disciplines	α	own	+	own	+	own	+	own
	Referral type	–		–		+	own	–	
	Reason for Ref.	–		–		+	own	–	
	Initiative	–		–		+	own	–	
	Hospital admitta	–		–		–		+	own
Therapy	Sort of therapy	–		–		+	own	–	
	Prescriptions	+	number	+	number	+	name, number dosage	–	
	New-old prescrip.	–		–		+	own	–	
Others	Labour certific.	α	number	+	own	–		–	
	Length of labour	α	days	–		–		–	
Disposition	Follow-up	–		–		+	own	–	

Existence:
+ Item included
– Item not included
ø Not pertinent
α Item obtained from a linked data set
∞ Item obtained from transformation of other items

Codification system:
(own) own codification system or description of the system used.

regular meeting with practice receptionists, etc. This is only possible on a small scale, allowing for classification and coding by the diagnosing practitioner.

Discussion

This review of data sets gives a view of opportunities and problems of developing harmonised data sets in primary care throughout Europe. The striking similarities found between the three data systems are promising, but the differences also show particular problems which must be paid attention to in the process of building up harmonised tools for obtaining comparable information on morbidity and interventions in primary care.

These are related not only to the differing content of ambulatory care between countries, but also to the stage of development of health care in member states and other factors such as available resources for data collection.

Objectives

Some of the major differences found can be explained in part by different objectives, and can therefore be assumed to follow deliberate planning. As an example, the epidemiology tailored data base of Nijmegen has almost no need for information on resource use, while ANAC-1 data is limited to this and ANAC-2 and NIVEL Study shows a balance between both.

The structure of the health care system has direct consequences for the organisation of data bases, and imposes differences upon the three data systems as well.

Data collection

Another difference is the data collection approach: a very small group of highly motivated practices in CMR, a large sample in the NIVEL Study, and a combination of routine collection (ANAC- 1) and sample of contacts (ANAC-2) in the SIG-7 system.

Furthermore, there are differences in the actors who require information: researchers (CMR and NIVEL Study), national planners (NIVEL Study) and regional/local planners and managers (SIG-7).

The relation between information requirements and feasibility may also be noted. High quality data collection will only be acceptable for a small group of collectors, or for a limited time period. Conversely, data with a loose quality control can be acceptable for general planning purposes while it will be useless for specific research.

Data/population

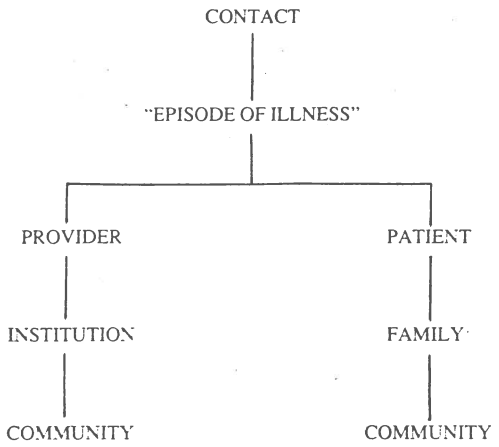
A major similarity between the systems is the use of population denominators (practice list or census data) although that the richness of this information is divergent. In this respect, the structure of health care plays a determining role, because in both countries every user is registered with one general practitioner, and specialist care is, generally, only available after referral. The differences found in the Spanish data are due to the fact that there is no accurate information about the patients on the individual General Practitioner (GP) list, and population (census) can only be used as a denominator as an aggregate at the level of health centres which match a town or neighbourhood. However, this is not only important in systems with practice lists. In 'free access' systems it may also be relevant to have information about catchment attraction of the GPs services.

Data/health problems

Another similarity can be found in the common data set of the three systems, which resulted from shared objectives. These similarities point to deliberate efforts to make the collected data comparable with others. Socio-demographic and morbidity data stand out as areas where the implementation of standardisation will be relatively simple to achieve. It must be kept in mind, however, that standards change over time. In particular, morbidity classification in primary care has been an area of rapid developments in the past decades, which will develop still further. There is a time lag between development and implementation, and as a consequence, existing data bases using standard classifications are at risk of becoming rapidly out of touch with these standards. This is particularly the case when data are to be collected over a larger period of time.

Greater standardisation differences appear on intervention (services) data. As seen in Table 3 the high prevalence of 'own' classifications (apart from the different level of detail given by each system) suggests that standardised classifications are

Figure 1. Units of aggregation of the information



far from being widely used. An added constraint in this is the use of some special coding for billing in specific European countries or agencies. Although some classification efforts have been made[3,4,12,13], it seems that a big effort in defining and using procedures standard classifications is still necessary.

A noteworthy difference between the systems is the possible aggregation levels in the three information systems (Figure 1). Because they identify episodes and contacts per patient, CMR and the Dutch National study have a richer aggregation perspective than the ANAC-2, where no analysis beyond contact level is possible. This difference can be traced to the aims of the projects and the related requirements. Information for some basic forms of utilisation review or practice administration can be satisfactorily collected without identifying episodes, but epidemiological monitoring and research needs to aggregate morbidity at episode or patient level. This clearly illustrates that even when the same harmonised data item is used, comparison of data can become impossible because of the strategy chosen.

Our conclusions suggest that these morbidity and intervention data sets can be a key tool in developing harmonisation on primary care information in Europe.

The three data systems have been shaped according to their objectives. The importance of the definition of objectives prior to actual data collection is self-evident, but seems to be easily overlooked in medical care. This implies that harmonisation does not necessarily mean developing a single *minimum data set* at European level[14]. It is an illusion to expect collected data to be multipurpose. Our work suggest the use of a common framework with a specific weight on particular *data types*, depending on the objectives and with a different level of development of concrete *data items*, balancing both the objectives and the data collection resources of each system. This can probably help to solve the ambiguity of the *flexible response* to new questions, and a consequent stable collection/processing of good quality data. Data bases will probably best develop step by step; but the level of expansion again depends upon the objectives. Careful planning through clear objectives, and the possibilities of standardisation do not alter the fact that the structure of health care will determine the comparability of data.

Careful *control of the quality of data* is essential, particularly when precise and detail information is needed. Quality control is costly in terms of contacts/patients/providers included in the data base, which in turn will have consequences for the relevance of the data base. Only for purposes where more global data suffice can secondary use of data be attempted. This is more feasible for planning and management than for research.

What appears to be evident is the standardisation needs of a framework for characterising and measuring morbidity and interventions, the *development and use of accepted classifications* for some data types (such as procedures).

The data collected is in some cases common, but the conditions under which they are collected differ. This causes particular problems in determining a *common denominator* to express the results. The practice population or the individual patient is the most explicit denominator, but it is not available in a number of European countries. The contact (visit) is on the other hand an easily available denominator, but lacks substance in interpreting data. It would contribute to the comparability and interpretation of data if a common, substantial denominator could be defined. The *'episode-of-illness'*[4,15] appears such a substantial conceptual denominator for registration systems. A clear unambiguous definition of this denominator would mean a very important step in standardisation of data base analysis.

References

1 Advanced Informatics in Medicine (AIM). Research and technology development on telematic systems in health care. Commission European Communities. Brussels, 1992.

2 WOOD M. MAYO F. MARSLAND D. Practice-based recording as an epidemiological tool. Ann Rev Public Health 1986:7:357-389.

3 International Classification of Health Problems in Primary Care-2-Defined. Oxford: Oxford University Press, 1983.

4 LAMBERTS H. WOOD M. eds. International Classification of Primary Care. Oxford: Oxford University Press, 1987.

5 Advanced Informatics in Medicine (AIM). AIM Workshop on Telematics in Primary Health Care. Commission European Communities. Brussels. 1991.

6 McACE consortium (AIM project A1009). European ambulatory care information systems: an exploratory study of harmonisation. Volume 2: Identifications of opportunities and problems in current health information systems, Commission European Communities. Brussels. 1990.

7 McACE consortium (AIM project A1009). European ambulatory care information systems: an exploratory study of harmonisation. Volume 1: Evaluation of Minimum Data Sets. Resource Use Management and Information Technology Support. Commission European Communities. Brussels. 1990.

8 Department of General Practice/Family Medicine. Morbidity figures from general practice. Data from four general practices 1978-1982. Nijmegen: Nijmeegs Universitair Huisartsen Instituut. 1985.

9 PORTELLA E. FERRUS L. CASAS JC. COBOS A. SIG/7, Sistema de Información para la planificación, gestión y evaluación de los servicios de Asistencia Primaria. Jornadas sobre Optimización en los Servicios Sanitarios de EADA. Barcelona, 12-14 Junio 1986.

10 FOETS M. VAN DER VELDEN J. VAN DER ZEE J. Morbidity and Interventions in General Practice. A cross-national survey in the Netherlands. Study design. Netherlands Institute of Primary Health Care (NIVEL). Utrecht (The Netherlands), 1986.

11 SPANJER M. General Practitioner in the Netherlands. Morbidity and interventions in General Practice, a cross-national survey. Netherlands Institute of Primary Health Care (NIVEL). Utrecht (The Netherlands), 1990.

12 Anatomical Therapeutic Chemical (ATC) classification index. Alphabetically sorted according to nonproprietary drug name. WHO Collaborating Centre for Drug Statistics Methodology. Norway, 1992.

13 International Classification of Process in Primary Care. Oxford: Oxford University Press, 1986.

14 CARRILLO E. BOLIBAR B. JUNCOSA S. PARKIN D. HUTCHINSON A. FISHER PJ. Conjuntos Mínimos de Datos en Atención Ambulatoria. Gaceta Sanitaria 1991;26/5:225-230.

15 GOULD M. AZEVEDO D. The content of adult primary care episodes. Public Health Reports 1982;97:48-57.

received	22-02-1993
revision asked by the editor	30-06-1993
accepted	30-06-1993

Consultations for women's health problems: factors influencing women's choice of sex of general practitioner

ATIE VAN DEN BRINK-MUINEN

DINNY H DE BAKKER

JOZIEN M BENSING

SUMMARY

Aim. *This study set out to examine the degree to which women choose to visit a woman doctor for women's health problems and the determinants of this choice. The differences between women and men doctors with regard to treating women's health problems were also studied.*

Method. *Data from the Dutch national survey of general practice were used. All group practices with both women and men general practitioners were selected. Analyses were restricted to consultations among women aged 15–65 years about menstruation, the menopause, vaginal discharge, breast examination and cervical smear tests.*

Results. *Given the size of their female practice population, women doctors saw considerably more women with women's health problems than did their male colleagues. Women were more likely to consult a woman general practitioner if she was more available (that is, working longer hours), and younger women were more likely than older women to choose women general practitioners. Sex differences in the treatment of women's health problems were small and mainly related to the verbal part of the consultation: counselling and providing information. The doctors' availability and their certainty about the working diagnosis explained differences in the verbal aspects of consultations. Women general practitioners had longer consultations than their male colleagues mainly because more health problems were presented per consultation.*

Conclusion. *In order to increase the possibility of patients choosing women general practitioners, policy should be directed towards the education of more women general practitioners and women general practitioners should be encouraged to work more days a week.*

Keywords: patient choice of doctor; women doctors; doctors' sex; women's health.

Introduction

ALL women are affected by menstruation and the menopause; some of them have problems with these physiological processes and consult their general practitioner. Many women undergo a vaginal examination in connection with these problems and other problems involving the genital system. Women also consult their doctor for sex-specific examinations such as breast examination and cervical smear tests, and for contraception and other sex-related issues. Earlier studies found that female patients, in general, have a preference for a doctor of their sex[1-3] and that they actually tend to choose a doctor of their own sex.[4-8] When consulting with problems of the female genital system, this preference is even stronger.[1,9,10] In one study, the preference for doctors of the same sex was strongest for items in which the complaint involved a sexual dysfunction and a thorough physical examination, for complaints of an intimate nature or where complete undressing was required.[2]

Why do women choose to visit a woman doctor for women's health problems? Women find it easier, or less embarrassing, to talk to a woman because of feelings of shame and fear, and taboos about genital problems.[1,9-11] Further, women are thought to have a better understanding of women's problems and are therefore easier to talk to.[10] Women doctors are also said to be more gentle when performing gynaecological examinations than men doctors.[11,12]

Patients' preference for a woman doctor could also be influenced by other factors, such as the general practitioner's experience and the patient's age and education. However, patient choice could be limited by the availability of women doctors. Women general practitioners are still in the minority, though their number has been increasing over recent years.[13-18] In addition, many of them work part-time and are therefore less frequently available than their mostly full-time male colleagues in the same group practices.[19]

Although much attention has been paid to women's health problems in the literature,[20-23] so far no data have been published about differences between women and men general practitioners in their treatment of these problems. Some studies indicate that counselling is an important factor in consultations,[14,24] particularly for gynaecological problems.[25,26] In consultations relating to women's health problems techniques such as reassurance and explanation have been advocated to reduce anxiety and to help patients relax.[27] As regards health problems in general, differences have been found in the treatment provided by women and men general practitioners.[19] Women doctors provided more counselling and advice on health education and lifestyle than male colleagues in the same group practices, while men more often gave information on health problems and treatment and offered more reassurance. The same study shows that women general practitioners admit to more uncertainty about their diagnoses than men general practitioners.

One could hypothesize that women and men general practitioners differ from each other in treating women's health problems. One reason could be that women doctors are more inclined to follow the two main principles of women's health care — consideration of the socialization and sexual identity of women and consideration of their social context.[28] Other principles of women's health care are showing respect for patients and encouraging them to cope with health problems and to take personal responsibility.[28] Following these principles can avoid medicalization of a problem. Advice on the integration of these principles into normal health care has been issued to the government.[29-31]

The number of women general practitioners is likely to increase in the future and this will have consequences for the accessibility and quality of care for women. The aims of this

A van den Brink-Muinen, MA, researcher; D H de Bakker, PhD, project leader; and J M Bensing, PhD, general director, Netherlands Institute of Primary Health Care (NIVEL), Utrecht, Netherlands.

Submitted: 21 January 1993; accepted: 28 July 1993.

study were, therefore, to determine to what degree women choose women general practitioners for women's health problems and the determinants of this choice, and whether women and men general practitioners differ in treating women's health problems and, if so, how this can be explained.

Method

The data were derived from the Dutch national survey of general practice,[32] a large scale study of morbidity and interventions in general practice. The study was conducted among 161 general practitioners working in 103 practices. Selection of participating doctors was based on a stratified random sample of all Dutch general practitioners (according to region, urbanization and distance to a general hospital). The general practitioners recorded detailed information about all patient contacts in a three-month period, with four groups covering one year (April 1987 to April 1988). The data recorded included patient characteristics, characteristics of the consultation (such as first or repeat consultation, length of consultation), problems presented, doctor's diagnoses (classified using *International classification of primary care*[33]) and services provided (diagnostic services, treatment, prescriptions, referrals). Counselling and information provided were recorded using the following categories: counselling passive/listening; counselling active/exploring; reassurance; information on health problems/disease; and information about treatment/medication/referral/operation/diet. The extent to which general practitioners were certain of their working diagnosis was recorded on a five-point scale for each problem presented (one, very uncertain to five, very certain). Data on the general practitioners' characteristics — sex, availability and experience — were obtained by written enquiry. Patients' characteristics, such as age, education and length of acquaintance with the doctor, were derived from national survey data.[32]

To ensure maximum uniformity in the data collection process all participating doctors were trained in the use of the diagnostic classification system. Written instructions with definitions were provided to keep at hand during the consultation.[34] Before the recording period started all the elements of the registration form were tested in each practice while a research assistant was present. During the recording period each practice was visited once a week by a research assistant to check the data as regards completeness and irregularities and to discuss problems that might have arisen. During the recording period the doctors were asked to formulate a diagnosis on the basis of a clinical vignette in order to get an idea of the extent to which participating general practitioners assign and name similar diagnoses. There was a reasonable degree of conformity (approximately 90%) indicating that this was a reasonably reliable method of registration. The services provided by the doctors were not validated.

For this study only data from group practices with both women and men general practitioners were considered (21 practices with 27 men and 23 women doctors) in order to balance patients' opportunity of choosing between a man and a woman general practitioner. This allowed distance to the practice to be excluded as a possible explanation for differences between women and men doctors, as well as the composition of the patient population, the characteristics of the area and the availability of other health services in the area.

Analysis

The analyses were restricted to women's health problems most commonly presented — consultations about menstruation, the menopause and vaginal discharge, and for breast examination and cervical smears. Only consultations with women patients aged 15–65 years (1654 consultations) were included, because women's health problems mainly arose in this age group (96.2% in this study).

Only health problems presented on the initiative of the patient alone or the patient and doctor together were selected in order to exclude screening consultations (5.3% of 1654 consultations were excluded) from the logistic regression analysis concerning preferences for a woman or man general practitioner. The two dependent variables were the general practitioner's sex and treatment provided — counselling (active or passive), information (on health problems or treatment), reassurance, vaginal examination, prescription and referral (to secondary care or to other primary health care professionals). The independent variables were general practitioner's sex, availability, experience and certainty about working diagnoses; and the patient's health problem, age, education and length of acquaintance with her doctor. The availability of the doctors was defined as the number of full time equivalents which they worked in the practice per week (0.1 full time equivalent is a morning or an afternoon). Experience was defined as the number of years a general practitioner had been working in general practice. Certainty about the working diagnoses was calculated for each general practitioner by taking the mean of certainty for all the problems presented. Patient education was divided into three categories by the level at which full-time education was finished: low (primary education), middle (secondary education) and higher (higher vocational training and university). The length of acquaintance of a patient with her doctor (and vice versa) was the number of years she had been registered with her present doctor.

In order to compare the numbers of women consulting a woman or a man general practitioner for women's health problems the consultations with each general practitioner were expressed both as numbers and as percentages of the listed women patients because the female patient population of men general practitioners was bigger than that of women general practitioners (60.4% versus 39.6% of women patients). However, not all consultations in the group practices studied were with the doctor with whom the patient was registered; it is quite possible for a woman listed with a man general practitioner to go to a woman general practitioner in the same group practice. Differences in patient ages[19] were taken into account by performing a direct standardization for the age distribution of all women aged 15–65 years in the practice population of the participating general practitioners (in five categories of 10 years).

The consultation length was recorded by the general practitioners using six categories. In the analysis the midmarks of these categories were used in order to calculate the mean consultation length.

The percentages and mean values presented were tested for statistical significance by a difference of proportions test and by a *t*-test; both test for independent random samples.[35] Multiple and logistic regression analyses were performed in order to explain differences between women and men general practitioners.

Results

Comparison of the two groups of general practitioners revealed no significant difference in their ages: the mean age of the 23 women doctors was 38.2 years and of their 27 male colleagues 41.8 years. Women and men general practitioners differed significantly in availability and certainty, but not in experience. Women general practitioners were less available (mean 6.1 versus 8.9 full time equivalents, $P<0.001$) and less certain about working diagnoses (mean 4.4 versus 4.6, $P<0.01$) than their male colleagues

Preference for a woman general practitioner

Women general practitioners were more often consulted for women's health problems than their male colleagues: 53.8% of all such consultations were with a woman general practitioner (Table 1). Given the size of their personal female practice population women doctors saw considerably more women's health problems, especially consultations about menopause problems and involving a cervical smear, than men doctors.

Women presented additional health problems in more of the consultations with women doctors than with men doctors (46.1% of 890 versus 40.9% of 763). More than half of the patients consulting their woman doctor about menopause problems (60.8%) and for a breast examination (55.4%) presented one or more other health problem in the same consultation; this was significantly more than in consultations with men doctors (37.8% and 42.3%, respectively; both $P<0.001$). In two thirds of the consultations for a cervical smear other health problems were presented (65.5% to women doctors and 64.4% to men); for menstruation problems the figures were 35.8% and 31.5%, respectively, and for problems concerning vaginal discharge 41.3% and 40.5%.

Logistic regression analysis was performed to find the determinants for the preference for a woman doctor. The chance a woman will visit a woman doctor about women's health problems is higher if the general practitioner is more available (odds ratio 1.08, $P<0.001$, for a doctor working 0.1 full time equivalents more). Younger women patients (in 10 year age bands) will more often choose a woman doctor than older patients (odds ratio 0.82, $P<0.001$).

Treatment

The services provided by women and men general practitioners for women's health problems are presented in Table 2. Women doctors spent time listening to their patients in significantly more consultations for a cervical smear and breast examination than men doctors. Active counselling was provided in significantly more consultations for menstruation problems with men doctors than women doctors while the reverse was true for consultations for a cervical smear. In most cases, men general practitioners provided reassurance in consultations more often than women general practitioners did. Information on health problems and their treatment was generally provided more often by men doctors than their female colleagues. No significant sex differences were found for performing vaginal examinations and referring patients, although overall more referrals to medical specialists were made in consultations with men doctors. Finally, men general practitioners prescribed medication in more consultations about menopause problems than women general practitioners, but no significant differences in prescribing were found for the other women's health problems.

In order to explain these differences logistic regression analysis was performed (Table 3). The chance of a patient receiving information or reassurance was higher when she visited a man

Table 1. Number of consultations for women's health problems and age standardized rate of consultation among women patients.

	Women GPs		Men GPs		
Type of problem	Number of consultations	Number per 1000 women patients (n = 10 566)	Number of consultations	Number per 1000 women patients (n = 16 112)	Ratio of rate women:men
Menstruation	380	34.7***	321	20.3	1.7:1
Menopause	51	5.4***	37	2.1	2.6:1
Vaginal discharge	208	18.8***	195	12.5	1.5:1
Cervical smear	177	16.9***	132	8.2	2.1:1
Breast examination	74	7.0**	79	4.9	1.4:1
Total	890	82.7	764	48.0	1.7:1

n = total number of women patients aged 15–65 years on GPs' personal lists. **$P<0.01$; ***$P<0.001$.

Table 2. Services provided in consultations for women's health problems.

	% of consultations									
	Menstruation		Menopause		Vaginal discharge		Cervical smear		Breast examination	
Service	Women GPs (n = 380)	Men GPs (n = 321)	Women GPs (n = 51)	Men GPs (n = 37)	Women GPs (n = 208)	Men GPs (n = 195)	Women GPs (n = 177)	Men GPs (n = 132)	Women GPs (n = 74)	Men GPs (n = 79)
Counselling										
Passive/listening	12.4	16.8	27.5	13.5	9.6	6.7	14.1***	4.5	17.6**	5.1
Active/exploring	16.8	25.5**	27.5	35.1	12.5	12.3	13.6***	4.5	17.6	12.7
Information										
Health problems/ disease	40.0	41.4	31.4	43.2	37.0	49.7***	14.1	12.9	29.7	35.4
Treatment/ medication/referral/ operation/diet	30.5	39.9***	21.6	40.5	37.5	44.1	12.4	9.8	23.0	21.5
Reassurance	16.1	24.3***	11.8	10.8	10.1	21.0***	7.3	15.2**	28.4	38.0
Vaginal examination	36.8	34.0	15.7	13.5	51.0	56.9	52.5	47.0	32.4	40.5
Prescription	44.7	43.9	45.1	73.0***	69.2	67.7	12.4	7.6	13.5	25.3
Referral	7.4	8.1	3.9	5.4	1.4	3.1	0.6	0.8	6.8	3.8

n = total number of consultations. **$P<0.01$; ***$P<0.001$.

Table 3. Odds ratios of the chance that a service will be provided by a woman general practitioner.

Service	Odds ratio		
	Uncontrolled	Controlled for health problem and patient's characteristics	Controlled for health problem and patient's and GP's characteristics
Counselling[a]	0.97	1.11	1.54**
Information[b]	0.76**	0.78	0.58***
Reassurance	0.53***	0.48***	0.38***
Vaginal examination	0.99	0.89	0.84
Prescription	0.87	0.90	1.06
Referral	1.08	1.33	0.76

[a]Passive/listening and/or active/exploring. [b]Health problems/disease and/or treatment/medication/referral/operation/diet. **$P<0.01$; ***$P<0.001$.

doctor for women's health problems than when she saw a woman general practitioner. Although the provision of information seems to be explained by patients' characteristics and health problems (column two), when the data are also controlled for the characteristics of the doctor another picture emerges (column three). Women general practitioners were more likely to provide counselling. The sexes did not differ significantly in terms of prescribing, referring or performing vaginal examinations.

The contribution of the various characteristics of the general practitioners is shown in Table 4. General practitioners who were more certain about the working diagnosis were more likely to provide counselling and information; doctors who were more available were less likely to provide information. More experienced doctors were more likely to provide counselling but less likely to provide information. However, experience is unlikely to be of great relevance because the odds ratios were all about 1.0. It should be noted that the strength of relationships with general practitioner characteristics will be overestimated, because the number of consultations per doctor is high and therefore standard errors will be smaller and as a result, relationships are more likely to achieve statistical significance.

Consultation length

The mean consultation time for women's health problems was 10.3 minutes with women doctors and 9.7 minutes with men doctors ($P<0.01$). This difference can mainly be ascribed to problems concerning the menopause and vaginal discharge: with women general practitioners consultations for these problems lasted a mean of 12.1 and 10.8 minutes, respectively, while for men general practitioners the mean times were 8.5 and 9.5 minutes, respectively (both $P<0.01$). There was no significant difference between the sexes for length of consultations for menstruation, breast examination and a cervical smear. Women seeing their doctor with menopause problems had the longest consultations (mean for all women 10.6 minutes), those attending for a breast examination the shortest (mean 9.6 minutes).

The general practitioner's sex alone makes only a small contribution to explaining the longer consultation time with women general practitioners (0.5% of variance explained, $P<0.01$). If patient characteristics, type of health problem and the number of health problems presented in one consultation are taken into account 17.8% of the variance is explained ($P<0.001$).

The availability and experience of the general practitioner, the length of acquaintance between doctor and patient and the number of problems the patient presents explain 20.0% of the variance in consultation length (Table 5). However, the general practitioner's sex and patient's education do not contribute to this explanation, nor do the general practitioner's certainty and patient's age. Part-time doctors had longer consultations, as did less experienced doctors. The less doctors were acquainted with the patient, the more time they took for the consultation. However, the strongest explanatory factor remains the number of problems presented by the patient.

Discussion

The system of data collection used in this study has advantages and disadvantages. On one hand, a great deal of accurate information is produced but, on the other hand, certain areas such as counselling and provision of information and reassurance allow room for interpretation on the part of the doctors. Whether counselling is recorded as active/exploratory or passive/listening will depend in part on the assessment of the doctor; provision of information may be so much part of their routine that it may not be recorded. These verbal and non-verbal aspects of communication need further investigation by means of independent observation methods, which ensure the validity of recording.

Another restriction is imposed by the group of general practi-

Table 4. Odds ratios of the chance that a service will be provided, by characteristics of general practitioner.

Service	Odds ratio by GP characteristic		
	Availability	Experience	Certainty
Counselling[a]	0.97	1.03**	6.74***
Information[b]	0.85***	0.97***	3.39***
Reassurance	0.90	0.98	1.71
Vaginal examination	0.95	0.99	1.59
Prescription	1.03	1.01	1.63
Referral	0.90	0.94	0.41

[a]Passive/listening and/or active/exploring. [b]Health problems/disease and/or treatment/medication/referral/operation/diet. **$P<0.01$; ***$P<0.001$.

Table 5. Standardized and unstandardized regression coefficients (beta and *B*) for length of consultation.

	Beta	*B*
GP characteristics		
Sex	–0.01	–0.08
Availability	–0.12***	–0.24
Experience	–0.09**	–0.05
Certainty	–0.03	–0.64
Patient characteristics		
Age	0.05	0.02
Education	0.05	0.39
Length of acquaintance	–0.11***	–0.73
Number of problems presented	0.35***	2.52

**$P<0.01$; ***$P<0.001$.

tioners selected for this study. Only group practices with women and men general practitioners were investigated, so that, in theory, patients had the chance of consulting a woman doctor. General practitioners in such practices may differ from other colleagues in practice style and hence the generalizibility of the results may be limited. This applies particularly to men general practitioners, because while fewer than half of all Dutch men general practitioners (42%) belong to a group practice, 80% of women general practitioners work in a group practice.[36]

These limitations should be borne in mind when considering the results of this study. More than half of the women studied visited a woman general practitioner for women's health problems (54%). This percentage is considerably higher than the figure of 43% for all health problems.[19] When the number of women in the practice population is taken into account, women doctors saw many more women with women's health problems than their male colleagues. The embarrassing nature of these health problems did seem to influence the choice of sex of general practitioner. Consulting a woman doctor seemed to diminish women's reserve about mentioning health problems in general; more women presented additional problems in consultations with women doctors than with men doctors.

Determinants of the preference for a woman doctor were the availability of the doctor and the age of the patient: women are more likely to consult a woman doctor if the doctor is more available, and younger women patients were more likely to choose a woman general practitioner than older patients. The issue of availability is an important finding. Only 14% of all Dutch general practitioners are women, many working part time[37] and women's preference in general practice is for part-time work.[36] The demand for women general practitioners is likely to increase among the future generation, given the preference of younger women for a doctor of their own sex. The implications for health policy are clear. To enable women to consult a woman doctor more women medical students have to be encouraged to become general practitioners. Additionally, vocational retraining should be organized for women general practitioners returning to work after raising a family.

Small sex differences were found in the treatment of women's health problems by general practitioners, mainly relating to the verbal part of the consultation. The small numbers of consultations, especially for menopause problems and breast examination, may explain why some differences did not reach statistical significance. In general, women doctors listened more often to their patients than men doctors (passive counselling); this was especially true with women who had problems with the menopause and with women undergoing an examination for breast or cervical cancer. Men general practitioners undertook more active counselling than women doctors with women consulting them for menstruation problems, while women general practitioners undertook more active counselling with women undergoing an examination for breast or cervical cancer. Information seemed to be more commonly provided by men than women doctors, as did reassurance. Communication and information are important issues, particularly from the point of view of women's health care, but it cannot be concluded from these findings that the principles of women's health care with regard to communication and provision of information are better met by women than by men general practitioners in the group practices investigated.

Another factor seen to be important in women's health care is the avoidance of medicalization where possible.[28] This would be apparent in prescribing and referring. With the exception of prescriptions for menopause problems, which were provided in significantly more consultations with men doctors than women, there were no sex differences in prescribing and referring. However, this exception is an interesting one, because women's health care emphasizes the menopause (and menstruation) as a normal physiological phenomenon that needs no medication in most cases.[28] It seems that women doctors support this idea more than men doctors. Men doctors more often gave information than women doctors, but because the content of the information was not investigated it cannot be assumed that this led to less medicalization.

How can these differences between women and men general practitioners in the treatment of women with women's health problems be explained? Counselling appears to be related to certainty about the working diagnosis: for both sexes combined more certainty coincided with more counselling. However, women doctors provided more counselling, especially passive counselling, but were less certain about their diagnoses, or were more willing to admit such uncertainty. This uncertainty is also demonstrated by a greater tendency to order laboratory tests and to ask the patient to make a repeat visit.[19] Whether this uncertainty should be considered as enhancing quality of care, or as increasing medicalization, cannot be determined from these data. Provision of information about health problems and their treatment is also related to more certainty. The more certain doctors are about the diagnosis, the more specific information they can provide. Less involvement in the practice also influences the extent of provision of information. This is probably a matter of time: part-time general practitioners plan and take more time per consultation and therefore they have more time to provide information to their patients. Another possible explanation is that counselling and provision of information are sex-related behaviours. Counselling can be considered as passive and affective behaviour, and may therefore be a specific characteristic of women.[28] Provision of information is regarded as active and instrumental behaviour and hence can be ascribed to men. Finally, reassurance was only related to sex: men general practitioners tried to reassure their women patients more often then their female colleagues. Perhaps this can be ascribed to the opposite sex of doctor and patient or it may indicate that reassurance is a typically male characteristic. Another explanation may be that women only reassure their patients if they are certain about the diagnosis. It can be concluded that to some degree certainty about the diagnosis and availability explain differences in the verbal aspects of consultations for women's health problems. However, the results do not imply that men or women supply a higher quality of care.

The length of consultation also deserves attention: women general practitioners had longer consultations for women's health problems than men general practitioners. In trying to find an explanation for this it was ascertained that the number of problems presented, in particular, influenced the consultation length. Since women general practitioners were more often consulted by patients presenting more than one health problem, this partly explains the sex difference. However, it also can be argued that women general practitioners encourage their patients to present additional problems. Other factors relevant to the length of consultation were the availability and experience of the general practitioner and the length of acquaintance between doctor and patient. The less general practitioners were involved in the practice, the less experienced they were and the less acquainted with a patient, the longer the consultation was. These factors explain only 20% of the variance in consultation length. Other factors must have contributed to the length of the consultation, for instance the content of communication, information and examination, but these were not investigated in this study.

Women more often consulted a woman than a man general practitioner for women's health problems and, in view of this preference, these problems especially should be a task for

women general practitioners. However, once in the consultation room there were few differences in the treatment of these problems, except in terms of communication. More detailed studies, particularly of the verbal and non-verbal aspects of consultations, are needed to determine whether there are differences between women and men general practitioners in the quality of care provided.

References

1. Fennema K, Meyer DL, Owen N. Sex of physician: patients' preferences and stereotypes. *J Fam Pract* 1990; **30:** 441-446.
2. Ackerman-Ross FS, Sochat N. Close encounters of the medical kind: attitudes toward male and female physicians. *Soc Sci Med* 1990; **14A:** 61-64.
3. Weyrauch KF, Boiko PE, Alvin B. Patient sex role and preference for a male or a female physician. *J Fam Pract* 1990; **30:** 559-562.
4. Hopkins EJ, Pye AM, Solomon M, Solomon S. The relation of patients' age, sex and distance from surgery to the demand on the family doctor. *J R Coll Gen Pract* 1968; **16:** 368-378.
5. Kelly JM. Sex preference in patient selection of a family physician. *J Fam Pract* 1980; **11:** 427-430.
6. Preston-Whyte ME, Fraser RC, Beckett JL. Effect of a principal's gender on consultation patterns. *J R Coll Gen Pract* 1983; **33:** 650-658.
7. Cooke M, Ronalds C. Women doctors in urban general practice: the patients. *BMJ* 1985; **290:** 753-755.
8. Graffy J. Patient choice in a practice with men and women general practitioners. *Br J Gen Pract* 1990; **40:** 13-15.
9. Haar E, Halitsky V, Sticher G. Factors related to the preference for a female gynaecologist. *Med Care* 1975; **12:** 782-790.
10. Nichols S. Women's preferences for sex of doctor: a postal survey. *J R Coll Gen Pract* 1987; **37:** 540-543.
11. Wigersma L. Kwaliteit van de hulpverlening door de huisarts bij genitale en seksuele klachten [Quality of care-giving by the general practitioner for genital and sexual complaints]. *Huisarts en Wetenschap* 1992; **35:** 53-56.
12. Oppenheimer C, Catalan J. Counselling for sexual problems. Ethical issues and decision-making. *Practitioner* 1981; **225:** 1623-1633.
13. Cooke M, Ronalds C. Women doctors in urban general practice: the doctors. *BMJ* 1985; **290:** 755-758.
14. Ogle KS, Henry RC, Durda K, *et al.* Gender-specific differences in family practice graduates. *J Fam Pract* 1986; **23:** 357-360.
15. Wakeford RE, Warren VJ. Women doctors' career choice and commitment to medicine: implications for general practice. *J R Coll Gen Pract* 1989; **39:** 91-95.
16. Hopper J. Full-time women general practitioners — an invaluable asset. *J R Coll Gen Pract* 1989; **39:** 289-291.
17. Kletke PR, Marder WD, Silberger AB. The growing proportion of female physicians: implications for US physician supply. *Am J Public Health* 1990; **80:** 300-304.
18. Maheux D, Dufort F, Béland F, *et al.* Female medical practitioners. More preventive and patient orientated? *Med Care* 1990; **28:** 87-92.
19. Bensing JM, van den Brink-Muinen A, de Bakker D. Gender differences in practice style: a Dutch study of general practitioners. *Med Care* 1993; **31:** 219-229.
20. Baart I, Baerveldt M (eds). *Dokteren aan vrouwen. De medicalisering van vrouwenlevens [Practice with women: medicalization of women's lives].* Amsterdam, Netherlands: SUA, 1986.
21. Horstman K, Brinkgreve C, van Lenning A, *et al.* Vrouwenkwalen? Feministische perspectieven op ziekte en gezondheid [Women's diseases? Feminist perspectives of illness and health]. *Tijdschr Vrouwenstudies* 1990; **11:** 175-183.
22. Meeuwesen L, Ruiter M, Westerbech-van Eeren J, *et al. Vrouwen en gezondheid [Women and health].* Hague, Netherlands: STEO, 1991.
23. Bransen E. Has menstruation been medicalised? Or will it never happen? *Soc Health Illness* 1992; **14:** 98-110.
24. Roter D, Lipkin S, Lortgaard A. Sex differences in patients' and physicians' communication during primary care visits. *Med Care* 1991; **29:** 1083-1093.
25. Petravage JB, Reynolds LJ, Gardner HJ, *et al.* Attitudes of women toward the gynecological examination. *J Fam Pract* 1979; **9:** 1039-1045.
26 Keane D, Woodward CA, Ferrier MB, *et al.* Female and male physicians. Different practice profiles. *Can Fam Physician* 1991; **37:** 72-81.
27. Haar E, Halitsky V, Stricker G. Patients' attitudes toward gynecological examination and to gynecologists. *Med Care* 1977; **15:** 787-795.
28. Ministerie van Sociale Zaken en Werkgelegenheid. Projectgroep vrouwenhulpverlening. *Vrouwenhulpverlening in de huisartsenpraktijk [Women's health care in general practice].* Hague, Netherlands: Ministerie van Sociale Zaken en Werkgelegenheid, 1986.
29. Ministerie van Sociale Zaken en Werkgelegenheid. Projectgroep vrouwenhulpverlening. *Eindadvies [Final advice].* Hague, Netherlands: Ministerie van Sociale Zaken en Werkgelgenheid, 1986.
30. Willemsen-Heddes ThM. Vrouwenhulpverlening en huisarts: een geïntegreerde visie [Women's health care and general practitioners: an integrated vision]. *Medisch Contact* 1990; **45:** 1470-1472.
31. Adviesgroep Vrouwenhulpverlening. *Advies functies van de vrouwenhulpverlening [Advice functions of women's health care].* Rijswijk, Netherlands: Ministerie van Welzijn en Volksgezondheid. 1990.
32. Foets M, van der Velden J, de Bakker D. *Dutch national survey of general practitioners: survey design.* Utrecht, Netherlands: NIVEL, 1992.
33. Lamberts H, Wood M (eds). *ICPC. International classification of primary care.* Oxford University Press, 1987.
34. Foest M, van der Velden J. *Een nationale studie van ziekten en verrichtingen in de huisartsprektijk. Basisrapport: meetinstrumenten en bijlagen [Dutch national survey of general practitioners. Base report: measuring instruments and appendices].* Utrecht, Netherlands: NIVEL, 1991.
35. Blalock HM. *Social statistics.* Toyko, Japan: McGraw-Hill, 1979.
36. Pool JB. *Cijfers uit de registratie van beroepen in de eerstelijnsgezondheidszorg 1991 [Numbers from registration of professions of primary health care 1991].* Utrecht, Netherlands: NIVEL, 1992.
37. Pool JB, Hingstman L. *Cijfers uit de registratie van beroepen in de eerstelijnsgezondheidszorg, deel I [Numbers from registration of professions of primary health care, part 1].* Utrecht, Netherlands: NIVEL, 1990.

Address for correspondence

Mrs A van den Brink-Muinen, NIVEL, Postbus 1568, 3500 BN Utrecht, Netherlands.

Molluscum contagiosum in Dutch general practice

S KONING

M A BRUIJNZEELS

L W A VAN SUIJLEKOM-SMIT

J C VAN DER WOUDEN

SUMMARY

Background. *While molluscum contagiosum is considered to be a frequently encountered disease, few data on its incidence are known.*

Aim. *The objective of this study was to describe the incidence of molluscum contagiosum in Dutch general practice and to assess the importance of venereal molluscum contagiosum.*

Method. *Data were taken from the national survey of morbidity and interventions in general practice, drawn from 103 practices across the Netherlands, with a study population of 332 300.*

Results. *The infection appeared to be common in childhood (cumulative incidence 17% in those aged under 15 years); the adult, sexually transmitted, form was rare. Incidence was higher between January and June than between July and December. Cases were unequally divided between recording practices, which is thought to have been caused by the occurrence of small epidemics.*

Conclusion. *The incidence of molluscum contagiosum in Dutch general practice was found to be 2.4 per 1000 person years. Molluscum contagiosum should still be considered as a mainly paediatric disease.*

Keywords: pox virus infections; skin diseases; morbidity; seasonal morbidity; differential diagnosis.

Introduction

MOLLUSCUM contagiosum is a self-limiting, viral cutaneous infection that has traditionally been regarded as a paediatric disease. Treatment with curettage, cryotherapy or electrocautery is considered easy. In the last few decades, it has become recognized that there are two different forms: the commonly known benign skin tumours in children and the sexually transmitted disease seen in young adults. Cobbold and MacDonald saw more patients with venereal than paediatric type molluscum contagiosum in 1970,[1] and suggested classification of molluscum contagiosum as a sexually transmitted disease. Most literature on molluscum contagiosum published since then refers to it as such. The incidence of the venereal form was shown to have increased dramatically from 1966 to 1983 in the United States of America,[2] as well as in England between 1971 and 1982.[3] Recently, a third group with molluscum contagiosum has been described: patients with the acquired immune deficiency syndrome (AIDS) or other immunosuppressive disorders and patients receiving immunosuppressive therapy.[4-7]

Even though molluscum contagiosum is commonly considered to be a frequently encountered disease,[1,5,6] few reliable data on its incidence are available.[5,6,8] Studies in a general or paediatric population are scarce and not recent.[9-13] They tend to describe epidemics, annual attack rates, or prevalence in a relatively small or local population. They frequently originate from warm climates, where the infection seems to be more common than in moderate climates. From general practice, only two brief reports were found.[14,15]

The aims of this study were to assess the incidence of molluscum contagiosum in Dutch general practice, by age, sex, region, season and degree of urbanization (population size). The hypothesis was that molluscum contagiosum in the general population was still much more common in children than in adults. In the Dutch health care system, general practitioners are the health care professionals that patients, including children, consult first and from whom patients need to be referred to obtain specialist care.[16] Therefore, they are good sources of information for studying morbidity in a general population.

Method

Data were analysed from the Dutch national survey of morbidity and interventions in general practice, carried out by the Netherlands Institute for Primary Health Care (NIVEL) in 1987 and 1988.[17] These data had been obtained from a randomly selected, non-proportionally stratified sample of 161 general practitioners in 103 practices, drawn from all registered Dutch general practitioners. Stratification took place to include a balance of practices across all regions of the Netherlands, and of rural and urban practices. The total practice population consisted of 332 303 patients, census data for whom had been recorded in advance.

The general practitioners were divided into four groups of approximately 40 each. Each group registered every contact between patient and practice during three months consecutively. The four registration periods were April–June, July–September, and October–December in 1987 and January–March in 1988, so that one whole year was covered and possible seasonal variations could be detected. The contacts were recorded on a specially designed contact registration form, on which were noted reasons for encounter, diagnoses and associated interventions. Reasons for encounter and diagnoses were subsequently coded, in accordance with a modified version of the *International classification of primary care*.

The national survey was based on an episode-oriented registration of morbidity, which meant that different consultations concerning the same health problem were linked to one episode of care. This enabled an accurate estimate of the incidence because provisional diagnoses were crystallized out into ultimate diagnoses. In the analysis only episodes in which the last diagnosis of the episode was molluscum contagiosum (*ICPC* subcode number S76.2) were considered.

In calculating incidence rates, the number of new episodes was

S Koning, MD, general practitioner trainee; M A Bruijnzeels, MA, data analyst; and J C van der Wouden, MA, PhD, research coordinator, Department of General Practice, Erasmus University, Rotterdam, Netherlands. L W A van Suijlekom-Smit, MD, paediatrician, Department of Paediatrics, Erasmus University, and University Hospital, Sophia Children's Hospital, Rotterdam, Netherlands.

Submitted: 15 October 1993; accepted: 1 March 1994.

used, the first consultation of which had taken place within the registration period. This number was divided by the total number of person years, that is, the period during which the patients were within the age group at risk multiplied by the number of patients.

The probability of a child visiting the general practitioner for molluscum contagiosum at least once before the age of 15 years, was calculated by conversion of the age-specific incidence rates into the cumulative incidence at age 15 years. The underlying assumption was that incidence rates are constant over time within each age category.[18]

Whenever possible, a distinction was made between venereal and non-venereal infection, by interpreting the registered localization of the lesions or the differential diagnosis.

Results

There were 236 episodes of molluscum contagiosum in the study year, of which 202 were new episodes. There were 271 consultation for the condition, thus the majority of episodes consisted of just one contact. Female patients accounted for 97 of the 202 new episodes (48.0%) and male patients for 105 (52.0%).

The number of new episodes and incidence in each age group are shown in Table 1. The majority of cases were recorded between the ages of six and 10 years. The difference in incidence between males and females was not statistically significant. This equal distribution among sexes was the same in any age group. The cumulative incidence up to the age of 14 years was 168 per 1000 children. This indicates that one out of six Dutch children aged 15 years have visited their doctor for molluscum contagiosum at least once in their life.

Eleven of 202 new episodes (5.4%) were recorded in patients aged 15 years and over, of which at least two and at most six can be considered to be the venereal form. This means that sexually transmitted molluscum contagiosum accounted for 1%–3% of all cases of molluscum contagiosum.

The incidence rates of molluscum contagiosum, by season, region, and degree of urbanization, are shown in Table 2. Of these three characteristics, region was the most distinguishing factor.

Forty one practices had made no diagnoses of molluscum contagiosum while 21 practices had each diagnosed one case (incidence of 1.4 per 1000 person years). Two or three cases had been diagnosed in 24 practices (2.8 per 1000 person years) and 17 practices had each diagnosed four or more cases (incidence of 5.7 per 1000 person years). Of the four practices that had recorded 58 cases altogether, three were situated in the central and one in the southern region.

Table 1. Incidence of molluscum contagiosum, by age.

Age (years)	No. of new episodes	Incidence (per 1000 person years)
0	0	0
1	7	6.2
2	16	15.2
3	10	9.5
4	8	7.7
5	15	14.4
6	24	22.8
7	22	21.1
8	25	24.4
9	19	18.8
10	25	25.0
11	7	6.8
12	6	5.8
13	3	2.7
14	4	3.3
15–24	7	0.5
>24	4	0.1
Total	202	2.4

Table 2. Incidence of molluscum contagiosum by season, region and degree of urbanization.

	Incidence (per 1000 person years)
Season	
April–June	2.8
July–September	1.9
October–December	1.8
January–March	3.3
Region	
North[a]	1.0
South[b]	2.7
Centre[c]	3.2
Urbanization (population)	
<30 000	2.3
30 000–50 000	2.5
>50 000	2.6

[a]Groningen, Friesland and Drenthe. [b]Noord-Brabant and Limburg. [c]All other provinces.

Discussion

The differential diagnosis of molluscum contagiosum includes keratoacanthoma, viral warts, granuloma and basal cell carcinoma. However, the condition is considered to be easily recognizable,[5,6,8] especially when multiple lesions are present. In the present study, diagnoses were usually made on clinical grounds, laboratory tests not being performed.

In a literature search, two brief reports on the incidence of molluscum contagiosum in general practice were found: in a large general practice in Belfast, Steele found four cases of molluscum contagiosum out of 2409 consultations,[14] which represented 2% of all dermatological cases. In the Hague, the Netherlands, Relyveld and colleagues found an incidence of four per 1000 patients per year, with a peak incidence between the ages of five and nine years.[15] Having been carried out in a big city, this study is not representative of rural areas; however, its result is comparable with the present findings.

No cases in children aged less than one year were observed in the study, which is in accordance with the international literature. This may reflect a passive immunity against the molluscum contagiosum virus that infants have obtained by maternal antibodies, or may indicate a long incubation period.[10]

The hypothesis that in general practice the paediatric type is more common than the venereal type of molluscum contagiosum, is strongly supported by the results: only between one and three per cent of the cases were of the venereal type. There are some limitations on the importance of this finding. First, the general practitioner may not recognize the genital pearly papules as molluscs. Secondly, a number of patients may be absent from these data because in the case of venereal disease, a Dutch patient can attend an outpatient department for sexually transmitted diseases directly without the need for prior referral from the general practitioner. No reliable incidence rates can be obtained from these clinics as molluscum contagiosum is not a reportable disease. The number of missing patients can be estimated to be small because the expected number of cases of venereal molluscum contagiosum occurring in the study population is 2.5 (calculated from the incidence per 100 000 of the population derived from English sexually transmitted disease clinics in 1982[19])

which is not higher than the actual number of two to six seen in the present study. That some dermatologists see relatively more cases of the venereal form than general practitioners might be explained by the probability that adults with molluscum contagiosum are more likely to be referred than children, and because of patients' direct attendance at outpatient departments for sexually transmitted diseases.

Several studies have associated molluscum contagiosum with various non-sexual modes of transmission: wrestling,[20] swimming,[10,13] intra-familial spread,[9,10] and sharing of towels.[21] Cultural differences may thus contribute to reported differences in incidence rates, peak ages, and sex distributions between countries. The finding that the incidence and prevalence in warmer climates are higher, and peak incidence is at a younger age than in cooler climates, is consistent in all studies. The peak incidence at ages six to 10 years found here is comparable with that of ages 10 to 12 years found in Scotland,[10] eight years in Japan,[13] and a mean age of six years in Alaska,[9] all in contrast with peak incidences at ages two to four years in Fiji, New Guinea and Congo.[10,12,22] The common explanation is that in these warmer countries transmission of the infection is facilitated by the moist conditions that are favourable to the virus,[21] as well as by the lightweight clothes and consequent easy skin-to-skin contact.[6] The finding that in the Netherlands the disease was encountered most in winter seems to be contradictory to this view. However, one has to consider the incubation period of the infection, which is not known precisely: estimates vary between two weeks and six months.[6] Assuming the longer incubation period to be correct would make this seasonal trend in incidence understandable.

The equal sex ratio found in the present study is in contrast to that in studies from Japan, Alaska and Fiji, where boys were affected more often.[9,10,13] However, neither in New Guinea, nor in the Hague,[12,15] were sex differences found. The most probable explanation is that between different countries and years, habits that are associated with the spread of the infection, such as swimming, are different for each sex.

The threefold difference in regional incidence is particularly striking, as neither climatic nor cultural differences between the three regions in the Netherlands play a considerable role. An important demographic feature of the northern region, its relatively low population density, cannot be the cause, because the degree of urbanization is not correlated with incidence. It may be that the skew in distribution of cases among practices brought about this result. In the past, several reports have been published on the occurrence of molluscum contagiosum in epidemics.[9,20,23] A few small epidemics of molluscum contagiosum may thus have caused the major regional differences in incidence found here.

This is the first study on the epidemiology of molluscum contagiosum in a large, national general practice population. Because molluscum contagiosum is not common, significant data on its incidence are difficult to obtain from small or local study populations. The main conclusions of this study are that the incidence of molluscum contagiosum in the Netherlands is 2.4 per 1000 person years, with a peak incidence between the ages of six and 10 years; the disease affects both sexes equally. Molluscum contagiosum is a mainly paediatric disease, the sexually transmitted form is uncommon. The incidence is higher between January and June than between July to December. Cases are unequally distributed among practices, possibly caused by small epidemics.

References

1. Cobbold RJC, MacDonald A. Molluscum contagiosum as a sexually transmitted disease. *Practitioner* 1970; **204:** 416-419.
2. Becker TM, Blount JH, Douglas J, Judson FN. Trends in molluscum contagiosum in the United States, 1966–1983. *Sex Transm Dis* 1986; **13:** 88-92.
3. Oriel JD. The increase in molluscum contagiosum. *BMJ* 1987; **294:** 74.
4. Cotton DWK, Cooper C, Barrett DF, *et al.* Severe atypical molluscum contagiosum infection in an immunocompromised host. *Br J Dermatol* 1987; **116:** 871-876.
5. Epstein WL. Molluscum contagiosum. *Semin Dermatol* 1992; **11:** 184-189.
6. Highet AS, Kurtz J. Molluscum contagiosum. In: Champion RH, Burton JL, Ebling FJG (eds). *Textbook of dermatology.* Oxford: Blackwell Scientific, 1992.
7. Schwartz JJ, Myskowski PL. Molluscum contagiosum in patients with human immunodeficiency virus infection. *J Am Acad Dermatol* 1992; **27:** 583-588.
8. Douglas JM. Molluscum contagiosum. In: Holmes KK, Mårdh P-A, Sparling PF, Wiesner PJ (eds). *Sexually transmitted diseases.* New York, NY: McGraw-Hill, 1990.
9. Overfield TM, Brody JA. An epidemiologic study of molluscum contagiosum in Anchorage, Alaska. *J Pediatr* 1966; **4:** 640-642.
10. Postlethwaite R, Watt JA, Hawley TG, *et al.* Features of molluscum contagiosum in the north-east of Scotland and in Fijian village settlements. *J Hygiene* 1967; **65:** 281-291.
11. Hawley TG. The natural history of mollumscum contagiosum in Fijian children. *J Hygiene* 1970; **68:** 631-632.
12. Sturt RJ, Muller HK, Francis GD. Molluscum contagiosum in villages of the West Sepik district of New Guinea. *Med J Aust* 1971; **2:** 751-754.
13. Niizeki K, Kano O, Kondo Y. An epidemic study of molluscum contagiosum. Relationship to swimming. *Dermatologica* 1984; **169:** 197-198.
14. Steele K. Primary dermatological care in general practice. *J R Coll Gen Pract* 1984; **34:** 22-23.
15. Relyveld J, Bergink AH, Nijhuis HGJ. Epidemiologische notities. Ulcus cruris, wratten en sterfsituatie. *Huisarts en Wetenschap* 1988; **31:** 266-267.
16. Boerma WGW, de Jong FAJM, Mulder PH. *Health care and general practice across Europe.* Utrecht, Netherlands: NIVEL, 1993.
17. Foets M, van der Velden J, de Bakker D. *Dutch national survey of general practice. A summary of the survey design.* Utrecht, Netherlands: NIVEL, 1992.
18. Kleinbaum DG, Kupper LL, Morgenstern H. *Epidemiologic research.* Belmont, CA: Wadsworth, 1982: 106-108.
19. Department of Health and Social Security. Sexually transmitted diseases. Extract from the annual report of the Chief Medical Officer of the Department of Health and Social Security for the year 1983. *Genitourin Med* 1985; **61:** 204-207.
20. Low RC. Molluscum contagiosum. *Edinburgh Med J* 1946; **53:** 657-659.
21. Brown ST, Nalley JF, Kraus SJ. Molluscum contagiosum. *Sex Transm Dis* 1981; **8:** 227-234.
22. Torfs M, Lambelin G. Considérations sur le molluscum contagiosum en milieu tropical. *Ann Soc Belg Med Trop* 1959; **39:** 703-709.
23. Murray MJ, Murray AB, Murray NJ, *et al.* Molluscum contagiosum and herpes simplex in Maasai pastoralists; refeeding activation of virus infection following famine? *Trans R Soc Trop Med Hyg* 1980; **74:** 371-374.

Acknowledgement

We wish to thank J van der Velden of the Netherlands Institute of Primary Health Care.

Address for correspondence

Dr J C van der Wouden, Department of General Practice, Erasmus University, PO Box 1738, 3000 DR Rotterdam, Netherlands.

Implementing Guidelines in General Practice. Evaluation of Process and Outcome of Care in Chronic Diseases

F.G. SCHELLEVIS,* J.TH.M. V EIJK,* E.H. VD LISDONK,‡ J. VD VELDEN§ and C. V WEEL‡

*Department of General Practice and Nursing Home Medicine/Institute for Research in Extramural Medicine, Vrije Universiteit Amsterdam, The Netherlands

‡Department of General Practice, Social Medicine and Nursing Home Medicine, University of Nijmegen, The Netherlands

§Netherlands Institute of Primary Health Care (NIVEL), Utrecht, The Netherlands

In a prospective longitudinal study over 21 months the performance of general practitioners and the disease status of their patients was measured during the formulation and implementation of guidelines on follow-up care. Data on 15 general practitioners and on 613 patients with hypertension, 95 with diabetes mellitus, 66 with chronic ischemic heart disease, 115 with chronic respiratory disease, and 17 with osteoarthritis were used for analysis. Performance measures were defined and for each disease a disease status indicator was used. The possibly modifying effect of compliance of the general practitioner, and of patient compliance with the guidelines was taken into account. The general practitioners formulated consensus guidelines on follow-up care and implemented these guidelines in their practices. The implementation was supported by peer review. The performance of general practitioners tended to conform more with the guidelines during the study period, especially with regards to actions that should be performed routinely. There were no major changes in the disease status indicators. Compliant hypertensive patients had a normotensive status more frequently than non-compliant patients. Diabetic patients were more likely to be normoglycaemic when they received care according to guidelines. None of the differences were statistically significant over time.

Key words: Guidelines, general practice, chronic diseases, quality of care, peer review.

INTRODUCTION

Formulating guidelines for optimal care is an important step in the process of measuring and improving the quality of care. In The Netherlands the development of guidelines started in 1982 with consensus meetings on controversial issues in specialist care by the 'Centraal Begeleidingsorgaan voor de Intercollegiale Toetsing' (CBO). Since 1989 guidelines ('standards') for optimal general practice care have been developed by the Dutch College of General Practitioners (NHG).

Until now, the process of formulating guidelines and their implementation in daily practice has been evaluated mainly by an indirect approach, by interviewing physicians or by using data on a highly aggregated national level [1–4], rather than directly on the level of individual doctors and their patients [5,6]. This paper reports on the evaluation of the implementation in general practice of guidelines for the follow-up care of patients with chronic diseases. Our study preceded the publication of the first Dutch standard [7]. Traditionally, three aspects of quality of care are considered: structure, process and outcome [8]. In this study we

Submitted 14 April 1994.

Correspondence: François G. Schellevis MD PhD, Institute for Research in Extramural Medicine, Vrije Universiteit, v.d. Boechorststraat 7, 1081 BT Amsterdam, The Netherlands.

evaluated aspects of process and outcome of care:

- did the actual care delivered by GPs tend to conform more to consensus guidelines for optimal care during their implementation, did it vary between GPs and between diseases?
- what was the course of disease status indicators of their patients during the implementation of guidelines, and was this course modified by patient characteristics and by compliance with the guidelines of GPs and of patients?

We assumed that GPs involved with the formulation and implementation of guidelines would increasingly perform according to these guidelines and that performance according to guidelines would result in a favourable course of disease status indicators in their patients.

METHODS

Design

Over 21 months from 1 January 1988, guidelines for optimal care were formulated, followed by their implementation in practice. During the whole period the GPs registered data on the actually delivered care, and of disease status indicators during all contacts with patients with selected chronic diseases. This evaluative study was of a prospective longitudinal descriptive nature.

Selection of GPs

The selection of practices and GPs has previously been described in detail [9]. In summary, 7 practices (with 15 GPs) were selected by convenience following their participation in the Dutch National Survey of General Practice [10]. The practices were located in the south-east part of the country. Two practices were staffed single-handedly, four had two GPs and one was a group practice with five GPs. Each GP had their own practice list. Five of the 15 GPs were women. Three practices were involved with vocational training for general practitioners, the others had no special relationship with a University Department of General Practice.

Selection of patients

The total list size of the practices at the start of the study amounted to 23,534 persons. The GPs identified in their practices all patients known to have one or more of the following diseases [9]:

- hypertension;
- diabetes mellitus;
- chronic ischemic heart disease (CIHD);
- chronic respiratory disease (asthma, chronic bronchitis, emphysema; CRD);
- osteoarthritis of knee and/or hip.

Baseline data [gender, age, and disease-specific data (date of diagnosis, doctor responsible for the follow-up care, blood pressure, and blood glucose)] were registered by the GPs at inclusion. For this study the patients had to satisfy two criteria: the diagnosis made before 1 January 1988, and a complete follow-up GP care during the study period. Table 1 lists the patient characteristics at baseline.

TABLE 1. Characteristics of five patient groups with chronic diseases at the start of the study

	Hypertension ($N = 613$)	Diabetes Mellitus ($N = 95$)	CIHD* ($N = 66$)	CRD ($N = 115$)	Osteoarthritis ($N = 17$)
Male (%)	31	33	49	58	18
Age (mean)	60	66	69	49	65
Initial BP (diastolic—mean—mmHg)	106				
Baseline BP (diastolic—mean—mmHg)	92		82		
Baseline blood glucose (mean—mmol/l)					
fasting		8.3			
non-fasting		10.9			

*CIHD=chronic ischemic heart disease; CRD=chronic respiratory disease; BP=blood pressure.

Process of formulating and implementing guidelines

Monthly meetings (except in July and in August) of 60–90 min were held with the participating GPs during the entire study period. These meetings were held in two, sometimes three subgroups. Attendance of all participating GPs was requested. When this was impossible, attendance of at least one representative per practice was urged. The first seven meetings in each subgroup of GPs were aimed at formulating consensus guidelines of optimal follow-up care for each of the five chronic diseases mentioned above. Because the subgroups did not discuss the five diseases in the same order, the first 7 months should be considered as one 'formulation phase'. For each disease the same procedure was followed. A summary of the 'state of the art' follow-up care and a draft version of proposed guidelines, both produced by the first author, were sent to the participants one week before each meeting. All subgroups received the same material. During the meeting the proposed guidelines were discussed and amended. A written report was prepared of each meeting and discussed and accepted in the next meeting. The presence of the first author at each meeting guaranteed that differences between the subgroups were discussed and that finally, identical guidelines were unanimously adopted. Table 2 (left column) lists the items of these guidelines for each of the five chronic diseases. All participating GPs explicitly stated their intention to act according to these guidelines and to motivate discrepancies between the guidelines and the actually delivered care.

During the subsequent meetings ('implementation phase') the GPs received written individual feedback on their actually delivered care. This information was derived from the data registered by the GPs themselves (see further). In each meeting discrepancies between the guidelines and the actual care regarding one disease were discussed using the method of peer review [11]. Again, a written report was made of these meetings, mailed to the participants, and discussed and approved at the next meeting.

The mean attendance rate of the individual GPs at the meetings was 68%. When computed on the practice level (attendance of at least one GP per practice) the mean attendance rate was 79%.

Measurements

Data on actually delivered care were recorded by the GPs on special forms during all consultations with the included patients over 21 months. These data included:

- diagnosis made during the consultation;
- performance of each of the following procedures (Yes/No): physical examination, blood pressure reading, measurement of body weight, blood glucose, serum creatinine, urine albumin excretion, ophthalmological examination (fundoscopy by the GP or referral to the ophthalmologist), influenza vaccination;
- making a follow-up appointment, defined in a term in weeks or months or by a specific date (Yes/No).

Diagnoses were coded by trained clerks according to the International Classification of Primary Care (ICPC) [12]. The diagnoses were clustered into disease episodes ('a problem or illness in a patient over the entire period of time from its onset to its resolution' [13]). At the first consultation of each episode it was indicated whether the episode was 'new' (disease never presented before) or 'old'.

Data on the disease status of the patients were registered on the same forms. These data included results of examinations and laboratory tests. Blood pressure was measured by the GPs with a sphygmomanometer or a digital manometer gauged at the start of the study. Diastolic blood pressure was read at the disappearance of the sounds (Korotkoff phase V). Blood glucose levels were generally measured in the office, using blood test strips and a reflectometer, or occasionally in a regional laboratory.

Outcome measures

For each of the five chronic diseases performance measures were defined, reflecting the extent of agreement of actual performance with performance according to the guidelines for each disease (Table 2, right column). The performance measures were first computed at the

TABLE 2. Guidelines on and measures of performance of general practitioners in the management of five chronic diseases

Guidelines	Performance measures
Hypertension	
• blood pressure measuring at each control visit	• percentage control visits in which blood pressure was measured
• urine albumin measurement once a year	• percentage of hypertensive patients whose urine albumin was measured during the study period
• follow-up appointment after each control visit	• percentage control visits finished by a follow-up appointment
Diabetes mellitus	
• taking recent history of signs and symptoms related to diabetes mellitus	—
• blood glucose measurement at each control visit	• percentage control visits in which blood glucose was measured
• body weight measurement at each control visit	• percentage control visits in which body weight was measured
• blood pressure measuring once a year	• percentage of diabetic patients whose blood pressure was measured during the study period
• influenza vaccination once a year	• percentage diabetic patients who were vaccinated against influenza during the study period
• serum creatinine measurement every three years	• percentage diabetic patients whose serum creatinine was measured during the study period
• ophthalmological examination every three years	• percentage diabetic patients who underwent ophthalmological examination during the study period
Chronic ischemic heart disease (CIHD)	
• taking recent history of signs and symptoms due to CIHD	—
• in case of hypertension: blood pressure measurement at each control visit	• percentage CIHD control visits in patients with hypertension in which blood pressure was measured
• in case of obesity: body weight measurement at each control visit	• percentage CIHD control visits in obese patients in which body weight was measured
Chronic respiratory disease (CRD)	
• taking recent history of signs and symptoms due to CRD	—
• lung examination at each control visit	• percentage control visits in which lungs were examined
• influenza vaccination once a year	• percentage CRD patients who were vaccinated against influenza each year
Osteoarthritis hip/knee	
• taking recent history of signs and symptoms due to osteoarthritis	—
• joint examination at each control visit	• percentage control visits in which joint examination took place

TABLE 3. Disease status indicators, and patient and care characteristics

Disease status indicators	Patient and care characteristics
Hypertension	
% of patients with diastolic blood pressure <95 mmHg	sex age initial blood pressure baseline blood pressure patient compliance received care: agreement with guidelines
Diabetes mellitus	
% of patients with blood glucose fasting <8.0 mmol/l or non-fasting <10.0 mmol/l	sex age baseline blood glucose patient compliance received care: agreement with guidelines
Chronic ischemic heart disease	
% of patients with diastolic blood pressure <95 mmHg	sex age baseline blood pressure patient compliance
Chronic respiratory disease	
incidence of exacerbations	sex age
Osteoarthritis	
incidence of joint related problems	sex age

patient level and then aggregated to the GP level.

Disease status indicators were defined for each of the five studied diseases and are listed in Table 3. For hypertension, diabetes mellitus, and CIHD these indicators were also mentioned as targets in the guidelines. For CRD and osteoarthritis episodes of exacerbations (acute bronchitis for CRD and joint related problems for osteoarthritis) were chosen as indicators. Exacerbations of chronic respiratory disease were defined as new episodes of acute bronchitis (ICPC code R78). Joint related problems were defined as new episodes of pain in knee/hip or myalgia (ICPC codes L13, L14, L15, or L18). Incidences were expressed in percentages of patients affected.

Potentially modifying variables

The potentially modifying variables were also listed in Table 2. Patient characteristics, such as age and diastolic blood pressure at the start of the study were dichotomised at the median of each disease group separately. Baseline blood glucose was dichotomised at 8.0 mmol/l for the fasting patient status or at 10.0 mmol/l for the non-fasting status. Initial diastolic blood pressure (measured in diagnosing hypertension) was dichotomised at 105 mmHg. Patient compliance was defined as the attendance at the minimum number of control visits according to the guidelines and dichotomised on the basis of the frequency distribution: 100% compliance vs less than 100%. A variable indicating received care was computed only for the cases of hypertension and diabetes mellitus, the number of cases in the other disease groups being too small. This variable reflected the agreement between the actual GP performance and the guidelines (observed vs expected) in each patient, viz the frequency of measurement of blood pressure in hypertensive patients and measurement of blood glucose in diabetic

patients. This variable was also dichotomised on the basis of the frequency distribution: 100% agreement vs less than 100%.

Analysis

Analysis was performed per disease. The entire study period was divided in periods and the results were presented as repeated measurements. For patients with diabetes five 4-month periods and for the others three 6-month periods were distinguished according to the guidelines on the frequency of follow-up visits. As mentioned above, the first 6–8 months can be regarded as the phase of formulation of guidelines, the following period as the implementation phase.

The performance measures of the GPs ($N = 15$) were presented as medians. Inter-doctor variation was expressed by computing the 33- and 67-percentile value.

The disease status indicators were presented as proportions of the patients. Subgroups were defined on the basis of the potentially modifying variables for bivariate analyses. Confidence intervals (CI) were computed at the 95% level. The analyses were carried out with SPSS-X and SPSS-PC.

TABLE 4. Performance measures of general practitioners ($N = 15$) in patients with hypertension. Median percentages and 33- and 67-percentiles) per 6-month period
Mean number of patients per GP: 40.9 (range 1–111)

	Measuring blood pressure		Making appointment	
Period	%	(33–67 perc)	%	(33–67 perc)
1	90.2	(84.7–97.0)	83.5	(72.2–95.7)
2	98.9	(97.1–100.0)	92.6	(90.9–99.3)
3	99.6	(98.2–100.0)	92.3	(85.1–100.0)

TABLE 5. Performance measures of general practitioners ($N = 15$) in patients with diabetes mellitus. Median percentages and 33- and 67-percentiles per 4-month period
Mean number of patients per GP: 6.3 (range 1–16)

	Measuring glucose		Measuring body weight	
Period	%	(33–67 perc)	%	(33–67 perc)
1	76.5	(66.7–81.0)	50.5	(31.8–52.8)
2	82.1	(68.8–85.7)	52.1	(30.0–54.2)
3	85.7	(78.6–87.5)	64.3	(35.0–66.7)
4	83.3	(77.1–83.3)	44.4	(12.5–47.2)
5	89.6	(80.0–91.8)	55.0	(25.8–56.3)

RESULTS

Performance

Hypertension. The GPs conformed increasingly to the guidelines during the study period with a decrease in variation between GPs in measuring blood pressure, and making an appointment (Table 4). This was reflected in a narrowing of the range between the 33 and 67 percentiles and in the results on the level of individual GPs: five GPs had a maximum score on both performance measures in the first period and these remained maximal, measuring blood pressure increased at least 10% in six GPs, and making a follow-up appointment increased at least 10% in five GPs.

Measurement of albuminuria was rarely performed during the study period: once in 60 of the 613 patients and twice or more in eight patients, never in the remaining patients.

Diabetes mellitus. The fraction of control visits in which blood glucose was measured increased steadily from 75 to 90% with a decrease in the range between the 33 and 67 percentiles (Table 5). The change in performance measures of individual GPs varied: two GPs had a maximum score in the first period which remained stable, and five GPs showed an increase of at least 10% between the first and the last period. The performance measures of the others remained stable at a lower level or decreased. Measuring body weight was less frequently performed and the increase was marginal: the measure of two GPs increased, of another two GPs decreased, and the others remained stable. All GPs but one measured the blood pressure of their diabetic patients twice or more during the study period. Only three GPs vaccinated all their diabetic patients against influenza. Serum creatinine was measured during the study period at least once in 24% of the diabetic patients. When extrapolated to 3 years (according to the guidelines), two GPs would be fully compliant in the measurement of serum

TABLE 6. Performance measures of general practitioners ($N = 15$) in patients with chronic respiratory disease. Median percentages and 33- and 67-percentiles per 6-month period
Mean number of patients per GP: 7.7 (range 1–22)

Period	Lung examination %	(33–67 perc)
1	92.1	(72.7–100.0)
2	98.6	(75.0–100.0)
3	93.8	(80.0–95.4)

creatinine of all their diabetic patients once in three years. Ophthalmological examinations were carried out in 26% of the patients at least once during the 21 months. One GP would reach an extrapolated measure of 100% of his diabetics undergoing an ophthalmological examination in 3 years.

Chronic ischemic heart disease (CIHD). The performance measures regarded only 41 of the 66 CIHD-patients, namely those with hypertension (N=17) and those with obesity (N=24). No clear trend was detectable in measuring blood pressure in hypertensive CIHD patients (mean percentages in period 1, 2 and 3: 100, 75, 80% of the visiting patients) and in measuring body weight in obese patients (27, 60 and 50% of the visiting patients, respectively). Aggregation on the GP level resulted in too small, and therefore unreliable, figures.

Chronic respiratory disease (CRD). Compliance with the guidelines regarding lung examination remained stable between 80 and 85% during the study period (Table 6) with a decrease of the interdoctor variation: five GPs remained on their maximum score, and four GPs showed an increase of at least 10%. Of the 115 patients 31% were vaccinated against influenza; the vaccination rate per GP ranged from 0 to 60% of the CRD patients.

Osteoarthritis. The agreement of GP performance with the guidelines on the follow-up of osteoarthritis patients, expressed in the fraction of control visits in which joints were examined, decreased during the study period from 72 to 59%. These figures are based, however, on limited numbers of patients and control visits. Therefore, aggregation to GP level has not been executed.

Disease status indicators

Hypertension. The proportion of hypertensive patients with a diastolic blood pressure below 95 mmHg did not change during the study period (Table 7). This course was not modified by sex, age, initial blood pressure, and baseline blood pressure. Whether the GP measured the blood pressure at each control visit or not did not modify the course of the number of normotensive patients (Table 7). There was a decreasing trend in the number of normotensive patients in the non-compliant patient group.

Diabetes mellitus. The proportion of diabetic patients with a normoglycaemic status fluctuated during the five 4-month periods with a marked dip in the fourth period (Table 8). This proportion tended to increase in females, older diabetics, and patients who were hyperglycaemic at baseline. In the subgroup in which the GP measured the blood glucose at each control visit the number of diabetics with a normoglycaemic

TABLE 7. Proportion of hypertensive patients with diastolic blood pressure <95 mmHg during three 6-months period in total, and controlled for GP care and patient compliance

		Received care: agreement with guidelines		Patient compliance	
Period	Total ($N = 613$) % (95% CI)	<100% ($N = 78$) % (95% CI)	100% ($N = 500$) % (95% CI)	<100% ($N = 183$) % (95% CI)	100% ($N = 414$) % (95% CI)
1	74 (70.0–78.0)	76 (63.4–86.4)	74 (69.2–77.8)	79 (69.6–87.1)	73 (68.2–77.2)
2	71 (66.4–74.6)	67 (54.0–78.7)	71 (66.5–75.2)	63 (52.2–73.3)	72 (67.7–76.5)
3	74 (70.2–78.2)	77 (62.5–87.2)	74 (69.5–78.1)	65 (54.1–74.6)	77 (72.2–80.9)

TABLE 8. Proportion of diabetics with normoglycaemia during five 4-month periods in total, and controlled for GP care and patient compliance

		Received care: agreement with guidelines		Patient compliance	
Period	Total ($N = 95$) % (95% CI)	<100% ($N = 50$) % (95% CI)	100% ($N = 44$) % (95% CI)	<100% ($N = 39$) % (95% CI)	100% ($N = 55$) % (95% CI)
1	47 (34.3–59.8)	61 (43.4–76.0)	27 (11.6–47.8)	40 (16.3–67.7)	49 (34.4–63.7)
2	44 (30.9–58.6)	39 (21.5–59.4)	50 (29.9–70.1)	39 (13.9–68.4)	46 (30.7–62.6)
3	49 (35.6–62.7)	47 (28.3–65.7)	52 (32.0–71.3)	53 (26.6–78.7)	48 (32.0–63.6)
4	32 (20.3–45.0)	22 (8.6–42.3)	39 (22.9–57.9)	29 (10.3–56.0)	33 (19.1–48.5)
5	47 (33.6–61.2)	39 (20.2–59.4)	55 (35.7–73.6)	53 (26.6–78.7)	45 (29.3–61.5)

status rose steadily (Table 8). Patient compliance in follow-up visits did not seem to modify the proportion of diabetics with normoglycaemia over time.

Chronic ischemic heart disease. The majority of CIHD patients in this study had a diastolic blood pressure below 95 mmHg already at baseline. Their number increased further during the study period from 87 to 91%. Subgroup analysis revealed a more pronounced increasing trend in women and older patients, and in the subgroup with a higher diastolic blood pressure at baseline. The results of the small number of compliant patients did not allow any conclusion.

Chronic respiratory disease. The overall incidence of exacerbations in patients with CRD did not change during the study period (19, 21 and 17% in the three periods). Analyses of subgroups showed higher incidences in the older patients.

Osteoarthritis. The 17 patients with osteoarthritis presented three new episodes of joint complaints to their GPs, one in the second semester and two in the last 6 months. These incidence rates did not allow any further analysis of subgroups.

DISCUSSION

Performance

The performance of GPs never reached full agreement with the consensus guidelines and it differed between the various performance measures, between the diseases, and between the GPs. There was an overall slightly positive trend during the 21-month period, which was reflected in an increase of the performance measure and/or a decrease in the variation between the GPs. Differences between periods were always within the 95% confidence interval values, indicating no significant change.

Some performance measures showed an increasing trend, others did not show a clear change. Those that increased reflected actions which had to be carried out routinely at each visit (blood pressure measurement, making appointment, measuring blood glucose) and had an immediate impact on the further management of the disease. Those that did not change during the study concerned actions that were to be carried out intermittently and/or had to do with risk factors with only a potential influence on the disease in the future: e.g. measurement of serum creatinine and ophthalmological examination, vaccination against influenza. Apparently, actions with direct consequences for the disease and its management were considered more important and were therefore performed more consistently during the implementation of the guidelines. This finding deserves more attention in research on quality of care and in projects on implementation.

The evaluation of actual GP performance from registered data bears the risk of under- or overestimation of the actual performance. A deficient registration regime is a possible source of underestimation. One could argue, however, that deficient registration represents a low

quality of care [14]. Overestimation would occur if the participating GPs registered procedures which they had not actually carried out. This was not checked, but it seems very unlikely that this occurred during the study period, since participation in the study was voluntary, there were no sanctions on non-compliant performance, the atmosphere during the meetings was friendly, and non-compliant performance led to more insight into one's own work. The GPs used the same form during the entire study period and for all patient groups. Although registration could have had an additional influence on performance, we assume that changes in performance, if any, were mainly induced by participation in the formulation and implementation of the guidelines and the follow-up meetings.

We have no clear explanation why only some GPs on some performance measures reached full compliance with the guidelines. It has been stated that "simply feeding back information on performance has almost no impact on changing clinical behaviour" [15]. Feedback most probably influences clinical practice in doctors who already agreed to review their practice [6]. In this project feedback of information was supported by peer review. More intensive feedback, e.g. on individual cases [16], or a longer implementation period might bring about more changes in performance.

Evaluation of the process of medical care should ideally be carried out by measuring performance of actions which can be considered to be related to the outcome of care. The value of measuring blood pressure in patients with hypertension and blood glucose levels in diabetic patients will not be disputed, but a meaningful performance measure reflecting optimal care for CIHD or osteoarthritis patients is not available. The availability of a meaningful performance measure would be an additional requirement for adequate guidelines [1].

Disease status indicators

Improvement of doctors' compliance with guidelines is generally expected to improve the outcome in patients [17], but there is little empirical evidence to support this. In this study the disease status indicators of patients with five common chronic diseases fluctuated during the study period without showing a clear trend. The results indicated improvement in some patient groups:

- in compliant hypertensive patients the number of normotensives increased;
- in diabetic patients who received care according to guidelines the number with a normoglycaemic status increased.

More detailed analysis of patients whose disease status indicators improved and those whose did not, could reveal other patient or care characteristics that modified the effects. Our findings should be regarded only as trends, because all the confidence intervals showed overlap. The modifying effect of patient compliance and received care could only be studied adequately for patients with hypertension and with diabetes mellitus for whom a simple quantitative measure could be computed. Similar trends might be visible in other patient groups if such measures were available.

Our results in diabetic patients are in accordance with a recent report on the influence of patient, doctor, practice and care characteristics on control in diabetes mellitus, where only 15% of the variation in glycolysated haemoglobin between patients could be explained [18]. In another study participation by GPs in standard setting for common childhood conditions showed improved respiratory function in their patients [19].

The disease status indicators used in this study were those upon which GPs usually base their management in daily practice. They reflect short-term results as in the case of diastolic blood pressure and blood glucose. For chronic diseases long-term outcome measures such as cardiovascular morbidity or mortality, diabetic retinopathy or respiratory function would be more preferable. However, this would require a far longer study period.

There could be an interaction between patient compliance and delivered care since delivery of optimal care is only possible in fully compliant patients. This interaction was not taken into account in our study. As none of the bivariate analyses showed differences it did not seem meaningful to carry out multivariate analysis.

It could be argued that the trends measured in this study can only be validly ascribed to the intervention if a control group of GPs and/or a

control period before intervention were included [20]. However, it was reasoned that neither approach would have provided a valid control. A control group of GPs would have consisted of doctors who continued to deliver their usual care during the same period without the influence of attending educational programs, reading publications or discussing problems on these subjects with colleagues. This would have created an artificial environment, which would no longer represent daily practice. A control group as well as a 'before and after' design would have included registration of actions during consultations on detailed research forms, without influencing the nature of delivered care. Such registration would undoubtedly have evoked the performance of actions mentioned on the forms, and thus the control group or control period would not have represented the usual care situation. The alternative of collecting information on delivered care from the patients' records would have been unsatisfactory because these are usually restricted to outcome measures relevant for the follow-up of the patients.

REFERENCES

1. Grol R, Development of guidelines for general practice care. *Br J Gen Pract* **43**: 146–151, 1993.
2. Hill M N, Levine D M and Whelton P K, Awareness, use, and impact of the 1984 joint national committee consensus report on high blood pressure. *Am J Publ Health* **78**: 1190–1194, 1988.
3. Vang J, The consensus development conference and the European experience. *Int J Tech Assess Health Care* **2**: 65–76, 1986.
4. Grol R, National standard setting for quality of care in general practice: attitudes of general practitioners and response to a set of standards. *Br J Gen Pract* **40**: 361–364, 1990.
5. North of England Study of Standards and Performance in General Practice, Medical audit in general practice. I: Effects on doctor's clinical behaviour for common childhood conditions. *Br Med J* **304**: 1480–1484, 1992.
6. Mugford M, Banfield P and O'Hanlon M, Effects of feedback of information on clinical practice: a review, *Br Med J* **303**: 398–402, 1991.
7. Anonymous, Standard of the Dutch College of General Practitioners on Diabetes mellitus type II [In Dutch]. *Huisarts Wet* **32**: 15–18, 1989.
8. Donabedian A, The quality of care. How can it be assessed? *JAMA* **260**: 1743–1748, 1988.
9. Schellevis F G, Lisdonk E vd, Velden J vd, Eijk JThM v and Weel C v, Validity of diagnoses of chronic diseases in general practice. *J Clin Epidemiol* **46**: 461–468, 1993.
10. Foets M, Velden J vd, Bakker D de, *Dutch national survey of general practice. Survey design*. Netherlands Institute of Primary Health Care, Utrecht, 1992.
11. Grol R, Mesker P and Schellevis F, Eds, *Peer review in general practice*. University Department of General Practice, Nijmegen, 1988.
12. Lamberts H and Wood M, *International Classification of Primary Care (ICPC)*. Oxford University Press, Oxford, 1987.
13. Classification committee of WONCA, An international glossary for primary care. In: *Classification committee of WONCA. International Classification of Health Problems in Primary Care (ICHPPC)-2-Defined*, (3rd Edn). Oxford University Press, Oxford, 1983.
14. Meyboom W A, Performance in general practice and patient record keeping [Dissertation] [In Dutch]. University of Nijmegen, Nijmegen, 1990.
15. Mitchell M W and Fowkes F G R, Audit reviewed: does feedback on performance change clinical behaviour? *J R Coll Phys London* **19**: 251–254, 1985.
16. Sommers L S, Sholtz R, Shepherd R M and Starkweather D B, Physician involvement in quality assurance. *Med Care* **22**: 1115–1138, 1984.
17. Haines A and Feder G, Guidance on guidelines [Editorial]. *Br Med J* **305**: 785–786, 1992.
18. Pringle M, Stewart-Evans C, Coupland C, Williams I, Allison S and Sterland J, Influences on control in diabetes mellitus: patient, doctor, practice, or delivery of care? *Br Med J* **306**: 630–634, 1993.
19. North of England Study of Standards and Performance in General Practice. II: Effects on health of patients with common childhood conditions. *Br Med J* **304**: 1484–1488, 1992.
20. Russell I and Grimshaw J, The effectiveness of referral guidelines: a review of the methods and findings of published evaluations. In: (Eds Roland M and Coulter A). *Hospital referrals*, pp. 179–211. Oxford University Press, Oxford, 1992.

Consultation rates and incidence of intercurrent morbidity among patients with chronic disease in general practice

F G SCHELLEVIS

E H VAN DE LISDONK

J VAN DER VELDEN

S H J L HOOGBERGEN

J TH M VAN EIJK

C VAN WEEL

SUMMARY

Background. *Information on frequency of consultation and presented morbidity among patients with chronic disease is relevant to the management of these patients in view of the increasing prevalence of chronic diseases.*

Aim. *This study set out to examine consultation rates and incidence of intercurrent morbidity in general practice in cohorts of patients with five common chronic diseases: hypertension, chronic ischaemic heart disease, diabetes mellitus, chronic respiratory disease and osteoarthritis.*

Method. *In seven practices with 15 general practitioners the records of all patients were screened for inclusion in the study. The data used for analysis were from 962 patients, whose diagnoses were made in agreement with diagnostic criteria, who were not under specialist care, and who were followed up for 21 months. A distinction was made between patients with one, or two or more of the five chronic diseases studied. For the single disease subgroups of patients with hypertension or diabetes two reference groups of people without a chronic disease, standardized for age and sex, were identified from the population in the same practices.*

Results. *Consultation rates were higher for patients with comorbidity than for patients with a single disease. Intercurrent diseases were presented more frequently to the general practitioner by patients with comorbidity than by patients with a single disease. Most intercurrent morbidity consisted of acute common diseases such as myalgia, upper respiratory tract infection and urinary tract infection. Patients with only hypertension or only diabetes had higher consultation rates than the corresponding reference group but did not have higher total incidence rates of intercurrent morbidity.*

Conclusion. *Patients with chronic disease consult their general practitioner frequently, and patients with more than one chronic disease consult even more frequently. The general practitioner has to deal with chronic disease and intercurrent acute disease in a single patient.*

Keywords: chronic disease; morbidity; consultation rates; associated conditions.

F G Schellevis, MD, PhD, general practitioner and J Th M van Eijk, PhD, professor, Department of General Practice and Nursing Home Medicine/Institute for Research in Extramural Medicine, Vrije Universiteit, Amsterdam, Netherlands. E H van de Lisdonk, MD, PhD, associate professor and C van Weel, MD, PhD, FRCGP, professor, Department of General Practice and Social Medicine, University of Nijmegen, Netherlands. J van der Velden, MD MPH, general practitioner and S H J L Hoogbergen, MD, student epidemiologist, Netherlands Institute of Primary Health Care (NIVEL), Utrecht, Netherlands.

Submitted: 6 April 1993; accepted: 7 October 1993.

Introduction

IN THE Netherlands chronic diseases are primarily managed by general practitioners. This care includes the diagnostic and therapeutic activities in the initial phase as well as the long term management of the disease.[1] It is expected that general practitioners will have to care for more patients with chronic disease in the near future owing to an increase in the number of elderly people[2] and to political measures emphasizing primary care. Insight into the workload generated by the care of patients with chronic diseases is relevant for the management of chronic diseases in general and for the organization of general practice in the future.

This study analysed consultation rates and intercurrent morbidity presented to the general practitioner by cohorts of patients with five common chronic diseases. In a previous study it was found that 7.6–40.3% of patients had combinations of two or more of the five chronic diseases (referred to as 'comorbidity').[3] Therefore, the influence of comorbidity on consultation rates and intercurrent morbidity (incident diseases presented to the general practitioner) justifies special attention. This study attempted to determine the difference in consultation rates between patients with a single chronic disease and patients with comorbidity, and the incidence and nature of intercurrent morbidity in these patient groups.

Intercurrent morbidity among patients with chronic diseases may be influenced by the phenomenon described by Berkson,[4] indicating a higher chance of diagnosing diseases in patients who are already receiving care than in patients who do not consult their general practitioner. For this reason reference groups without a chronic disease were included in the analysis.

Method

Study cohorts

The selection of practices and patients has previously been described in detail.[5] In summary, seven practices (15 general practitioners) were selected following their participation in the Dutch national survey of general practice.[6] The total practice population consisted of 23 534 people at the start of the study (1 January 1988).

The general practitioners identified all patients in their practices known to have at least one of the following diseases: hypertension, chronic ischaemic heart disease, diabetes mellitus, chronic respiratory disease (asthma, chronic bronchitis, emphysema) and osteoarthritis of knee and/or hip. A total of 1989 patients were identified. Background data were collected for each patient on the diagnostic procedures performed in dia-

gnosing the chronic disease, the date of the diagnosis and the physician responsible for the follow-up care. Five cohorts were defined on the basis of the following criteria: diagnosis made before 1 January 1988, diagnosis in agreement with the diagnostic inclusion criteria of *ICHPPC-2-defined*[7] and patient not receiving follow-up care for the chronic disease from a specialist at the start of the study. This final requirement ensured that all morbidity during the study period was presented to the general practitioner, because in the Netherlands medical specialists can only be consulted after referral by a general practitioner. Patients who left the practice during the 21-month study period were excluded.

Application of these criteria left a total of 962 patients who were included in the analysis. The characteristics of these patients are shown in Table 1. Each cohort was divided into single disease and comorbidity subgroups.

Reference groups

Two separate reference groups of people without a chronic disease were constructed from the population of the same practices for the single disease subgroups of the cohort with hypertension and that with diabetes using data from the Dutch national survey of general practice. Hypertension and diabetes were chosen because in the Netherlands these patients are usually included in a surveillance scheme, thus allowing the usual number of surveillance visits per year to be subtracted from the total number of consultations in order to obtain consultation rates that can be compared with those of the reference groups.

People with any of the five chronic diseases under study were excluded, as well as those with any other non-vertebral osteoarthritis, stroke, peripheral vascular disease, rheumatoid arthritis, any malignant neoplasm, and dementia. The reference groups were standardized for age and sex relative to the relevant single disease subgroups. The reference group for the hypertension subgroup consisted of 14 623 people, the diabetes reference group of 15 847 people.

Data collection

During the 21-month study period all consultations with the 962 patients in the five disease cohorts were registered by the general practitioners on special research forms. For each consultation the general practitioner recorded one or more diagnosis at the highest diagnostic level appropriate.[8] These diagnoses were coded by trained clerks according to the *International classification of primary care (ICPC)*.[9] In cases of more than one consultation for the same diagnosis, the consultations were clustered by the clerks into episodes of disease ('a problem of illness in a patient, over the entire period of time from its onset to its resolution'[7]). The diagnosis for the episode was characterized by the diagnosis of the last registered consultation during the episode, as is usual in general practice morbidity studies.[10,11] Whether the episode was 'new' (never presented before) or 'old' (already existing at the start of the study period) was indicated on the research form at the first consultation for each episode. For this study only episodes of disease that had started during the study period were included.

Data on the consultations and intercurrent morbidity of the reference groups were obtained from the Dutch national survey of general practice. The data for the reference groups were collected during the three months preceding the data collection for the study cohorts (October–December 1987). Determination of the number of consultations and construction of episodes of disease were carried out as described for the study cohorts.

Table 1. Background characteristics of the five patient cohorts.

Chronic disease	Number of patients[a]	% of males	Mean age (range) in years	% with single disease
Hypertension	549	*34.8*	60 (22–92)	*86.3*
Chronic ischaemic heart disease	183	*61.7*	67 (34–87)	*73.8*
Diabetes mellitus	119	*42.0*	65 (29–88)	*68.9*
Chronic respiratory disease	252	*59.9*	45 (3–86)	*88.9*
Osteoarthritis	80	*32.5*	69 (39–87)	*68.8*

[a]Sum is greater than total of 962 owing to comorbidity among some patients.

Measures

The numbers of consultations and of episodes of disease were rescaled to rates per annum. The consultation rate was taken as the total number of consultations (face-to-face contact with the general practitioner at the practice or at home) per year, irrespective of the presented morbidity. Consultation rates are presented as means.

Intercurrent morbidity reflects episodes of new diseases presented to the general practitioner. The total incidence rate of intercurrent morbidity was taken as the total number of episodes per 1000 patients per year. Intercurrent morbidity was also studied at the level of *ICPC* chapter headings and at the level of dia-gnoses.

Reliability and validity of study cohort data

In one practice, with two general practitioners, the registration was interrupted for three months owing to reorganization of the practice. Correction for this interruption was made in the calculation of the consultation rates and the incidence rates for each cohort.

The accuracy of the number of consultations recorded by general practitioners was checked for a 2% sample of the patients by comparison with the practice patient records. Of all the consultations covered by the charts 70% appeared to be present in the study database. The consultations which were not reported were mainly for repeat prescriptions and those during evenings and weekends. No correction was made for this underreporting.

In order to determine the agreement between the 15 general practitioners in the diagnostic labelling of diseases, each general practitioner was asked to make diagnoses for each of 30 written case histories. The mean inter-observer agreement was 90%.

Analysis

Univariate description analysis was carried out to calculate consultation rates and total incidence rates for the different subgroups and the reference groups. Confidence intervals of the means are presented at the 95% level. Data analysis was performed with *SPSS-X/SPSS-PC*.

Results

Consultation rates

The consultation rates were higher in all the comorbidity subgroups than in the corresponding single disease subgroups (Table 2). The largest differences were found for patients with chronic respiratory disease and those with osteoarthritis, where the consultation rates in the comorbidity subgroups were 51% and 52% higher, respectively, than in the single disease subgroups. However, the comorbidity subgroups in these cohorts were both small (27 and 25 patients, respectively). The confidence intervals were large for all estimated means for both the single disease and

Table 2. Mean consultation rates in the five disease cohorts and two reference groups.

	Mean consultation rate (95% CI)		
	Single disease group	Reference group	Comorbidity group
Hypertension	4.7 (4.4 to 5.0)	3.0 (2.9 to 3.1)	5.4 (4.3 to 6.6)
Chronic ischaemic heart disease	5.5 (4.7 to 6.2)	–	6.0 (4.8 to 7.3)
Diabetes mellitus	5.7 (4.9 to 6.4)	2.8 (2.7 to 2.9)	6.5 (4.7 to 8.4)
Chronic respiratory disease	4.9 (4.3 to 5.5)	–	7.4 (5.0 to 9.8)
Osteoarthritis	4.2 (3.2 to 5.2)	–	6.4 (4.9 to 8.0)

CI = confidence interval.

comorbidity subgroups, indicating high individual variation in consultation rates.

Patients in the reference groups without a chronic disease had lower mean consultation rates than the respective single disease subgroups (Table 2).

Intercurrent morbidity

In all five cohorts patients with comorbidity had higher total incidence rates of intercurrent morbidity than the corresponding patients with a single disease (Table 3). This difference varied from 8% for diabetic patients to 74% for patients with osteoarthritis. The total incidence rate for the hypertension reference group was 42% higher than for the corresponding single disease group. However, the difference in rates between the single-disease diabetic patients and their reference group was only 4%.

Analysis of the incidence rates of intercurrent morbidity at the level of *ICPC* chapter headings among the five disease cohorts showed the highest rates in the chapters K (circulatory), L (musculoskeletal), R (respiratory) and U (urology). No important differences in this pattern were found between the subgroups within each of the five cohorts. In the two reference groups the highest rates were found for the chapters D (digestive), L (musculoskeletal), R (respiratory) and S (skin).

At the level of diagnoses the five diseases with the highest incidence rates were determined for each single disease subgroup. Only eight acute diseases were found in the top five for all five cohorts: myalgia, upper respiratory tract infection, acute bronchitis, urinary tract infection, influenza, ear wax, pneumonia and sinusitis. Comparison between the single disease and comorbidity subgroups showed a trend of higher incidence rates for these acute diseases in the comorbidity subgroups of patients with hypertension, diabetes and chronic ischaemic heart disease. The reference groups had lower incidence rates of these acute diseases than the corresponding single disease subgroups.

Table 3. Total incidence rates of intercurrent morbidity in the five disease cohorts and two reference groups.

	Total no. of episodes per 1000 patients per year		
	Single disease group	Reference group	Comorbidity group
Hypertension	1247	1776	1645
Chronic ischaemic heart disease	1794	–	2021
Diabetes mellitus	1618	1683	1755
Chronic respiratory disease	1784	–	2489
Osteoarthritis	1504	–	2619

Discussion

The mean annual consultation rate for all patients in general practice in the Netherlands is 3.2.[12] The mean consultation rate for patients with a single chronic disease in this study varied from 4.2 to 5.7; for patients with more than one of these diseases it varied from 5.4 to 7.4. Having more than one of the studied chronic diseases results in more consultations per year than having only one. The consultation rates do not increase linearly with the number of chronic diseases. The consultation rates in the two reference groups were lower than the corresponding single disease subgroups with hypertension and diabetes. Assuming that patients with hypertension and diabetes visit their general practitioner two to four times a year for regular control of their chronic disease, the number of consultations for other reasons appears to be lower than for patients without a chronic disease. This finding confirms previous results for patients with hypertension.[13] Patients with only hypertension or only diabetes probably present their other problems during their control visits whenever possible. The results presented here only indicate trends in consultation rates, since the confidence intervals of the means show considerable overlap between subgroups.

A morbidity study of the entire population of four general practices in the Netherlands found a total incidence rate of 1681 episodes of disease per 1000 patient years.[10] The figures for intercurrent morbidity for the single disease subgroups and the reference groups in this study are of a similar magnitude. The total incidence rates were lower in the single disease subgroups than in the comorbidity subgroups. In patients with diabetes this difference was small but it was substantial in the cohorts of patients with chronic respiratory disease and with osteoarthritis.

The results presented here confirm an earlier finding that patients with chronic diseases also present 'common diseases' to the general practitioner.[14] It could be argued that patients with chronic diseases would save their problems until the next control visit. This would lead to increased morbidity without increased consultation rates in patients with chronic diseases compared with those without chronic diseases, according to the phenomenon described by Berkson.[4] However, patients with only hypertension or only diabetes consulted their general practitioner more frequently than the reference groups without these diseases, and did not present more intercurrent morbidity. On the contrary, patients without a chronic disease had higher rates of total intercurrent morbidity.

The incompleteness of the study database, representing only 70% of the consultations listed on the patient charts, could not be resolved satisfactorily. The fact that the participating general practitioners had to complete research forms for consultations of only some of their patients made it difficult for them to do so in all cases. This incompleteness results in an underestimation of the consultation rates, but correction by simply adding 30% in all groups would ignore possible differences between subgroups with regards to the completeness of their data. Therefore, the observed figures have been presented without correction. This seems to be warranted for two reasons. First, many of the missing consultations were for repeat prescriptions. Secondly, intercurrent morbidity was analysed at the level of episodes of disease, which decreased the importance of a missed consultation.

In interpreting the results it should be remembered that there is an overlap in the comorbidity subgroups from the different

cohorts: patients with more than one chronic disease are present in more than one comorbidity subgroup. Comorbidity is defined here as a combination of two or more of five chronic diseases. The proportion of all chronic diseases in a general practice population represented by the five studied diseases depends on the definition of chronicity used. If the summated prevalence of the diseases used for exclusion in the construction of the reference groups is taken as 100%, the five chronic diseases studied represent about 70% of this. Including other chronic diseases, like gastrointestinal and mental health problems, to form additional study cohorts would change the results obtained.

Patients with chronic diseases have high consultation rates which are even higher when these patients have more than one chronic disease. However, patients with only hypertension or only diabetes do not have high total incidence rates of intercurrent morbidity, suggesting that Berkson's phenomenon[4] plays only a minor role in the results for these patients. Patients with chronic diseases also present common diseases to the general practitioner. This emphasizes the important role of general practitioners in the management of all diseases in a single patient.

References

1. Hasler J, Schofield T (eds). *Continuing care: the management of chronic disease.* 2nd edition. Oxford University Press, 1990.
2. Olshansky SJ, Ault AB. The fourth stage of the epidemiologic transition: the age of delayed degenerative diseases. *Milbank Q* 1986; **64:** 355-391.
3. Schellevis FG, van der Velden J, van de Lisdonk EH, *et al.* Comorbidity of chronic diseases in general practice. *J Clin Epidemiol* 1993; **46:** 469-473.
4. Berkson J. Limitations of the application of fourfold table analysis to hospital data. *Biometrics* 1946; **2:** 47-53.
5. Schellevis FG, van de Lisdonk E, van der Velden J, *et al.* Validity of diagnoses of chronic diseases in general practice. *J Clin Epidemiol* 1993; **46:** 461-468.
6. Foets M, van der Velden J, de Bakker D. *Dutch national survey of general practice. Survey design.* Utrecht, Netherlands: Netherlands Institute of Primary Health Care, 1992.
7. Classification Committee of WONCA in collaboration with the World Health Organization. *ICHPPC-2-defined (International classification of health problems in primary care).* 3rd edition. Oxford University Press, 1983.
8. Weed LL. Medical records, medical education and patient care. Cleveland, OH: Cleveland Press of Case Western University. 1969.
9. Lamberts H, Woods M (eds). *ICPC. International classification of primary care.* Oxford University Press, 1987.
10. van den Hoogen HJM, Huygen FJA, Schellekens JWG, *et al* (eds). *Morbidity figures from general practice. Data from four general practices 1978-1982.* Nijmegen, Netherlands: University Department of General Practice, 1985.
11. Lamberts H. *In het huis van de huisarts. Verslag van het transitieproject [Report of the transition project].* Lelystad, Netherlands: Meditekst, 1991.
12. Groenewegen PP, de Bakker DH, van der Velden J. *Nationale studie naar ziekten en verrichtingen in de huisartspraktijk. Basisrapport verrichtingen in de huisartspraktijk.* [*Dutch national survey of general practice. Report on interventions in general practice*]. Utrecht, Netherlands: Netherlands Institute of Primary Health Care (NIVEL), 1992.
13. van Weel C. Does labelling and treatment for hypertension increase illness behaviour? *Fam Pract* 1985; **2:** 147-150.
14. Beale N, Searle M, Woodman J. Use made by patients of chronic disease surveillance consultations in general practice. *Br J Gen Pract* 1992; **42:** 51-53.

Address for correspondence

Dr F G Schellevis, Department of General Practice and Nursing Home Medicine, Vrije Universiteit, vd Boechorststraat 7, 1081 BT Amsterdam, Netherlands.

Cholesterol management in Dutch general practice

A comparison with national guidelines

Trudy van der Weijden[1], Jack BF Hutten[2], Bart J Brandenburg[1], Richard PTM Grol[1], and Koos van der Velden[3]

[1]Centre for Quality of Care Research, Universities of Maastricht and Nijmegen, Dept. General Practice, [2]Interuniversity Centre for Social Science Theory and Methodology, Utrecht University, [3]Netherlands Institute for Primary Health Care, Utrecht, The Netherlands.

Weijden T vd, Hutten JBF, Brandenburg BJ, Grol RPTM, Velden K vd. Cholesterol management in Dutch general practice. A comparison with national guidelines. Scand J Prim Health Care 1994;12:281–8.

Objective **– To examine cholesterol diagnosis and treatment by Dutch general practitioners (GPs) in the period before publication of national guidelines, in order to develop implementation strategies based on discrepancies found between daily practice and the guidelines.**
Design **– Data of the 'Dutch National Survey of General Practice', in which GPs were involved in extensive consultation registration, were used. Patients were included for analysis if serum cholesterol, or the ICPC-code lipid metabolism disorder, or cholesterol-lowering treatment was registered.**
Setting **– General practice.**
Participants **– 161 GPs, 177 practice-nurses.**
Outcome measures **– Reasons for consultation, diagnoses, therapy, inter-doctor variation.**
Results **– The main discrepancies between daily practice and the guidelines concerned indications for cholesterol measurement, repeated measurements to diagnose hypercholesterolaemia, and attention for diet advice. A remarkable inter-doctor variation in diagnosis, and less so in treatment, was also found.**
Conclusion **– The inter-doctor variation justifies the publication of the standard guidelines. Implementation strategies should aim at indications for cholesterol testing, repeating measurements for diagnosis, and advice on diet.**

Key words: **cholesterol, general practice, guidelines, implementation.**

Trudy van der Weijden, MD, Department of General Practice, University of Maastricht, Postbus 616, 6200 MD Maastricht, The Netherlands.

The publishing of the results of important cholesterol intervention studies in the 1980s (1,2) stimulated consensus meetings in various countries. In addition there are developments such as the introduction of new cholesterol-lowering medication and the portable capillary blood-testing device for cholesterol measurement. These developments (3) prompted the Dutch College of General Practitioners to set a well-balanced (4) standard for hypercholesterolaemia, which was published in November 1991 (5). Reservedness in screening and drug therapy characterizes this standard (Fig. 1).

Meanwhile the cholesterol topic is still contro-

Figure 1. The cholesterol guidelines of the Dutch College of GPs

Screening:
Selective case finding; men and women, 18–65 yrs, with one of the following risk factors; signs of familial hypercholesterolaemia (xanthoma, xanthelasmata/arcus senilis before the age of 40), CHD is patient history, CHD in sibling or parent while younger than 60, hypertension, diabetes mellitus, familial hyperlipidaemia in the family.
Diagnosis:
The mean of 3 total cholesterol tests ≥ than 6.5 mmol/l. Determination of HDL and triglycerides only in case cholesterol lowering drugs are being considered.
Management:
< 6.5 mmol/l: general advice about low-fat diet
> 10.0: consultation of an internist
6.5–10.0: diet therapy, with support, for 6 months; referral to dietist if unsuccesful.
Drugs may be considered if, after 6 or 12 months of diet therapy, serum cholesterol is:
6.5–7.9 mmol/l and ≥ 2 risk factors (as defined under screening)
8.0–10.0 mmol/l and ≥ 1 risk factor(s)
Target level: 6.5 mmol/l

versial (6-8), and there is no consensus in the literature as to which screening strategy is to be preferred in general practice (9).

Not much is known about the usual cholesterol care by Dutch GPs. A description of the usual cholesterol care, including the inter-doctor variation, in the period prior to publication of the guidelines, serves various purposes; it may show problems with the feasibility of the guidelines. It might also increase insight into deficiencies in the provided care and points of attention for implementation of the standard (10).

In this study we present findings on the following questions:

1. Who were the patients, in the period prior to publication of the guidelines, whose serum cholesterol was tested by the GP, and how was hypercholesterolaemia diagnosed? How was the inter-doctor variation?
2. How were patients with hypercholesterolaemia treated by their GP in the period prior to publication of the guidelines? How was the inter-doctor variation?

Material and methods

Data of the 'Dutch National Survey of General Practice' (11) were used to answer these questions. In this survey, 161 GPs and 177 practice-nurses, working in 103 Dutch practices serving 335000 patients, registered all doctor-patient or nurse-patient contacts during a period of three months. The survey, lasting from April 1987 to March 1988, consisted of four consecutive registration periods of three months each to account for seasonal influences. Selection of participating GPs was based on a stratified (according to region, urbanization, and distance to a general hospital) random sample.

Data recorded include patient characteristics, characteristics of the consultation, reason for consultation, diagnosis, and interventions (diagnostic tests, non-drug and drug treatment, referral). Different health problems presented in one consultation were registered separately. Each health problem contained a maximum of three reasons for consultation, one diagnosis, and two differential diagnoses. Reasons for consultation and diagnoses were coded in ICPC-codes.

The GPs could mark on their registration form that they requested a lipid spectrum, including total cholesterol and/or triglycerides and/or lipoproteins and/or free fatty acids. Checking the laboratory-forms showed that all lipid spectrum applications included at least total cholesterol.

A distinction was made between patients who consulted their GP for the first time in relation to reasons that resulted in a request for a lipid spectrum ('new' patients), and those who had already consulted their GP before the registration-period, with problems related to cholesterol diagnosis or intervention ('known' patients).

To answer the first question all contacts in which a lipid spectrum was determined were se-

Figure 2. The structure of the 2 cholesterol files.

National Survey:	161 GPs with 335,000 enlisted patients	
	↓	
Registration period:	166,000 patients contacted their GP during a period of 3 months	
	↓	↓
Files:	**File 1:**	**File 2:**
	1380 patients (with 1462 contacts) whose lipid spectrum was determined	877 patients (with 1098 contacts) with hypercholesterolaemia
	↓	↓
	528 'new' + 852 'known'	210 'new' + 667 'known'

lected (file 1). File 1 consists of 1380 patients (1462 contacts).

Regarding question 2 a second, separate selection was made consisting of:

- all patients of file 1 with hypercholesterolaemia*;
- all patients in whom the ICPC-code lipid metabolism disorder (T93) was registered as reason for consultation or diagnosis;
- all patients for whom a cholesterol-lowering drug was prescribed.

This file 2 includes 877 patients (1098 contacts) with hypercholesterolaemia.

The structure of the files is summarized in Fig. 2.

The amount of inter-doctor variation was studied by means of the coefficient of variation (CV) (12). The CV (dividing the standard deviation by the mean) is a measure for relative variability. To explore the determinants of this variation, multiple linear regression analysis was used. GPs in file 2 who had less than three patient contacts on the subject (n=59) were excluded from this

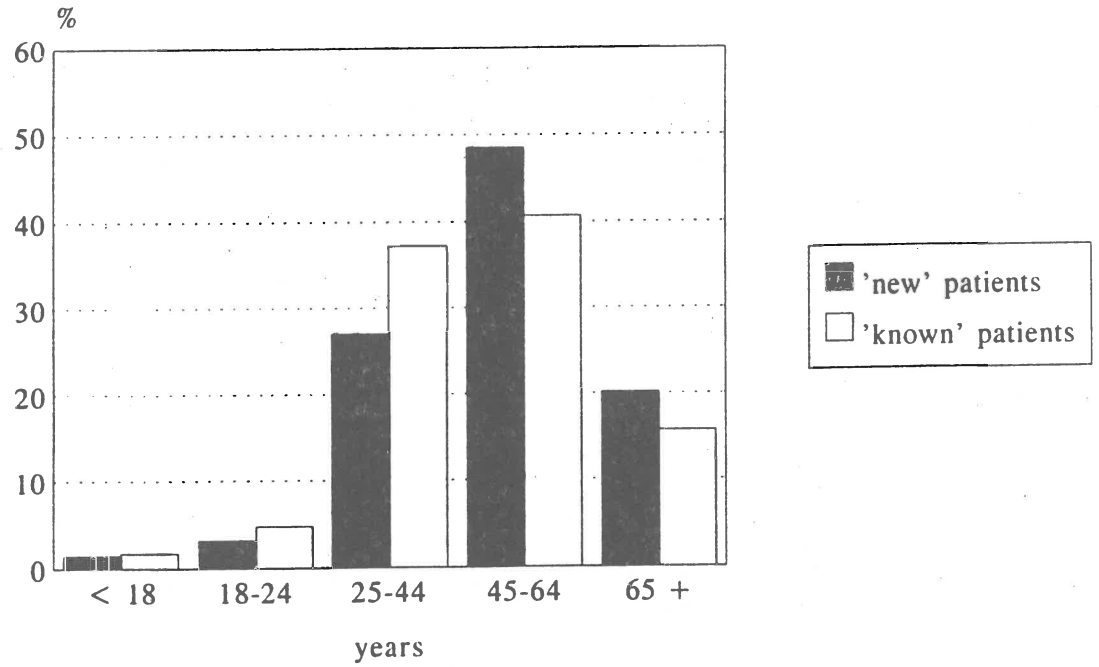


Figure 3. Age-distribution of 'new' and 'known' patients. Percentages.

Table I. Top 10 reasons for consultation and top 10 diagnoses for patients in whom serum cholesterol was tested (percentages)

'new' patients (n=528)		'known' patients (n=852)	
A. Reason for consultation			
1. General weakness/tiredness	18.6	Blood pressure measurement	20.4
2. Symptoms thorax/rib	8.1	Blood test (metabolism)	8.3
3. Blood pressure measurement	6.6	General weakness/tiredness	6.6
4. Complete medical examination	6.3	Complete medical examination	5.8
5. Blood test (metabolism)	5.5	Partial medical examination (metabolism)	5.4
6. Headache (excl. sinus)	4.5	Lipid metabolism disorder	3.1
7. Symptoms leg/thigh	4.0	Results test/procedures	3.1
8. Vertigo/dizziness	3.6	Observation/health education/diet (metabolism)	2.0
9. palpitations/awareness of heartbeats	3.2	Generalized abdominal pain/cramps	1.8
10. Partial medical exam (metabolism)	3.2	Symptoms leg/thigh	0.9
B. Top 10 diagnoses			
1. Uncomplicated hypertension	6.3	Uncomplicated hypertension	25.0
2. No disease	6.1	Lipid metabolism disorder	15.6
3. Lipid metabolism disorder	5.5	Complete medical examination	4.9
4. Neurasthenia	4.9	No disease	4.7
5. General weakness/tiredness	4.4	Diabetes mellitus	3.5
6. Complete medical examination	3.8	No diagnosis	2.1
7. Feeling anxious/nervous/inadequate	3.0	Feeling anxious/nervous/inadequate	2.0
8. Acute stress/situational disturbance	2.1	Neurasthenia	1.8
9. Diabetes mellitus	2.1	Depressive disorder	1.5
10. Angina pectoris	1.7	General weakness/tiredness	1.3

analysis. Correlation coefficients were calculated to explore substitution between different kinds of therapy.

Results

Diagnostics

A cholesterol test was ordered in 0.8% of all the patients (1380/166,000) who had visited their GP during the registration period of three months.

Demographic characteristics

The patients for whom a lipid spectrum was requested were on average 50 years old. Half the patients were between 45 and 64 years old. The sex distribution of the patients involved was almost symmetrical. According to the guidelines, the age-criteria for screening of cholesterol are between 18 and 65 years. Of the 'new' patients 21.5% did not meet these age criteria (Fig. 3). This percentage was considerably higher for older women (25.4%) than for older men (17.4%).

Reasons for consultation

The ten reasons for consultation most frequently presented by the patients are listed in Table IA.

For 'new' patients a lipid spectrum was most frequently – 33.9% of all reasons – indicated on non-specific symptoms such as general weakness/tiredness, headache, vertigo/dizziness, palpitations/awareness of heartbeats. Contacts concerning health check-ups (complete and partial medical examination) accounted for 9.5% of all the reasons for consultations.

Of the 'known' patients the non-specific indications (general weakness/tiredness and generalized abdominal pain/cramps) were of lesser

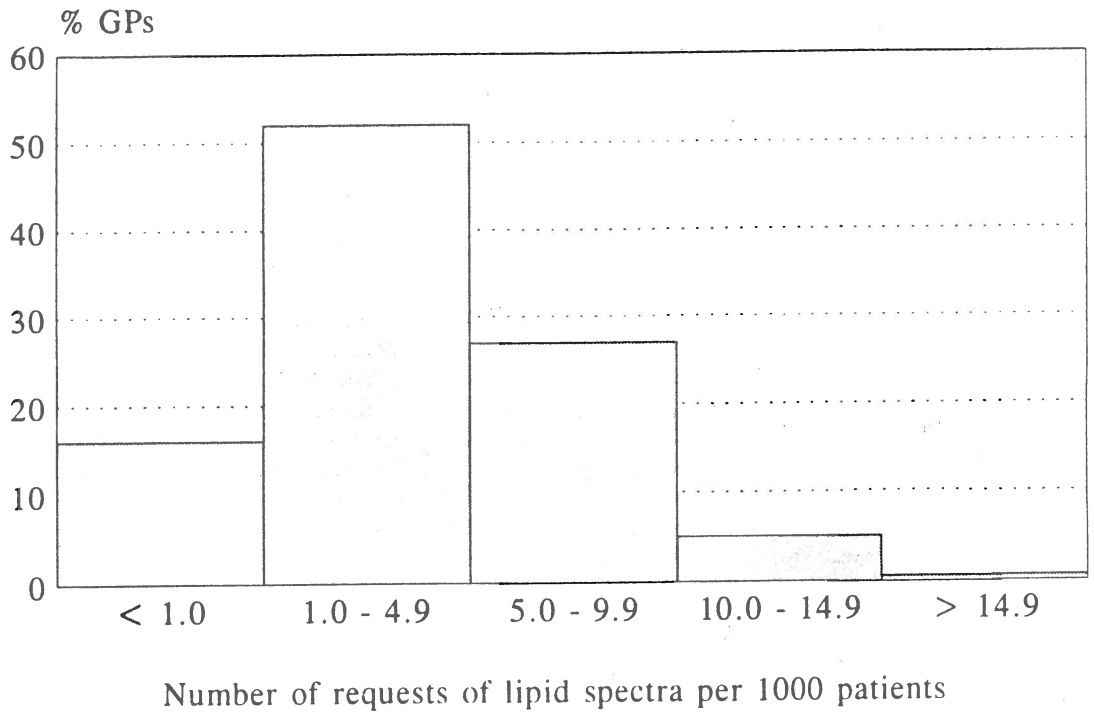


Figure 4. Inter-doctor variation in number of requests of lipid spectra per 1000 patients. Percentage GPs to number of requests per 1000 patients.

magnitude (8.4% of all the reasons for consultations).

Cholesterol measurements in this group of patients were more often indicated in control of hypertension (blood pressure measurement), lipid metabolism disorder, or health-check-ups.

Diagnoses

In 5% of the patients serum cholesterol was determined by more than one measurement during the registration period of three months.

Table IB lists the top 10 diagnoses of 'new' and 'known' patients with the reasons for consultation as mentioned in Table IA. In the first group hypertension, no disease, and lipid metabolism disorder were most common, but did not occur much more than the other diagnoses. The other diagnoses, except diabetes mellitus and angina pectoris, were mainly psychiatric in nature.

In the group of 'known' patients hypertension and lipid metabolism disorder were by far (40.6%) the most common diagnoses.

Inter-doctor variation

The number of requests for lipid spectra varied between the GPs (Fig. 4). Half of the GPs requested 1 to 5 measurements per 1000 patients over

Table II. Linear regression analysis on number of cholesterol requests explained by GP and practice characteristics. Standardized regression coefficients (beta).

	Beta	significance
GP's age	-0.004	0.96
GP's sex (male)	-0.017	0.89
Mean age of patients	0.023	0.79
Percentage women patients	0.022	0.83
Size of practice population	-0.240	0.01
Degree of urbanization	0.091	0.32
Type of practice		
Health centre	0.087	0.31
GP working solo	0.226	0.01

- adjusted R square: 0.07
- the GP with 27.4 requests per 1000 patients was left out in this analysis

the period of three months. The GP with the highest number of measurements requested 27.4 per 1000 enlisted patients. The average was 4.1 lipid spectra per 1000, with a standard deviation of 3.5 (CV = 0.85).

Doctor and practice characteristics were explored to explain this inter-doctor variation (Table II). GPs working solo requested cholesterol tests more often, while GPs with large patient lists requested less often.

Therapy

Non-drug therapy

The GPs advised and/or informed 42% of the patients; 39% of 'known' and 51% of 'new' patients.

Diet therapy was given to 15% of 'known' and 13% of 'new' patients. The GPs referred three percent of the patients to a dietician.

Drug therapy

Cholesterol-lowering drugs were prescribed to 25% of 'known' and 1% of 'new' patients. Table III gives an overview of prescribed drugs. Bile acid binders were prescribed mostly, followed by nicotinic acid, and fibrates. Over 10% of women under the age of 50 were prescribed lipid lowering drugs.

Inter-doctor variation

Inter-doctor variation in therapy was large for the prescription of fibrates (CV = 3.28). There were also variations in the prescription of nicotinic acid (CV = 2.86) and in referral to a dietician (CV = 2.31). Inter-doctor variation in therapy could not be explained by any of the doctor and practice characteristics mentioned in Table II.

No correlation was found between the degree of prescribing diet and the degree of prescribing cholesterol-lowering drugs (r= -0.18, p>.05), or other combinations.

Discussion

Methods

In the US and Canada telephone surveys (13-16) and medical chart audits (17-19) were conducted to determine the management of hypercholesterolaemia by GPs. It is quite possible that these research methods yield primarily socially desirable information or expose only part of reality, respectively. Determination of usual care based on intensive consultation registration seems to overcome these shortcomings.

On the other hand, these data give very little insight into the anamnestic part of the consultation, as well as management of Familial Hypercholesterolaemia for which no specific ICPC-code exists. Cholesterol screening is supposed to be executed in an anticipatory way, by case finding. The trigger for cholesterol screening might be the visit to the GP rather than the patient's reason for that visit. Still, considering the fact that reasons for consultation relating to different health problems were registered separately, as well as the large size of the data set, and the unmistakable indications for screening such as medical examination, conclusions can be drawn from the data.

Unfortunately, the norm for hypercholesterolaemia was not standardized between the different laboratories. Analysis of file 2 does not leave us with hard data but gives a description of management by GPs of the concept of hypercholesterolaemia.

Table III. Drug therapy for men and women pre- and postmenopausal (age criterion 50 years) (percentages)

	♂n=470	♀≤50 n=122	♂>50 n=285	total n=877
No medication	61	68	44	57
Bile acid binders	4	1	7	5
Fibrates	13	9	16	14
Nicotinic acid	8	3	13	9
Thyroxine	–	1	–	0.1
Cardiovascular medication	13	7	17	13
Other medication	9	16	16	13

Diagnostics

Indications for serum cholesterol measurement in 'new' patients consisted mainly of non-specific symptoms, not valid according to the guidelines. Less than 15% of all the indications were possibly indicated according to the guidelines, i.e. 6.6% concerning blood pressure measurement (possibly hypertension) and 8.1% concerning symptoms of thorax/rib (possibly angina pectoris or more likely fear of cardiovascular disease?).

Serum cholesterol was measured more than once in only five percent of the patients during the three months' period. It seems unlikely that serum cholesterol was measured more than once in the other cases before or after the registration period.

The number of non-specific diagnoses is remarkable, especially in 'new' patients.

Therapy

We do not know what the contents and quality of the advice or information given to the patient were, nor how they were given. Recently a low rate of lifestyle advice in general practice (20) as well as a lack in communication skills regarding hypercholesterolaemia (21) have been reported.

The designers of the guidelines attach great importance to extensive patient information at the moment hypercholesterolaemia is diagnosed for the first time. At the time of the National Survey, patient information was given in only half of those cases. Diet advice was given to only 14% of patients. It is possible that some diet advice was not registered but was given in another consultation outside the registration period. Despite this potential bias it is still a low number considering the fact that diet therapy is, according to the guidelines, the foundation of cholesterol-lowering therapy, deserving attention in every cholesterol-related consultation.

Only three percent of the patients were referred to a dietician during the registration period. It is not known in how many cases diet advice had already been given by a dietician, so there might be underestimation.

Over 10% of pre-menopausal women with hypercholesterolaemia were treated with lipid-lowering drugs. Considering the guideline that Familial Hypercholesterolaemia is the only indication for drug treatment of pre-menopausal women, this is a fairly high percentage. In the years when the 'National Survey' was conducted, the HMG coenzyme-A reductase inhibitors were not on the market. It would be interesting to know how the prescription of cholesterol-lowering drugs has been changed, now that the HMG coenzyme-A reductase inhibitors are available.

No evidence could be found for a substitution-effect between advice or diet therapy and the prescription of drugs.

Conclusion

Managing hypercholesterolaemia was not a clear-cut task for Dutch GPs at the time of the National Survey. The large inter-doctor variation justifies publication of guidelines for cholesterol management in general practice.

Discrepancies between daily practice and the guidelines may point at potential problems with feasibility of the guidelines. Possible barriers to change, which should be taken into account in implementing the Standard, are situated mainly in the field of indications for screening, diagnosis of hypercholesterolaemia, informing the patient when hypercholesterolaemia is diagnosed, and diet therapy as the foundation of cholesterol-lowering treatment. These are quite similar to points of attention recently assessed in the US (22).

Physicians' attitudes and motivation, rather than availability of practice guidelines, seem to relate more to actual preventive performance (23). Further research on the implementation of cholesterol guidelines in general practice is recommended. Much improvement can be achieved in this area.

Acknowledgements

We are grateful to the members of the Advisory Committee, M. Hofstra, M. den Hollander, W. Meyboom, A. Groeneveld, and F. König, for their critical advice and support.

References

1. Lipid Research Clinics Program. The Lipid Research Clinics Coronary Primary Prevention Trial results. I. Reduction in incidence of coronary heart disease. JAMA 1984;251:351-64
2. Frick MH, Elo O, Haapa K, Heinonen OP, Heinsalmi P, Helo P et al. Helsinki Heart Study: Primary prevention trial with gemfibrozil in middle-aged men with dyslipidaemia. Safety of treatment, changes in risk factors, and incidence of coronary heart disease. N Eng J Med 1987;317:1237-45.
3. Erkelens DW. Herziening Consensus Cholesterol. [Updating of the Cholesterol Consensus] Ned Tijdschr Geneeskd. 1991;135:2337-40.
4. Rutten G, Laan J van der. Hypercholesterolaemia: setting a Dutch national standard. Br J Gen Pract 1992;42:411-4.
5. Binsbergen JJ van, Brouwer A, Drenth BB van, Haverkort AFM, Prins A, Weijden T van der. NHG-standaard cholesterol. [Standard cholesterol of the Dutch College of General Practitioners] Huisarts Wet 1991;34:551-7.
6. Smith GD, Pekkanen J. Should there be a moratorium on the use of cholesterol lowering drugs? BMJ 1992;304:431-4.
7. Dunnigan MG. The problem with cholesterol. No light at the end of this tunnel? BMJ 1993;306:1355-6.
8. Law MR, Thompson SG, Wald NJ. Assessing possible hazards of reducing serum cholesterol. BMJ 1994;308:373-9.
9. Weijden T van der. Wetenschappelijke achtergronden bij de NHG-standaard cholesterol. Een literatuuroverzicht. [Scientific background of the cholesterol guidelines. A literature review.] Huisarts Wet 1992;35:90-6.
10. Zwaard A, Zijlstra S, Grol R. Kwaliteits- en deskundigheidsbevordering rond NHG-standaarden. [Quality and expertise assurance on behalf of the standards of the Dutch College of General Practitioners] Huisarts Wet 1989;32:501-4.
11. Foets M, Velden J van der, Bakker DH de. Dutch National Survey of General Practice. Survey Design. Utrecht, NIVEL 1992.
12. Blalock HM. Social statistics. Mcgraw-Hill, Tokyo, 1979.
13. Langner NR, Hasselback PD, Dunkley GC, Corber SJ. Attitudes and practices of primary care physicians in the management of elevated serum cholesterol levels. Can Med Assoc J 1989;141:33-8.
14. Shea S, Gemson DH, Mossel P. Management of high blood cholesterol by primary care physicians: diffusion of the National Cholesterol Education Program Adult Treatment Panel Guidelines. J Gen Intern Med 1990;5:327-34.
15. Kofron PM, Råstam L, Pirie PL, Holder W, Luepker RV. Physician practice for cardiovascular disease risk-factor reduction in six upper midwestern communities. J Fam Pract 1991;32:49-55.
16. Rosser WW, Palmer WH. Dissemination of guidelines on cholesterol. Effect on patterns of practice of general practitioners and family physicians in Ontario. Can Fam Physician 1993;39:280-4.
17. Madlon-Kay DJ. Family physician recognition and treatment of severe hypercholesterolemia. J Fam Pract 1987;24:54-6.
18. McBride PE, Pacala JT, Dean J, Plane MB. Primary care residents and the management of hypercholesterolemia. Am J Prev Med 1990;6:71-6.
19. Hudson JW, Keefe CW, Hogan AJ. Cholesterol measurement and treatment in community practices. J Fam Pract 1990;31:139-44.
20. Silagy C, Muir J, Coulter A, Thorogood M, Yudkin P, Roe L. Lifestyle advice in general practice: rates recalled by patients. BMJ 1992;305:871-4.
21. Dahlgren LO, Diwan VK, Tomson G, Wahlström R. On the variations in conceptions among primary care physicians regarding hypercholesterolaemia: a phenomenographic analysis. Scand J Prim Health Care 1992;10:316-20.
22. McBride PE, Plane MB, Underbakke G. Hypercholesterolemia: the current educational needs of physicians. Am Heart J 1992;123:817-24.
23. Fix KN, Oberman A. Barriers to following National Cholesterol Education Program guidelines. An appraisal of poor physician compliance. Arch Intern Med 1992;152:2385-7.

Received May 1993
Accepted February 1994

Patients with suspected meningitis: a study in general practice

Rinco Koorevaar, Marc A Bruijnzeels, Johannes C van der Wouden, Emiel van der Does, Koos van der Velden, Lisette WA van Suijlekom-Smit

To assess the management of patients with suspected meningitis by general practitioners, we used data from the Dutch National Survey of Morbidity and Interventions in General Practice. In this study, involving 161 general practitioners with a practice population of 335,000 persons, all patient contacts in general practice and all hospital admissions were registered. Additional information was gathered by interviewing the GPs involved. We selected patients with a provisional diagnosis of meningitis by the GP and/or a hospital diagnosis of meningitis. Of the 17 patients with the provisional diagnosis of meningitis by the GP eight had a final diagnosis of meningitis (predictive value of the provisional diagnosis: 46%). In the majority of patients with another final diagnosis the GP reported meningeal irritability and lowered consciousness, but this was not confirmed in hospital. Of all ten patients with a final diagnosis of meningitis eight had the correct provisional diagnosis of meningitis by the GP (sensitivity of the provisional diagnosis meningitis: 80%). We conclude that it is often difficult to diagnose meningitis in general practice. Inevitably, patients will be referred with a provisional diagnosis of meningitis which cannot be confirmed in hospital.

Eur J Gen Practice, 1995; 1: 21-4.

Meningitis is one of the most severe infections, but does not occur frequently in general practice. The incidence of meningitis is highest among children younger than two years of age.[1] The case fatality rate of bacterial meningitis varies between 3 to 30% and long-term neurological sequelae develop in as many as one third of all survivors.[2] An early and correct diagnosis of meningitis is important for the early admittance of the patient to hospital and the immediate initiation of antibiotic treatment.[3,4] Diagnostic delay is likely to result in an unfavourable outcome.[5-7] The diagnosis can be difficult in infants and in the elderly, and in the early stage of the disease.[8] None of the symptoms found in patients with meningitis are pathognomonic for meningitis. Fear of missing the diagnosis exists in general practice. From a quality assurance point of view it will be helpful to see what happens in the diagnostic trajectory. We studied the presentation, management and outcome of patients presenting with meningitis or suspected meningitis in general practice by using data of a large, nation-wide morbidity study.

Patients and methods

Between April 1st, 1987, and March 31st, 1988, the Dutch national survey of morbidity and interventions in general practice was carried out by the Netherlands institute of primary health care (NIVEL).[9] For this survey 103 general practices (161 general practitioners), divided into four groups, recorded all contacts with patients during one of the four successive three months registration periods (contact registration). The practices were a stratified random sample across the whole of The Netherlands. The general practice study population consisted of 335,000 persons and was representative of the population of The Netherlands. Data were recorded for each consultation, including type of contact, reasons for consultation, working hypothesis and differential diagnosis, certainty of the working hypothesis, diagnostic procedures, treatment, referral and follow-up appointment. If a patient was admitted to hospital during the registration period (and was discharged not longer than one month after the registration period ended) the general practitioner filled in a registration form including the discharge diagnosis. If a patient went directly to a hospital without seeing a GP, a hospital registration form was also filled in.

The data on the registration forms were coded by a researcher using the International Classification of Primary Care (ICPC).[10] In order to trace all patients with meningitis we searched whether ICPC-code N71 was considered as a diagnosis by the general practitioner on the contact re-

Rinco Koorevaar, *MD.*
Marc A Bruijnzeels, *MA Data-analyst.*
Johannes C van der Wouden, *MA PhD, Research coordinator.*
Emiel van der Does, *MD PhD, Professor of General Practice.*
Department of General Practice, Room Fe319, Erasmus University Rotterdam, PO Box 1738, 3000 DR Rotterdam, The Netherlands.
Koos van der Velden, *MPH, Epidemiologist.*
Netherlands Institute of Primary Health Care (NIVEL), PO Box 1568 ,3500 BN Utrecht, The Netherlands.
Lisette WA van Suijlekom-Smit, *MD PhD, Paediatrician.*
Department of Paediatrics, Erasmus University and University Hospital/Sophia Childrens Hospital Rotterdam, PO Box 1738, 3000 RN Rotterdam, The Netherlands.
Correspondence to: Dr Johannes C van der Wouden.

Submitted: July 14th, 1994; accepted in revised form: November 1st, 1994.

gistration form or as a discharge diagnosis on the hospital registration form. We checked the original forms to exclude other possible diagnoses belonging to the same ICPC-category. Patients with septicaemia as part of meningococcal meningitis were also included. Episodes of meningitis were only studied if they started within the registration period. In order to complete the medical history we visited the general practitioners of the patients involved. We checked the patient's record held by the GP and the hospital discharge letters and interviewed the general practitioner about the patient contacts using a structured questionnaire. In the patients with a hospital discharge diagnosis we considered the discharge diagnosis to be the final diagnosis. If the patient was not referred we decided upon a final diagnosis after analysing all data, including the course of the disease.

Results

In the contact registration 17 patients were given a provisional diagnosis of meningitis by the general practitioner; in the hospital registration two more patients with a hospital diagnosis of meningitis were found, who had been sent to hospital by their GP with another provisional diagnosis. No patients with meningitis went directly to a hospital without consulting a GP. Table 1 shows the reported meningitis symptoms and diagnoses in general practice and in hospital of patients with a working hypothesis of meningitis by the GP but with another final diagnosis. Table 2 gives the clinical parameters of the patients with a final diagnosis of meningitis.
Information by the general practitioners, data from the patient charts and the hospital discharge letters were obtained for all patients except one (patient F).

Of the 17 patients with a working hypothesis of meningitis four were not referred to hospital. Two of them had a provisional diagnosis of mumps-meningitis and neck rigidity was found in both patients; one also had a headache and a transient delirium. These patients received a final diagnosis of mumps-meningitis. Another patient received the provisional diagnosis of viral meningitis but an X-ray of the maxillary sinus showed signs of sinusitis. Retrospective analysis of the disease course of another patient made the diagnosis of meningitis most unlikely. In this patient we made the final diagnosis: infection of unknown origin. None of these four unreferred patients developed sequelae. Thirteen patients (76%) were referred to hospital. Only one of these patients was not admitted. After examination in the outpatient department, the final diagnosis was hypertonia of the neck musculature. All 12 admitted patients suffered from infections: six had meningitis, six an infection of another type. Special attention was given to the reported meningitis symptoms of the referred patients with a final diagnosis other than meningitis. In the majority of this group of patients the GP reported meningeal irritability and a change in the level of consciousness but these symptoms were not confirmed in hospital.

Two patients were traced in the hospital registration with a discharge diagnosis of meningitis who were admitted with another provisional diagnosis by the GP. One patient with a provisional diagnosis of mumps was referred because he vomited and refused to drink; mumps-meningitis and -pancreatitis was diagnosed. The other patient with a provisional diagnosis of fever of unknown origin was referred because he was apathic and hypotonic. Analysis of the presentation to the GP of these two patients made it clear that symptoms pointing to meningitis were present.

A classical presentation of meningitis (two or more of the following symptoms present: meningeal irritability, disturbed consciousness, headache, petechiae) was found by the general practitioners in six of the ten meningitis patients. None of the patients in this study was treated with antibiotics before referral to hospital. Three of the patients (O, Q, S), all referred and with a final diagnosis of men-

Table 1. Patients with a final diagnosis other than meningitis (n=9): symptoms of meningitis in general practice and hospital.

Patient	Sex, age	Symptoms of meningitis found by the GP	Working hypothesis of the GP	Symptoms of meningitis found in hospital	Final diagnosis
A	F8	meningeal irritability	meningitis	not referred	infection of unknown origin
B	F29	head-/neckache	viral meningitis	not referred	sinusitis
C	M2	lethargic, neckache	meningitis	none	broncho-pneumonia
D	M43	meningeal irritability	meningitis	none	pneumonia
E	F10	lethargic, headache	meningitis	none	upper respiratory tract infection
F	M1	?	meningitis	meningeal irritability	upper respiratory tract infection and meningeal irritability
G	F5	meningeal irritability, lethargic, neckache	viral meningitis	none	viral infection
H	F0,5	meningeal irritability	meningitis	meningeal irritability	viral infection and meningeal irritability
I	M3	meningeal irritability	meningitis	none	hypertonic neck musculature

Table 2. Patients with a final diagnosis of meningitis (n=10): clinical parameters.

Patient	Sex, age	Symptoms of meningitis found by the GP	Working hypothesis of the GP	Symptoms of meningitis found in hospital	Final diagnosis
J	M3	meningeal irritability	mumps-meningitis	-	meningitis (mumps)
K	M6	meningeal irritability headache, delirium	mumps-meningitis	-	meningitis (mumps)
L	M5	headache	mumps	meningeal irritability	mumps-meningitis and -pancreatitis
M	F21	meningeal irritability headache	meningitis	meningeal irritability, lethargic	meningitis (viral)
N	F31	head-/neckache	viral meningitis	meningeal irritability	meningitis (viral)
O	M5	meningeal irritability headache, apathetic	meningitis	meningeal irritability	meningitis (viral)
P	M0,5	meningeal irritability lethargic	meningitis	meningeal irritability	meningitis (*Haem infl*)
Q	F1	apathetic, hypothonic	infection of unknown origin	meningeal irritability, lethargic	meningitis (*Haem infl*)
R	M17	meningeal irritability lethargic	meningitis	meningeal irritability, lethargic	meningitis (meningococceal)
S	M2	meningeal irritability lethargic, petechiae	meningitis	meningeal irritability, petechiae	meningitis (meningococceal)

ingitis, developed sequelae (chronic headaches, hemiparesis, disturbance of equilibrium). In all of these cases the patient was referred by the GP on the day the problem was presented to him. No death due to meningitis was found.

Discussion

A survey of the literature did not reveal any study about patients suspected of meningitis in general practice. We found four Danish studies that reported on the provisional diagnosis of the GP in patients with a hospital diagnosis of meningitis.[3,8,11,12] These studies used hospital data and are therefore not comparable to our study. They reported a correct diagnosis by the GP in 50-95% of the referred patients with meningitis. In our study, in two of the ten patients with confirmed meningitis, the GP had another working hypothesis, but these patients were referred because the GP considered further evaluation necessary. In the Danish studies the predictive value of the provisional diagnosis of meningitis by the GP in the referred patients was 9% and 40%, in our study 6/13 = 46%. These studies also indicate that it is difficult to diagnose meningitis correctly in general practice.

The data of the National Survey provides a unique opportunity to look at cases of meningitis in general practice. The large study population makes it possible to study a disease that is quite rare in general practice. The study design enabled us to trace the patients with meningitis or suspected meningitis both referred to hospital and not referred.

In three patients with mumps clinical signs of meningitis were found; one of them had also signs of pancreatitis. Because of the usually benign character of mumps-meningitis admittance to hospital is not mandatory.[13] The two patients in this study who showed the typical signs of a mumps-meningitis were not referred by their general practitioners and recovered completely without sequelae.

In 53%, 9 of the 17 patients with a provisional diagnosis meningitis by the general practitioner, the final diagnosis turned out to be different. In the majority of these patients the GP reported meningeal irritability and/or disturbed consciousness but this was not confirmed in hospital. Variation of symptoms in time could explain this difference. Another explanation could be the interdoctor variation in examining ill patients, especially for the presence of meningeal irritability. We can conclude that it is difficult to diagnose meningitis correctly in general practice. Hence, it is inevitable that patients will be referred with a provisional diagnosis meningitis which is not confirmed in hospital. ■

Acknowledgement

The authors thank Dr Arthur M Bohnen for his helpful comments.

References

1 Klein JO, Feigin RD, Mc Cracken Jr GH. Report of the task force on diagnosis and management of meningitis. *Pediatrics* 1986; 78 (Suppl): 959-82.

2 Saez-Llorens X, Ramilo O, Mustafa MM, Mertsola J, McCracken Jr GH. Molecular pathophysiology of bacterial meningitis: current concepts and therapeutic implications. *J Pediatrics* 1990; **116**: 671-84.

3 Mathiassen B, Thomsen H, Landsfeldt U. An evaluation of the accuracy of clinical diagnosis at admission in a population with epidemic meningococcal disease. *J Int Med* 1989; **226**: 113-6.

4 Sell SH. Long term sequelae of bacterial meningitis in children. *Pediatr Infect Dis* 1983; 2: 90-3.

5 Swartz MN, Dodge PR. Bacterial meningitis- a review of selected aspects. *N Eng J Med* 1965; **272**: 725-31, 779-87, 954-63, 1003-9.

(References continued on page 24)

(Patients with suspected meningitis: a study in general practice, continuation of page 23)

6 Gary N, Powers N, Todd JK. Clinical identification and comparative prognosis of high-risk patients with Haemophilus influenzae meningitis. *Am J Dis Child* 1989; **143**: 307-11.
7 Kaplan SI, Smith EO, Wills C, Feigin RD. Association between preadmission oral antibiotic therapy and cerebrospinal fluid findings and sequelae caused by Haemophilus influenzae type b meningitis. *Pediatr Infect Dis* 1986; **5**: 626-32.
8 Romer FK. Difficulties in the diagnosis of bacterial meningitis. *Lancet* 1977; **2**: 345-7.
9 Bensing JM, Foets M, Velden J van der, Zee J van der. De nationale studie van ziekten en verrichtingen in de huisartspraktijk. *Huisarts Wet* 1991; **34**: 51-61.
10 Lamberts H, Wood M, editors. ICPC. International classification of primary care. Oxford: Oxford University Press, 1987.
11 Nielsen B, Sorensen HT, Nielsen JO. Children admitted for observation for suspected meningitis. Problems in diagnosis in general practice. *Scand J Prim Health Care* 1988; **6**: 229-32.
12 Sorensen HT, Moller-Petersen J, Krarup HB, Pedersen H, Hansen H, Hamburger H. Diagnostic problems with meningococcal disease in general practice. *J Clin Epidemiol* 1992; **45**: 1289-93.
13 Ratzan KR. Viral meningitis. *Med Clin North Am* 1985; **69**: 399-413.

ANGIOTENSIN CONVERTING ENZYME INHIBITOR ASSOCIATED COUGH: A POPULATION-BASED CASE-CONTROL STUDY

L. E. VISSER,[1] B. H. CH. STRICKER,[2,4]* J. VAN DER VELDEN,[3] A. H. P. PAES[1] and A. BAKKER[1]

[1]Department of Pharmacoepidemiology & Pharmacotherapy, Utrecht Institute of Pharmaceutical Sciences, Utrecht, [2]Netherlands Centre for Monitoring of Adverse Reactions to Drugs, Rijswijk, [3]Netherlands Institute for Primary Health Care, Utrecht and [4]Department of Epidemiology and Biostatistics, Pharmacoepidemiology Unit, Erasmus University, Rotterdam, The Netherlands

(Received 16 June 1994; received for publication 6 December 1994)

Abstract—The objectives of this study were to determine the risk for coughing as an adverse reaction to angiotensin converting enzyme (ACE) inhibitors under everyday circumstances in a large population and to study whether this adverse effect is more common in women. A population-based case-control study was used. The study was set in the practices of 161 Dutch general practitioners (GPs), in which all consultations, morbidity, mortality, medical interventions and prescriptions were registered during 4 consecutive 3-month periods in 4 consecutive groups of 40–41 GPs. The subjects were 2436 patients with incident coughing and up to 3 controls per case were obtained (total group: 7348 controls), matched for GP and a contemporary consultation in the same 3 months. All cases and controls were 20 years or older and had no notification of respiratory infections, influenza, tuberculosis, asthma, chronic bronchitis, emphysema, congestive heart failure, sinusitis, laryngitis, haemoptysis or respiratory neoplasms during the 3-month period. The results showed that cases were 3.6 times as likely as controls to have been exposed to ACE inhibitors (95% CI: 2.4–5.5) but after adjustment for potential confounders the odds ratio was 2.5 (95% CI: 1.6–3.9). The crude odds ratio for males was 2.7 (95% CI: 1.4–5.1) and for females 4.2 (95% CI: 2.4–7.5). The adjusted odds ratio for males was 1.8 (95% CI: 0.9–3.5) and for females 2.7 (95% CI: 1.5–4.8). Cases were 2.7 (95% CI: 1.3–5.9) and 3.9 (95% CI: 2.3–6.5) times as likely as controls to have been exposed to captopril and enalapril, respectively but the adjusted odds ratio for enalapril, 2.3 (95% CI: 1.4–3.9) was higher than for captopril, 1.8 (95% CI: 0.8–3.8). In conclusion we can say that the risk for coughing is increased two- to threefold among ACE inhibitor users. Although the odds ratio for females is higher than the odds ratio for males, the difference is modest.

Coughing ACE inhibitors Case-control study Pharmacoepidemiology Enalapril Captopril

INTRODUCTION

Since anecdotal reports began to appear in the literature in the early 1980s, linking captopril and coughing, this association has been well documented. Coughing is probably related to the pharmacological effect of the angiotensin converting enzyme (ACE) inhibitors. All ACE inhibitors have since been documented to induce this side effect. The precise mechanism is still unclear but several have been proposed, such as bronchial hyper-reactivity, increased

*All correspondence should be addressed to: Dr B. H. Ch. Stricker, Pharmacoepidemiology Unit, Department of Epidemiology and Biostatistics, Erasmus University, Dr Molewaterplein 50, 3015 GE Rotterdam, The Netherlands.

cough reflex or alterations in prostaglandins, bradykinin or substance P. Many predisposing factors have also been reported to be related to the development of coughing, e.g. age, dose, duration of therapy, smoking status and use of β-blockers. Isolated case reports and a number of small or uncontrolled studies suggest coughing to be sex related, with women more likely to experience this side effect [1–18]. This finding has been challenged by others however [19, 20].

The reported frequency of coughing to ACE inhibitors varies widely in the literature, depending on the populations under study and the methodological approaches used. Two large-scale postmarketing studies with unselected patient populations yielded much lower incidence estimates than did the studies in hospital or tertiary clinics [21, 22]. In some studies, the frequency of coughing to ACE inhibitors was expressed as an incidence rate whereas in others cumulative incidences or prevalence estimates were used. Most studies were uncontrolled or inadequately controlled. In some follow-up studies, for instance, users of calcium-antagonists and thiazide diuretics were used as control groups [10, 12, 13, 23–25]. Whether this was a good choice is questionable, because the indications for use of these drugs may differ from the reason for use of ACE inhibitors. Moreover, both pharmacological groups have been suggested to be protective against ACE inhibitor-induced coughing [26, 27]. There appears to be only one study in which along with a frequency estimate a risk estimate was also calculated [25]. This study, however, comprised a relatively small patient population.

In the light of these observations, we conducted the present population-based case-control study in order to obtain a risk estimate of coughing to ACE inhibitors under everyday circumstances in a large population. Second, we studied whether this adverse effect is more common in women.

METHODS

Data source

From 1 April 1987 to 1 April 1988 a National Survey of General Practice was conducted in the Netherlands. A random sample of the total population of general practitioners (GPs) in the Netherlands was taken, stratified by degree of urbanization, by region, and by distance from the hospital. This resulted in the participation of 161 established GPs, covering a catchment population of approx. 335,000 persons (52% females). All consultations, morbidity data, prescriptions and other medical interventions were registered during 3 months in 4 consecutive groups of 40–41 GPs. The database consisted of 168,021 consulting patients with a total of 361,018 consultations [28]. The following information contained in the National Survey has been taken into consideration for the purpose of this study: GP and patient identification codes; patient age and sex; reasons for consultation (symptoms and diagnoses) as well as concurrent diseases and data on prescribed drugs. Morbidity data were coded according to the International Classification of Primary Care (ICPC) [29] and prescription data according to the Anatomical Therapeutic Chemical (ATC) classification scheme, as recommended by the WHO [30]. At the time of data collection the only available ACE inhibitors in The Netherlands were captopril and enalapril.

Design

In order to determine the risk for coughing as an adverse reaction to ACE inhibitors, a case-control study was performed. Potential cases comprised all patients aged 20 years and older, who consulted a GP for coughing for the first time during the 3-month period (incident cases). Hence, case patients were counted once even if they had had two incident episodes of coughing during the study period. Three matched controls for every case were randomly selected from the patients who had had at least one consultation during the 3-month period with the same GP as the case and who had not reported coughing during that period. Cases and controls were excluded if any of the following diseases were registered as additional diagnosis or as underlying disease: respiratory infections; influenza; tuberculosis; asthma; chronic bronchitis; emphysema; congestive heart failure; sinusitis; laryngitis; haemoptysis or respiratory neoplasms.

Data analysis

To ensure that the period of possible exposure to the drug of interest was identical in cases and controls and to control for seasonal variations in morbidity and prescriptions, the date of first consultation for coughing of the case was used as a reference point for the determination of

exposure in cases and controls. For each last prescription of an ACE inhibitor prior to the reference date, a treatment course was calculated from the number of prescribed tablets divided by the prescribed number of tablets per day. The length of the treatment course was multiplied with a factor of 1.1 with a minimum of 2 days and a maximum of 14 days in order to control for residual drug effects and moderate undercompliance. Subjects were defined as users if the reference date fell between the start and stop date of the estimated treatment course.

The odds ratio of exposure to ACE inhibitors among patients with coughing, compared to control patients, was calculated as a crude odds ratio with a 95% confidence interval (95% CI). Only matched analyses were performed in this study. For each potential co-factor a crude exposure odds ratio was calculated. The following variables were considered as potential confounders: age; gender; number of prescriptions; number of consultations; use of nonsteroidal anti-inflammatory drugs (NSAID), β-blockers [31], thiazides [27] or calcium-antagonists [26] and smoking status. Factors that caused a change in the crude odds ratio of use of ACE inhibitors of at least 5% were included in the multivariate model, applying conditional logistic regression [32] with the EGRET statistical package [33]. Subanalyses were performed for users of captopril and enalapril, by applying the same multivariate models. Furthermore, we investigated whether this side effect was dose-dependent, using the same conditional model with 3 dose levels. A separate analysis was performed with a case definition encompassing all acute respiratory infections.

RESULTS

There was a total of 1602 patients in the study base who had been prescribed an ACE inhibitor during the 3-month period. 698 had been prescribed captopril (44%) and 904 enalapril. In this group, coughing was reported by 155 patients giving a period prevalence of 9.7%. In the rest of the population (166,419 patients), coughing was registered in 10,907 patients (period prevalence: 6.6%), which would be compatible with a risk difference of 3.1%. Of the total number of patients with coughing ($n = 11{,}062$), 2436 were incident cases without a concurrent respiratory disease or congestive heart failure. The control group consisted of 7348 patients without concomitant respiratory illness or heart failure. Although it was not possible to find 3 controls for every case, all cases had at least 2 and the large majority of cases 3 matched controls.

A comparison of cases and controls is shown in Table 1. Several differences emerged which could explain differences in exposure frequency between cases and controls, irrespective of a real association between ACE inhibitors and coughing. Cases and controls were similar in age and there were slightly more women in the control group. Cases received significantly more prescriptions and had more consultations per person during the 3-month period than the control group. Patients without coughing were more likely to be taking β-blockers and NSAID. Numbers of prescriptions for calcium-antagonists and thiazides were not significantly different between cases and controls. Patients with coughing were more likely to be smokers. Factors were included in the model to study their effect upon the crude estimate.

Of the cases, 47 (1.9%) were identified as users of ACE inhibitors, against 39 (0.5%) of the controls. The results of the logistic regression analyses without (crude) and with (adjusted) controlling simultaneously for confounding variables are presented in Table 2.

Table 1. Comparison of cases and controls matched for GP

	Cases ($n = 2436$)	Controls ($n = 7348$)	Odds ratio (95% CI) or p-value
Mean age in years (SD)	46.7 (17.7)	46.9 (18.3)	$p = 0.714$
Females (%)	1431 (59%)	4557 (62%)	0.87 (0.79–0.96)
Mean No. prescriptions	2.14	1.46	$p < 0.0001$
Mean No. consultations	2.46	2.36	$p = 0.0016$
ACE inhibitors*	47 (1.9%)	39 (0.5%)	3.6 (2.4–5.5)
NSAID*	166 (6.8%)	737 (10%)	0.65 (0.5–0.7)
β-blockers*	89 (3.7%)	429 (5.8%)	0.6 (0.5–0.7)
Thiazides*	31 (1.3%)	91 (1.2%)	1.0 (0.7–1.5)
Ca-antagonists*	13 (0.5%)	64 (0.9%)	0.6 (0.3–1.1)
Smoking	62 (2.5%)	35 (0.5%)	5.5 (3.6–8.4)

*Number of cases and controls who were exposed to the drug on the index day.

Table 2. Crude and adjusted odds ratios (OR) for exposure to ACE inhibitors

	Overall		Females		Males	
	Crude OR	Adjusted OR	Crude OR	Adjusted OR	Crude OR	Adjusted OR
All ACE inhibitors	**3.6 (2.4–5.5)**	**2.5 (1.6–3.9)**	**4.2 (2.4–7.5)**	**2.7 (1.5–4.8)**	**2.7 (1.4–5.1)**	1.8 (0.9–3.5)
Captopril	**2.7 (1.3–5.9)**	1.8 (0.8–3.8)	2.4 (0.8–7.1)	1.5 (0.5–4.5)	**3.1 (1.02–9.5)**	2.3 (0.7–7.1)
Enalapril	**3.9 (2.3–6.5)**	**2.3 (1.4–3.9)**	**5.3 (2.6–10.8)**	**3.4 (1.7–7.0)**	**2.4 (1.1–5.4)**	1.7 (0.7–3.9)

Exposure odds ratios with a value of unity outside the 95% confidence interval are printed in **bold**.

Cases were 3.6 times as likely as controls to have been exposed to ACE inhibitors (95% CI: 2.4–5.5) but after adjustment the odds ratio was 2.5 (95% CI: 1.6–3.9). When results were analysed according to sex a crude odds ratio of 4.2 (95% CI: 2.4–7.5) was found for females which declined to 2.7 (95% CI: 1.5–4.8) after adjustment. For males, the crude odds ratio was 2.7 (95% CI: 1.4–5.1) and 1.8 (95% CI: 0.9–3.5) after adjustment. Patients with coughing were 2.7 (95% CI: 1.3–5.9) and 3.9 (95% CI: 2.3–6.5) times as likely as control patients to have been exposed to captopril and enalapril, respectively but after adjustment only the odds ratio for exposure to enalapril remained significant. Subanalyses for captopril and enalapril according to sex yielded significant crude and adjusted values for females on enalapril but not on captopril (Table 2). Although the crude exposure odds ratio was significant for captopril in males, adjusted odds ratios were nonsignificant in males for both captopril and enalapril. A dose–response relationship could not be observed for captopril as the crude odds ratio increased from 2.0 (95% CI: 0.3–12.2) in those treated with doses of 25 mg or less per day, via 2.2 (95% CI: 0.7–6.9) in those treated with 25–50 mg per day, to 2.1 (95% CI: 0.4–12.2) for those patients treated with doses of more than 50 mg per day. For enalapril, the crude odds ratio was 3.3 (95% CI: 1.5–7.0) in those treated with a dose of 10 mg or less per day, 5.3 (95% CI: 2.2–12.7) in patients on 10–20 mg per day, and 4.5 (0.8–27.2) in those using doses of more than 20 mg per day.

A separate analysis of the aforementioned data with a case definition including acute respiratory infections revealed similar results and is therefore not demonstrated here.

DISCUSSION

In this study, we assessed the risk for coughing in relation to treatment with ACE inhibitors in a large population under everyday circumstances. The results suggest that the risk for coughing is increased two- to threefold among patients exposed to ACE inhibitors. Although our findings could be compatible with the previously made suggestion that coughing may occur more frequently to enalapril [2, 18, 19, 34, 35], especially as the adjusted odds ratio for exposure to captopril was nonsignificant. On the other hand, there is no doubt that captopril may cause coughing and we think that if a difference between the odds ratios of exposure to enalapril and captopril exists, it is only modest. In our study, the exposure odds ratio for females was higher than that for males as the odds ratio of 2.7 and 3.4 in the adjusted analyses for all ACE inhibitors and for enalapril, respectively was not demonstrated in males. This is in line with earlier suggestions [1–18]. Why the exposure odds ratio in women is higher than in men is speculative. There are several possible explanations. First, women might visit their GP more readily and might thus be more easily recognized or diagnosed as coughing. As we controlled for the number of consultations and prescriptions, however, this explanation is not very likely. Second, women might be more sensitive to bradykinin, prostaglandins or substance P, for instance by a difference in receptor population or sensitivity, or develop higher levels to ACE inhibitors. Third, there could be other known or unknown gender-related co-factors that might explain this difference and which were not controlled for in this study. Although the difference seems to be fairly modest, the consequences for public health may be substantial as the exposure prevalence to ACE inhibitors increases.

Although potential biases must be considered in interpreting the results of any case-control study, we think that most of these were adequately dealt with. Selection bias was not very likely to occur for two reasons. First, all consultations were registered, all cases who met the entry criteria were included in the study and selection of the control subjects was performed at random and independent of drug use. Second, during the classification of symptoms and diseases, GPs were not aware of the

research hypothesis and no selection was made by the GP. Although it must be emphasized that the fact that not all patients will consult their GP because of coughing means that some selection inevitably occurs, it does not necessarily mean that this will bias the risk estimate. To control for this potential bias, the reference group consisted of patients who had consulted the GP in the same 3-month period. This reference group was also chosen to ensure that patients were alive and present in the catchment area and thus eligible for visits to the GPs and for filling a prescription, in order to prevent a spurious overestimation of the odds ratio. Recall bias could not have played a role as the case-control study made use of data which were already gathered before disease onset. Another type of information bias might occur, however, if GPs were more likely to diagnose coughing in users of ACE inhibitors than in other patients. Information bias could also be the result of patients who presented themselves to their GP with captopril or enalapril-associated coughing after reading the data sheet. As adverse reactions encountered during the National Survey were registered as such and because none of the patient visits for coughing were registered as adverse reactions, this seems unlikely as a substantial source of information bias. Because many GPs were not yet aware of the association between ACE inhibitors and coughing in 1987, cases of coughing may have been misclassified as acute respiratory disease, bronchitis, asthma or worsened cardiac failure. As patients with such diagnoses were also excluded from the control group, this does not necessarily jeopardize a valid estimation of the odds ratio. This assumption is endorsed by the fact that a re-analysis with a case definition consisting of coughing or acute respiratory infections did not substantially change our risk estimates. Misclassification might, however, lead to an underestimation of the incidence of respiratory adverse effects to ACE inhibitors, especially in view of the aforementioned fact that not all patients go to their GP because of coughing. Misclassification of disease in control patients might occur if they suffered from coughing, but did not present this complaint to the GP. Even if this occurred, however, it is unlikely that many controls were coughing during the short time window in which they were matched to cases. Moreover, coughing during the same period as the matched case without consulting the GP will introduce a bias towards the null hypothesis and thus lead to a conservative estimation of the odds ratio. Criteria for defining exposure should also be evaluated. Subjects were only considered as exposed if the reference date fell between the start and stop date of the last prescription which had been filled prior to the occurrence of coughing. Subjects who filled a prescription of an ACE inhibitor outside the time frame between this reference date and the beginning of the data collection, may have been misclassified as nonexposed. This nondifferential misclassification will lead to a conservative estimation of the odds ratio. Subjects may also have been misclassified as nonexposed, if they received a prescription of an ACE inhibitor from a medical specialist. As the percentages of patients under treatment of a cardiologist or a specialist in internal medicine was low and not significantly different between cases (2.7%) and controls (3.3%) ($p = 0.18$), such an effect is probably negligible. An important point is the control for confounding bias which lowered our estimates of the odds ratios in the multivariate analysis. The adjusted odds ratios remained statistically significantly away from unity, except in males. It should be noted that the crude odds ratio estimates were influenced most by the adjustments for gender, number of consultations, number of prescriptions and use of NSAID and β-blockers. In our study, patients who coughed had received significantly less prescriptions of β-blockers. This is not compatible with an earlier study suggesting that β-blockers predispose to ACE inhibitor-induced coughing [31]. The meaning of the low crude odds ratio for use of NSAID and β-blockers is unclear. Some NSAID, e.g. sulindac, have been advocated in the treatment of coughing to ACE inhibitors which would be in line with a low odds ratio. A lower exposure prevalence of β-blockers in cases has also been noted by others [3]. A likely explanation is that some medical practitioners attributed coughing to chronic obstructive lung disease or asthma, conditions in which β-blockers and NSAID are contra-indicated. This would explain why the exposure prevalence in cases was lower than in controls but also that—even though cases and controls with these concurrent diagnoses were excluded—some misclassification may have occurred. Although in our study smoking appeared to be related to the outcome of interest, there was no relationship with exposure to ACE inhibitors. Hence, smoking was not considered as a confounding variable in this study.

It should be emphasized, however, that the registered number of smokers in the database was far below the expected percentage in The Netherlands. The latter is probably explained by the fact that few GPs register the smoking status of their patients. Due to the short study period of the National Survey, a duration–effect relation could not be studied. For those persons who got their first prescription before the data collection started, it was not possible to determine for how long they had been on ACE inhibitor treatment. As coughing may appear after 1 day as well as after 17 months [36], however, it may be questioned whether a duration–effect relationship would have been demonstrated. In our study, a dose–response relationship between coughing and captopril was not demonstrated, which is compatible with the literature [2, 3, 9, 25]. Users of a daily dose below 10 mg of enalapril, however, had a slightly lower risk than users of higher daily doses. There was no evidence for a protective effect of calcium-antagonists or thiazides, as has been suggested elsewhere [26, 27].

It is very difficult to compare studies of the frequency of ACE inhibitor-induced coughing in the literature. This is exemplified by the fact that the reported frequency of coughing varies from 0.5 to 39% of recipients [37]. This has several reasons. First, the risk estimate is given as a straightforward frequency in some studies but as point prevalence, period prevalence, incidence rate or cumulative incidence in others. Also, the length of the study period varies. Second, while some studies rely on spontaneous reporting, in other studies patients were specifically asked for coughing. Third, several studies comprise case series or small uncontrolled studies with retrospective review of medical records, whereas the two largest studies so far were uncontrolled. Although there were also some controlled studies, these were small and not always representative of the use of ACE inhibitors under everyday circumstances. Fourth, the setting differs in studies from patients hospitalized in tertiary clinics to data gathered in an outpatient setting. Fifth, none of the studies was controlling for the number of consultations and prescriptions per patient. For these reasons, we performed this case-control study in a large population under everyday circumstances. The latter is an especially important feature of the actual situation in postmarketing surveillance.

In conclusion, we demonstrated that the risk for coughing is increased two- to threefold in users of ACE inhibitors and that this increase in mainly seen in women. In view of the fact that other studies demonstrated the same under different circumstances, this gender-related effect justifies further study concerning the mechanism.

REFERENCES

1. Inman WHW. Enalapril-induced cough. **Lancet** 1986; 2: 1218.
2. Coulter DM, Edwards IR. Cough associated with captopril and enalapril. **Br Med J** 1987; 294: 1521.
3. Yeo WW, Ramsey LE. Persistent dry cough with enalapril: incidence depends on method used. **J Hum Hypert** 1990; 4: 517–520.
4. Stoller JK, Elghazawi A, Mehta AC, Vidt DG. Captopril-induced cough. **Chest** 1988; 93: 659–661.
5. Webb D, Benjamin N, Collier J, Robinson B. Enalapril-induced cough. **Lancet** 1986; 2: 1094.
6. Kaufman J, Casanova JE, Riendl P, Schlueter DP. Bronchial hyperreactivity and cough due to angiotensin-converting enzyme inhibitors. **Chest** 1989; 95: 544–548.
7. Hume AL, Murphy JL, Lauerman SE. Angiotensin-converting enzyme inhibitor-induced cough. **Pharmacotherapy** 1989; 9: 88–90.
8. Morice AH, Brown MJ, Higenbottam T. Cough associated with angiotensin converting enzyme inhibition. **J Cardiovasc Pharmacol** 1989; 13 (Suppl 3): S59–S62.
9. McEwan JR, Choudry N, Street R, Fuller RW. Change in cough reflex after treatment with enalapril and ramipril. **Br Med J** 1989; 299: 13–16.
10. Yeo WW, Foster G, Ramsay LE. Prevalence of persistent cough during long-term enalapril treatment: controlled study versus nifedipine. **Q J Med** 1991; 80: 763–770.
11. Poole MD, Postma DS. Characterization of cough associated with angiotensin converting enzyme inhibitors. **Otolaryngol Head Neck Surg** 1991; 105: 714–716.
12. Yeo WW, Maclean D, Richardson PJ, Ramsey LE. Cough and enalapril: assessment by sponteanous reporting and visual analogue scale under double-blind conditions. **Br J Clin Pharmacol** 1991; 31: 356–359.
13. Os I, Bratland B, Dahlof B, Gisholt K, Syvertsen JO, Tretli S. Female sex as an important determinant of lisinopril-induced cough. **Lancet** 1992; 339: 372.
14. Strocchi E, Valtancoli G, Ambrosioni E. The incidence of cough during treatment with angiotensin converting enzyme inhibitors. **J Hypert** 1989; 7 (Suppl 6): S308–S309.
15. Moore N, Noblet C, Joannides R, Ollagnier M, Imbs JL, Lagier G. Cough and ACE inhibitors. **Lancet** 1993; 341: 61.
16. Lefebvre J, Poirier L, Lacourciere Y. Prospective trial on captropril-related cough. **Ann Pharmacother** 1992; 26: 161–164.
17. Carre A, Vasmant D, Elmalem J, Thiery P. Tolerability of ramipril in a multicenter study of mild-to-moderate hypertension in general practice. **J Cardiovasc Pharmacol** 1991; 18 (Suppl 2): S141–S143.
18. Strocchi E, Valtancoli G, Ricci C, Malini PL, Bassein L, Ambrosioni E. Postmarketing studies of subjective side effects; a case for strict methodological criteria and careful analysis of data. **Pharmacol Res** 1992; 25 (Suppl 1): 79–80.
19. Goldszer RC, Lilly LS, Solomon HS. Prevalence of

cough during angiotensin converting enzyme inhibitor therapy. **Am J Med** 1988; 85: 887.
20. Simon SR, Black HR, Moser M, Berland WE. Cough and ACE inhibitors. **Arch Intern Med** 1992; 152: 1698–1700.
21. Inman WHW, Rawson NSB, Wilton LV, Pearce GL, Speirs CJ. Postmarketing Surveillance of enalapril. I: results of prescription event-monitoring. **Br Med J** 1988; 297: 286–289.
22. Cooper WD, Sheldon D, Brown D, Kimber GR, Isitt VL, Currie WJ. Post-marketing surveillance of enalapril: experience in 11,710 hypertensive patients in general practice. **J R Coll Gen Pract** 1987; 37: 346–349.
23. Os I, Bratland B, Dahlof B, Gisholt K, Syvertsen JO, Tretli S. Lisinopril or nifedipine in essential hypertension? A Norwegian multicenter study on efficacy, tolerability and quality of life in 828 patients. **J Hypert** 1991; 9: 1097–1104.
24. Woo J, Woo KS, Kin T, Vallence-Owen J. A single-blind, randomized, cross-over study of angiotensin-converting Enzyme Inhibitors and triamterene and hydrochlorothiazide in the treatment of mild to moderate hypertension in the elderly. **Arch Intern Med** 1987; 147: 1386–1389.
25. Sebastian JL, McKinney WP, Kaufman J, Young MJ. Angiotensin converting enzyme inhibitors and cough. **Chest** 1991; 99: 36–39.
26. Fogari R, Zoppi A, Tettamanti F, Malamani GD, Tinelli C, Salvetti A. Effects of nifedipine and indomethacin on cough induced by angtiotensin converting enzyme inhibitors: a double-blind, randomized, cross-over study. **J Cardiovasc Pharmacol** 1992; 19: 670–673.
27. Kamei J, Kasuya Y. The effect of hydrochlorothiazide on the enhanced coughing associated with treatment with enalapril. **Eur J Pharmacol** 1992; 213: 137–139.
28. Foets M, Velden J van der, Bakker DH de. Survey design. Dutch National Survey of General Practice, Utrecht, NIVEL; 1992.
29. Lamberts H, Woods M. **International Classification of Primary Care.** Oxford: Oxford University Press; 1987.
30. Anonymous. **Anatomical Therapeutic Chemical (ATC) Index: Including Defined Daily Doses (DDD) for Plain Substances.** Oslo: World Health Organization; 1987.
31. Puolijoki HJ, Nieminen MM, Siitonen LO, Lahdensuo AH, Reinikainen PM. Is a simultaneous beta-blocker therapy a risk factor for enalapril-induced cough? **Respiration** 1989; 55: 127–128.
32. Kleinbaum DG, Kupper LL, Muller KE. **Applied Regression Analysis and Other Multivariate Methods.** Boston: PWS-Kent Publishing Company; 1988.
33. Egret. Statistics and Epidemiology Research Corporation, Seattle, WA.
34. Town GI, Hallwright GP, Maling TJB, O'Donnell TV. Angiotensin converting enzyme inhibitors and cough. **N Z Med J** 1987; 100: 161–163.
35. Croog SH, Levine S, Testa Ma, *et al.* The effects of antihypertensive therapy on the quality of life. **N Engl J Med** 1986; 314: 1657–1664.
36. Just PM. The positive association of cough with angiotensin-converting enzyme inhibitors. **Pharmacotherapy** 1989; 9: 82–87.
37. Israili ZH, Hall WD. Cough and angioneurotic edema associated with angiotensin-converting enzyme inhibitor therapy. A review of the literature and pathophysiology. **Ann Intern Med** 1992; 117: 234–242.
38. Hood S, Nicholls MG, Gilchrist NI. Cough with angiotensin converting enzyme inhibitors. **N Z Med J** 1987; 100: 6–7.
39. Gibson GR. Enalapril-induced cough. **Arch Intern Med** 1989; 149: 2701–2703.
40. Lunde H, Hedner Th, Samuelsson O, Lotvall J, Andren L, Lindholm L, Wiholm B. Dyspnoea, asthma and bronchospasm in relation to treatment with angiotensin converting enzyme inhibitors. **Br Med J** 1994; 308: 18–21.
41. Adams JW, Hard PB. Adverse effects of ACE inhibitors. **Ann Intern Med** 1993; 118: 314–315.
42. Reisin L, Schneeweiss A. Complete spontaneous remission of cough induced by ACE inhibitors during chronic therapy in hypertensive patients. **J Hum Hypert** 1992; 6: 333–335.
43. Fuller RW. Cough associated with angiotensin converting enzyme inhibitors. **J Hum Hypert** 1989; 3 (Suppl 1): 159–161.
44. Berkin KE, Ball SG. Cough and angiotensin converting enzyme inhibition. **Br Med J** 1988; 296: 1279.
45. Sesoko S, Kaneko Y. Cough associated with the use of captopril. **Arch Intern Med** 1985; 145: 1524.
46. Hallwright GP, Maling TB, Town GI. Enalapril and cough: case report. **N Z Med J** 1986; 99: 66–67.
47. Semple PF, Herd GW. Cough and wheeze caused by inhibitors of angiotensin converting enzyme. **N Engl J Med** 1986; 314: 61.

Do out-of-office laboratory tests affect diagnoses in general practice?

Joost OM Zaat[1], François G Schellevis[1], Jacques ThM van Eijk[1] and Koos van der Velden[3]

[1] Department of General Practice and Nursing Home Medicine, Institute for Research in Extramural Medicine, Vrije Universiteit, Amsterdam, [2] Netherlands Institute of Primary Health Care (NIVEL), Utrecht, The Netherlands.

Zaat JOM, Schellevis FG, van Eijk JThM, van der Velden K. Do out-of-office laboratory tests affect diagnoses in general practice? Scand J Prim Health Care 1995:13;46–51.

Objective – **To find out whether the GP diagnosis changed by out-of-office laboratory test results and whether his diagnosis became more certain.**
Design – **Descriptive study.**
Setting – **Dutch survey of morbidity and interventions in general practice: stratified random sample of 161 GPs with a total list of 335 000 patients.**
Subjects – **2 081 episodes of illness with at least one consultation with clinical chemistry, haematology, or serology tests and at least one follow-up consultation.**
Main outcome measurements – **Change in ICPC component or chapter between the consultation in which a laboratory test was ordered and the follow up contact; change in exact ICPC code in cases with important diseases (infectious diseases, haematological disorders, endocrine abnormalities, auto-immune processes and malignancies (n=330)); change in certainty of a diagnosis and change in somatic/psychosocial orientation.**
Results – **After laboratory tests done in the first consultation the ICPC component changed in 46% of the diagnoses. Of the diagnoses made in first consultations without laboratory tests 41% changed in the follow up consultation. The diagnosis after laboratory tests was the same as before in 51% of the consultations with important diseases. Certainty about a diagnosis increased significantly after laboratory tests ($p < 0.001$). An abnormal laboratory result did not affect the clinical certainty of the general practitioner or the percentage of altered diagnoses.**
Conclusion – **The usefulness of tests should be assessed not only in terms of the number of diagnoses changed or of the percentage of abnormal results, but also in terms of the changed certainty concerning a diagnosis.**

Key words: **laboratory test use, diagnostic process, general practice.**

JOM Zaat MD, PhD, EMGO Institute/ Department of General Practice and Nursing Home Medicine, Vrije Universiteit, Van de Boechorststraat 7, 1081 BT Amsterdam, The Netherlands.

Dutch GPs are traditionally restrictive in their use of laboratory tests. They order one or more tests in only 4% of all patient consultations (1), while practitioners in Belgium order tests in 7% of their consultations (2), and family physicians in the US in 28% of all outpatient consultations (3). Little is

known about the diagnostic impact of all these laboratory tests in everyday practice. In view of the role of laboratory tests in general practice we wanted to know if diagnoses change or gain in certainty after test use. We analysed, in a descriptive study, the data from the Dutch National Survey of Morbidity and Interventions in General Practice (4) in order to answer the following questions:

- how often did the diagnosis of the GP change after laboratory tests and how often could this have happened due to results of laboratory tests?
- did the certainty about his diagnosis change after receiving the results of laboratory tests?
- did the opinion of the GP concerning the somatic or psycho-social background of the problem change due to laboratory tests?
- was a change in diagnosis influenced when one or more tests gave abnormal test results?

Methods

Design

The design and data collection of the Dutch National Survey are described elsewhere (5). A random stratified sample of 161 general practitioners working in 103 practices with a total list of 335 000 patients was drawn from all 6 400 Dutch GPs. Data were recorded for all patient consultations (both surgery consultations and home visits) during three consecutive months between 1 April 1987 and 31 March 1988.

We selected for this study a group of illness episodes that met the following criteria:

- episodes with two or more consultations, haematological, clinical-chemical, or serological laboratory tests being performed in at least one consultation;
- at least one follow-up consultation after the one in which tests were performed;
- the first consultation had to be for presentation of a new or recurrent problem to the GP. Recurrent problems were defined as old problems that were not active at the start of the registration period.

The diagnostic 'impact' of a laboratory test early in an episode can be different from test use in later consultations. A distinction was therefore made between episodes with tests performed during the first consultation ('consultations with laboratory tests'), and episodes in which tests were performed during the second or later consultation ('consultations without laboratory tests').

Measuring instruments

All diagnoses by the GP were centrally coded according to an adapted version of the ICPC by specially trained research assistants (6). The ICPC chapters 'eye' (F), 'ear' (H), 'nervous system' , 'skin' (S), 'urology' (U), 'pregnancy' (W), disorders of the female sex organs' (X), and 'disorders of the male sex organs' (Y) were combined into the category 'others' because these chapters covered only a small number of episodes with laboratory tests. Chapter Z (social problems) has no separate diagnosis component and was therefore not included in the analysis.

Change of diagnosis was studied in two ways. In a global one we counted shifts between the component's *symptoms* (component 1) and *diagnoses* (component 7) within and between ICPC chapters as changes, disregarding shifts within these components. In order to investigate the changes in diagnoses in more detail (within the diagnosis component of an ICPC chapter) we selected the episodes with diagnoses in the second consultation for which clinical-chemical or haematological tests could have been important. The diagnoses selected in this second consultation were infectious diseases (with the exception of upper respiratory tract or lower urinary tract infections), haematological disorders, metabolic disturbances, autoimmune disease, and malignancy of the GI tract. Laboratory tests may be important to establish the diagnosis in such problems. We compared the diagnosis in this second consultation with the orginal diagnosis in the consultation before a laboratory test was done.

In each consultation the GP rated the certainty of his diagnosis on a five-point scale on the contact-registration form (very uncertain to very certain). Whether it was a somatic or psycho-social complaint was to be scored similarly (purely somatic, somatic with some psycho-social aspects, somatic with definitive psychosocial causes, psychosocial with some somatic aspects, pure pyscho-social).

Results of all laboratory tests from hospital or general practice laboratories were classified as normal or abnormal, based on the reference values of each particular laboratory.

Analysis

Non-parametric tests for paired observations (Wilcoxon) were used to test differences in certainty and opinions concerning somatic causes of complaints between the consultations. A significance level of 5% was chosen.

Results

The events in 1857 first consultations with laboratory tests and their following second consultations were analysed, while 1322 episodes without laboratory tests in the first consultation were analysed.

Change of diagnosis

The diagnosis changed to another ICPC component in 46% of the 1857 first consultations with tests. The percentage of changed diagnoses varied considerably between ICPC chapters (Table I). The component showing the greatest shift (73%) was component 1, 'symptoms' from chapter A (General). The yield from laboratory tests in this group seems high, but of the 152 patients whose diagnoses from the first consultation belonged to this component, 41 had no other diagnosis in the second consultation; 20 diagnoses moved to 'psychic disorders' (10 to the symptoms and 10 to the diagnoses component, e.g. depressive state); 11 diagnoses moved to the diagnoses in chapter B 'blood' (anaemia, etc.); 6 to endocrine diagnoses, 10 to the diagnoses in chapter R, 'respiration' (mainly upper airways infections); and the others were distributed over the remaining ICPC chapters.

The diagnosis changed in the follow-up consultation of 41% of the 1,322 first consultations without laboratory tests (Table I). The percentage of changed diagnoses also varied considerably.

The important diagnoses (infectious diseases, malignancy of the GI tract, autoimmune disease, metabolic or hematological disorders) were established in 330 second consultations (18%). Of these, 51% had the same ICPC code in both consultations. Laboratory tests may therefore have supplied important new information in half of these cases, i.e. in 9% of the total number of diagnoses.

Table I. Percentages of changed diagnoses in consultations with and without laboratory tests requested in the first consultation.

		WITH TESTS	WITHOUT TESTS
ICPC	Diagnosis in consultation	Changed diagnoses in 2nd 2nd consultation	Changed diagnoses in 2nd consultation
A-1	symptoms general	73 (152)*	68 (41)
A-7	diagnoses general	52 (160)	58 (79)
B-7	diagnoses haematology	36 (110)	47(17)
D-7	diagnoses gastroenterology	36 (165)	34 (119)
K-7	diagnoses cardiovascular	37 (108)	31 (55)
L-7	diagnoses musculo-skeletal	36 (94)	33 (82)
P-1	symptoms psychic	65 (91)	61 (59)
P-7	diagnoses psychic	48 (127)	36 (45)
R-7	diagnoses respiratory	33 (178)	29 (314)
T-7	diagnoses endocrinological	31 (134)	30 (20)
	Others	34 (538)	26 (491)
	Total	46 (1857)	41 (1322)

* In parentheses: the total number of diagnoses made during the first consultation for that particular category

Table II. Percentages of certain and uncertain diagnoses in first consultations with laboratory tests (absolute numbers)

Diagnosis 1st conconsultation→ / Diagnosis 2nd consultation ↓	Uncertain	Doubtful	Certain	Total 2nd consultation
Uncertain	25	8	4	(117)
Doubtful	24	26	9	(222)
Certain	51	66	88	(1023)
Total 1st consultation	100 (220)	100 (416)	100 (726)	(1326)

Changes in certainty and somatic/psychic score

Certainty concerning the diagnosis increased after laboratory tests (Table II). Only one quarter of the 220 uncertain diagnoses in the first consultation with laboratory tests remained uncertain in the second one; the diagnoses were now certain in somewhat more than half the cases ($p<0.001$). In contrast with this gain in certainty, the diagnoses in the consultations without tests became less certain during the second consultation than they were in the first one ($p<0.001$).

With respect to the somato-psychic scale the group with laboratory tests showed a tendency to consider diagnoses as more somatic in the second consultation but this was not significant ($p=0.21$) (Table III). GPs considered the diagnoses to be somatic more often in first consultations without laboratory tests than in consultations with laboratory tests (65% vs 55%).

Abnormal test results

There were 1,110 (16%) abnormal test results out of 7,176 tests. In 603 consultations (32% of the consultations) the GP received one or more abnormal result(s). After abnormal tests, 44% of diagnoses changed, versus 43% after normal results.

Certainty about the diagnoses increased after laboratory tests for consultations with abnormal and with normal results (both $p<0.001$). More diagnoses were generally considered to be somatic when test results were abnormal than when they were normal (72% *vs* 57%, $p=0.001$).

Discussion

There are few studies of the diagnostic impact of laboratory tests. In a hospital population, laboratory tests led to a change of diagnosis in 22% of consultations and to a 'better understanding' of the illness in 20% (7). Other studies also showed that only a small percentage of the problems with diagnoses was affected by supplementary diagnostic tests (8, 9). In general practice, for example, the erythrocyte sedimentation rate appeared to be mainly a tool for confirmation of a diagnosis, rather than a test for detection of unexpected pathology (10, 11).

The results of our study show that more than half the diagnoses did not change with respect to their ICPC component after receiving test results. The fact that the diagnosis did change in 46% of the consultations can probably not be considered to be (entirely) due to the tests performed. It made little difference whether a result was abnormal; moreover, similar percentages of diagnoses of consultations without tests also changed in the course of follow-up consultations. The 'net' profit from tests generally appears to be small: the diagnosis was clear in the first consultation in half the cases of 'important diagnoses'. Laboratory tests apparently served to confirm the diagnosis.

Some investigators consider that the desire for certainty is the main reason for the use of many and unnecessary tests (12). There are indications that lack of certainty as a characteristic of physicians is related to the use of laboratory tests (13-15). Comments such as "general practitioners feel more assured after laboratory tests" cannot be made on the basis of the data described here. We

Table III. Opinion concerning somatic or psychic background of the diagnoses after the 1st and 2nd consultation in first consultations with laboratory tests (absolute numbers in parentheses)

Diagnosis 1st consultation→ / Diagnosis 2nd consultation↓	Somatic	Somatic/ psychic	Psychic	Total 2nd consultation
Somatic	87	42	25	(1063)
Somatic/psychic	9	32	22	(228)
Psychic	4	26	53	(218)
Total in 1st consultation	100 (1030)	100 (293)	100 (186)	(1509)

can only state that the diagnosis in the second consultation was more certain after laboratory testing. Whether this certainty was correct or influenced patient care we cannot say.

However, our approach is not without problems. Laboratory tests are only a small part of the diagnostic process and it is difficult with our approach to decide whether a test result indeed contributed to a change of a diagnosis. Various other factors can also have contributed to this change. Our study does not include detailed information about either history-taking or physical examination. The question about the contribution of test results therefore cannot be answered conclusively.

It is possible that the GP consciously adapts the diagnosis in a consultation to the desire to carry out an investigation (16): a diagnosis of iron-deficiency anaemia may not be truly considered, but nevertheless the physician requests a haemoglobin estimation, and therefore diagnoses anaemia – with a high degree of uncertainty – prior to obtaining the result of the test. If the result is normal, the physician will return to the real, first diagnosis (e.g. fatigue or work problems). In this instance the recorded 'diagnosis' was changed, but in the physician's mind the diagnosis had remained the same. This bias could not be taken into account and could well be one of the explanations for changed diagnoses. However, our central coding process minimized at least the inter-doctor variation in the registration of their diagnoses.

In general, the benefit from laboratory tests seems to be that diagnoses gain in certainty, while frequently not changing. There is a need for further studies of the value of tests in evaluating the assurance of physician and patient.

Acknowledgement

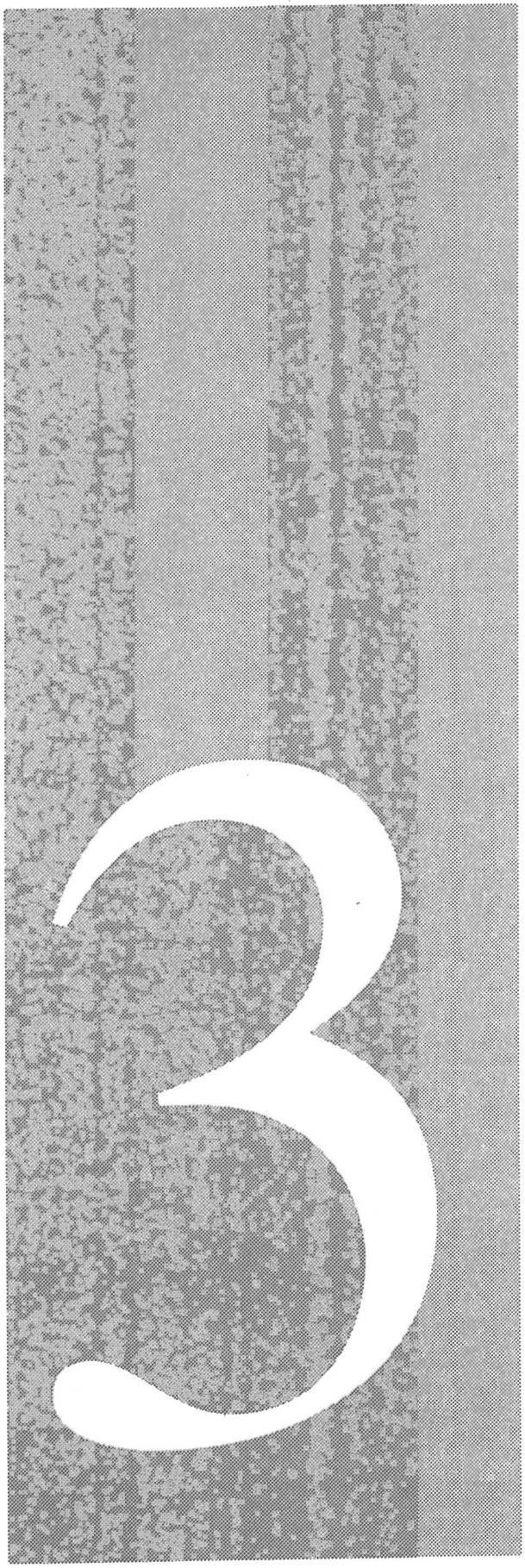
The Dutch National Survey on Morbidity and Interventions in General practice was supported by the Ministry of Welfare and Health and Cultural Affairs.

References

1. Kluijt I, Zaat JOM, van der Velden J, van Eijk JThM, Schellevis FG. The use of clinical-chemical, haematological, and serological laboratory tests in general practice. Huisarts Wet 1991; 34: 67-71. [In Dutch, summary in English]
2. De Maeseneer J. General practice; a survey. Gent: Rijksuniversiteit Gent, 1989 [in Dutch, summary in English]
3. Rosenblatt RA, Cherkim DC. The structure and content of family practice. Current status and future trends. J Fam Pract 1982; 15: 681-722.
4. Bensing JM, Foets M, van der Velden J, van der Zee J. The Dutch national survey of morbidity and Interventions in general practice. Huisarts Wet 1991; 34: 51-61. [in Dutch, summary in English]
5. Foets M, van der Velden J. Dutch national survey of morbidity and interventions in general practice. Summary of the study design. Utrecht: NIVEL, 1992.
6. Lamberts H, Woods M, eds. International Classification of primary care. Oxford: Oxford University Press, 1987.
7. Wertman BG, Sostrin SV, Pavlova Z, Lundberg GD. Why do physicians order laboratory tests? JAMA 1980; 243: 2080-2.

8. Stilwell JA, Young D, Cunnington A. Evaluation of laboratory tests in hospitals. Ann Clin Biochem 1980; 17: 281-6.
9. Hampton R, Harrison MJG, Mitchell JRA, Prichard JS, Seymour C. Relative contributions of history-taking, physical examination, and laboratory investigation to diagnosis and management of medical outpatients. BMJ 1975; ii: 486-9.
10. Dinant GJ, Knottnerus JA, van Wersch JWJ. Discriminating ability of the erythrocyte sedimentation rate: a prospective study in general practice. Br J Gen Pract 1991; 41: 365-70.
11. Grønlie M, Hjortdahl P. The erythrocyte sedimentation rate; its use and usefulness in primary health care. Scand J Prim Health Care 1991; 9: 97-102.
12. Kassirer JP. Our stubborn quest for diagnostic certainty: a cause of excessive testing. N Engl J Med 1989; 320: 1489-92.
13. Epstein AM, Begg CB, McNeil BJ. The effect of physicians' training and personality on test ordering for ambulatory patients. Am J Public Health 1984; 74: 1271-3.
14. Ornstein SM, Markert GP, Johnson AH, Rust PF, Afrin LB. The effect of physician personality on laboratory test ordering for ambulatory patient. Med Care 1988; 26: 536-43.
15. Zaat JOM, van Eijk JThM. General practitioners' uncertainty, risk preference and their use of laboratory tests. Med Care 1992; 30: 846-54.
16. Howie JGR. Diagnosis: the Achilles heel. J R Coll Gen Pract 1972; 22: 310-5.

Received August 1993
Accepted May 1994



3

SUPPLY OF CARE

WORKLOAD AND JOB SATISFACTION AMONG GENERAL PRACTITIONERS: A REVIEW OF THE LITERATURE

PETER P. GROENEWEGEN[1] and JACK B. F. HUTTEN[2]

[1]Netherlands Institute of Primary Health Care, P.O. Box 1568, 3500 BN Utrecht and
[2]Interuniversity Center for Sociological Theory and Methodology, State University of Utrecht, The Netherlands

Abstract—The workload of general practitioners (GPs) is an important issue in health care systems with capitation payment for GPs services. This article reviews the literature on determinants and consequences of workload and job satisfaction of GPs. Determinants of workload are located on the demand side (list size and composition of the patient population) and the supply side (organization of the practice and personal characteristics of the GP). The effects of workload and job satisfaction on workstyle and quality of work have been reviewed. The length of consultations or booking intervals seems to be an important restriction for workstyle and quality of work.

Key words—general practice, workload, job satisfaction, quality of care

1. INTRODUCTION

There is an ever increasing literature on the correlates of workload of general practitioners (GPs). Most of the literature deals with health care systems with capitation payment for general practitioner services, like in Great Britain and in the Netherlands. One gets the impression that workload is less often perceived as a problem in fee for service systems, where market forces on demand and supply work more strongly than in capitation systems (or salary systems). The policy problem of the justification of income in relation to the amount of workload of GPs seems to be less relevant in fee for service systems. Because of the different problem-context of workload studies in capitation systems like the United Kingdom and the Netherlands and in fee for service systems (in the present review mainly the United States), it may be important to know from what kind of health care system a particular study originates. In the text we have indicated other systems than the U.K. and the Netherlands.

The background to a number of workload studies is the question whether the level of capitation payment justly reflects the amount of work that is generated by different groups of patients [1–3]. Doctors with a comparable list size might have a different workload due to the composition of the practice population.

Another important background to workload studies is the supposed relation between workload and the quality of care of GPs. A high workload might lead to stress and time pressure on the GPs and this in turn might result in shorter consultations, a lower quality of communication between GP and patient, detecting fewer of the patient's problems and generally a lower quality of care [4–9].

Finally, some authors study workload of GPs because of issues in planning and management, such as the ideal size and composition of the primary care team [10], task delegation [11], repeat consultations [11–13] and ideal booking intervals [14].

In this article we review the existing literature on workload, job satisfaction and related issues of general practitioners. Our aim is to sketch an integrated model of the determinants and effects of variations in GPs workload. There are, however, some restrictions. We will not review separate descriptive items, such as the level of the number of consultations in different studies [15]; we are mainly interested in relations, e.g. between list size and the number of consultations per patient. The extensive literature on after-hour calls will not be discussed.

In Section 2 we discuss definitions of workload. The third section describes the relations that have been found between list size, characteristics of the patients on the list and workload. This results in a very simple model in which variations in general practitioner workload are determined only by factors on the demand (or patient) side. In Section 4 the literature on the relations between workload, the ways GPs manage their workload or cope with it and job satisfaction will be reviewed. This results in a somewhat more complicated model that takes into account 'objective' as well as 'subjective' aspects of workload, moderated by the way GPs manage or cope with their workload. Finally, in Section 5 we extend our review to the effects of the before mentioned variables on the workstyle and content of work of general practitioners.

2. DEFINITIONS OF WORKLOAD

Undertaking research into the workload of GPs presupposes some definitions and measurement procedures. We distinguish between 'objective' and 'subjective' workload. We will discuss objective workload first.

Objective workload is generally defined in terms of the amount of time that certain activities consume or the frequency that certain activities take place. The answer to the question which activities to include in definition and measurement of workload depends on what one sees as important aspects of the work of GPs, it depends also on the specific research question and background of the study and of course on the pragmatics of available data and ease of data collection.

In a general sense the work of GPs can be split up into patient related activities and other activities. In an editorial Ball [2] distinguishes between two important aspects of the work of GPs within the patient related activities: cure and care, and adds that quantitative measures of workload are easier for the curative aspects than for the caring aspects. Branthwaite and Ross [16] use more or less the same dimensions: clinical aspects and psychosocial aspects of work. Fleming's report on the Practice Activity Analysis [17] distinguishes two patient related elements of work: patient services (such as repeat prescriptions and telephone calls) and consultations (office visits and home visits), and three groups of other activities: continuing education, health service administration and practice administration.

A study by the DHSS [18] brings in another dimension: the activities while the GP is on duty. While on duty the GP might be involved in patient related activities, in other professional activities such as keeping up with the literature or in non-professional activities. The question is of course how on duty hours have to be counted in studies of workload. The actual measurement of workload is in terms of frequency of activities or the amount of time they take and workload is defined as higher with a greater frequency of activities or with a greater amount of time they take. The next scheme gives an overview of the measures of workload used in empirical studies (the numbers refer to the list of references at the end of this article).

Inclusion of length of consultations and booking intervals in this table might seem a bit strange. The length of consultations could be seen rather as a consequence of workload than as an independent indicator of it. Booking intervals could be seen as a result of decisions to manage workload. However, because in the literature they are used on the same level as the other measures of workload, they are included here. We will come back to this in the fourth section where the ways GPs cope with their workload will be discussed.

The more subjective aspects of workload range from rather direct questions whether physicians feel overworked [19] to measures of job stress [7, 16] and job satisfaction [6, 7, 16]. Makin a.o. [7] measure job stress by a 32 item scale developed from interviews with GPs. The items were factor analyzed in four factors: interruptions, emotional involvement, administrative involvement and home/work interface, and routine medical work as sources of job stress. Branthwaite and Ross [16] used 50 items related to sources of satisfaction and pressure, developed from interviews with GPs. The items on sources of pressure cluster in five factors: uncertainty and insecurity, isolation, poor relationships with other doctors, disillusionment in perceived role and awareness of changing demands. The items on sources of satisfaction could be reduced to three factors: psychosocial aspects of work, clinical aspects and management of the practice. Moreover, they used one item to indicate overall satisfaction ('on the whole, I am happy with the work I do').

Apart from job stress, Makin a.o. [7] have also measured job satisfaction with a 9-item Likert scale, previously used in a more extended form in research on other occupational groups, and with a single item measuring overall job satisfaction. Grol a.o. [6] operationalize job satisfaction in terms of how often GPs experience positive or negative emotions with different aspects of their work. They used four positive emotions (contentedness, challenge, self-esteem and well-being) and four negative emotions (lack of time, frustration, tenseness and doubts). They asked GPs 'how often do these feelings occur' when (a) helping patients with diagnosable physical affections; (b) helping patients with psychosomatic complaints and psychosocial problems; and (c) involved in extra activities in addition to consulting hour encounters and home visits? Factor analysis over all emotions and the three aspects of work revealed two factors: positive feelings and negative feelings.

Data collection methods vary from mailed questionnaires where GPs were asked to estimate their workload in terms of hours worked (see e.g. Ref. [4]), to studies where GPs had to record their workload for some time (see e.g. Refs [1, 17] and [19] for a combination of a questionnaire and recording in practice), and to studies where an external observer

Table 1. Measures of workload of general practitioners

Kind of activity	Operational variable	Reference
Patient related activities	No. of consultations	5, 10, 11
	No. of visits	10, 11
	No. of consultations per patient	1, 11
	No. of visits per patient	4, 11
	Ratio office consultations/home visits	4, 11
	Hours office consultations	4
	Hours home visits	4
	Hours office consultations + home visits	17
	Hours patient services	17
	Hours private practice	17
	Length of consultations	5, 8, 14
	Booking intervals	4, 5, 8
Other activities	Hours practice administration	4, 17
	Hours education etc.	4, 17
	Hours health service administration	17

was used to measure the GP's distribution of time (see Ref. [32]).

3. LIST SIZE, PATIENTS' CHARACTERISTICS AND WORKLOAD

The most important article on the relation between list size and workload is Butler and Calnan's 'The economy of time in general practice' [4]. Their point of departure is that 'general practice is demand led'; the workload of GPs is determined mainly by the demand of their patients and the influence of the GP on his own workload is limited. If this line of reasoning which is also at the background of capitation payment of GPs, is true, list size should be the principle determinant of workload:

LIST SIZE ⟶ WORKLOAD

There is indeed a statistically significant relation of list size with different indicators of workload, but correlations are not very strong. Positive correlations have been found with the number of consultations [5, 19] and with indicators of number of hours worked [4, 19] and in particular the time spend in office consultations. The relatively weak correlations with the number of hours worked indicate deviations from linearity and the influence of other factors on workload.

Negative correlations have been found with the number of consultations per patient and with the length of consultations [4, 5, 19–21]. While GPs with a larger list size spend more time in practice activities, the consultations are generally shorter and the number of consultations and home visits per patient are lower. The variation is higher for GPs with smaller list sizes, suggesting that in these cases other factors also influence the length of consultations and their number per patient; for GPs with larger list sizes the variation within categories of the list size is smaller [19].

The relation between list size and indicators of workload such as the number of hours spend in practice activities is not linear over the whole range of values of list size. The shape of the curve is an inverse *J*, indicating that the workload is not proportionally higher in the higher categories of list size. Interestingly enough the shape of this relation might depend on the system of remuneration. Mechanic [22] in his 1975 study of physicians in fee for service and prepaid primary care settings in the United States, found that prepaid physicians react to a greater number of patients by treating more patients in the same time, while fee for service physicians tend to increase their number of practice hours.

List size is only one indicator for the pressure of demand in general practice. Research into the determinants of utilization of primary medical care [23] suggest that there is a large variation between sociodemographic groups in the use of services and insofar as general practices differ in composition of the patient population this might result in different levels of workload with the same list size. This idea is incorporated in the differential capitation fees in the United Kingdom and has been discussed in The Netherlands where for publicly insured patients a flat capitation fee exists [3].

Apart from list size there are therefore other influences on workload:

Marsh and McNay [10] analyzed the influence of characteristics of the practice population on the workload of GPs. Women have more consultations and home visits, so do young children and the elderly (in particular home visits). Turn-over of the practice population might have a small effect on workload, because newly entered patients consume somewhat more of the GPs time. People with a higher social class background have somewhat more consultations, but lower social class people have somewhat more home visits. Unemployed men consult their doctor more often than employed men; the length of unemployment has no clear linear relationship with consulting a GP, but people who are for five years or more out of work have a higher chance of consulting the GP [24].

Effects on total workload depend of course on the representation of these groups in the practice and on differences between practices in these respects. Wilkin a.o. [19] found no effect on indicators of total workload from the percentage of children aged 0–4 years; a higher percentage of elderly (65 years and older) and of people in social class 1 or 2 was associated with longer hours of work in patient contact. Fleming [1] poses the question whether the existing differentiation in capitation fees in Great Britain justly reflects differences in workload as indicated by the number of consultations per patient. The sex of the patient is the most important determinant of workload. The workload associated with young children and elderly patients is underestimated in the current fee schedule. However, the effects seem to be marginal because practices do not differ greatly in the percentage of women on their list and the higher workload associated with young children is compensated for by a lower workload for children of 5–14 years and men of 24 years and older.

Not only the number of consultations and the time spend in patient contact are influenced by patient characteristics, but also the length of consultations. Morrell [25] found that the length of consultations is influenced by age and social class of the patient, the health problem presented and whether or not a new problem is presented. Westcott [14] found no difference between consultations with male and female patients, but shorter consultations for patients aged 15–29 years, longer consultations for patients aged 45–64 years and for patients with 'psychoneurotic' problems and a tendency towards shorter consultations with patients from the lower social classes. Andersson and Mattsson [26] in a small study of Swedish salaried GPs report longer consultations for women, elderly patients and patients with psychosocial complaints. Repeat consultations take more time than first consultations. They also report a large amount of variation between GPs, while practice composition did not differ greatly between GPs. "Personal style and different working manners of the

doctors may thus explain the main reason for this variation" [26, p.132], is their conclusion.

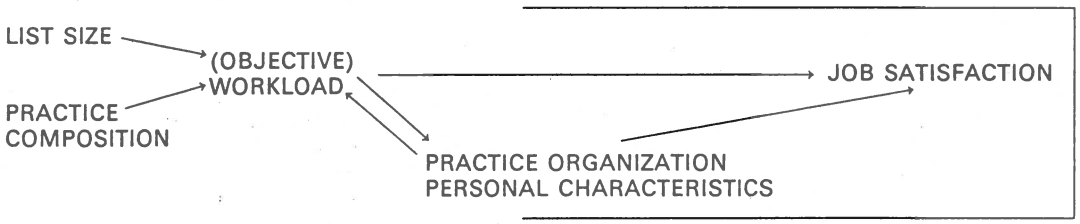
Verhaak [27] studied more than 2500 videotaped consultations and found that consultations of women and older patients last longer. This difference is related to the fact that women and elderly people more often present more than one complaint in the consultation, which corresponds to the findings of Roland and others [8] in the 5 min condition in their study of experimentally manipulated booking intervals. Psychosocial problems and consultations where patient education takes place take more time.

In summary we can state that the workload of GPs is influenced by list size as well as by the characteristics of the practice population served and the health problems presented. As is indicated by the relatively weak correlations demand related factors are not the only determinants of workload in general practice. As will be further explored in the next section workload is not just something that happens to general practitioners, but at least in part it is something that GPs themselves can influence.

4. WORKLOAD, POSSIBILITIES TO MANAGE WORKLOAD AND JOB SATISFACTION

The possibilities for managing workload can be divided theoretically into two groups. The first group relates to the organizational aspects of practice, such as single-handed or group practice, the use of a booking system, secretarial support, availability of a practice-nurse or other possibilities to delegate tasks and so on. The second group relates to personal characteristics of the general practitioner. Some of these characteristics may have a direct effect on the objective workload; others on the subjective experience of workload in terms of feelings of overload and job satisfaction. Examples of these personal characteristics are age and experience of the general practitioner, but also aspects of personality and attitudes towards work.

We can therefore extent our simple scheme as follows:



In this section we will review the literature on the relations between (objective) workload, practice organization and personal characteristics (the possibilities to manage workload) and job satisfaction.

Apart from the organizational conditions and personal characteristics workload also depends on the workstyle of general practitioners in the consultation room. General practitioners can manage their workload by using different strategies to end consultations, indirectly by their way of communicating with patients and more directly by proposing a repeat consultation, referring the patient or giving a prescription. We will conceptualize these 'management strategies' as a consequence of workload in terms of workstyle, having a feedback effect on initial workload. This will be discussed in Section 5.

Practice organization

As far as the organizational conditions of practice are concerned Wilkin [19] found that single handed practices have a higher workload in the sense that they have a higher average consultation ratio and a longer average consultation time (but there is no significant difference in other indicators of workload such as number of consultations per week, time in surgery and time in patient contact). Practices employing a nurse have a significantly higher workload on all indicators except for the average length of consultations. However, the question of cause and effect is in place here, because a Dutch study [28] of task delegation to a practice secretary (not even a registered nurse) has already indicated a substantial potential time gain. Having an appointment system (which almost all general practitioners have) coincides only with longer consultation time. It should be mentioned that these are the result of bivariate comparisons; the influence of list size and practice composition has not been controlled for.

Appointment systems and an organization of deputizing services are more often found in practices with a larger list size [29]. In group practices the appointment system through a practice receptionist will to a certain extent level out differences in initially demand related workload between partners [30]. Bridgstock [29] also found that appointment systems make it possible to spend more time per consultation, without, however, increasing the (estimated) total number of hours worked. Empirical work showing the extent of task delegation and its relation on the one hand with characteristics of practice organization and scale and workload on the other hand is scarce (apart from the literature on nurse practitioners in the United States) [31]. The research of Breslau and others [31] in 70 primary care teams in the United States indicates that the extent of task delegation (patient-care tasks: history taking, physical exam, treatment management and follow-up; technical routinized tasks: weighing and measuring) is greater for technical, routinized tasks and is greater in modern practice settings. Modern practice settings are defined as primary care teams that are organized as part of larger organizations as opposed to traditional teams that are independent single handed or group practices. In traditional teams the amount of delegation of patient care tasks is related to physician' attitudes and the ratio of physicians to paramedics, while the extent of delegation of technical routinized tasks is related only to the level of skills of the paramedics. Delegation in modern practice settings could not be

predicted from these variables. Dutch research [28] shows that the acceptance of delegation by patients is higher for the more technically oriented aspects of the task of GPs.

Although the evidence is very scattered, the conclusion of Branthwaite and Ross [16] could be retained as a hypothesis that practice organization in the form of larger partnerships makes it possible to structure the workload.

Personal characteristics

As far as the relations between personal characteristics and workload are concerned, Wilkin a.o [19] found that female GPs have less consultations, spend less time in practice and in direct patient contact. They also found that female GPs show less feelings of being overworked. But this can be a consequence of the fact that women have significant smaller personal list size than men [19]. Wilkin [19] found no relation between the age of GPs and their workload. The study of Armstrong o.a. [21] showed also that, although younger GPs have smaller list size than older GPs, age has no influence on the number of consultations and the number of home visits. Boots [32], however, found that more experienced (and thus on the average older) GPs spend less time on diagnostic aspects of their work and more on advising and counseling their patients.

Calnan [33] investigated the relation between the work orientation of GPs and their workload. More medically oriented GPs (as opposed to more socially oriented GPs) have a higher workload in terms of hours worked. And Branthwaite and Ross [16] found that a higher satisfaction with the clinical aspects of work coincides with more hours in consultations. The more socially oriented GPs in Calnan's study have larger consultation times. This finding is consistent with Verhaak's finding [27] that the personal communication style of GPs influences the length of consultations. An open style of communication (which is characteristic of general practice oriented GPs as opposed to clinically oriented GPs—a distinction very close to Calnan's) results in longer consultations, partly because patients have the opportunity to express more complaints in general and more psycho-social complaints in particular. Boots [32] distinguishes between GPs with a broader task perception and those with a smaller which is also close to Calnan's and Verhaak's distinctions. He found that the allocation of time of GPs with a broader task perception is less influenced by the pressure of the number of patients the GP sees per day.

Job satisfaction

The relations between workload and job satisfaction, and between practice organization and personal characteristics and job satisfaction have not been extensively studied. Calnan [33] found that the more socially oriented GPs have a higher job satisfaction. Although Makin [7] did not directly study the relations between workload and job satisfaction, their findings are indicative in this field. They found that the most important sources of stress in GPs are: interruptions of the daily routine, emotional involvement, administrative workload, and routine work. These sources of stress (except for routine work) are negatively related to overall satisfaction.

McCranie [34] conclude that, on the whole, the family physicians in their study in the United States were very satisfied with their work and careers. Reasons for relative dissatisfaction were: time pressure in practice, the treatment of emotional problems beyond their training, financial costs of their practice, the paper work and perceived interference of external regulations in the patient–physician relationship. Problems related to workload, as time pressure and the amount of paper work, where more mentioned by family phycians who had been in practice for a longer time and those who worked in solo-practices. Mawardi [35], also in a U.S. study, states that time, the lack of leisure time and time pressures of work, present the biggest source of stress in medical practice.

In a pilot study by Rankin a.o. [36] the emotional state of GPs was recorded hourly. The main cause of dissatisfaction was the uncertain state of the work: the factors of hassle and lack of time together accounted for more than the half of the reasons given for a drop in mood. Furthermore their results suggest that GPs get the most pleasure in the work for which they are trained and less pleasure in dealing with social problems which may take a long time and have unsatisfactory outcomes.

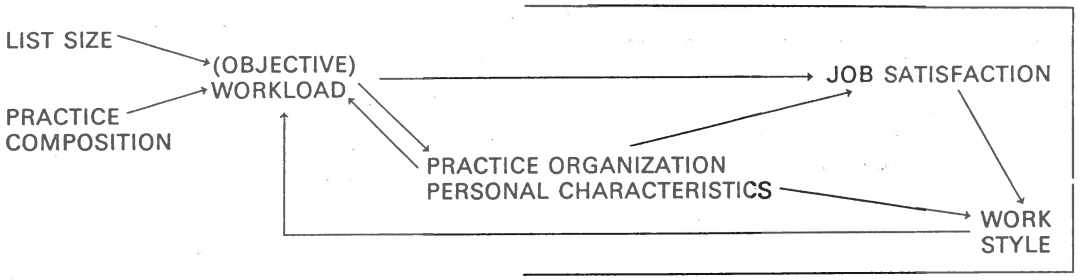
Breslau [37] argue that varied and complex tasks will lead to greater job satisfaction than routine and repetitive tasks. That can be seen as an explanation for their finding that primary care physicians in modern organizational work settings in the U.S. had a lower satisfaction with the work activity, with coworkers and with income, than in more traditional office practices.

In conclusion we can say that the relations between (objective) workload, possibilities to manage workload, and job satisfaction have not been studied systematically and as conclusive as the more simple relations between list size and practice composition, and workload. Practice organization (partnership, appointment system, delegation) seems to be a way to manage workload and as part of the coping possibilities of GPs might be hypothesized also to influence the relation between (objective) workload and job satisfaction. As far as personal characteristics are concerned the work orientations or task perceptions of GPs seem to be an important issue because they relate both to the allocation of time in general practice and to job satisfaction. Work orientation or task perception might be the more important because it is also one of the determinants of work style and possibly through the feedback of workstyle on (objective) workload.

Combining the insights from Section 3 on the influence of demand characteristics on workload and from this section on the influence that the general practitioner himself can exert on his workload, a more theoretical distinction in the work and workload of GPs might be the distinction between that part of the workload that the GP has no influence on and the part that is influenced by the GP [33].

5. EFFECTS ON WORKSTYLE

Part of why workload and job satisfaction is an important topic in health services research is that given levels of workload and job satisfaction might affect the workstyle of general practitioners and the quality of their work. On the other hand aspects of workstyle are closely interrelated with the way GPs manage their workload and with the GP or supply induced workload itself. Relevant aspects of workstyle are: the kind of interventions that take place in consultations (e.g. doctors who do a lot of minor surgery themselves vs those who do not), the communication style of the GP, referrals, prescriptions and repeat consultations. In this part of the review our emphasis is not on explaining differences in e.g. referral rates between GPs; we only look at these aspects of workstyle as far as they are part of a more or less integrated model of workload and job satisfaction. The analytical scheme developed throughout this review can now be extended with workstyle:



We have found only two references to research that investigates the relation between job satisfaction and workstyle.

The relation between job satisfaction and workstyle in general has been investigated by Grol [6]. They have looked at the quality of care as exemplified by prescriptions, referrals, the medical aspects of care and the psycho-social aspects of care. They used both administrative data (on referrals and prescriptions) and direct observation according to a protocol (medical and psycho-social aspects of care). Satisfaction was measured in terms of positive and negative feelings (see Section 2). Positive feelings of the GP coincide with a more open communication style and more attention to psychosocial aspects of the complaints of the patients, but also with a higher number of referrals. Negative feelings coincide with more prescriptions and less explanation to the patient.

Melville [38] has investigated the relation between job satisfaction and the quality of prescribing. Job satisfaction was measured with a five item Likert type scale (Cronbach's alpha 0.92). The quality of prescribing was measured by looking at six groups of drugs that were under debate at the time, either because of serious adverse reactions or possible interactions or because of the risk of addiction or dependency. It turned out that the hypothesis that GPs with lower levels of job satisfaction have a lower quality of prescribing, was confirmed for four out of six groups of drugs, while in two cases there was no significant relationship. Another interesting finding was that GPs with lower levels of satisfaction more often prescribe drugs without actual face to face contact with patient [39].

The findings of Melville are only partially confirmed by Haaijer's study of prescribing in general practice [40]. GPs' satisfaction with their work (measured and analyzed as four separate items) did not relate to the volume and cost of prescribing, but one of the items—the estimate of the percentage of patients presenting trivial complaints—did relate to a lower quality of prescribing.

Apart from this research on the direct relationship between job satisfaction and workstyle, there is research on the relationship between workload and work style, especially on length of consultation and workstyle. In 'the 5 min consultation' Roland [8] experimentally manipulated the booking intervals of five GPs in one group practice. The three conditions were booking intervals of 5, $7\frac{1}{2}$ and 10 min. The 5 min condition resulted in shorter consultations. The effects on work style are in the number of problems in general and psychosocial problems in particular that were detected by the GP. Contrary to the authors' expectations there was no difference between the conditions in the field of prescriptions, examinations, referrals and repeat consultations. GPs were not more stressed, but patients were less satisfied in the 5 min condition. A replication of this study [9] only showed improved communication (GPs spent more time in explaining the problem to the patients) and higher patient satisfaction in the condition of longer booking intervals.

These results are supported by a study by Howie and his colleagues [41, 42]. In 'measurement of stress as it affects the work of the general practitioner' [42] they put forward two hypotheses: The first is that doctors behave differently when surgery workload and other loads reach a level which causes tension or strain; the second is that 'pressured' behavior from the doctor is less satisfying to the patient, less effective clinically and more costly in terms of referrals and prescriptions. Until now only the first results of their project have been published and the hypotheses have not yet been fully tested. What they did find [41] was a relation between the average length of consultations and the handling of psychosocial problems. GPs were categorized as 'fast', 'intermediate' or 'slow' on the basis of a recording of the time they spent in all consultations. The effects on workstyle were investigated for consultations with patients with respiratory illness. Aspects of workstyle investigated were the extent to which psychosocial problems were explored and the prescription of antibiotics. They found no difference between the three groups in the detection of psychosocial problems. There was, however, a

difference between GPs who usually have short or intermediate consultations on the one hand and those with usually long consultations on the other hand. The latter GPs more often explore the psychosocial problems to a greater extent. No direct relation with the prescription of antibiotics was found.

Hartzema and Christensen [43], in a study in the U.S., found that one of the nonmedical factors associated with prescribing volume is the number of patients processed per clinic hour in a primary care HMO setting.

In an older study in Aberdeen, Richardson [44] found a positive relation between consultation rate and the rate of repeat consultations. Therefore they conclude that GPs can influence their workload, because repeat consultations are substantially under the doctors own control.

Practice organization and personal characteristics of GPs relate to aspects of workstyle. GPs with a general practice orientation, as opposed to a clinical orientation, have a more open style of communication [27].

GPs in partnerships and especially in health centres refer less patients [45, 46]; so do GPs with a broader task perception [47]. The relation between list size and referral rate is ambiguous, presumably because of the separate relations between list size and consultation rate, and consultation rate and referral rate [48]. Wilkin and Smith [49] in their review of the literature conclude that the evidence on list size provides little support for the view that high referral rates might result from pressure of work. Boots [32] in his time study of general practice even found that a smaller number of patients per unit of time coincided with more referrals.

Haaijer's summary of the literature on the relation between list size and prescribing in general practice [40] confirms the conclusion of Wilkin and Smith on referrals. In her empirical analyses she did not find any clear relations either.

Summarising this section, we can say that studies, explicitly devoted to the analysis of the effects of workload (or indirect indicators of workload, like list size) and job satisfaction on workstyle and quality of work are scarce. Moreover, the studies on job satisfaction and quality of work only look at part of the model we have developed in the course of this review. For this reason as well as given the actual findings in these studies, no firm conclusions can be drawn, although there are indications of a positive relation of satisfaction and quality of work. The quality of the doctor–patient relationship also seems to be affected by workload. Longer consultations give more room to the patients and lead to higher patient satisfaction. In the broader literature on aspects of workstyle, such as communication style, referrals and prescriptions, workload is only one out of many possibly influencing variables. It is difficult to isolate the influence of workload or workload related variables in these studies.

6. CONCLUSIONS

We have organized the literature in this review around increasingly complex analytical models of the determinants and effects of GPs' workload. Starting with the relation between list size and workload, we have subsequently extended the discussion with the influence of differences in the composition of the GPs' practices. Together list size and practice composition indicate the demand led part of workload. It is important to distinguish this demand led part of workload from the part that is influenced by the GP him/herself. This distinction has to be taken into account in the measurement of aspects of workload by dividing consultations or visits according to the initiative for the contact, the patient or the GP. First contacts in an episode of care could be defined as patient initiated and repeat consultations as physician initiated. This distinction presupposses data collection at the level of contacts or episodes of care through a recording of activities in general practice. But these kind of studies are seldomly done on a large enough scale to enable elaborate multivariate analyses of the interrelations between the variables in the more complex models.

The more complex models look at the relations between workload and job satisfaction or the subjective experience of workload and the possibilities to manage workload. These relations have not been studied systematically. Practice organization seems to have a relation to the possibilities to manage workload. Work orientation or task perception might be as important. The attempts of GPs to manage their workload might become visible in the length of consultations. Length of consultation is on the one hand influenced by the pressure of patient initiated demand and by the management decisions of the GP (e.g. in setting booking intervals) and on the other hand the length of consultation could be seen as a restriction on the GPs' possibilities to exert a certain style of practice and to reach standards of quality.

Work style is affected by workload through the restrictions of time, but at the same time influences workload through the decisions to explore deeper the patient's problems or to end the consultation with a referral or prescription. Although these mutually influencing relations can be distinguished analytically, it is very difficult to analyse them empirically.

As was already mentioned in the introduction, workload is mainly an issue in health care systems with capitation payment for GPs' services. The question is whether GPs' reactions to workload differ according to the system of professional reimbursement. The relations between list size and number of hours worked (weakly positive) and consultation rate and length of consultation (weakly negative) suggest that in capitation systems larger list size or higher patient load coincide with GPs 'processing' more patients per unit of time. Mechanic's study [22] which compares capitation and fee for service, shows a different reaction to workload in fee for service systems: higher patient load coincides with more hours of work at constant length of consultations. One could add to this a hypothesis on salaried systems. Because in salaried systems the number of hours worked is fixed by contract and there is no competition for patients (as there is at least to some extent in capitation systems and to a large extent in fee for service systems), one would expect to find increases of waiting lists as a typical reaction to higher patient load in salaried systems. These,

hypothetical, different reactions to workload warrant further study of workload under different conditions of remuneration.

REFERENCES

1. Fleming D. M. The case for differential capitation fees based on age in British general practice. *Br. Med. J.* **297,** 966–968, 1988.
2. Ball J. G. Workload in general practice. *Br. Med. J.* **277,** 868–870, 1978.
3. Wal G. van der Voor een differentiering binnen het abonnementshonorarium. Huisarts en Wetenschap. **27,** suppl., 68–71, 1984.
4. Calnan M. and Butler J. R. The economy of time in general practice: an assessment of the influence of list size. *Soc. Sci. Med.* **26,** 435–441, 1988.
5. Knight R. The importance of list size and consultation length as factors in general practice. *J. R. College gen. Practitioners* **37,** 19–22, 1987.
6. Grol R., Mokkink H., Smits A. Work satisfaction of general practitioners and the quality of patient care. *Family Practice* **2,** 128–135, 1985.
7. Makin P. J., Rout U. and Cooper C. L. Job satisfaction and occupational stress among general practitioners; a pilot study. *J. R. College gen. Practitioners* **38,** 303–306, 1988.
8. Roland M. O., Bartholomew J., Courtenay M. J. F. The 'five minute' consultation: effect of time constraint on verbal communication. *Br. Med. J.* **292,** 874–876, 1986.
9. Ridsdale L., Carruthers M., Morris R. and Ridsdale J. Study of the effect of time availability on the consultation. *J. R. College gen. Practitioners* **39,** 488–491, 1989.
10. Marsh G. N. and McNay R. A. Factors affecting workload in general practice-II. *Br. Med. J.* **269,** 319–321, 1974.
11. Marsh G. N. and McNay R. A. Team workload in an English general practice-I. *Br. Med. J.* **269,** 315–318, 1974.
12. Carney T. A. A comparison of workload and morbidity recording by partners in a group practice. *J. R. College gen. Practitioners* **30,** 271–277, 1980.
13. Carney T. A. Characteristics of general practitioners with high and low recall rates. *Family Practice* **4,** 183–191, 1987.
14. Westcott R. The length of consultations in general practice. *J. R. College gen. Practitioners* **27,** 552–555, 1977.
15. Thomas K., Birch S., Milner P., Nicholl J., Westlake L. and Williams B. Estimates of general practitioners workload: a review. *J. R. College gen. Practitioners* **39,** 509–513, 1989.
16. Branthwaite A. and Ross A. Satisfaction and jobstress in general practice. *Family Practice* **5,** 83–93, 1988.
17. Fleming D. M. Workload review. *J. R. College gen. Practitioners* **32,** 292–297, 1982.
18. Department of Health and Social Security (1987). See also the comments to this study on the implications for remuneration by Ridsdale L. General practitioner workload: research and policy. *J. R. College gen. Practitioners* **38,** 390–391, 1988.
19. Wilkin D., Hodgkin P. and Metcalfe D. Factors affecting workload: is received wisdom true? In *The Medical Annual, 1986. The Yearbook of General Practice* (Edited by Pereira Gray D. J.). Wright, Bristol, 1986.
20. Fleming D. M. Consultation length (letter to the editor), *J. Royal College gen. Practitioners* **37,** 179, 1987.
21. Armstrong D. and Griffin G. A. Patterns of work in general practice in the Bromley Health District. *J. R. College gen. Practitioners* **37,** 264–266, 1987.
22. Mechanic D. The organization of medical practice and practice orientations among physicians in prepaid and non-prepaid primary care settings. *Med. Care* **13,** 189–204, 1975.
23. Muller C. Review of twenty years of research on medical care utilization. *Hlth Services Res.* **21,** 129–144, 1989.
24. Yuen P. and Balarajan R. Unemployment and patterns of consultations. *Br. Med. J.* **298,** 1212–1214, 1989.
25. Morrell D. C., Evans M. E., Morris R. W. and Roland M. O. The 'five minute' consultation: effect of time constraint on clinical content and patient satisfaction. *Br. Med. J.* **292,** 870–873, 1986.
26. Andersson S-O. and Mattsson B. Length of consultations in general practice in Sweden: views of doctors and patients. *Family Practice* **6,** 130–134, 1989.
27. Verhaak P. F. M. Bewerkelijkheid van huisartsconsulten met verschillende klachten of patiënten, uitgedrukt in tijd. *Tijdschrift voor Sociale Gezondheidszorg* **64,** 558–562, 1986.
28. De Haan J. *De doktersassistente; Delegeren van taken in een Huisartspraktijk*. Lelystad, Meditekst, 1986.
29. Bridgstock M. General practitioners' organization and estimates of their workload. *J. R. College gen. Practitioners* **26,** suppl. 1, 16–24, 1976.
30. Freeman G. K. Receptionists, appointment systems and continuity of care. *J. R. College gen. Practitioners* **39,** 145–147, 1989.
31. Breslau N., Wolf G., Novack A. H. Correlates of physician' task delegation in primary care. *J. Hlth soc. Behav.* **19,** 374–384, 1978.
32. Boots J. M. J. *Het werk van de huisarts: resultaten van een functie analyse in* 93 *huisartspraktijken*. Maastricht, Rijksuniversiteit Limburg, 1983.
33. Calnan M. Images of general practice: the perceptions of the doctor. *Soc. Sci. Med.* **27,** 579–586, 1988.
34. McCranie E. W., Hornsby J. L. and Calvert J. C. Practice and career satisfaction among residency trained family physicians: a national survey. *J. Family Practice* **14,** 1107–1114, 1982.
35. Mawardi B. H. Satisfaction, dissatisfaction, and causes of stress in medical practice. *JAMA* **241,** 1483–1486, 1979.
36. Rankin H. J., Serieys N. M. and Elliot-Binns C. P. Determinants of mood in general practitioners. *Br. Med. J.* **294,** 618–620, 1987.
37. Breslau N., Novack A. H. and Wolf G. Working settings and job satisfaction: a study of primary care physicians and paramedical personnel. *Med. Care* **16,** 850–862, 1978.
38. Melville A. Job satisfaction in general practice: implications for prescribing. *Soc. Sci. Med.* **14A,** 495–499, 1980.
39. Furthermore, there is some but not very strong evidence that face to face consultations lead to less prescribing compared to indirect consultations; compare Haaijer-Ruskamp F. M., Stewart R. and Wesseling H. Does indirect consultation lead to overprescribing in general practice? *Soc. Sci. Med.* **25,** 43–46, 1987.
40. Haaijer-Ruskamp F. M. *Het voorschrijfgedrag van de huisarts*. Thesis, University of Groningen, 1984.
41. Howie J. G. R., Porter A. M. D. and Forbes J. F. Quality and the use of time in general practice: widening the discussion. *Br. Med. J.* **298,** 1008–1010, 1989.
42. Porter A. M. D., Howie J. G. R. and Levinson A. Measurement of stress as it affects the work of the general practitioner, *Family Practice* **2,** 136–146, 1985.
43. Hartzema A. G. and Christensen D. B. Nonmedical factors associated with the prescribing volume among family practitioners in an HMO. *Med. Care* **21,** 990–1000, 1983.
44. Richardson I. M., Howie J. G. R., Durno D. A study of general practitioner consultations in Northeast Scotland. *J. R. College gen. Practitioners* **23,** 132–142, 1973.

45. Wijkel D. Lower referral rates for integrated health centres in the Netherlands. *Hlth Policy* **6**, 185–198, 1989.
46. Wijkel D. Encouraging the development of integrated health centres: a critical analysis of lower referral rates, *Soc. Sci. Med.* **23**, 35–41, 1986.
47. Dopheide J. P. Verwijzingen door de huisarts: enkele determinanten van het verwijscijfer van solo-huisartsen op het verstedelijkt platteland en in forensengemeenten. *Gezondheid Samenleving* **3**, 141–151, 1982.
48. Posthuma B. H. and van der Zee J. Tussen eerste en tweede lijn, deel II: over praktijkgrootte en productiecijfers van huisartsen, Utrecht, NHI, 1978.
49. Wilkin D. and Smith A. Explaining variation in general practitioner refferals to hospital. *Family Practice* **4**, 160–169, 1987.

PROFESSIONAL REIMBURSEMENT AND MANAGEMENT OF TIME IN GENERAL PRACTICE

AN INTERNATIONAL COMPARISON

MICHAEL CALNAN,[1] PETER P. GROENEWEGEN[2] and JACK HUTTEN[3]

[1]Centre for Health Services Studies, The University of Kent, Canterbury, Kent CT2 7NZ, U.K., [2]Netherlands Institute of Primary Health Care, Utrecht and [3]Interuniversity Center for Sociological Theory and Methodology, Utrecht, The Netherlands

Abstract—A hypothetical model was proposed for explaining the relationship between general practitioners' system of payment and the amount of time spent in patient and non-patient work. It was hypothesized that GPs reactions to higher workload vary according to the payment system. In this paper we compare two health care systems which have both mixed systems of payment of GPs. In England and Wales up until April 1990 GPs are partly paid by capitation (approx 45% of their income), partly by allowance (38% of their income) and for a much smaller part fee for service (18% of their income). In the Netherlands GPs are paid by capitation for the publicly insured patients (63% of the average practice list) and fee for service for the privately insured patients. We expect (among other things) a stronger, positive relationship between list size and hours worked in the Netherlands and a comparably strong, negative relationship between list size and booking intervals in the Netherlands and in England and Wales.

Drawing on data collected from national surveys of GP workload in the Netherlands and England and Wales these propositions were examined. The results of this comparative analysis showed some support for the propositions in that the relation between list size and number of hours worked in patient related activities is stronger in the Dutch setting than in England and Wales, and about the same strength for the relationship between list size and booking intervals.

Key words—general practitioners, payment system, workload, United Kingdom, Netherlands

1. INTRODUCTION

Professional reimbursement and workload of general practitioners (GPs) are two related topics that are in the centre of attention of health policy makers in the United Kingdom and in The Netherlands. In both countries the reimbursement system of GPs has been under debate during the past years [1]. In the U.K. the new contract greatly increases the capitation part of total payments and decreases the allowance part. In the Netherlands plans are made to come to a uniform payment system, based on a mix of capitation and fee for service. Behind these changes are ideas about the possibilities to influence the behaviour of GPs through the system of payment.

Workload of GPs, as a topic of policy debate, is mainly discussed in health care systems with (at least partly) capitation payment [2]. In capitation systems the amount of income of GPs is directly related to the number of patients on the list of a GP and not to the number of services, as in fee for service systems. It seems that workload is less often perceived as a problem in fee for service systems, where market forces on demand and supply work more strongly than in capitation systems (or salaried systems). At the background of the debate on workload in capitation systems is the question whether the level of capitation payment justly reflects the amount of work that is generated by different groups of patients [3].

The relation between professional reimbursement system and workload or management of time (workload is often expressed in terms of hours spent on different activities and length of consultations) is not often investigated. The main reason is that these investigations in most cases require international comparisons. In most countries there is only one dominant system of payment, or, if there are more, comparative studies are difficult, because of in-built selection of GPs and/or patients. A notable exception is Mechanic's [4] analysis of the different reactions of paediatricians in pre-paid practice and in fee for service practice to differences in workload. Combined with ideas gained from a review of the literature on workload [2] this formed the basis of a hypothetical model to explain how GPs manage differences in the levels of workload under different remuneration systems. The aim of this paper is to examine the implications of this model for explaining GPs behaviour in England and Wales and The Netherlands.

2. BACKGROUND AND RESEARCH QUESTIONS

We consider three basic professional reimbursement systems. The first is capitation payment; GPs receive a fixed amount of money per patient per year, irrespective of the number of service GPs render to their patients. The second is fee for service payment;

GPs receive a fee for every item of service that is separately valued according to some contract. Finally, there is salaried service; GPs are on the pay roll of some organization funded by public or private sources and they receive a fixed amount of money for a fixed number of hours work.

These basic systems contain different incentives that influence the behaviour of GPs [5]. We are here only interested in one aspect of GPs' behaviour, their management of time. The question is how these payment systems influence the management of time in general practice.

To answer this question we make some simple assumptions. GPs can be modelled as rational human beings who allocate scarce resources in such a way as to maximize utility.

Utility is a very broad concept, containing everything that enhances well being and social approval. To use the concept of utility in an informative way, one has to specify these broad utility arguments, or maximants, with reference to the specific problem under investigation. In the context of a model of the allocation of time under different systems of professional reimbursement, we pose the income per unit time invested and the amount of free time as the only relevant utility arguments. That is not to say that GPs have no other aims in life than to gain money and to have leisure time, but that for this specific problem and for the time being it suffices to use only these utility arguments.

The main resource of GPs is their time. General practice does not require very expensive and scarce equipment or buildings, as do some medical specialties (although one could also formulate hypotheses about investments in equipment under different systems of payment).

The allocation of time to work and leisure is restricted by the GPs anticipation of the reactions of patients and by institutional arrangements in contracts. Capitation payment implies fixed patient lists and therefore more administrative barriers for patients to change GPs than in fee for service systems. Also local groups of GPs tend to adhere to informal rules of conduct which restrict overt competition for each other's patients. But still, in capitation systems patients have the freedom to join the list of another GP. In salaried systems patients usually only have the choice to go private and to pay for services themselves. Institutional arrangements for the allocation of time are most clear in salaried systems, where there are fixed working hours for GPs. Capitation contracts sometimes contain arrangements for surgery hours and out of hours duties.

Given these premises, we predict the following responses of GPs to a high patient load as compared to a low patient load:

- in capitation systems the total number of hours worked by GPs is relatively unaffected by higher patient load. A higher patient load is typically managed by seeing more patients in the same time period. At a given list size the only way to increase utility is by working less hours, thereby increasing the monetary gain per unit of time and at the same time increasing number of leisure hours. An increase in list size means more income, irrespective of the number of services rendered and thus of time invested; GPs will have a tendency not to increase their number of hours worked proportionately to the increase in list size.
- In fee for service systems a higher patient load leads to a higher number of hours worked, while the time invested per patient remains constant. In fee for service systems there is, in contrast to capitation, no fixed patient list. The only way to increase income is to render more services; this increases at the same time the number of hours worked proportionately. The possible tendency to increase services with a better balance between payment and time invested is counteracted by more consumer influence.
- In salaried systems the total number of hours worked and income are fixed. A higher patient load does not affect the number of hours worked, nor the time spent per patient. The effect of a higher patient load is increasing waiting lists.

Described here are hypothetical patterns of behaviour. Observed behaviour in actual health care systems will deviate from this for at least two reasons. First of all, GPs behaviour is more complex than the way we have modelled it here; GPs have other aims in their (professional) lives and are subject to other restrictions in real life situations. But our simple model serves to clarify the principle. Secondly, most health care systems are not pure types, but form a mix of different kind of payment systems. The question now is where general practice in The Netherlands and in England and Wales fit into this hypothetical model. Can we use it to make predictions about patient load and management of time of GPs in these countries?

England and Wales and The Netherlands have in common that the payment system of GPs is partly based on capitation, but with respect to the other part the two countries differ [6]. The payment system of GPs in England and Wales under the pre-1989 contract consists of three parts:

- a capitation fee per patient on the list of a GP, differentiated by age of the patient,
- allowances; basic practice payment and allowances for seniority and for having attended vocational training (fixed payments),
- fee for service, e.g. for vaccinations and preventive services.

The average percentage of income is approx 45 for capitation, 38 for allowances and 17 for fees. The

relatively large part of allowances in the payment system justifies the classification of the system as a mix of capitation and salaried service (although not in the legal sense, but in the sense of the incentives influencing the behaviour of GPs) [7].

The payment system of GPs in the Netherlands consists of two parts:

- capitation payment for the publicly insured patients on the list of a GP,
- fee for service payment for the privately insured patients.

Publicly insured patients are all employees below a certain income ceiling and their dependents, old age pensioners who used to be publicly insured before retirement and certain categories of social welfare dependents. They make up 65% of the population. GPs always have a mix of publicly and privately insured patients, but the actual percentages depend e.g. on the affluence of the community. The payment system of Dutch GPs can be classified in terms of the basic systems of professional reimbursement as a mix of capitation and fee for service.

This classification of the payment of GPs in England and Wales and in The Netherlands, combined with the hypothetical model of the reactions of GPs to differences in patient load, leads us to the following predictions about differences between the two systems:

- the relations between patient load and GPs' management of time will be more or less the same in both systems, because of the common element in both systems;
- the (positive) relation between list size and number of hours worked will be stronger in The Netherlands than in England and Wales;
- the (negative) relation between list size and length of consultations will be approximately equally strong or weak, because both in ideal type salaried systems and fee for service systems length of consultations is unaffected by patient load;
- the (positive) relation between list size and time it takes to get an appointment with one's GP, will be weaker in The Netherlands than in England and Wales.

3. DATA AND METHODS

To test these predictions, we have used two datasets. The first is the dataset from the workload studies of the Health Services Research Unit of the University of Kent at Canterbury; the second stems from the Dutch National Survey or Morbidity and Interventions in general practice.

Data collection and sample

The data on England and Wales have been collected in a national sample survey of unrestricted principals in autumn 1984. Sample size was 2104; the response rate was 67%. Non-response analysis revealed no serious biases. All data were collected through self-completed questionnaires [8].

The Dutch National Survey consists of several linked data sets: recording of all doctor–patient contacts during three months in 1987/88, a self-completed questionnaire and a diary kept during one full week. Contacts were recorded by 168 GPs, questionnaires were completed by 161 of them and 155 of them kept a diary. The GPs in this study were selected from three independent, stratified random samples from all Dutch GPs. Non-response analysis showed over-representation of younger GPs, female GPs and partnerships [9].

Definition of variables

Our predictions contain the following variables: personal list size, number of hours worked, length of consultations and time to get an appointment. In the England and Wales data, personal list size was arrived at for single handed GPs by just taking their list size; and for partnerships by dividing them in those with personal lists and those with a free flow system. For those with a free flow system, the answer to the following question was used: "If you were to change to a personal list of patients, what size would it have to be to give you the same workload as you have now?" In the Dutch data practice list size has been established by setting up an age/sex-register of the practices. For single handed practices, personal list size equals practice list size. For partnerships, the practice list size has been distributed over the partners according to their full-time equivalents of involvement in the practice. Trainees and temporary locums or assistants who have a structural share in the practice, have been treated in the same way as partners, i.e. they have been assigned a personal list size according to the full-time equivalents of involvement of the practice.

As far as hours of work are concerned, we restrict ourselves to activities within the practice and we distinguish between the total hours of work and hours spent in patient-related activities. For England and Wales, we have used the answers to the question: "Approximately how many hours did you spend in the last full week (when not on leave) on each of the following activities within the practice?" Total hours of work is the sum of all categories, while the number of hours in patient-related activities is arrived at by summing the hours of surgery consultations with NHS patients and the hours of home visits with NHS patients including travelling. For the Dutch situation, we used the diaries kept by the GPs. Total number of hours worked was arrived at by summing the time spent on activities within the practice on weekdays within office hours (8.00 hr a.m. until 5.00 hr p.m.) and hours spent in patient-related activities by summing the time for consultations in the practice and for home visits.

The length of consultations was not directly measured in the England and Wales data. As a proxy, the normal booking interval was used ("If you do operate an appointments system, what is the normal booking interval?"). In the Dutch data, both the actual length of consultations and the booking intervals are available. We have used the latter for the sake of comparability ("How much time do you set aside per patient in your appointments system?").

The time it takes to get an appointment was measured as the answer to the question (for England and Wales): "If a patient telephoned last Monday morning to request a non-urgent appointment, when would he have been able to see his doctor (in a partnership: any doctor)?" and (for the Netherlands): "Within how many days patients can usually get an appointment?" In both cases, we use the percentage answering 'on the same day'. This was used as a measure, admittedly crude, of accessibility.

Comparability of data

The two studies took place on different points in time and used different methods of data collection. This poses the question whether or not the data are still comparable enough to test our predictions. The central point here is to note the objective of our study. If we were interested in comparing levels of workload and time management of GPs in both countries, differences in data collection and timing of the studies would have prohibited a meaningful comparison. However, we are not interested in levels, but in relations between variables. It is the comparison of the strength and form of the relations between list size and time management, that we are interested in. In this case, the differences in method and timing of data collection are not problematic.

Method of analysis

To test our predictions, we break down the average number of hours worked, the average booking intervals and the percentage able to get an appointment on the same day by categories of list size for visual inspection and test of linearity (F-ratio test for the difference between η^2 and R^2). Our predictions are about differences between England and Wales and the Netherlands in the strength of the relations between list size and selected dependent variables. Technically, this means a test between two models, one with no interaction between country and list size and one with an interaction term of country and list size. The first model assumes parallel regression lines for England and Wales and the Netherlands; the second model assumes significantly different slopes of the two regression lines. These models have been tested against each other using GLIM (Generalized Linear Interactive Modelling [10]). The test statistic F is computed as the difference of the sums of squares of the two models, divided by the scale parameter.

4. RESULTS

One of the *general* predictions of the hypothetical model was that the relation between patient load and GPs management of time will be more or less the same in both the system in the Netherlands and in England and Wales. This prediction received some support (see Table 1) in that for both countries, there was a similar pattern of statistical relationships between list size and the indices of time. In both countries, the relationship between list size and hours worked in contact with patients and practice based activities was positive, and the relationship between list size and consultation length negative. However, in both countries, there was *no* statistical significant relationship between list size and length of time it takes to get an appointment, although this lack of relationship may be an artefact of the rather crude way the concept was measured.

List size and hours worked

The first of the specific predictions suggested that the relationship between list size and number of hours worked will be stronger in the Netherlands than in England and Wales. Figures 1(a) and (b) show these relationships and show that for both countries, there was a positive linear relationship. Statistical tests (F-ratio) show for both data sets, there was no significant deviation from linearity—see Table 2.

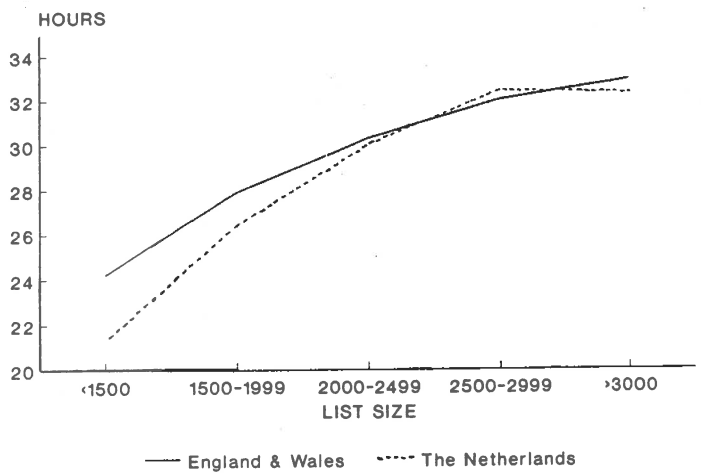
However, as is suggested in Fig. 1(a), the relationship between list size and hours worked in patient related activities, appears to be slightly stronger in the Netherlands than in England and Wales. Table 3 shows that in this case there is a significant difference between the model that assumes parallel regression lines in both countries and the interaction model. The model parameters show that the slope of the regression line for the Dutch data is steeper than for the

Table 1. Correlations between list size, hours worked, consultation length and length of time to get an appointment (Pearson's correlations)

	England and Wales List size	Netherlands List size
Hours per week in patient contact	0.24 ($N = 1364$)*	0.57* ($N = 155$)
Hours per week in practice based activities	0.19 ($N = 1366$)*	0.61* ($N = 155$)
Consultation length (booking internal)	−0.19 ($N = 1173$)*	−0.39* ($N = 159$)
Length of time to get an appointment	−0.02 ($N = 1133$)	−0.05 ($N = 159$)

*$P < 0.01$.

(a)



(b)

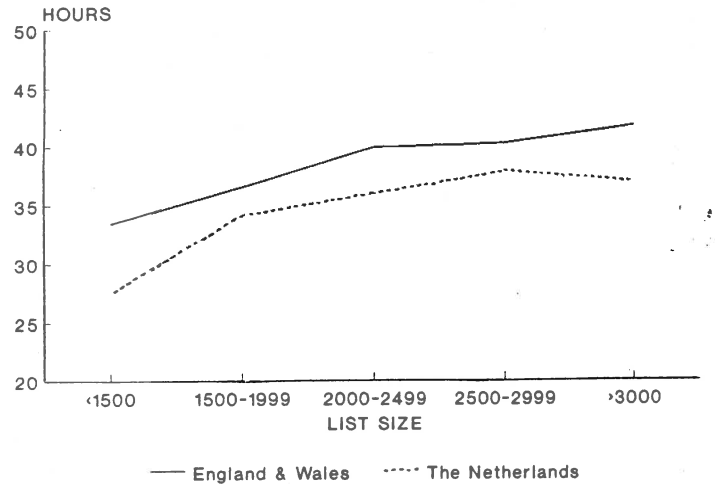


Fig. 1. (a) Hours spent in patient contact. (b) Hours spent in practice-based activities.

English and Welsh data. Hence, there was some support for our prediction although there was little difference between the countries in the relationship between total hours worked in practice based activities and list size. In this case the interaction model is no significant improvement on the parallel regression lines model.

List size and consultation length

Figure 2 shows the relationship between list size and consultation length (booking interval) for both countries. In both countries, there was a marked, inverse linear relationship between list size and booking interval (statistical tests showed in both data sets, there was no significant deviation from linearity—see Table 2). It was predicted that there would be little difference between the countries in terms of the strength of the relationship between list size and booking interval. The grouped data, shown in Fig. 2, suggest a stronger relation in the Netherlands. However, the test, based on list size as a continuous variable, of the interaction model against the parallel regression lines model showed no significant difference.

List size and percentage who can see a doctor the same day

The final prediction suggested that in England and Wales, the relationship between list size and the time it takes to get an appointment with one's GP, will be stronger in England and Wales than in the Netherlands. However, in both countries there was no significance between the two variables and therefore no evidence of a linear relationship (Fig. 3). Thus, there was no support for this prediction (see also Table 3).

5. DISCUSSION

The aim of this paper was to examine the relationship between different arrangements for financial reimbursement of general practitioners and the way they managed different levels of workload. Starting

Table 2. Analyses of variance for list size (in five categories) and (I) hours in patient contact; (II) practiced based activities; (III) booking interval; (IV) length of time it takes to see a doctor

	England and Wales					The Netherlands				
	Sum of squares	*df*	Mean square	*F*	Significance	Sum of squares	*df*	Mean square	*F*	Significance
(I) List size and hours in patient contact										
Between groups	8560.8799	4	2140.2200	22.8067	0.0000	1251.9692	4	312.9923	8.2020	0.0000
Linearity	7893.7646	1	7893.7646	84.1179	0.0000	1192.4508	1	1192.4500	31.2484	0.0000
Dev. from linearity	667.1154	3	222.3718	2.3696	0.0690	59.5183	3	19.8394	0.5199	0.6692
(II) List size and hours in practice based activities										
Between groups	8217.8965	4	2054.4741	12.9722	0.0000	824.0726	4	206.0181	6.470	0.0001
Linearity	7653.5549	1	7653.5549	48.3256	0.0000	709.4762	1	709.4762	22.3122	0.0000
Dev. from linearity	564.3416	3	188.1139	1.1878	0.3131	114.5964	3	38.1988	1.2013	0.3115
(III) List size and booking interval										
Between groups	322.9514	4	80.7378	14.2235	0.0000	130.7329	4	32.6832	10.1304	0.0000
Linearity	312.0103	1	312.0103	54.9664	0.0000	122.1606	1	122.1606	37.8645	0.0000
Dev. from linearity	10.9411	3	3.6470	0.6425	0.5818	8.5724	3	2.8575	0.8857	0.4501
(IV) List size and % able to see doctor on same day										
Between groups	4.0933	4	1.0233	4.7302	0.0009	1.2551	4	0.3138	2.3877	0.0536
Linearity	0.5314	1	0.5314	2.4561	0.1173	0.3785	1	0.3785	2.8803	0.0917
Dev. from linearity	3.4620	3	1.1873	5.4883	0.0010	0.8766	3	0.2922	2.2235	0.0878

off from the three basic systems of remuneration of GPs—capitation, fee for service, salaried service—we have hypothesized typical reactions of GPs in terms of allocation of time to a higher workload. Ideally, these hypotheses should be tested in a comparison of the 'pure' types of remuneration, controlling for all kinds of other relevant differences. For a number of reasons this is usually not possible in the field of research on remuneration systems. 'Pure' types of remuneration are relatively rare; if there would be two 'pure' types within one health care system, there will also be other relevant differences (e.g. in insurance status of patients); and in comparing different health care systems, there will even be more relevant differences, e.g. cultural differences in perceptions of health and health care. In a situation like this, one can either refrain from doing research or acknowledge the methodological difficulties and try to fill in a new piece of the large puzzle about the effects of professional remuneration systems. We think the latter approach is more fruitful. In doing so we have added a new question in the broader field of effects of remuneration systems, the question of differences in allocation of time under different levels of workload in different payment systems. We do not claim to give definitive answers to this question; the available data sets with all their weaknesses are being used for the exploration of new hypothesis, that should be tested more rigorously in the future.

A comparison was made between a system with a mix of salary and capitation (England and Wales: pre 1990-contract) and a system with a mix of capitation and fee for service (Netherlands). This comparison is of particular importance for the National Health Service as the recent changes in financial arrangements for general practitioners in England and Wales actually involve a shift towards a mix of capitation and fee for service payments. In other words, what is the effect on doctors workload and the management of time of having a fee for service component compared with a salaried component. The answer to this question, at least according to this evidence, is that the fee for service element tends to show, as was predicted, that an increase in workload leads to longer working hours in patient contact, although not an increase in hours worked in other practice based activities. This suggests that general practitioners

Table 3. Test of significance for difference between parallel regression lines (model 1) and interaction between country and list size (model 2), for the relation between list size and (I) hours in patient contact; (II) practice based activities; (III) booking interval; (IV) length of time it takes to see a doctor

	Sum of squares	*df*	Mean square	Scale parameter	*F*	*df*	Significance
(I) List size and hours in patient contact							
model 1	119718	1501	619.6	79.4	7.8	1,1500	$P < 0.01$
model 2	119098	1500					
(II) List size and hours in practice based activities							
model 1	199091	1501	320.6	132.6	2.4	1,1500	N.S.
model 2	198770	1500					
(III) List size and booking interval							
model 1	7225,7	1329	9.44	5.44	1.7	1,1328	N.S.
model 2	7216,3	1328					
(IV) List size and % able to see doctor on same day							
model 1	268,9	1289	0.00	0.21	0.00	1,1288	N.S.
model 2	268,9	1288					

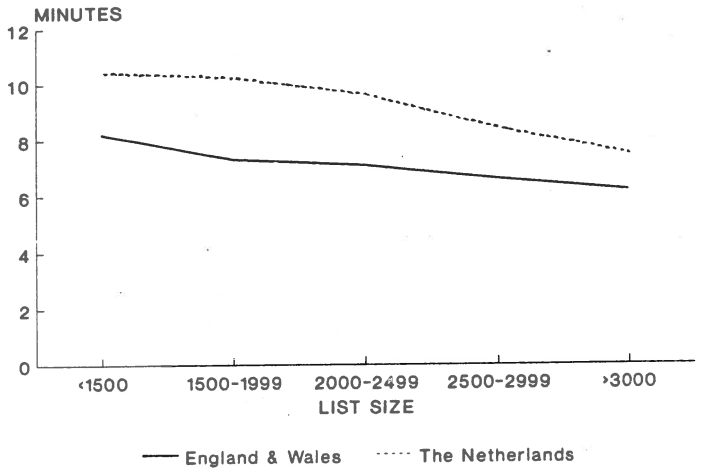


Fig. 2. Average consultation length.

working in systems with a stronger fee for service element may tend to invest more time in patient contact than in other activities. These other activities partly concern administrative duties, but partly also cooperation within primary care and professional activities like post-graduate education. It was also predicted that, because of the common feature of capitation payment in both systems, the relationship between list size and booking interval would be similar. This prediction was shown to be corroborated. The relationship turned out to be similar, indeed. Other research on the relation between list size and management of time in England shows a similar pattern [11], although in this particular study, the strength of the relationships was weaker amongst the lower list size bands. The results suggest that British GPs are more free to allocate their time, which may reflect the extent to which a payment method influences the behaviour of GPs. The final part of the analysis examined the relationship between list size and accessibility (as measured in this study) which was weak under both systems and the salaried element found in the system in England and Wales did not generate queuing or delay.

Ideally this analysis of the impact of different systems of financial reimbursement should involve a comparison between countries with markedly different systems. The systems found in the Netherlands and England and Wales have some similarities and the introduction of the new contract in England and Wales with the reduction in income from allowance have made them even more similar. Thus, it might be possible to predict from the analysis of the Netherlands data, what the implications of the recent changes will be in England and Wales. Of course in doing so, it has to be taken into account that our analysis was based on data from two studies with different methodologies and sample sizes, stressing the tentative character of our conclusions. One of the conclusions which emerges from this analysis will be

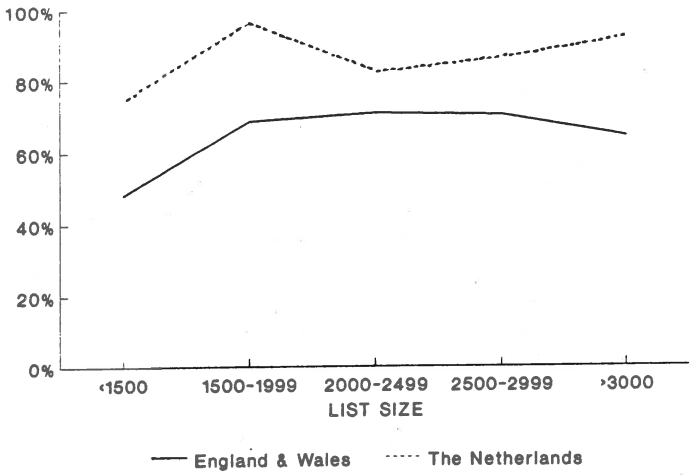


Fig. 3. Appointment the same day.

that the new contract might lead to a change in the relationship between list size and the management of time, with general practitioners who are faced with a heavier workload increasing their hours in patient contact. However, how will general practitioners manage the increase in time spent in patient contact? Judging from the evidence presented here, it might mean less time spent on other activities, as the results showed there was no significant difference between the countries in the strength of the relationship between list size and hours spent on all activities and list size and consultation length. It has been suggested [12] that the new contract will lead to an overall increase in list size and according to our evidence if this occurs there might be a deterioration in the quality of care not only because of a reduction in consultation length [13] but also a reduction in the number of hours general practitioners spend on non-patient contact activities which may mean less time spent on post-graduate education as well as time spent in liaising with other professions. This problem may be exacerbated in the proposals in the new White Paper on the health service in the U.K. in that doctors have been given the option of a more managerial role in terms of controlling use of resources. However, this apparent problem of increasing doctors involvement in management as well as increasing hours in contact with patients may be resolved by delegation of routine tasks, such as health promotion activities to nurses [14].

The analysis also has implications for the proposed changes in the GP-contract in the Netherlands. The proposed new contract for the Netherlands, consisting of a uniform (i.e. the same for publicly and privately insured patients) system of capitation combined with fee for service, is very close to U.K.'s new contract, although the fees will undoubtedly concern other services. Both systems are therefore converging to the same kind of mix, with the capitation part serving to guarantee continuity of care and the fee for service part stimulating active GPs. However, in both systems, then, with larger list size, non-patient related activities will be neglected. Fee for service runs the risk of increasing quantity of care more than quality of care [15]. A counterbalance to this can be found in stimulating postgraduate education and peer review. However, postgraduate education and peer review are typical examples of non-patient related activities. It could be speculated that the allowances in the U.K.'s old contract served to cover these non-patient related activities. Mixed systems of payment are in the centre of attention at this moment [16]—hybridization creates a stronger breed—but it is our feeling that the positive aspects of fixed allowances or salaries are not enough taken into account [17].

REFERENCES

1. Ham C., Robinson R. and Benzeval M. *Health Check; Health Care Reforms in an International Context*. King's Fund Institute, London, 1990.
2. Groenewegen P. P. and Hutten J. B. F. Workload and job satisfaction among general practitioners; a review of the literature. *Soc. Sci. Med.* **32,** 1111–1119, 1991.
3. Fleming D. M. The case for differential capitation fees based on age in British general practice. *Br. Med. J.* **297,** 966–968, 1988. Ball J. G. Workload in general practice. *Br. Med. J.* 868–870, 1978. Wal G. van der. Voor een differentiering binnen het abonnementshonorarium. *Huisarts en Wetenschap* **27,** Suppl. 68–71, 1984.
4. Mechanic D. The organization of medical practice and practice orientations among physicians in prepaid and non-prepaid primary care settings. *Med. Care* **13,** 189–204, 1975.
5. Glaser W. A. *Paying the Doctor; Systems of Remuneration and their Effects*. The Johns Hopkins Press, Baltimore, 1970; Rosen B. Professional reimbursement and professional behavior: emerging issues and research challenges. *Soc. Sci. Med.* **29,** 455–462, 1989.
6. Groenewegen P. P., van der Zee J. and van Haaften R. *Remunerating General Practitioners in Western Europe*. Avebury, Aldershot, 1991.
7. In an article in *The Receptionist* on the different parts of income of GPs in the U.K. it was formulated thus: "It could be likened to receiving a 'salary' for making himself available for a substantial amount of time and agreeing to look after the patients under his care." (*The Receptionist*, programme 1, issue 1, p. 13, 1984).
8. Calnan M. and Butler J. R. The economy of time in general practice: an assessment of the influence of list size. *Soc. Sci. Med.* **26,** 435–441, 1988.
9. Foets M. and van der Velden J. *Een National Studie van Ziekten en Verrichtingen in de Huisartspraktijk; Basisrapport Meetinstrumenten en Procedures*. NIVEL, Utrecht, 1990.
10. Aitkin M., Anderson D., Francis B. and Hinde J. *Statistical Modelling in GLIM*. Clarendon Press, Oxford, 1989.
11. Wilkin D. and Metcalfe D. H. H. List size and patient contact in general medical practice. *Br. Med. J.* **289,** 1501–1505, 1984.
12. Morrell D. The new general practitioner contract in the NHS-review—is there an alternative. *Br. Med. J.* **298,** 1005–1007, 1989.
13. Roland M. O., Bartholomew J., Courtenay M. J. F. *et al.* The 'five minute' consultation: effect of time constraint on verbal communication. *Br. Med. J.* **292,** 874–876, 1986. Ridsdale L., Carruthers M., Morris R. and Ridsdale J. Study of the effect of time availability on the consultation. *J. R. Coll. Gen. Pract.* **39,** 488–491, 1989. Williams S. J. and Calnan M. Convergence and divergence: assessing criteria of consumer satisfaction across primary, dental and hospital care settings. *Soc. Sci. Med.* **33,** 707–716, 1991.
14. Calnan M. *Preventing Coronary Heart Disease: Prospects, Policies and Politics*. Routledge, London, 1991.
15. Forbes J. F. Public policy and primary medical care in the U.K.; an economic perspective on promoting better health. *Fam. Pract.* **5,** 27–241, 1988.
16. Newhouse F. P. Pricing behavior and imperfections in the medical care market place. *Second World Congress on Health Economics*, Zürich, 1990.
17. Potrykus C. A salaried service for family doctor? Health Visitor. *Hlth Visitor* **64,** 216, 1991.

LIST SIZE, COMPOSITION OF PRACTICE AND GENERAL PRACTITIONERS' WORKLOAD IN THE NETHERLANDS

PETER P. GROENEWEGEN,[1] JACK B. F. HUTTEN[2] and KOOS VAN DER VELDEN[1]

[1]Netherlands Institute of Primary Health Care (NIVEL), P.O. Box 1568, 3500 BN Utrecht, The Netherlands and [2]Interuniversity Center for Sociological Theory and Methodology (ICS), State University of Utrecht, The Netherlands

Abstract—Workload of general practitioners plays an important role in discussions about list size and remuneration in health care systems with fixed patient lists and capitation payments, such as in the Netherlands and in the United Kingdom.

Against the background of the fairness of differences in income level between GPs the question is posed to what extent differences in list size reflect differences in workload and to what extent differences in patient characteristics influence workload. Both list size and practice composition relate to the demand led character of general practice.

Data collected in the National Study of Morbidity and Interventions in General Practice are used. Central to this study is a three month recording of all contacts of 161 general practitioners (and their locums, assistant GPs and trainees) in the Netherlands. For each practice a patient register has been made to relate contacts to the practice population. The participating GPs kept a detailed diary covering 24 hr a day during one week.

As indicators of workload several contact rates, hours worked in practice per week (in direct patient care and in other activities) and average length of office consultations are used.

Demand related characteristics have the strongest relation to the number of hours worked by GPs, particularly the number of hours spent in patient-related activities. Rates of contacts, with the exception of the office contact rate, are not related to list size, but mainly to practice composition. The average length of consultations is negatively related to list size and some characteristics of the practice population.

Key words—workload, general practice, list size, the Netherlands

1. INTRODUCTION

Workload of general practitioners (GPs) is an important issue in health care systems with capitation payment for GPs' services. Capitation payment implies a fixed amount of money per patient on an annual basis—sometimes differentiated by age groups in the practice population—irrespective of the amount of services actually rendered by the GPs. In relation to workload the question is posed whether or not the level of capitation payment justly reflects the amount of work generated by the practice population [1–3]. Both, factors on the 'demand' or 'patient' side and factors situated on the 'supply' side determine the level of workload of GPs [4]. The total number of patients on their list and composition of the practice population indicate the so called 'demand-related' influence on workload [4, 5]. Particular characteristics of the practice population might be relevant because some groups of patients, for instance women and the elderly, contact their GP more often or longer than others. 'Supply-related' determinants of workload are e.g. organizational aspects of the practice and personal characteristics of the GP.

A 'fair' capitation system assumes a linear relation between list size and workload of GPs: when GPs have larger lists, they have to work harder and therefore earn more. However, it is conceivable that GPs with a comparable list size have a different workload due to differences in the composition of the practice, e.g. in age distribution. Capitation systems like the one in the U.K. take these differences into account by setting different capitation fees for different age groups. In the Dutch capitation system this kind of differentiation does not exist: fees are only determined by list size. This makes the question relevant whether or not list size is the main determinant on the 'demand' side of the GP's workload. In this paper we will answer this question. First we provide some background information on the Dutch remuneration system and specify the research question. Next we describe the data and methods used. The results are presented in the fourth section. The paper ends with a discussion of the results.

2. BACKGROUND TO THE DUTCH CAPITATION SYSTEM AND RESEARCH QUESTIONS

In January 1988 6275 GPs practised as principals in the Netherlands. Another 349 worked as an assistant to a principal and a number of GPs had a trainee in their practice. The number of inhabitants per principal was 2345; including assistants to a principal this number is 2221 [6]. Actual list size is a bit lower because some groups of the population do not use the normal GP services, e.g. those in military service, institutionalized people and so on. Single handed practice is the dominant form of organization (55%

of the GPs). Thirty-one percent of the GPs are in partnerships of two GPs and 7% are in group practices of three or more GPs. Finally 8% of the GPs work in an integrated health care centre (where they work together with community nurses, social workers and in most cases physiotherapists). Nearly all GPs in the Netherlands work in independent practice. Only 5% works on a salaried basis, mainly in health care centres.

There are two groups of patients based on their insurance status: publicly insured patients (63% of the population) and privately insured patients (37% of the population). Publicly insured patients are all employees (and their dependents) with a salary below a certain income ceiling, old age pensioners who used to be publicly insured before retirement and certain categories of social welfare dependents. The rest of the population is privately insured with one of the many private (profit or nonprofit) insurance companies. The payment system of GPs differs according to these two groups of patients. Publicly insured patients are formally on the list of GPs. The GP receives a fixed amount of money for each patient on the list, irrespective of the number of consultations or services. Privately insured patients are not formally on the list of a GP, but they usually consult the same GP. They pay directly to the GP for the consultation and/or services rendered. Whether or not they will be refunded depends on their insurance policy. Private insurance policies offer coverage of GP services to varying degrees: from complete coverage of GPs' services through various schemes with deductibles to no coverage at all. Hospital and specialist care are always covered by private insurance policies. All GPs have both privately and publicly insured patients, but the actual mix differs from practice to practice.

The system of payment for GPs services has been debated for the last 5–10 years. The debate centres around two topics. The first is the distinction between publicly and privately insured patients. There is agreement now that this distinction will be abolished by the introduction of a uniform system of insurance. The second is about the system of payment for GPs services under a uniform system of insurance. The original position of the GPs organization was that payment should be capitation based, possibly differentiated according to the amount of work associated with different groups of patients. During the last few years attention has shifted to a mixed system consisting of a lower capitation fee and additional fees for consultations and separate services, comparable to the system used in Denmark [7]. It seems probable that capitation will be part of the payment system of Dutch GPs, although in a different form than it is nowadays.

As indicated in the introduction, an assumption underlying undifferentiated capitation payment is that list size is the main determinant of workload of GPs. In the literature evidence is found for the influence of list size on workload. For U.K.'s National Health Service positive correlations are reported of list size with the number of consultations [8, 9] and with indicators of the number of hours worked and in particular the time spent in office consultations [5, 9]. Negative correlations are found with the number of consultations per patient and with the length of consultations [5, 8–11]. On the other hand, it is clear that other factors than list size influence the amount of work of GPs. Wilkin a.o. [9] did find that a higher percentage of elderly and of people from lower social class background in the practice population is associated with more time in direct patient care. However, they did not present a multivariate analysis including list size. It is a common finding in health care utilization research that women use GPs' services more frequently, as do the very young children and the elderly [12]. In the U.K., people with a higher social class background have somewhat more consultations, but lower class people have somewhat more home visits [13]. Unemployed men consult their GP more often than employed men [14]. In the Netherlands, publicly insured patients consume more medical services than privately insured patients [15, 16]. Besides consultation rates the length of consultations is influenced by these characteristics of patients [17–19]. Finally, there is some evidence that the location of practice influences the workload of GPs [5], partly as a result of a higher demand for services among urban populations, but possibly also as a result of better accessibility of services [20].

The actual influence of these characteristics on the workload in general practice depends on the amount of variation between practices. The relative influence of list size compared to other demand related factors has not been assessed in the literature. Therefore, in this article the question is posed: What is the relative influence of personal list size on one hand and of practice composition on the other hand on the workload of Dutch general practitioners?

3. DATA AND METHODS

The study presented here is part of a large project, the National Survey of Morbidity and Interventions in General Practice [21]. Central to this project is a three month recording of all contacts of 161 general practitioners (and their trainees, locums and assistants) in the Netherlands. In this article data on the 161 independent GPs and 7 permanent locums/assistant GPs are analyzed.

Selection of participating GPs is based on a stratified (according to region, urbanization and distance to a general hospital), random sample. Next to the 98 randomly selected GPs, 5 volunteering GPs and 58 GPs who were partners of these originally selected GPs participated. Data recorded include patient characteristics, characteristics of the consultation (e.g. first or repeat consultation, length of

consultation, time of the day), complaints of the patient, diagnosis, and intervention (diagnostic services, treatment, prescriptions, referral). For all practices ($N = 103$) an age/sex register has been set up to be able to relate the contacts to the practice population. This register contains information about 335,000 patients. Out of the participating general practitioners 155 kept a detailed diary covering 24 hr a day during one week (including the weekend) within the three months recording period.

Dependent variables

In the literature different indicators of workload are used [4]. They can be divided in three groups: indicators based on rates of contact of patients, indicators based on the time GPs spent in different activities and the average length of consultations. We will use indicators from each of these groups.

(a) Rates of contact. The number of contacts, during the 3 months of registration, divided by the personal list size of the GP, expresses the relative workload, given the size of the practice population. Contacts can be handled by the practice secretaries or by the GPs themselves. It is also important to distinguish between office consultations and home visits of GPs. Both patients and GPs can be the initiator of the consultation. If the relative number of contacts initiated by patients is larger, the workload generated by the practice population will be larger. Therefore, four different kind of contact rates will be used.

—Total contact rate: all contacts in the practice, handled by GP or practice secretary;
—office contact rate: all office consultations handled by the GP;
—home visit rate: all home visits conducted by the GP;
—patient initiated contact rate: all contacts that, according to the GP, were initiated by the patients themselves.

Contacts of trainees and temporary locums or assistant GPs were added to the contacts of their principal GP because they had no personal lists.

(b) Time spent in practice activities. The diaries kept during one week, were used to compute the time GPs spent in practice. Hours on call/on duty were only taken into account as far as they included actual work. This variable is labelled total hours worked in practice. A distinction was made between two groups of activities namely direct patient care and other activities such as practice administration, education and consultations with colleagues. Hours spent in both kinds were separately analyzed. Inclusion of all hours on duty, irrespective of actual practice activities, would do justice to an important aspect of workload of GPs. On the other hand, this would make a comparative analysis of the relation between list size and hours worked difficult. Large differences in the number of hours worked could be caused by the fact that some GPs were on duty during the week they kept the diary, while others were not [22].

(c) Average length of consultations. Time spent on each individual patient is the third kind of indicator. Particular groups of patients need more time in a consultation than others. To control for differences in travelling time involved in home visits, we have computed the average length of *office* consultations only.

Independent variables

The independent variables are list size and aspects of the practice composition.

(a) List size. The total number of patients per practice was established through the patient register that was set up for each practice. This includes privately as well as publicly insured patients. The practice list size was transformed into an indicator of the personal list size. In the case of single handed practices personal list size equals practice list size. Patients in partnerships or health care centres were asked which GP they usually visit. Patients who did not report one particular GP were distributed over the partners according to their full-time equivalents of involvement in the practice, established at the beginning of the registration period. Trainees and temporary locums or assistants have not been included in the measurement of personal list size. Permanent assistant GPs were treated in the same way as partners, i.e. they have been assigned a personal list size according to full time equivalents of involvement in the practice.

(b) Practice composition. Based on earlier research [4], the following indicators of practice composition are used in the analyses:

—percentage of women;
—percentage of elderly people (over 75 yr of age);
—percentage of young children (under 5 yr of age);
—percentage of publicly insured patients;
—percentage of unemployed people;
—percentage of people with a foreign nationality;
—percentage of low educated people (people over 25 yr of age with primary education only).
—degree of urbanization is included because of the possible differences between urban and rural areas as mentioned before.

The relative influences of list size and practice composition on indicators of workload are tested in a two-step multiple regression analysis. In the first step list size is separately included in the equation. In the second step aspects of practice composition are added. By comparing the explained variances (R^2) of the equations, we can test whether the second equation provides significantly better predictions than the first one. If so, the standardized regression coefficients are considered to indicate the variables with the largest contribution.

4. RESULTS

4.1. The workload of the GPs

Table 1 shows the main descriptive statistics of the variables used in the analyses. The average total contact rate during the three months of registration is 1.18. This means that on the average each patient on the list contacts the GP's practice once in three months. However, there is a large variation between GP's practices in this respect. In general, the number of office contacts is five times as high as the number of home visits. Most contacts are initiated by the patients.

The 155 GPs who completed the diary work 48 hr. As indicated before, hours on duty in which no actual practice activities were noted, were excluded. There is a large variation: some GPs work only one day a week in the practice, others for more than 75 hr. About 71% of that time is spent in direct patient care. The average length of office consultations is 8.4 min.

The average personal list contains almost 2000 patients. Fifty-two percent of the practice population consists of women. There is only small variation between practices in this respect. The age distribution of the practice populations is indicated by the percentage of very old people (75 yr and older) which averages 5%, and the percentage of very young people (under 5 yr of age) which averages 7%. Sixty-one percent is publicly insured, 5% unemployed, 2% of foreign nationality and 16% of lower educational level.

The correlations between the indicators of workload are given in Table 2. Of course the correlations between the indicators of the same group are high. The total contact rate correlates significantly with the other rates. Especially the office contact rate is strongly associated with the rate of contacts initiated by the patients. No significant relation is found between the home visit rate and the office contact rate. This might indicate that a higher demand for office consultations is not compensated by less home visits or *vice versa*.

The correlations between indicators of the allocation of time show that the total working time per week is more strongly related to direct patient care than to other activities in the practice like administration and education. It is interesting to see that the amount of time spent in direct patient care correlates negatively with the number of hours in non-patient-related activities: the more time GPs spent in patient related activities, the less time they spent in other activities.

The correlations between indicators in different groups are sometimes contrasting. The total contact rate is not significantly related to any indicator in the other groups. The office contact rate has a negative relation with the hours spent in practice, especially in direct patient care. On the other hand, the number of home visits per patient on the list, has a positive relation with the hours spent in direct patient care per week. So, controlling for list size, GPs who provide more home visits work more hours per week in direct patient care. Furthermore the correlations show that the average length of office consultations is negatively associated with the home visit rate.

4.2. List size as a determinant of workload

In this section the zero order correlations between personal list size and the indicators of GP's workload will be explored. In Figs 1–3 personal list size is divided into seven groups. For each group the average score on the indicators is computed. The correlation coefficients are based on the unclassified list sizes.

Table 1. Descriptive statistics[1]

	Mean	Min	Max	SD	*N*
Rates of contact[2]					
Total contacts	1.18	0.65	1.99	0.27	168
Office contacts	0.62	0.35	1.29	0.14	168
Home visits	0.14	0.02	0.38	0.71	168
Contacts initiated by patients	0.56	0.32	1.02	0.14	168
Working hours (per week)[3]					
Total hours worked	48.1	8.0	78.5	11.2	155
Hours in direct patient care	33.7	6.0	55.8	9.8	155
Hours in other practice activities	14.4	0.0	21.3	3.7	155
Average length of consultations[4] (min)	8.4	3.2	14.2	1.9	168
Personal list size	1988	220	3889	762.5	168
Practice composition					
% women	52.1	42.3	67.8	3.5	168
% elderly	5.2	0.7	17.0	2.9	168
% 0–4 year	6.8	2.7	19.0	2.6	168
% public insured	60.8	31.3	83.6	10.1	168
% unemployed	4.9	1.1	17.1	2.4	168
% foreign nationality	2.2	0.9	12.9	2.3	168
% low educated	15.7	4.3	35.9	6.2	168

[1]Due to the stratified character of the sample these descriptive statistics can not be seen as true estimates of population parameters. A weighing procedure has been designed to compute true estimates, but for the purpose of this study, the analysis of relations between list size and practice composition and indicators of workload, this is not necessary.
[2]Number of contacts during 3 months divided by personal list size.
[3]Only hours in which practice activities were reported, were included.
[4]Only office contacts were included.

Table 2. Intercorrelations between the indicators of workload

	Total contact rate	Office contact rate	Home visit rate	Patient initiated contact rate	Total hours worked	Hours in patient care	Hours in other activities
Office contact rate	0.42**						
Home visit rate	0.40**	0.01					
Patient initiated contact rate	0.32**	0.83**	0.13				
Total hours worked	−0.03	−0.15*	0.19*	−0.10			
Hours in direct patient care	−0.07	−0.17*	0.25**	−0.08	0.74**		
Hours in other activities	0.05	0.00	−0.04	−0.05	0.52**	−0.20*	
Average length of office consultations	−0.08	0.06	−0.14*	0.01	0.10	0.01	0.12

1-tailed significance: $*P < 0.05$. $**P < 0.000$.

When the rate of all contacts in a practice (including contacts handled by practice secretaries) is considered, there is no significant relation with list size. However, in the case of the office contact rates (handled by GPs themselves) a negative relation with list size is found. GPs with larger practices have a lower rate of contacts initiated by patients.

Figure 2 shows that GPs with a larger personal list size work more hours per week. Larger lists coincide with more hours spent in direct patient care. On the other hand, time spent in other activities has no relation with list size.

List size is negatively related to the length of office consultations (see Fig. 3). However, GPs with very large lists (with more than 3000 patients) spent more time on each patient than GPs with 2250–3000 patients.

4.3. The influence of list size and practice composition

To determine the relative influence of list size and the characteristics of the practice population on indicators of GP's workload, two-step regression analyses were carried out. In the first step list size is the only independent variable. In the second step list size and characteristics of practice composition (see Section 3) are included. The coefficient of determination (R^2) of these two equations will be compared and the difference will be tested. The contact rates are considered in Table 3.

Personal list size turns out to be a bad predictor of the contact rates of GPs. The first equation explains little or no variance. The inclusion of aspects of practice composition significantly improves the coefficients of determination. However, the coefficients of most of the equations are still low. Only in the case of home visits about one third of the variance is explained by 'demand-related' determinants of workload. The standardized regression coefficients in the second equation indicate the variables with the strongest contribution. GPs in urban areas have a lower contact rate during three months: even when we control for differences in practice composition and list size. This might be an effect of easier access to alternative sources of health care in urban areas. The age composition of the practice population is an important determinant of contact rates. GPs who have a higher percentage of children under 5 years of age on their list record a higher rate of contacts, especially of office contacts. Home visit rates on the other hand are higher in practices with a larger percentage of elderly. The extra workload in this respect is at least partly compensated by a negative relation of the percentage elderly with the office contact rate. Other factors with significant contributions are the practice composition by gender (in the case of the office contact rate) and education (total contact rate).

In Table 4 a comparable analysis is presented for the number of hours worked. List size explains one fourth of the variation in the number of hours worked per week. Although the inclusion of aspects of practice composition improves the explained variance, the difference between a regression equation consisting of list size and an equation in which practice composition is added is not significant. This means that list size is the most important factor in determining hours worked per week. Time spent in

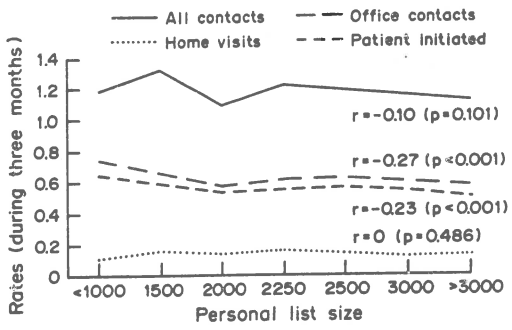


Fig. 1. Contact rates.

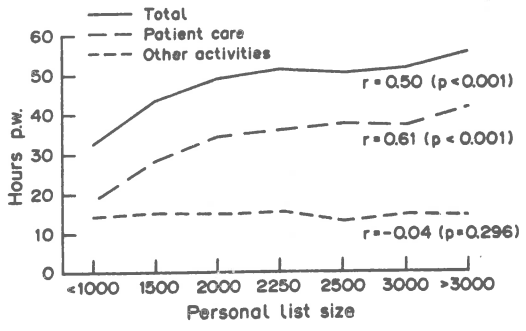


Fig. 2. Hours worked in practice.

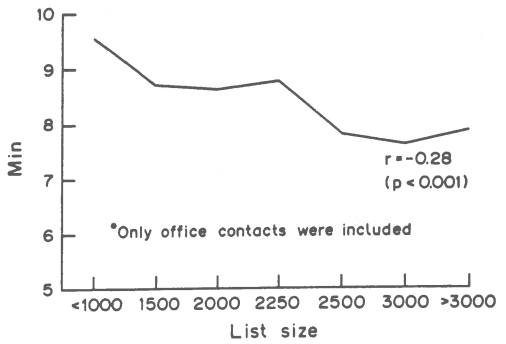


Fig. 3. Average length of consultations.

direct patient care is significantly better predicted when aspects of practice composition are included in the regression equation. The same can be said about the time spent in other activities, like practice administration, post-graduate education and so on. However, the low coefficient of determination in the second equation indicates that factors situated on the 'demand side' have only a very small influence on the hours spent in these activities.

Almost half of the variance in time spent in direct patient care is explained by all 'demand-related' variables. The standardized regression coefficients in the second equation show that list size is still the most important factor when time spent in patient care is concerned. Furthermore, a higher percentage of elderly in the practice population coincides with a larger number of hours spent in direct patient care. This might be a result of GPs visiting elderly people more often at home.

The influence of 'demand-related' variables on the amount of time spent in other activities is rather low: the percentage of elderly and of publicly insured patients are the only two with a significant negative relation. In the case of the elderly this might be caused by GPs needing to spend more time in patient care and a negative relation between time in patient care and in other activities. The influence of the percentage publicly insured patients might be explained by the fact that less administrative procedures have to be carried out for publicly insured patients. Privately insured patients e.g. have to be billed for every consultation.

In Table 5 the results of the two regression analyses on the average length of office consultations are shown.

Only 8% of the variation in the average length of office consultations is determined by list size. When the aspects of the practice composition are taken into account the explained variance increases significantly. GPs with larger lists have on the average shorter consultations than GPs with smaller lists. A larger percentage of unemployed people on a GP's list is associated with longer office consultations. The percentage of publicly insured patients is nega-

Table 3. Results of two regression equations on contact rates: standardized regression coefficients (beta's) and explained variance (R^2)

	Total contact rate	Office contact rate	Home visit rate	Patient initiated contact rate
Equation 1				
List size	−0.10	−0.27*	0.00	−0.23*
Explained variance (R^2)	0.01	0.07	0.00	0.05
Equation 2				
List size	−0.11	−0.17*	−0.02	−0.11
% young children	0.24*	0.30*	0.08	0.30*
% elderly	0.09	−0.19*	0.54*	−0.07
% women	0.12	0.19*	−0.09	0.14
% low educated	0.42*	0.16	0.07	0.09
% unemployed	0.13	0.09	0.07	0.11
% foreign nationality	−0.13	−0.03	−0.05	−0.02
% publicly insured	0.09	−0.09	−0.05	−0.04
Urbanization	−0.23*	0.10	−0.12	0.17
Explained variance (R^2)	0.24	0.22	0.32	0.19
Difference in R^2	0.23	0.15	0.32	0.14
Significance	0.000	0.000	0.000	0.001

*$P < 0.05$.

Table 4. Results of two regression equations on hours worked: standardized regression coefficients (beta's) and explained variance (R^2)

	Total hours worked p.w.	Hours spent in patient care	Hours spent in other activities
Equation 1			
List size	0.50*	0.61*	−0.04
Explained variance (R^2)	0.25	0.37	0.00
Equation 2			
List size	0.55*	0.64*	−0.00
% young children	0.02	0.05	−0.03
% elderly	0.08	0.35*	−0.33*
% women	−0.02	−0.04	0.02
% low educated	0.02	−0.14	0.21
% unemployed	0.11	−0.02	0.17
% foreign nationality	−0.01	0.05	−0.08
% publicly insured	−0.21	0.06	−0.39*
Urbanization	0.05	0.15	−0.12
Explained variance (R^2)	0.30	0.48	0.12
Difference in R^2	0.05	0.11	0.12
Significance	0.260	0.000	0.019

*$P < 0.05$.

Table 5. Results of two regression equations on average length of *office* consultations: standardized regression coefficients (beta) and explained variance (R^2)

	Average length of office consultations
Equation 1	
List size	−0.28*
Explained variance (R^2)	0.08
Equation 2	
List size	−0.16*
% young children	0.11
% elderly	0.06
% women	0.07
% low educated	−0.09
% unemployed	0.30*
% foreign nationality	−0.01
% publicly insured	−0.24*
Urbanization	−0.03
Explained variance (R^2)	0.18
Difference in R^2	0.10
Significance	0.013

*$P < 0.05$.

tively related to the average length of office consultations.

In these regression analyses only independent effects of list size and indicators of practice composition on aspects of workload are described. However, it is conceivable that there are also interaction effects between list size and indicators of practice composition, in the sense that e.g. the percentage of elderly on the GP's list has a different effect on workload in larger practices compared to smaller ones. Interaction effects have been explored by reanalyzing the data separately for GPs with personal list sizes below 2000 patients and above 2000 patients. As far as contact rates are concerned, total contact rate is lower for smaller practices in urbanized areas. In the case of larger practices urbanization has no influence on total contact rates. The office contact rate for larger practices is not affected by practice composition, but it is for smaller practices. No interaction effects have been detected for the other contact rates.

In the group of GPs with smaller practices practice composition has no influence on total hours worked per week and hours spent in patient care. For these smaller practices the number of hours spent in non-patient-related activities is not related to list size nor to practice composition. However, in larger practices each of the indicators of the number of hours worked by a GP is affected by practice composition.

With regard to the average length of office consultations no interaction effects have been found.

5. DISCUSSION

A basic assumption underlying undifferentiated capitation payment of GPs is that list size is the main determinant of their workload. In this article we have tested this assumption by comparing the influence of list size and characteristics of the composition of the GPs' lists on different indicators of workload.

Calnan and Butler [5] tried to assess the degree that the workload of GPs is demand led. They found that list size was particularly important in explaining variations in the number of hours spent in surgery consultation and surgery consultation rates. The research presented here corroborates their findings. List size turned out to have the strongest influence on the number of hours spent in patient care and amongst the contact rates list size has a significant (and negative) relation with the office contact rate. Calnan and Butler also report only a weak association of list size and consultation length as measured by the booking intervals. They conclude that list size is an insensitive measure of patient demand in the sense that it does not take into account the prevalence or severity of illness in the practice population. In our analysis we went one step further by including indicators of the practice composition. Although these are not direct measures of the prevalence and severity of disease in the practice population, they are indirect measures because they have been chosen to represent groups of the population that make higher demands on GPs' services.

Our analysis shows that list size is the main determinant of the total number of hours worked by GPs. The relation is stronger if we only look at the hours in patient-related activities. Composition of the practice also has a significant influence on these indicators of workload. Time spent in non-patient-related activities is only influenced by some aspects of the practice composition. Approximately half of the variation in hours in direct patient care is explained by 'demand-related' characteristics of the practice. Only 10% of the variance in the non-patient-related activities is explained.

In contrast to what is found in the (mainly British) literature, we only find a negative relation between list size and office contact rate, but not for the total contact rate and home visit rate. These rates are more influenced by characteristics of the practice composition, but the explained variance is still low. When we look at the total contact rate, which also includes contacts that have been handled by the practice secretary (such as repeat prescriptions), it seems that larger list size in the Netherlands does not affect the accessibility of the practice. However, the office contact rate with the GP is negatively related to list size (but not very strongly), suggesting that delegation of work towards the practice secretary is a coping strategy for the higher demands of a larger list size in the Netherlands.

Except for the home visit rate, the total hours worked and particularly for the hours spent in direct patient care, demand-related variables are not important determinants of workload as judged from the coefficients of determination in the regression analyses. It may be assumed that the GP's own decisions on how to allocate his time and the organization of the practice (including delegation towards practice secretaries) are of more importance.

Flat capitation rates reflect the GP's involvement in direct patient care relatively well; differentiation of capitation fees according to practice composition improves that situation and also better reflects the activities of GPs in different kind of contacts, particularly for home visits.

However, when one also values non-patient-related activities, not only administrative in kind, but also related to post-graduate education and peer review and cooperation within the primary health care team, a system like the pre-1990 contract in the U.K. with a fixed allowance, apart from capitation would do more justice to the GP's workload.

REFERENCES

1. Ball J. G. Workload in general practice. *Br. Med. J.* **1**, 868–870, 1978.

2. Fleming D. M. The case for differential capitation fees based on age in British general practice. *Br. Med. J.* **297**, 966–968, 1988.
3. Wal G. van der. Voor een differentiering binnen het abonnementshonorarium. *Huisarts Wetenschap* **27**, suppl., 68–71, 1984.
4. Groenewegen P. P. and Hutten J. B. F. Workload and job satisfaction among general practitioners: a review of the literature. *Soc. Sci. Med.* **32**, 1111–1119, 1991.
5. Calnan M. and Butler J. R. The economy of time in general practice: an assessment of the influence of list size. *Soc. Sci. Med.* **26**, 435–441, 1988.
6. Dam F. van, Pool J. and Hingstman L. *Figures from the Information System on Professionals in Dutch Primary Health Care for 1988.* NIVEL, Utrecht, 1989.
7. Flierman H. A. and Groenewegen P. P. Het Deense voorbeeld: een gemengd systeem van honorering per verrichting en abonnement. *Huisarts Wetenschap* **32**, 212–218, 1989.
8. Knight R. The importance of list size and consultation length as factors in general practice. *J. R. Coll. gen. Pract.* **37**, 19–22, 1987.
9. Wilkin D., Hodgkin P. and Metcalfe D. Factors affecting workload: is received wisdom true? In *The Medical Annual, 1986. The Yearbook of General Practice* (Edited by Pereira Gray D. J.). Wright, Bristol, 1986.
10. Armstrong D. and Griffin G. A. Patterns of work in general practice in the Bromley Health District, *J. R. Coll. gen. Pract.* **37**, 264–266, 1987.
11. Fleming D. M. Consultation length (letter to the editor). *J. R. Coll. gen. Pract.* **37**, 179, 1987.
12. Zee J. van der. *De Vraag Naar Diensten van de Huisarts.* Utrecht, NHI, 1982.
13. Marsh G. N. and McNay R. A. Factors affecting workload in general practice—II. *Br. Med. J.* **4**, 319–321, 1974.
14. Yuen P. and Balarajan R. Unemployment and patterns of consultations. *Br. Med. J.* **298**, 1212–1214, 1989.
15. Vliet R. C. J. A. van and van de Ven W. P. M. M. Consumptieverschillen tussen ziekenfonds- en particulier verzekerden nader onderzocht. *Gezondheid Samenleving* **7**, 81–94, 1986.
16. Mootz M. The need for health care utilization and its relation to insurance coverage for medical expenses. *Gezondheid Samenleving* **6**, 232–237, 1985.
17. Morrell D. C., Evans M. E., Morris R. W. and Roland M. O. The 'five minute' consultation: effect of time constraint on clinical content and patient satisfaction. *Br. Med. J.* **292**, 870–873, 1986.
18. Westcott R. The length of consultations in general practice. *J. R. Coll. gen. Pract.* **27**, 552–555, 1977.
19. Verhaak P. F. M. Bewerkelijkheid van huisartsconsulten met verschillende klachten of patiënten, uitgedrukt in tijd. *Tijdschrift Sociale Gezondheidszorg* **64**, 558–562, 1986.
20. Joseph A. E. and Phillips D. R. *Accessibility and Utilization: Geographical Perspectives on Health Care Delivery.* Harper & Row, New York, 1984.
21. Bensing J. M., Foets M., van der Velden J. and van der Zee J. De nationale studie van ziekten en verrichtingen in de huisartspraktijk, achtergronden en methoden. *Huisarts Wetenschap* **34**, 51–61, 1991.
22. Incidental differences in the number of hours worked due to hours on duty can be reduced by only including hours worked during (more or less arbitrarily defined) office hours (8:00 hr a.m. till 5:00 hr p.m.). However, this does not affect our findings. In another analysis this more restricted definition of hours worked has been used. Compare Calnan M., Groenewegen P. P. and Hutten J. B. F. Professional reimbursement and management of time of general practitioners, an international comparison (submitted).

Gender Differences in Practice Style: A Dutch Study of General Practitioners

JOZIEN M. BENSING, PHD, ATIE VAN DEN BRINK-MUINEN, MA,
DINNY H. DE BAKKER, PHD

The differences between female and male general practitioners (GPs) were studied regarding three different factors: 1) Do female GPs see more female patients than their male colleagues in the same practice?; 2) Are female GPs confronted with different types of health problems from their male colleagues?; and 3) Do female GPs provide different services to their patients? Data from the Dutch National Study on Morbidity and Interventions in General Practice were used. All practices in this study with both female (n = 23) and male (n = 27) GPs were selected. This resulted in detailed data on 47,254 consultations, 62% of which were with female patients. The three research questions all received an affirmative response: 1) female patients tend to choose female general practitioners; 2) female GPs see different health problems from their male colleagues, and that is only partly because the patient so chooses; and 3) besides the expected differences in female-specific problems, there is a clear GP-gender effect in the presence of 'social' and 'metabolic' problems in the female GP's consultations. Some differences in the provision of services between male and female GPs occurred, with female GPs spending more time on their patients and having a stronger tendency to provide continuity of care. In addition to a gender effect (both physician and patient) a part-time effect in most issues studied was observed. Key words: gender; general practice; patients' preference; morbidity; interventions; part-time working. (Med Care 1993; 31:219-229)

The classical dyad in the medical encounter consists of a male physician and a female patient. Approximately 60% of all patients who visit Dutch primary health care practitioners are women.[1] This seems to be a fairly universal phenomenon.[2-7] The male-to-female ratio for general practitioners (GPs) shows a completely different picture, and—in the Netherlands at least—an even more unbalanced one: in 1991, 87% of independently established Dutch GPs were men, and only 13% were women.[8] A greater participation of female doctors in general practice can be expected as a result of the sharp increase of female medical students in the last 10 years—a recognized phenomenon in the western world.[9-16]

Because the number of female medical students is growing, gender issues have at-

From the Netherlands Institute for Primary Health Care (NIVEL), Utrecht, the Netherlands.

This research was supported by the Dutch Ministry of Welfare, Health and Cultural Affairs.

The preliminary results of this study were presented as a paper on the Third European Conference on Health Services Research, London, 1991.

Address correspondence to: Mrs. Jozien M. Bensing, PhD, NIVEL, Postbus 1568, 3500 BN Utrecht, The Netherlands.

This article was originally submitted for consideration on 3-26-92. It was accepted for publication on 9-15-92 following the completion of all necessary reviews and/or revisions.

tracted attention in the medical field. Nowadays, there is a general call for more female GPs to enable patients to choose between a male and a female GP and to enhance the quality of care for specific health problems.[17–19] The underlying assumption is that women prefer a female GP and that some health problems are more easily presented to and differently cared for by a female GP. It is interesting to test the legitimacy of such a call from a scientific point of view. An examination of the literature shows that for each of these issues some questions remain unanswered and some results have yet to be confirmed.

Patients' preference for a specific type of physician is generally measured by asking for the patients preference in a questionnaire.[20–24] In most of these studies, patients seem to have a preference for a physician of the same sex,[22–25] although in one study a general preference for female physicians has been reported.[26] However, some studies also suggest that, until recently, there was a general preference for male physicians among male and female patients.[19,22] Female patients often report that they prefer female doctors for sex-specific health problems,[20,22,26] for embarrassing or intimate problems,[23,27] or for behavioral problems;[22] at this moment it is unclear whether female patients prefer female doctors for other types of health problems, or perhaps more generally, regardless of the specific types of health problems. Sometimes the validity of these questionnaire-based measures of preference as measures of the real preference for male or female physicians is questioned, and it is suggested that they are more likely measures of a general sex bias.[22,28] This validity problem is not encountered in studies in which patients preference is measured by looking at the patient's actual consultation behavior. In such studies, it is assumed that the patient shows his or her preference for a male or female physician by making an appointment with a physician of his or her choice. In these studies, female patients generally chose female physicians.[2,7,17,28,30] The generalizibility of the results of most of these studies is dubious however, because the study took place in only one general practice with male and female physicians,[7,17] or because data from larger data-sets are analyzed, without being corrected for the availability of female and male physicians.[11] The latter is a problem because of the uneven distribution of female and male GPs, while the most important factor in choosing a GP is geographical accessibility.[19] A new study is necessary with the participation of an equally available and sufficient number of female and male GPs.

The literature on the *prevalence of various problems* presented to male and female GPs, or the type of diagnoses made by male and female GPs is much more scarce. There is some evidence from morbidity surveys that female physicians are more often visited for genitourinary problems and preventive procedures,[2,7] In one British study, some differences in morbidity pattern were found between male and female GPs, but the overall conclusion of this study was that "there seems to be no recognizable pattern in differences in morbidity seen by women and men doctors other than for gender associated conditions."[2(p 755)] However, this conclusion is based on one study only; more research on different data sets is necessary to confirm this conclusion.

Still less knowledge exists on *gender differences in the services provided.* In a recent American study, female doctors were shown to spend more time with their patients, especially with their female patients,[31] a result that was not confirmed in a British study;[11] female doctors seem to have a more favorable attitude to psychosocial factors in patient care, patient education, and health counseling,[9] although it is not yet clear whether this attitude is also reflected in actual practice behavior. One recent Canadian study, after adjustment for other factors, demonstrated that women provided more counseling and psychotherapy,[16] and also

ordered more laboratory tests, a result that was also found in England.[11] In this study, however, we concluded that the percentage of women in the 15-to-44 age group accounted for most of the differences between men and women doctors rather than the sex of the doctor. Again, it would seem desirable to find out if these results can be replicated.

Accordingly, the present study has been designed, taking into account what is known from literature. Three research questions have been formulated:

1. Do female patients have a preference for female GPs?;
2. Are female GPs confronted with other types of health problems when compared with their male colleagues?;
3. Are female physicians different from their male colleagues with respect to the services that they provide to their patients?

Methods

We used data from the NIVEL National Study of Morbidity and Interventions in General Practice, a large nationwide study among 161 GPs working in 103 practices.[32,33,34] We based selection of participating GPs on a stratified (according to region, urbanization and distance to a general hospital), random sample of all Dutch GPs. The GPs registered detailed information about all patient contracts in a 3-month period, with four groups (together) covering 1 year (April 1987 to April 1988). Data recorded included patient characteristics, characteristics of the consultation (e.g. first or repeat consultation, length of consultation, time of the day), problems presented and physician's diagnoses (classified in the International Classification of Primary Care), and services provided (diagnostic services, treatment, prescriptions, referrals). To ensure maximal uniformity in the data-collection process, all participating doctors were trained in the use of the classification systems. A written instruction with definitions was provided to keep at hand during the consultation.[32] Before the registration period started, all elements of the registration form were tested in each individual practice while a research assistant was present. During the registration period, a research assistant visited each practice to check the data for completeness and irregularities and to discuss problems that might have arisen. The doctors received feedback on their practice profile compared to the 'average' practice profile in the National Study on Morbidity and Interventions in General Practice, a service that was highly appreciated by the doctors and provided an extra opportunity for data control.

For this study we used select portions of the National Study's data set. We selected all group practices with male and female GPs (21 group practices with 27 male and 23 female GPs) to equalize patients' opportunity to choose between a male and a female GP. In this way, distance to the practice location can be ruled out as a possible explanation for differences between male and female GPs, as well as other relevant factors such as the composition of the practice population, the particular characteristics of the neighborhood, and the availability of other health services in the surroundings. There was no significant difference between the age distribution of male and female GPs; the mean age was 38.2 years for male GPs and 41.8 years for. As part-time jobs are common for female physicians but can also be an alternative explanation for differences found between male and female physicians, the group practices have been grouped into practices with female partners working less than a 0,6 full-time-equivalent (n = 10) and practices with female partners working more than a 0,6 full-time-equivalent (n = 11). For reasons of intelligibility these groups are called "part-time working," respectively "full-time working."

In this article, (which is part of a larger study) we used only the routine consultation data for the analyses; home visits, emergencies, and special consultations were ex-

cluded. Altogether 47,254 patients from the selected group practices were examined by the doctors; 29,322 of these patients were women (62.0%). To distinguish between GP and patient gender effects, the results are presented for both patient sexes. In this way it is also possible to study the different types of dyads (male-male, male-female, female-male, female-female). The age-sex register of the patients consulting male and female GPs is presented in Table 1. As there are some differences in the age-sex register between male and female GPs and as age is an important factor in morbidity figures, a direct standardization for age was performed for all analyses presented. Furthermore, an additional standardization for morbidity was performed in analyzing the last research question. The influence of GPs gender will be analyzed under different conditions: 1) with part-time versus full-time working female physicians; and 2) with male versus female patients.

Results

Gender Preferences in Patient Contacts

Contacts with female patients comprise 55.3% of the total workload of male physicians and 71.1% of the female physician's workload. In Figure 1 the result of gender preferences in terms of the sex distribution of patient contacts is presented for each of the 21 group practices with male and female GPs. If there had been no gender preferences, the male/female patient ratio would have been equal for male and female GPs and, as a consequence, each bar in Figure 1 would have had zero length. In none of the group practices is there an over-representation of female patients with male doctors; in contrast, in all but one of the 21 practices, the female GPs examine a significantly higher proportion of female patients as compared with their male colleagues in the same practice. In the only practice where there is no such difference, the female GP works only 1 day per week. The differences are larger in those practices where the female GP works 0.6 full-time equivalent or more. Female patients more often choose a female GP than do male patients if and when they have the opportunity to choose one; this tendency is stronger, as there is more opportunity.

Gender Differences in the Problems Presented

The morbidity pattern of the International Classification of Primary Care (ICPC) chapters for male and female GPs is presented in Figure 2. A detailed Table is pre-

TABLE 1. Age Distribution of Consulting Patients, by GP's and Patients' Sex and Female GP's Involvement in Practice

	♀GP < 0.6 FTE				♀GP ≥ 0.6 FTE			
	♂Patient		♀Patient		♂Patient		♀Patient	
Age (years)	♂GP	♀GP	♂GP	♀GP	♂GP	♀GP	♂GP	♀GP
0–14	11.8	14.9[a]	6.5	6.6	13.6	20.4[b]	11.3[b]	9.1
15–44	43.2	48.4[a]	52.0	57.7[b]	48.7[a]	45.8	50.4[b]	59.0
45–54	13.9	10.8	11.9	11.6	12.6	12.3	11.8	11.3
55–64	15.9[b]	9.9	13.1[b]	10.2	11.6	10.6	12.1[b]	9.0
≥65	15.2	16.0	16.5[b]	13.9	13.5[b]	10.8	14.4[b]	11.6
N	5157	1310	7701	3360	7621	3844	8887	9374

♀, female; ♂, male.
FTE, full-time equivalents.
[a] $P < 0.001$
[b] $P < 0.01$

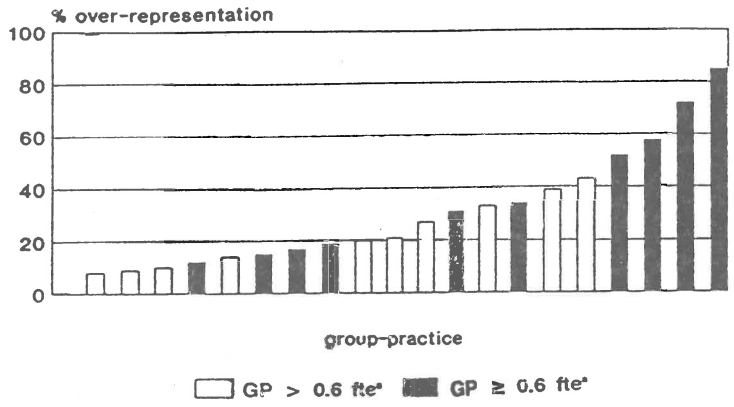


FIG. 1. Over-representation of female patients with female doctors in 21 group practices.

sented in Appendix A. Part-time working female physicians see nearly twice as many gynecologic problems compared with their male colleagues (8.9% vs. 4.7%); musculoskeletal problems are seen less frequently by these female GPs (20.6% against 24.1%). The other differences between the morbidity patterns of male and female GPs from this group are not statistically significant. When we look at the morbidity pattern presented to full-time working female GPs compared with that presented to their male colleagues, some considerable differences can be noted: female GPs see more gynecologic problems (7.5% vs. 4.3%), pregnancy/family planning (6.7% vs. 3.9%), social problems (3.0% vs. 1.8%), and endocrinologic/metabolic problems (3.1% vs. 2.0%). Conversely, female GPs see fewer musculoskeletal problems (23.8% vs. 27.0%), respiratory problems (13.2% vs. 15.9%), and problems of the male genital system (0.6% vs. 1.1%). The over-representation of gynecologic and pregnancy/family planning reasons for encounter can be ascribed to the higher percentage of female patients with female phy-

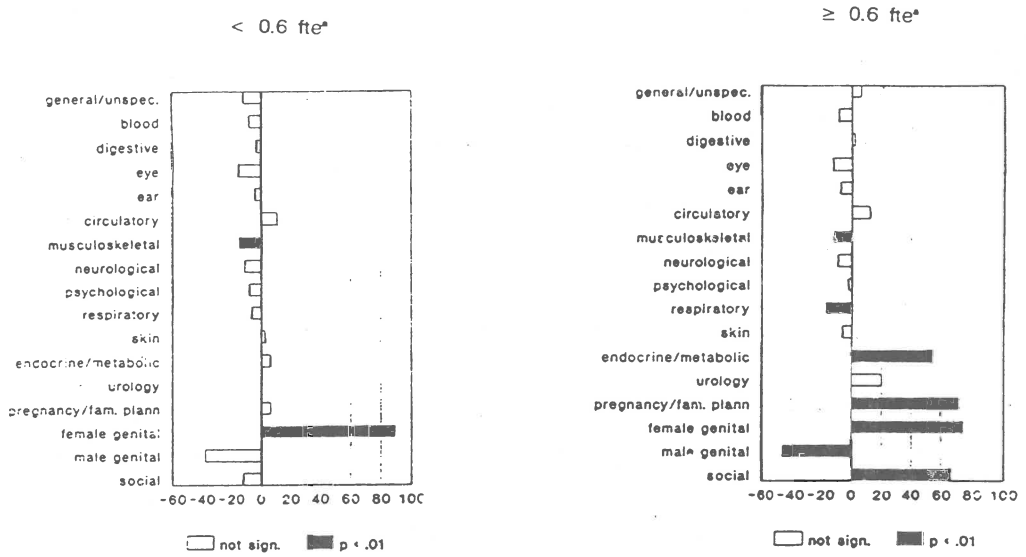


FIG. 2. Over-representation of Reasons for Encounter presented to female and male GPs by female GP's involvement in practice.

sicians, but social problems are presented more often to female physicians by both male (2.8% vs. 1.9%) and female (3.1% vs. 1.8%) patients. The same can be said of the endocrinologic/metabolic problems (2.9% vs. 1.7% for the male patients; 3.1% vs. 2.4% for the female patients).

Gender Effects in Provided Services

In Table 2 some consultation characteristics are presented for female and male GPs under two conditions (full-time and part-time; male and female patients). Despite a direct standardization for patients' age, and morbidity, some differences remain in the consultation characteristics. On average, female physicians spend more time with their patients than their male colleagues (not presented in table): 32.7% (vs. 25.7%) of their consultations last longer than 10 minutes (Note: the average consultation length in the Netherlands is 8.1 minutes[35]). In group practices with part-time working female physicians there are also more long consultations than in group practices with full-time working female GPs (34.0% vs. 24.8%). In addition, female patients also have longer consultations than men (30.8% vs. 23.2%). As a combined effect consultations of more than 10 minutes occur 2.3 as often among part-time working female GPs with female patients as compared with full-time working male GPs with male patients.

The way a GP manages his or her consultations is indicated by three variables: 1) who took the initiative for the consultation?; 2) is it a first or a repeat consultation of a problem presented in earlier consultations?; and 3) is the consultation concluded with a new appointment? In the full-time group a consistent picture emerges: male GPs have more first consultations, more consultations on the patient's initiative, and fewer consultations that end in a new appointment. These results pertain to both male and female patients. In the part-time group no such picture is seen: there is only a difference between male and female physicians in the amount of consultations on the patient's initiative; this is higher among female physicians, regardless of the sex of the patient.

Female physicians are less sure about the diagnosis than their male colleagues, regardless of their involvement in practice and regardless of the sex of their patient; part-time working female GPs consider more complaints in their psychosocial context as their male colleagues, a result that is not found among full-time working GPs.

TABLE 2. Consultation's Characteristics by GP's and Patients' Sex and Female GP's Involvement in Practice (%)

	♀GP < 0.6 FTE				♀GP ≥ 0.6 FTE				
	♂Patient		♀Patient		♂Patient		♀Patient		Overall %
	♂GP	♀GP	♂GP	♀GP	♂GP	♀GP	♂GP	♀GP	
Time of consultation >10 min.	28.2	29.6	34.6	41.8[a]	17.9	24.5[a]	21.9	33.4[a]	28.0
First consultation	44.6	45.2	41.9	43.8	49.6[a]	45.8	48.4[a]	44.9	45.6
Patient-initiated	62.7	71.1[a]	63.1	70.6[a]	74.1[a]	69.6	75.7[a]	73.3	69.1
Further appointment	69.5	68.2	72.6[a]	70.2	58.6	66.8[a]	57.7	66.0[a]	65.7
Psychological aspect	23.9	35.8[a]	29.0	41.1[a]	31.6	31.5	33.8	33.8	31.2
Unsure/doubt	9.3	16.4[a]	9.6	15.9[a]	10.7	14.8[a]	10.3	14.0[a]	15.4
N	5157	1310	7701	3360	7621	3844	8887	9374	47254

♀, female; ♂, male.
FTE, full-time equivalent.
[a] $P < 0.001$.

Table 3 presents an overview of services provided by male and female GPs for their male and female patients. Regardless of the sex of the patient, full-time working women doctors ordered more laboratory tests, wrote down fewer prescriptions, and performed fewer technical-medical interventions. They did more passive and active counseling, but registered fewer reassurances as compared with their male colleagues. They also gave less information. Generally the same picture emerges when we compare part-time working female GPs with their male colleagues, with one notable exception: this particular group of female physicians is in general much more active in providing information, general health education, and lifestyle advice as compared with their male colleagues.

Discussion

This study has some limitations. Because of the study's design, only group practices with male and female physicians have taken part in the study. It is possible that a self-selecting group of physicians is drawn to such group practices; the majority of Dutch GPs work in single-handed practice. This limits the generalizibility of the study. It is also possible that some of the differences that we found are the result of an agreement between the partners in the group practices (for instance agreements on antenatal care). One final warning is necessary about the validity of some of the measures for the services provided, especially the "softer" services: information giving, counseling, and general health education. While, in general, there is little doubt whether a GP has taped a patient's ankle or not, in some instances no such clear distinction can be made in case of (for instance) information giving. Despite the training, the written definitions provided, and the very careful process of data collection, some physician-specific inaccuracies are possible. An observational study is

TABLE 3. Provided Services by GP's and Patients' Sex and Female GP's Involvement in Practice (%)

	♀GP < 0.6 FTE				♀GP ≥ 0.6 FTE				
	♂Patient		♀Patient		♂Patient		♀Patient		Overall
	♂GP	♀GP	♂GP	♀GP	♂GP	♀GP	♂GP	♀GP	%
Internal diagnostics	76.2	74.0	73.2	72.1	77.2	76.1	75.6	74.2	75.0
External diagnostics	6.1	6.6	8.8	11.7[a]	5.6	7.3[a]	7.5	8.9[a]	7.8
Counselling:									
Passive/listening	14.2	18.9[a]	16.6	18.7[b]	9.7	17.4	10.4	15.9[a]	14.2
Active/exploring	17.4	15.5	20.1[a]	15.5	12.9	20.5	13.5	17.3[b]	16.4
Reassuring	15.7	13.4	17.6[a]	13.0	17.5	10.4	18.4[a]	11.0	15.3
Information about:									
Health problems/treatment	38.6	51.4[a]	38.3	45.6[a]	49.4[a]	38.6	46.6[a]	34.3	41.8
General health education	3.2	5.3[a]	3.5	5.6[a]	4.3	3.8	4.1	4.3	4.1
Wait and see	5.1	4.9	4.2	4.2	5.9	5.0	5.6[b]	4.7	5.0
Lifestyle advice	3.0	5.2[b]	2.4	4.1[b]	3.3	2.7	1.8	2.4[b]	2.7
Medical treatments	9.5[a]	6.3	5.6	4.9	8.9	7.6	6.6[a]	4.9	6.7
Prescription	41.2	41.0	44.0[a]	39.3	42.5[b]	39.7	46.6[a]	43.2	43.1
Medication without prescription	1.3	4.5[a]	1.9	4.3[a]	3.7	2.8	3.1	2.9	2.9
Referral primary care	4.4	4.3	3.6	2.9	3.9	4.7	3.4	4.0	3.8
Referral medical specialist	6.5	5.7	4.9	5.3	6.2	5.8	4.6	5.1	5.4
Consultation	1.6	2.9[b]	1.2	2.3[a]	1.5	1.7	0.8	1.1	1.3
N	5,157	1,310	7,701	3,360	7,621	3,844	8,887	9,374	47,254

♀, female, ♂, male.
FTE, full-time equivalent.
[a] $P < 0.001$.
[b] $P < 0.01$.

better equipped to measure this group of variables.

Despite these limitations, this study reveals some interesting results. To start with the main conclusion, gender issues indeed seem important in general practice.

First, it is clear from our data that compared with male patients female patients tend to choose a female GP if and when they have the opportunity to do so; this tendency is strongest with patients of full-time working female GPs.

Second, female GPs see a different morbidity pattern from their male colleagues; for part-time working female GPs this is mainly to be found in the disproportionate amount of gynecologic problems; for the full-time GPs there is a much more variegated picture. In addition to the expected over-representation of female-specific health problems (family planning/pregnancy, gynecology) and under-representation of male-specific health problems (male genital system), these female physicians also see more social and endocrine problems and less musculoskeletal and respiratory problems. Moreover, these differences pertain to both patient sexes. Male patients as well as female patients tend to present more social problems and endocrine problems when confronted with a female GP. The former is interesting, because it includes relational problems; the latter category is interesting because it includes metabolic and eating disorders. It is not clear from our data which is the cause and which is the result of this phenomenon. Perhaps male patients who are willing to discuss their social and/or metabolic problems with a GP tend to choose a female GP rather than a male GP. Conversely female GPs by their attitude and questions stimulate patients to discuss social and metabolic problems. In all probability, both situations occur as doctor and patient tend to socialize each other. Of course the same can be said about the underrepresentation of musculoskeletal and respiratory problems with female physicians; these types of health problems are more popular in the male physician's consultation room. The GP's gender is important in general practice; GPs seem to attract not only patients of the same sex, but also specific types of health problems, regardless of the sex of the patient. These types of health problems correspond with the existing sex-biases or stereotypes: "masculine" health problems (around the male genital system and the musculoskeletal system, which incorporates health problems arising from sports-related accidents, job injuries, and so on) are overrepresented with the male GPs; "feminine" health problems (around gynecology and family planning/pregnancy, but also around human relationships, food habits, and so on) are overrepresented with the female GPs. With part-time working female GPs only the differences in sex-specific morbidity become apparent; with full-time working GPs there also seems to be a more subtle or indirect gender effect. This may explain why only marginal differences around these types of health problems are found in earlier studies, where no distinction was made between full-time and part-time working female GPs.[7,11]

The third conclusion that can be drawn from this study is that female GPs and male GPs seem to develop a different working style. As noted in previous studies,[31] the female GPs from this study tend to have longer consultations, especially when they have a part-time job and when confronted with female patients.

The service profile of female GPs is similar to what is known from other studies,[11,16] with female GPs doing more counseling and ordering more laboratory tests, but writing fewer prescriptions and doing fewer technical-medical interventions. The situation of information giving is a bit complicated: part-time working female GPs seem to be more active in a variety of information-giving activities than full-time female GPs; the full-time female GPs in our study registered even less information giving than their male colleagues. Why this is not clear at the mo-

ment. Results from other studies indicate that female GPs have a higher number of information-giving utterances.[31,36] One possible explanation is that, in these studies, no distinction was made between full-time and part-time working female GPs. Our study clearly indicates some important differences within the group of female physicians between those with a full-time and those with a part-time job. As many women doctors prefer a part-time job, it is possible that in those studies the part-time effect is measured instead of the gender effect. Another possible explanation is that there could be a trade-off effect between information giving and counseling, which becomes most prominent under time pressure (of all subgroups studied, the full-time working female physicians proved to have—by choice or by force—the tightest appointment schedule). Information giving is considered active and instrumental behavior, whereas counseling is considered passive and affective behavior. This is consistent with feminist literature in which male behavior is dominated by active and instrumental types of behavior, whereas female behavior is dominated by passive and affective behavior. A last explanation that cannot be ruled out pertains to the validity of the registered communication variables as has been stated before. Because of the puzzling results, more research on this issue is necessary, partly with other research methods (preferably observation methods).[31,36]

From the way they manage their consultations it seems that full-time female GPs have a special interest in the continuity of care (relatively many repeat consultations visits, many visits on the GPs initiative, and many follow-up appointments). This tendency towards continuity of care can also be interpreted as a consequence of the female GP's greater uncertainty about the exact diagnosis of the patient's health problems, or, of course, the female GP's greater willingness to admit such uncertainty. This is similar to the female GP's stronger tendency to order laboratory tests. All in all, a picture emerges of a female GP who is or admits more often than her male colleague that she is not sure about what exactly is wrong with the patient, and consequently orders laboratory tests and asks the patient to come back for a repeat visit (with which the patient complies). Whether this particular work style must be considered precise and quality enhancing, or instead as insecure and medicalizing, can not be determined from these data, but is an interesting topic for further research.

A last result that has to be discussed is the unexpected but highly relevant distinction between part-time and full-time working female physicians. Part-time working female GPs seem to spend more time on their patients, time that seems to be largely spent on information giving and counseling, two important types of behaviour in preventive and psychosocial care. Possibly as a result, they are more sensitive to the psychosocial aspects of the patient's health problems. Conversely, with part-time working female GPs, patient's preferences are a bit less marked as compared with their full-time working colleagues, possibly because of their restricted availability; there is also less continuity in care, as expressed in proportion of repeat visits, contacts on GP's initiative, and further appointments. This study clearly demonstrates the relevance of part-time versus full-time working as a research topic in itself.

References

1. Groenewegen PP, de Bakker, DH, van der Velden J. Verrichtingen in de huisartspraktijk. Utrecht: Nivel, 1992.

2. Cooke M, Ronalds C. Women doctors in urban general practice: the patients. Brit Med J 1985;290:753.

3. Verbrugge LM, Steiner R. Physician Treatment of men and women patients. Sex bias or appropriate care? Med Care 1981;19:609.

4. Weyrauch KF, Boiko PE, Alvin B. Patient sex role and preference for a male or female physician. J Fam Practice 1990;30:559.

5. Gray J. The effect of the doctor's sex on the doctor-patient relationship. J R Coll Gen Pract 1982; 32:167.

6. Challacombe CB. Do women patients need women doctors? The Practitioner 1983;227:848.

7. Preston-Whyte ME, Fraser RC, Beckett JL. Effect of a principal's gender on consultation patterns. J R Coll Gen Pract 1983;33:654.

8. Pool JB, Hingstman L. Cijfers uit de registratie van beroepen in de eerstelijnsgezondheidszorg, deel I. Utrecht: Nivel, 1990.

9. Maheux B, Dufort F, Béland F, et al. Female medical practitioners. More preventive and patient oriented? Med Care 1990;28:87.

10. Kletke PR, Marder WD, Silbergere AB. The growing proportion of female physicians: implications for US physician Supply. Am J Public Health 1990;80:300.

11. Cooke M, Ronalds C. Women doctors in urban general practice: the doctors. Br Med Journal 1985; 290:755.

12. Hooper J. Full-time women GPs—an invaluable asset. J R Coll Gen Pract 1989;39:289.

13. Rosenthal MM. Perspectives on women physicians in the USA through cross-cultural comparison: England, Sweden, USSR. Int J Women's Studie 1979;2:528.

14. Ogle KS, Henry RC, Durda K, et al. Gender-specific differences in family practice graduates. J Fac Practice 1986;23:357.

15. Wakeford RE, Warren VJ. Women doctors' career choice and commitment to medicine: implications for general practice. J R Coll Gen Pract 1989;39:91.

16. Keane D, Woodward CA, Ferrier BM, et al. Female and male physicians. Different practice profiles. Can Fam Phys 1991;37:72.

17. Kelly JM. Sex preference in patient selection of a family physician. J Fam Practice 1980;11:427.

18. Hayden J. Women in general practice. Time to equalize the opportunities. Br Med J 1991;303:733.

19. Gray J. The effect of the doctor's sex on the doctor-patient relationship. J R Coll Gen Pract 1982; 32:167.

20. Haar E, Halitsky V, Sticher G. Factors related to the preference for a female gynaecologist. Med Care 1975;12:782.

21. Ahmad WIU, Kernohan EEM, Baker MR. Patients' choice of general practitioner: influence of patient's fluency in English and the ethnicity and sex of the doctor. J R Coll Gen Pract 1989;39:153.

22. Fennema K, Meyer DL, Owen N. Sex of physician: patients' preferences and stereotypes. J Fam Practice 1990;30:441.

23. Ackerman-Ross FS, Sochat N. Close encounters of the medical kind: attitudes toward male and female physicians. Soc Sci Med 1980;14A:61.

24. Weyrauch KF, Boiko PE, Alvin B. Patient sex role and preference for a male or female physician. J Fam Practice 1990;30:559.

25. Linn LS, Cope DW, Leake B. The effect of gender and training of residents on satisfaction ratings by patients. J Med Educ 1984;59:964.

26. Nichols S. Women's preferences for sex of doctor: a postal survey. J Gen Pract 1987;37:540.

27. Zare N, Sorenson JR, Heeren T. Sex of provider as a variable in effective genetic counselling. Soc Sci Med 1984;19:671.

28. Weisman CS, Teitelbaum MA. Physician Gender and the physician-patient relationship: recent evidence and relevant questions. Soc Sci Med 1985;20:1119.

29. Graffy J. Patient choice in a practice with men and women general practitioners. Br J Gen Pract 1990;40:13.

30. Hopkins EJ, Pye AM, Solomon M, et al. The relation of patients' age, sex and distance from surgery to the demand on the family doctor. J R Coll Gen Pract 1968;16;368.

31. Roter D, Lipkin, Korsgaard A. Sex differences in patients' and physicians' communication during primary care visits. Med Care 1991;29:1083.

32. Foets M, van der Velden J. Een Nationale Studie van Ziekten en Verrichtingen in de Huisartspraktijk. Basisrapport: Meetintsrumenten en bijlagen. Utrecht: Nivel, 1991.

33. Bensing JM, Foets M, van der Velden J, et al. De Nationale Studie van ziekten en verrichtingen in de huisartspraktijk: achtergronden en methoden. Huisarts en Wetenschap 1991;34:51.

34. Groenewegen PP, Hutten JBF, van der Velden K. List size, composition of practice and general practitioners' workload in the Netherlands. Soc Sci Med 1992;3:263.

35. van der Zee J. De Nationale Studie en het overheidsbeleid ten aanzien van de huisarts. Med Contact 1990;45:611.

36. Hall JA, Palmer RH, Orav EJ, et al. Performance quality, gender and professional role. A study of physicians and nonphysicians in 16 ambulatory care practices. Med Care 1990;28:489.

Appendix Percentage[a] of total number of consultations with reason for encounter[b], by GP's sex and female GP's involvement in practice

	GP < 0.6 FTE			GP ≥ 0.6 FTE[a]		
	♂GP	♀GP	ratio ♀/♂	♂GP	♀GP	ratio ♀/♂
General and unspecified	11.5	10.2	0.89	9.5	10.1	1.06
Blood	1.3	1.2	0.92	1.2	1.1	0.92
Digestive	8.3	8.0	0.97	8.9	9.1	1.02
Eye	2.7	2.3	0.85	3.3	2.9	0.88
Ear	4.9	4.7	0.96	5.3	4.9	0.92
Circulatory	12.0	13.3	0.11	9.2	10.3[c]	1.12
Musculoskeletal	24.1[d]	20.6	0.85	27.0[d]	23.8	0.88
Neurological	5.5	4.9	0.89	5.7	5.2	0.91
Psychological	6.3	5.8	0.92	4.9	4.8	0.98
Respiratory	14.3	13.4	0.94	15.9[d]	13.2	0.83
Skin	11.2	11.5	1.03	13.9	13.1	0.94
Endocrine, metabolic	3.3	3.5	1.06	2.0	3.1[d]	1.55
Urology	2.2	2.2	1.00	1.5	1.8	1.20
Pregnancy, family planning	6.4	6.8	1.06	3.9	6.7[d]	1.72
Female genital system	4.7	8.9[d]	1.89	4.3	7.5[d]	1.74
Male genital system	0.8	0.5	0.62	1.1[d]	0.6	0.54
Social	2.6	2.3	0.88	1.8	3.0[d]	0.67
N	12858	4670		16508	13218	

♂, male; ♀, female.
FTE, full-time equivalent.
[a] As one consultation can have more than one reason for encounter the total adds up to more than 100%.
[b] Classified in the International Classification of Primary Care (ICPC).
[c] $P < 0.01$.
[d] $P < 0.001$.

THE INFLUENCE OF SUPPLY-RELATED CHARACTERISTICS ON GENERAL PRACTITIONERS' WORKLOAD

PETER P. GROENEWEGEN[1] and JACK B. F. HUTTEN[2]

[1]Netherlands Institute of Primary Health Care-NIVEL, P.O. Box 1568, 3500 BN Utrecht, The Netherlands and [2]Interuniversity Centre for Social Science Theory and Methodology, Utrecht University, The Netherlands

Abstract—The workload of general practitioners (GPs) is usually defined in terms of the number of hours worked (divided in time spent on different practice tasks), rates of contact (office consultation and home visit rates) and length of consultations. They are influenced by two groups of factors: demand-related influences and supply-related influences. Demand-related influences refer to the list sizes of GPs and the composition of the practice population. Supply-related influences refer to the way GPs themselves manage their workload. In this article the relative influence of demand- and supply-related variables on the workload of Dutch GPs is assessed.

The data for this analysis has been collected as part of the Dutch National Survey of Morbidity and Interventions in General Practice. We draw on four data sources: a three months recording of all contacts between GPs and their patients, a census of the practice population of the GPs, a mailed questionnaire among GPs and a one week diary kept by the GPs. The population consists of 168 GPs.

The number of hours spent by GPs on practice activities is mainly determined by demand-related characteristics. List size and the percentage of elderly on the list are positively related to the time spent on direct patient care. Running a free flow consultation hour is the only factor on the supply side with an additional effect. GPs supervising a trainee and those with a larger percentage of elderly and publicly insured patients on their list spent more hours on other activities such as practice administration, deliberation and reading medical literature.

List size and the percentage of elderly on the list have a negative influence on the office contact rate, while the percentage of low educated patients on the list and the number of practice secretaries per GP have a positive impact. Furthermore, GPs without a free flow consultation hour and those working in health centres tend to have smaller office contact rate than the others. Home visit rates are smaller when the practice secretaries provide a higher percentage of consultations in the practice, in single handed practices and in the case of female GPs. However, the percentage of elderly on the list is the main determinant of the home visit rate. The average length of consultations is not substantially affected by either supply- or demand-related characteristics.

Key words—workload, practice organization, general practice, The Netherlands

1. INTRODUCTION

Research on the workload of GPs is motivated by the debate on the payment system of GPs on the one hand and by concerns about implications of workload variations for the content and quality of care on the other hand [1]. It is e.g. argued that a 'fair' remuneration system must reflect differences in workload: GPs with a higher workload should have a higher income than those with a smaller workload [2]. Moreover, workload is assumed to influence the performances of GPs in their daily practice. A higher workload indicates that less time is available for each patient which might affect the quality of care [3, 4].

The total workload of GPs is generated from two sources: the demand for care by patients (demand related part) and other practice activities (supply related part). The first part refers to the number of consultations handled by the GP and the time spent on direct patient care. Other practice activities are divided into practice organization (practice administration, deliberation with practice secretaries or colleagues) and medical professional activities such as participation in refresher courses, peer review and reading medical literature.

Both parts are influenced by factors situated at the demand and supply side. The 'demand-led character' of general practice is studied by Calnan and Butler [5]. They used list size as the indicator of demand. A small positive relation is found between list size and workload. Groenewegen *et al.* extended this analysis by adding other indicators at the demand side, viz. indicators of the composition of the practice population [2]. Overall, the variation in workload indicators explained by demand-related variables increases when indicators of the composition of the practice population are added to the influence of list size as such.

This is, however, still only one side of the coin. On the other side GPs in The Netherlands, like in the U.K., work mainly in independent practice. They have a contract with the public health insurance

funds to treat the publicly insured patients on their list. The independent character of general practice gives them freedom to run their practice as they prefer. This implies that they make their own decisions about the allocation of their time and the management of their practice. GPs themselves have some influence on the level of their workload by the delegation of consultations to other workers in the practice or through the organization of their practice, e.g. through their appointment system. It is therefore reasonable to suppose that both demand-related and supply-related variables influence the workload of GPs. The aim of this paper is to find out to what extent supply-related variables have an additional influence on the workload of GPs.

2. BACKGROUND AND RESEARCH QUESTIONS

Concerns regarding a 'fair' payment system and the quality of care in general practice are the main backgrounds to study the workload of GPs. Dutch GPs are partly paid by capitation. That is, for each publicly insured patient on their list GPs receive a fixed amount of money, irrespective of the number of services they render. Capitation fees are not differentiated according to the age of the patient, like they are in the U.K. and Italy [6]. Publicly insured patients are approx. 65% of the Dutch population and each GP has a mix of publicly and privately insured patients; the latter pay their GP a fee for each consultation.

Payment systems might be analysed from two different angles. The first is the way payment systems, as instruments of health policy, influence the behaviour of physicians. From this point of view, the assumption of capitation payment might be that physicians invest in their patients' health by providing preventive care to decrease future workload. The second angle looks at the payment system from the point of view of the relation between payment level and current workload or the principle of rewards according to effort. It is from the latter angle that we look at the payment system in this article. From this point of view the assumption behind a flat capitation fee is that all patients require more or less the same amount of work from the GP. Therefore, list size is, by implication, seen as the main determinant of workload. But, is this really so and to what extent does list size determine workload? Not all patients require the same amount of care. Research in health services utilization shows that some groups in the population use more services than others and if these groups would be unequally distributed across practices, an equally large patient list might generate a different workload depending on the composition of the practice population. Differences in income in a capitation system would be justified if list size is indeed the major determinant of workload.

The relation between list size on the one hand and the content and quality of care on the other hand is an important background to study the workload of GPs. One of the policy measures to increase the quality of care in general practice is the reduction of the list sizes of GPs. This also assumes that list size is an important determinant of the workload of GPs. However, not only list size seems to be relevant but also the way the GPs allocate the available amount of time. A reduction of the number of patients on the lists does not necessarily mean that GPs spend more time on direct patient care: the extra available time can also be spent on other practice activities or leisure time [3]. So, also regarding the relation between workload and the performance of the GPs in daily practice, it is important to investigate the influence of all possible factors on the workload of GPs.

In an earlier paper [2] we have analysed the relative influence of list size and composition of the practice on workload of GPs. List size turned out to be the most important determinant of the total number of hours worked and of the number of hours spent on direct patient care. Practice composition is the most important determinant of the consultation rates, mainly of the home visit rate, of the number of hours spent in non-patient related activities and of the length of consultations. The amount of variance in workload explained by demand-related variables, i.e. list size and practice composition, is highest for the number of hours spent on direct patient care. Still, only half of the variance in hours in direct patient care is explained by demand-related variables. For the other indicators of workload, variance explained by demand-related variables is lower. Hence enough room is left for supply-related variables to exert their influence.

Three types of supply-related factors can be distinguished. Firstly, personal characteristics of the GP; especially gender and age are discussed in the literature. Female GPs have been found to have a lower workload, but also a smaller personal list size [7]. Age of the GP is not related to workload [7], but younger GPs have smaller list sizes suggesting that more experienced GPs can easily cope with a larger potential workload [8]. The second type of supply-related factors are the organizational aspects of the practice that increase or decrease the possibilities to manage a certain amount of demand for care by the patients. Single handed practices have been found to have higher consultation rates and longer consultations [7]. Larger partnerships provide the opportunity to structure the workload [9, 10] through the appointment system and a good deputizing service in evening hours and weekends [11]. Task delegation to a practice secretary is a potential time saver [12].

Thirdly, the involvement of GPs in different kinds of activities can be mentioned. Certain GPs are involved in specific activities that increase their workload in terms of hours worked. Other activities outside general practice, such as a teaching job at medical school, have to be left out when analysing workload in general practice. However, having a trainee in one's practice, and running a dispensary or

doing home deliveries—mainly in rural areas among Dutch GPs—are activities that might increase workload.

Against this background our research question is whether these groups of supply-related variables increase or decrease the workload of GPs, given the influence of demand-related variables. For those supply-related variables that indicate extra activities, we expect to find a positive influence on the workload of GPs. The interpretation is that GPs who engage in extra activities have a higher workload than would be expected on the basis of their list size and composition of practice. For those supply-related variables that indicate the possibilities to manage workload, we expect to find a negative relation. The interpretation is that GPs who have more possibilities to manage the demand of care in their practice, have a lower workload than would be expected on the basis of their list size and composition of practice.

3. DATA AND METHODS

The study presented here is part of a large project, the National Survey of Morbidity and Interventions in General Practice [13]. Central to this project is a three month recording of all contacts of 161 general practitioners (and their trainees, locums and assistants) in the Netherlands. In this article data on the 161 independent GPs and 7 permanent locums/assistant GPs are analysed.

Selection of participating GPs is based on a stratified (according to region, urbanization and distance to a general hospital), random sample. Next to the 98 randomly selected GPs five volunteering GPs and 58 GPs who were partners of these originally selected GPs participated. Data recorded include patient characteristics, characteristics of the consultation (e.g. first or repeat consultation, length of consultation, time of the day), complaints of the patient, diagnosis, and intervention (diagnostic services, treatment, prescriptions, referral). For all practices ($N = 103$) an age/sex register has been set up to be able to relate the contacts to the practice population. This register contains information about 335.000 patients. Out of the participating general practitioners 155 kept a detailed diary covering 24 hr a day during one week (including the weekend) within the three months recording period. Non-response analysis shows no main differences between the GPs who completed the diary and all participating GPs in the study. Comparison to the total population of GPs in the Netherlands, younger GPs, female GPs and GPs working in partnerships are over-represented. This means that our data are not fully representative for all Dutch GPs. However, this does not affect the studied relations between the demand- and supply-related factors on the one hand and workload of GPs on the other hand.

Dependent Variables

In the literature different indicators of workload are used [1]. They can be divided in three groups: indicators based on rates of contact of patients, indicators based on the time GPs spent on different activities and the average length of consultations. We will use indicators from each of these groups.

(a) Rates of contact

The number of contacts, during the three months of registration, divided by the personal list size of the GP, expresses the relative workload, given the size of the practice population. Contacts can be handled by the practice secretaries or by the GPs themselves. It is also important to distinguish between office consultations and home visits of GPs. Furthermore, both patients and GPs can be the initiator of the consultation. If the relative number of contacts initiated by patients is larger, the workload generated through the practice population will be larger. Therefore, four different kind of contact rates will be used.

- —Total contact rate: all contacts in the practice, handled by GP or practice secretary;
- —office contact rate: all office consultations handled by the GP;
- —home visit rate: all home visits conducted by the GP;
- —patient initiated contact rate: all contacts that, according to the GP, were initiated by the patients themselves.

Contacts of trainees and temporary locums or assistant GPs were added to the contacts of their principal GP because they had no personal lists.

(b) Time spent in practice activities

The diaries were used to compute the time GPs spent in practice. Hours on call were only taken into account as far as they included actual work. Inclusion of all hours on duty, irrespective of actual practice activities, would do justice to an important aspect of workload of GPs. On the other hand, this would make a comparative analysis of the relation between list size and hours worked difficult. Large differences in the number of hours worked could be caused by the fact that some GPs were on duty during the week they kept the diary, while others were not. Furthermore, a distinction was made between two groups of activities namely direct patient care and other activities such as practice administration, education and consultations with colleagues.

(c) Average length of consultations

The GPs reported the actual length of each consultation, in minutes, on the registration form. To control for differences in travelling time involved in home visits, we have computed the average length of *office* consultations only.

Independent Variables

The independent variables are list size, aspects of the practice composition and supply-related variables.

(a) List size

The total number of patients per practice was established through the patient register that was set up for each practice. This includes privately as well as publicly insured patients. The practice list size was transformed into an indicator of the personal list size. In the case of single handed practices personal list size equals practice list size. Patients in partnerships or health care centres were asked which GP they usually visit. Patients who did not report one particular GP were distributed over the partners according to their full-time equivalents of involvement in the practice, established at the beginning of the registration period. Trainees and temporary locums or assistants have not been included in the measurement of personal list size. Permanent assistant GPs were treated in the same way as partners, i.e. they have been assigned a personal list size according to full-time equivalents of involvement in the practice.

(b) Practice composition

Practice composition is measured by means of the percentage of certain patient groups on the personal lists of the GPs. According to earlier research [1], the following patient groups need to be considered: women, elderly people (over 75 years of age), young children (under 5 years of age), publicly insured patients, unemployed people, people with a foreign nationality and low educated people (people over 25 years of age with primary education only). Besides, degree of urbanization is included as indicator for the practice composition [14].

(c) Supply-related variables

The factors situated on the 'supply-side' are already discussed in the introduction. The related variables are:

(1) GP characteristics:

—Practice experience (number of years in practice);
—GP's gender (0 = female, 1 = male).

(2) Organizational aspects:

—rota services not only for the weekends, but also for evening hours on weekdays (0 = no, 1 = yes);
—operating free flow consultation hours, next to an appointment system (0 = no, 1 = yes);
—mode of practice; two dummy variables are included in the analyses: single handed practice (0 = no, 1 = yes) and health centre (0 = no, 1 = yes);
—number of practice secretaries per GP in the practice;
—range of activities usually delegated to practice secretaries: a scale of 13 items is used (Cronbach's $\alpha = 0.79$) [15];
—proportion of consultations actually handled by practice secretaries.

(3) Involvement in specific activities:

—doing home deliveries (0 = no, 1 = yes);
—having a trainee in the practice (0 = no, 1 = yes);
—running a dispensary for pharmaceutical prescriptives (0 = no, 1 = yes).

To analyse the influence of the supply-related variables on the indicators of workload, we will start with a simple description of the means of the workload indicators for categories of the supply-related variables. Next the zero-order Pearson correlations between the supply-related variables on the one hand and between supply- and demand-related factors on the other hand will be described. Descriptive statistics of the dependent variables have been reported in an earlier article [2].

The relative influences of demand- and supply-related variables on indicators of workload are assessed in a two-step multiple regression analysis. In the first step demand-related variables are included in the equation. In the second step supply-related variables are added. By comparing the explained variances (R^2) of the equations, we can test whether the second equation provides significantly better predictions than the first one. The standardized regression coefficients of the second equation are considered to indicate the variables with the largest contribution.

4. RESULTS

Table 1 describes the workload of GPs grouped according to the supply-related variables, without taking into account differences in list size and practice composition. Female GPs have a higher office consultation rate, a lower home visit rate, work less hours in total and in direct patient care and have on the average longer consultations. Less experienced GPs have a higher office contact rate, a higher patient initiated contact rate and longer consultations.

Nearly all Dutch GPs participate in rota services for the weekends and weekday evenings. A few only participate in rota services for the weekends. The latter have a higher total contact and home visit rate.

All GPs in our sample run appointment schemes for their consultations. Almost one third also have free flow consultation hours. They have a higher office consultation rate, a higher patient initiated consultation rate and work more hours per week.

GPs in single handed practice work more hours per week in total and in patient related activities, while GPs in integrated health centres have a lower total consultation rate and a lower home visit rate.

In practices with more than one practice secretary per GP the home visit rate is lower. The same goes for GPs who delegate a broader range of tasks to their practice secretary. Those GPs whose practice secretaries handle a higher percentage of the consultations, have a higher total contact rate, a lower office contact rate and a lower patient initiated contact rate.

GPs who do home deliveries have a higher total contact rate. Having a trainee in one's practice results in more hours worked in total and in more hours spent in other activities. GPs who run a dispensary have a higher total contact rate, a lower office contact rate, a lower patient initiated contact rate and spent less hours on direct patient care.

The supply-related variables are correlated amongst each other, as shown in Table 2. Female GPs are underrepresented in the group of GPs with more years of experience. This mirrors the increase of the number of female GPs in the past decade [16]. GPs with more years of practice experience more often have a trainee. Female GPs more often work in partnerships and less in single handed practices. They are less often involved in doing home deliveries. No rota services for weekday evenings is associated with working in single handed practices, with doing deliveries and running a dispensary, and with a larger percentage of consultations handled by the practice secretary. Together this refers to the more rural practices.

Running a free flow consultation hour is not related to any of the other supply related variables. GPs in single handed practices have more practice secretaries per GP, but the range of tasks delegated is smaller. GPs in health centres also have more practice secretaries per GP and they also delegate a broader range of activities. On the whole, more practice secretaries per GP coincides with a broader range of activities delegated. GPs who have a trainee delegate a broader range of tasks to the practice secretary. Those who run a dispensary delegate a smaller range of tasks, but a higher percentage of consultations is actually handled by the secretary.

Table 1. Indicators of workload of GPs according to supply-related variables

	Total contact rate	Office contact rate	Home visit rate	Patient initiated contact rate	Total hours worked per week	Hours spent in patient care	Hours spent in other activities	Average length of office consultations	*N*
No. of years in practice									
<6 yr	1.26	0.68	0.14	0.62	44.8	31.6	13.1	9.20	32
⩾6 yr	1.17	0.61*	0.14	0.55*	49.1	34.4	14.7	8.13*	123
Gender of the GP									
Male	1.18	0.61	0.15	0.55	49.4	35.2	14.2	8.23	140
Female	1.21	0.67*	0.11*	0.58	40.4*	24.8*	15.6	9.26*	28
Rota services evenings									
No	1.45	0.56	0.20	0.47	50.3	34.8	15.4	8.63	8
Yes	1.18*	0.62	0.14*	0.56	48.3	33.9	14.4	8.28	150
Free flow consultations									
No	1.15	0.60	0.14	0.54	46.7	32.6	14.1	8.45	103
Yes	1.22	0.66*	0.14	0.59*	50.4*	35.5	15.0	8.32	65
Mode of practice									
Single handed	1.19	0.60	0.15	0.55	51.3	37.9	13.4	8.33	57
Partnership	1.21	0.63	0.14	0.56	46.6	31.7	14.9	8.32	93
Health centre	0.99*	0.62	0.09*	0.54	46.5*	31.4*	15.1	9.01	18
No. of Practice secretaries per GP									
⩽1	1.18	0.61	0.15	0.56	48.0	33.8	14.2	8.35	119
>1	1.19	0.64	0.12*	0.56	48.2	33.2	15.0	8.52	49
Range of activities delegated									
<14	1.22	0.61	0.16	0.55	46.5	32.4	14.2	8.09	80
⩾14	1.18	0.63	0.12*	0.57	49.9	34.9	15.0	8.52	77
% of consultations by practice secretaries									
<29	1.06	0.65	0.15	0.60	48.4	34.0	14.4	8.43	84
⩾29	1.30*	0.58*	0.13	0.51*	47.8	33.3	14.5	8.37	84
Doing home deliveries									
No	1.15	0.63	0.13	0.57	47.4	33.3	14.0	8.40	107
Yes	1.30*	0.61	0.16	0.54	49.6	34.3	15.3	8.22	53
Having a trainee									
No	1.20	0.63	0.14	0.57	47.2	33.8	13.4	8.37	124
Yes	1.20	0.60	0.13	0.54	51.5*	33.6	17.9*	8.40	35
Running a dispensary									
No	1.12	0.63	0.13	0.58	48.7	34.8	14.0	8.37	128
Yes	1.37*	0.58*	0.16	0.50*	46.2	30.3*	15.9	8.49	40

*$P < 0.05$.

Table 2. Pearson correlations between supply-related factors

	1	2	3	4	5	6	7	8	9	10	11
1. No. of years in practice											
2. Gender of GP (male)	0.16*										
3. Rotaservices evenings (yes)	−0.15	−0.09									
4. Free flow consultation (yes)	−0.04	0.03	0.01								
5. Single handed practice (yes)	0.14	0.15*	−0.19*	0.02							
6. Health centres (yes)	−0.05	0.05	0.08	−0.04	−0.25**						
7. No. of practice secretaries per GP	−0.03	0.22**	−0.01	−0.02	0.24**	0.22**					
8. Range of activities delegated	0.01	−0.00	0.10	0.08	−0.22*	0.33**	0.14				
9. % of consultations by secretaries	0.10	0.03	−0.23**	−0.01	0.02	−0.14	0.21**	0.01			
10. Doing home deliveries (yes)	0.09	0.18*	−0.26**	−0.08	0.02	−0.15	0.11	−0.09	0.29**		
11. Having a trainee (yes)	0.26**	0.09	−0.02	−0.12	−0.04	0.07	0.15	0.21*	0.12	−0.02	
12. Running a dispensory (yes)	0.01	0.03	−0.33**	−0.07	0.01	−0.19*	0.06	−0.27**	0.46**	0.42**	−0.13

Significance: $*P < 0.05$, $**P < 0.01$.

These consultations have obviously to do with repeat prescriptions.

Table 3 shows the correlations between supply- and demand-related factors. Male GPs, more experienced GPs (who are more likely to have a trainee in their practice) and GPs working in a single handed practice have more patients on their personal lists. List size is positively related to the number and range of activities of the practice secretaries.

We see that some personal GP characteristics coincide with the composition of their practice population. Older GPs have a larger proportion of elderly (which also coincides with a larger proportion of low educated patients) and a smaller proportion of young children on their personal list and female GPs have relatively more female patients. Besides, female GPs are more likely to work in urban areas.

Doing home deliveries and running a dispensary are almost exclusively activities of GPs in rural areas. Therefore, GPs who run a dispensary have a smaller percentage of patients with a foreign nationality on their personal lists. They also have a smaller percentage of women and unemployed persons in their practice population. Practices in big cities are more likely to participated in a rota system for the evenings. Health centres are more frequently situated in urban areas with a younger population and relatively more patients with a foreign nationality. The personal lists of GPs working in single handed practices contain a higher percentage of elderly.

In a multivariate analysis the influence of the supply-related variables has been assessed, taking the influence of list size and practice composition into account by first running a regression analysis with the demand-related variables only and then adding the supply-related variables (see Table 4). Adding the supply-related variables in the regression equation changes the coefficients of the demand variables only slightly. The influence of list size increases, especially in the explanation of contact rates; the influence of urbanisation disappears in the explanation of the total contact rate. There is a negative relation between personal list size on the one hand and the total contact rate, office contact rate and the patient initiated contact rate on the other hand.

Table 3. Pearson correlations between supply- and demand-related factors

	List Size	% Young children	% Elderly	% Women	% Low educat	% Unemployed	% Foreign	% Publ. insured	Urbanization
No. of years in practice	0.30**	−0.38**	0.16*	−0.17*	0.25**	−0.01	0.03	−0.01	−0.10
Gender of GP (male)	0.50**	−0.04	0.07	−0.50**	0.07	−0.17*	−0.10	−0.02	−0.16*
Rotaservices evenings (yes)	−0.04	0.03	−0.14	0.10	−0.31**	0.06	0.03	−0.05	0.21**
Free flow consultation (yes)	−0.01	0.00	−0.01	0.09	0.03	0.07	0.17*	0.03	0.05
Single handed practice (yes)	0.20**	−0.24**	0.35**	0.04	0.23**	0.04	0.03	−0.08	0.13
Health centres (yes)	−0.04	0.24**	−0.24**	0.01	−0.27**	−0.07	0.32**	−0.15	0.19*
No. of practice secretaries per GP	0.30**	0.20**	−0.12	−0.05	−0.03	−0.08	0.08	−0.02	−0.16*
Range of activities delegated	0.23**	0.11	−0.27**	0.08	−0.21**	0.11	0.15	0.01	0.08
% of consultations by secretaries	0.19*	−0.10	0.12	−0.20*	0.35**	−0.03	−0.21**	0.18*	−0.39**
Doing home deliveries (yes)	0.09	0.03	0.06	−0.21**	0.06	−0.11	−0.09	−0.02	−0.23**
Having a trainee (yes)	0.20**	−0.03	−0.11	−0.05	0.09	0.08	0.01	0.12	−0.10
Running a dispensory (yes)	−0.16*	0.15*	0.08	−0.27**	0.16*	−0.31**	−0.40**	−0.08	−0.45**

Significance: $*P < 0.05$; $**P < 0.01$.

Table 4. Results of two regression equations on (a) contact rates and (b) hours worked and average length of office consultations: standardized regression coefficients (beta's) and explained variance (R^2); (1) model including demand-related factors only; or (2) model including both demand- and supply-related factors

	(a)								(b)							
	Total contact rate		Office contact rate		Home visit rate		Patient initiated contact rate		Total hours worked		Hours spent on patient care		Hours spent on other activities		Average length of consultations (min)	
	1	2	1	2	1	2	1	2	1	2	1	2	1	2	1	2
List size	−0.11	−0.23*	−0.17*	−0.26*	−0.02	−0.01	−0.11	−0.28*	0.55*	0.58*	0.64*	0.62*	−0.00	0.07	−0.16*	−0.17
% Young children	0.24*	0.16*	0.30*	−0.09	0.08	0.06	0.30*	0.27*	0.02	−0.07	0.05	0.02	−0.03	−0.13	0.11	0.01
% Elderly	0.09	0.05	−0.19*	−0.19*	0.54*	0.50*	−0.07	−0.12	0.08	0.07	0.35*	0.29*	−0.33*	−0.26*	0.06	0.10
% Women	0.12	0.16	0.19*	0.17	0.09	0.00	0.14	0.17	−0.02	−0.06	−0.04	−0.02	0.02	−0.06	0.07	0.02
% Low educated	0.42*	0.35*	0.16	0.35*	0.07	0.12	0.09	0.30*	0.02	0.04	−0.14	−0.11	0.21	0.20	−0.09	−0.16
% Unemployed	0.13	−0.02	0.09	0.02	0.07	0.08	0.11	0.03	0.11	0.08	−0.02	−0.01	0.17	0.13	0.30*	0.25*
% Foreign nationality	−0.13	−0.07	−0.03	−0.09	−0.05	−0.00	−0.02	−0.02	−0.01	−0.00	0.05	0.05	−0.08	−0.06	−0.00	0.03
% Publicly insured	−0.18	−0.15	−0.09	−0.14	−0.05	−0.10	−0.04	−0.13	−0.21	−0.22*	0.06	0.05	−0.39*	−0.39*	−0.24*	−0.18
Urbanization	−0.23*	−0.01	0.10	0.06	−0.12	−0.14	0.17	0.06	0.05	0.07	0.15	0.11	−0.12	−0.03	−0.03	0.04
No. of years in practice		−0.13		−0.10		−0.02		−0.14		−0.23*		−0.11		−0.19*		−0.12
Gender of GP		0.09		0.04		0.19*		0.17		−0.02		0.08		−0.14		−0.06
Rota services evenings		−0.02		0.03		−0.13		0.05		−0.01		0.00		−0.00		−0.13
Free flow consultations		0.11		0.18*		−0.01		0.09		0.20*		0.14*		0.12		−0.05
Single handed practice		−0.09		−0.06		−0.19*		0.01		0.05		0.09		−0.05		−0.00
Health centres		−0.19*		−0.19*		−0.13		−0.25*		−0.02		−0.02		−0.00		0.01
No. of pr. sec. per GP		0.12		0.19*		0.00		0.11		−0.05		−0.07		0.01		0.08
Range of act. deleg.		0.06		0.14		−0.13		0.16		0.05		0.03		0.03		0.13
% of consult. by p.s.		0.47*		0.14		−0.21*		−0.27*		−0.10		−0.03		−0.11		0.07
Doing home deliveries		0.14		0.14		0.06		0.09		0.09		0.04		0.07		−0.06
Having a trainee		0.01		−0.02		0.00		−0.01		0.17*		−0.04		0.31*		0.07
Running a dispensary		−0.05		−0.17		−0.17		−0.20		0.10		−0.00		0.15		0.11
R^2 (1)		0.24		0.22		0.32		0.19		0.30		0.48		0.12		0.18
R^2 (2)		0.52		0.40		0.44		0.39		0.39		0.53		0.22		0.25
Diff. R^2		0.28		0.18		0.12		0.20		0.09		0.05		0.10		0.07
Sign (F)		0.000		0.000		0.014		0.000		0.136		0.410		0.183		0.478

*$P < 0.05$.

With regard to GPs' gender only the regression coefficient of the relation with the home visit rate is significant. The original, bivariate differences (see Table 1)—female GPs have a higher office contact rate, work less hours and have longer consultations—disappear in a multivariate analysis controlled for demand-related and other supply-related variables. Only the home visit rate remains lower for female GPs. More experienced GPs work less hours per week and spend less time in non-patient related activities. The other bivariate conditions—lower office contact rate, patient initiated contact rate and shorter consultations—disappear.

Controlling for demand-related and other supply-related variables, participating in rota services also for the evenings on weekdays is not related to workload. Those who also operate a free flow consultation hour have a higher office consultation rate and work more hours in total and in direct patient care. GPs in single handed practices have a lower home visit rate, although in the bivariate analysis (reported in Table 1) they have a higher home visit rate compared to GPs in health centres. The latter have a lower total contact rate, office consultation rate and patient initiated contact rate, but their home visit rate—which is lower in the bivariate analysis—does not differ significantly from the other practice modes. The variables related to the role of the practice secretaries are not related to the number of hours worked. The availability of practice secretaries is positively related to the office contact rate. When practice secretaries actually handle more consultations, the total contact rate is higher and the home visit rate and the patient initiated contact rate are lower.

Having a trainee results in longer hours in total and in more hours spent on practice organization. Involvement in the other two extra activities—running a dispensary and doing home deliveries—is in the multivariate regression analysis not related to the workload indicators.

Table 4 also shows the coefficients of determination of the regression equation including only demand-related variables, and of the equation also including supply-related variables. Adding the supply-related variables significantly improves the coefficients of determination for the contact rates. Although there is also some improvement in the explained variance of the hours worked and length of consultations, this is not significant.

5. DISCUSSION

In this article we have analysed the influence of variables that relate to the GP and his/her practice organization on indicators of workload. The influence of list size and composition of practice (demand related variables) was taken into account. New to the analysis in this paper is first that both demand-related and supply related influences on workload have been studied, and secondary that they have been analysed in a multivariate way, which is important because of the correlations between variables.

We expected to find positive relations between indicators of workload and supply-related variables that indicate extra activities of GPs. This turned out only to be the case for having a trainee, but not for running a dispensary and doing home deliveries. It seems that the higher workload of GPs who run a dispensary is largely met by their practice secretaries, who handle e.g. a large number of contacts concerning repeat prescriptions.

Supply-related variables that indicate possibilities for managing the workload are supposed to be related to a lower workload. GPs who only work with an appointment system have a lower workload compared to those who have free flow consultations as well. Reducing the accessibility of the practice seems to be a rather easy way to regulate one's workload. In this study, having a single handed practice does not coincide with a higher workload: controlling for the other variables single handed GPs have less home visits and do not differ on the other workload variables. Availability of practice secretaries and the actual number of consultations handled by the practice secretary keep the practice accessible (higher total contact rate and office contact rate). These aspects of delegation have no influence on the number of hours worked by GPs.

Gender is the only personal characteristic of GPs related to the home visit rate, while the workload in terms of hours worked of more experienced GPs is lower than of less experienced GPs.

As far as the relative influence of demand- and supply-related characteristics on the workload of GPs is concerned, it is concluded that the number of hours spent by GPs on practice activities is mainly determined by demand-related characteristics (list size and practice composition). The increase of the coefficient of determination after adding supply-related variables is not significant. The rates of contact, however, are more strongly influenced by supply-related characteristics. The average length of consultations is not substantially affected by either supply- or demand-related characteristics.

It may be concluded that the allocation of time by GPs at the general level of working hours and the division between direct patient care and other activities is influenced by aggregate characteristics of the practice population, mainly list size. At the level of consultations aggregate characteristics of neither the practice population nor the organization of practice are of any influence. The length of consultations seems to be more directly influenced by what happens in the consultation, i.e. the combined influence of the complaints presented in the consultations and the actions taken by the GPs. This combined influence could better be studied at the appropriate level of consultations instead of at the aggregate level of average length of consultations.

The rate of contacts, i.e. the number of contacts divided by the practice population, is influenced by the organization of the practice, and most strongly by the percentage of consultations that is handled by the practice secretary. However, the range of activities delegated to the practice secretary does not have an independent influence on the contact rates. An interpretation of this finding might be that the practice secretary mainly gives administrative support, e.g. in handling repeat prescriptions.

In this article we have conceptualized GPs' workload in terms of hours worked, length of consultations and rates of contact [17]. The relation between list size and allocation of time is by and large comparable between the Netherlands and the U.K., only the relation with the number of hours spent in patient related activities is somewhat stronger in the Netherlands [18]. Although variations between GPs in the number of hours worked are strongly related to list size, the same relation is not as clear at the aggregate level of health care systems. In comparing European countries there is no clear relation between crude estimates of average working hours and density of GPs [6]. In his European study of referrals, Fleming [19] finds a positive relation between the number of consultations per week and the density of GPs. However, after calculating consultation rates the relation disappears [20].

One of the backgrounds of workload studies, as mentioned in the Introduction, are the debates on remuneration. Workload of GPs is mainly studied in health care systems with capitation payment of GPs. In the context of these debates our study shows that differentiating the flat capitation fee that is being used now in the Netherlands would be an improvement. The workload consequences of differences in practice composition would then to some extent be compensated for. However, substantial differences between GPs would remain, especially in the field of contact rates. This could be compensated for by introducing fees per consultation alongside capitation, as is the case in Denmark [21]. This would, however, increase the administrative costs of the system. There is evidence that the introduction of fees per consultation in combination with capitation (and a relatively small portion of fees for separate services within consultations) does not lead to an increased number of consultations [22]. Pure fee for service systems, however, might lead to an increased number of consultations. Fleming's study [19] shows that the health care systems with fee for service payment for GPs are amongst the highest in average number of consultations; also when consultation rates are calculated on the basis of his figures, the fee for service systems rank amongst the highest six countries (out of a total 11 European countries) [23].

The length of consultations does not clearly relate to aggregate characteristics on demand or supply side. However, given the large variation in the average number of consultations per week in Fleming's study, ranging from 220 in the former Federal Republic of Germany to 60 in Norway, average length of consultations will also differ; even if we take into account that total working hours also differ. The average length of consultations in countries with salaried GPs such as Sweden is two to three times as long as in the U.K. and The Netherlands [18, 24].

In conclusion we would like to point to the necessity of research into the workload of GPs at two levels. The first is studies into the length of consultations at the lowest level of aggregation; the second is internationally comparative studies on working hours and rates of contacts in a sufficiently large number of countries to disentangle influences of density of GPs, systems of remuneration and other system characteristics.

REFERENCES

1. Groenewegen P. P. and Hutten J. B. F. Workload and job satisfaction among general practitioners: a review of the literature. *Soc. Sci. Med.* **32,** 1111, 1991.
2. Groenewegen P. P., Hutten J. B. F. and van der Velden J. List size, composition of practice and general practitioners' workload in the Netherlands. *Soc. Sci. Med.* **34,** 263, 1992.
3. Butler J. R. and Calnan M. *Too Many Patients? A Study of the Economy of Time and Standards of Care in General Practice*. Aldershot, Avebury, 1987.
4. Howie J. G. R., Porter A. M. D. and Forbes J. F. Quality and the use of time in general practice: widening the discussion. *Br. Med. J.* 1008, 1989.
5. Calnan M. and Butler J. R. The economy of time in general practice: an assessment of the influence of list size. *Soc. Sci. Med.* **26,** 435, 1988.
6. Groenewegen P. P., van der Zee J. and van Haaften R. *Remunerating General Practitioners in Western Europe*. Aldershot, Avebury, 1991.
7. Wilkin D., Hodgkin P. and Metcalfe D. Factors affecting workload: is received wisdom true? In *The Medical Annual, 1986. The Yearbook of General Practice* (Edited by Pereira Gray D. J.). Wright, Bristol, 1986.
8. Armstrong D. and Griffin G. A. Patterns of work in general practice in the Bromley Health District. *J. R. College gen. Practitioners* **37,** 264, 1987.
9. Branthwaite A. and Ross A. Satisfaction and job stress in general practice. *Family Practice* **5,** 83, 1988.
10. Freeman G. K. Receptionists, appointment systems and continuity of care. *J.R. College gen. Practitioners* **39,** 145, 1989.
11. Bridgstock M. General practitioners' organization and estimates of their workload. *J.R. College gen. Practitioners* **26,** suppl.1, 16, 1976.
12. de Haan J. *De doktersassistente; delegeren van taken in een huisartspraktijk*. Lelystad, Meditekst, 1986.
13. Foets M., van der Velden J. and de Bakker D. *Dutch National Survey of General Practice, a Summary of the Survey Design*. Utrecht, NIVEL, 1992.
14. The degree of urbanization has been included in the set of demand-related characteristics. Urban dwellers are less well off in terms of health and health related behaviour (compare e.g. Verheij R. *De grote-stadspraktijk*. Utrecht, NIVEL, 1993; Garretsen H. F. L. and Raat H. *Gezondheid in de vier grote steden*. Den Haag, Wetenschappeliijke Raad voor het Regeringsbeleid, 1989).
15. Foets M., Stokx L., Hutten J. and Sixma H. *De huisartsenquete van de nationale studie: datareductie door schaalconstructie*. Utrecht, NIVEL, 1991.

16. Kortenhoeven D. *Vrouwelijke artsen en vestiging als huisarts*. Utrecht, Bohn, Scheltema & Holkema, 1990.
17. The subjective side—how do GPs experience their workload and how is subjective workload related to carrying capacity—has not been studied. Compare Branthwaite and Ross (Ref. [9]); Makin P. J., Rout U. and Cooper C. L. Job satisfaction and occupational stress among general practitioners; a pilot study. *J. R. College gen. Practitioners* **38,** 303, 1988; Dierendonck D. van, Sixma H. and Groenewegen P. P. De opgebrande huisarts. *Medisch Contact* **47,** 331, 1992.
18. Calnan M., Groenewegen P. P. and Hutten J. B. F. Professional reimbursement and management of time in general practice: an international comparison. *Soc. Sci. Med.* **35,** 209, 1992.
19. Fleming D. M. *The European Study of Referrals from Primary to Secondary Care*. Thesis Publishers, Amsterdam, 1993.
20. We calculated consultation rates by multiplying the average number of consultations per week by 52 and divided this by the number of inhabitants per GP (as a proxy of list size). Overall there is no clear relation between this proxy of list size and consultation rate. However, the eight countries falling in the range between 1400 and 2100 inhabitants per GP show a slightly positive relation between list size and consultation rate.
21. Flierman H. A. *Changing the Payment System of General Practitioners*. Utrecht, NIVEL, 1991.
22. Krasnik A., Groenewegen P. P., Pedersen P. A. *et al.* Changing remuneration systems: effects on activity in general practice. *Br. Med. J.* **300,** 1698, 1990.
23. Sandier, on the basis of only six health sysetms, does not find evidence for a higher number of visits in systems with fee for service payment for GPs. Sandier S. Health service utilization and physician income trends. *Hlth Care Financing Rev.* annual supplement, 33–48, 1989.
24. Andersson S-O. and Mattson B. Length of consultations in general practice in Sweden: views of doctors and patients. *Family Practice* **6,** 130, 1989.

APPENDIX

LIST OF PUBLICATIONS FROM THE NATIONAL SURVEY IN DUTCH

1. Theses
2. Books and reports
3. Articles
4. Other publications

APPENDIX

1. Ph.D. THESES *(partial or largely based on data from National Survey)*

- Flierman H. **Changing the payment system of general practitioners** (Thesis). Utrecht: NIVEL, 1991.
- Kersten T. **De invloed van de huisarts in de tweedelijn**. (Proefschrift) Utrecht: NIVEL 1991.
- Nijland A. **De praktijkassistente in de huisartspraktijk: progressie in professionalisering** (proefschrift). Meppel: Krips Repro 1991.
- Zaat J. **De macht der gewoonte** (proefschrift). Amsterdam: Thesis Publishers 1991.
- Schellevis F. **Chronic diseases in general practice. Comorbidity and quality of care**. (Thesis). Utrecht: Pascal, 1993.
- Tijhuis MAR. **Social networks and health** (Thesis). Utrecht: NIVEL, 1994.
- Delnoy DMJ. **Physician payment systems and cost control** (Thesis). Utrecht: NIVEL, 1994.
- Van Suijlekom-Smit LWA., Crone-Kraaijeveld E. **Het zieke kind, een zorg voor huisarts en kinderarts** (Proefschrift). Erasmus Universiteit Rotterdam 1994
- Van der Wouden JC. **Taakopvattingsonderzoek bij huisartsen** (proefschrift). Rotterdam: EUR 1993.

2. BOOKS AND REPORTS

- Foets M., Van der Velden J., Van der Zee J. **Een Nationale Studie van Ziekten en Verrichtingen in de Huisartspraktijk. Projectvoorstel**. Utrecht: NIVEL 1986.
- Foets M., Van der Velden J., Van der Zee J. **Morbidity and interventions in general practice. A cross-national survey in the Netherlands**. Utrecht: NIVEL, 1986.
- Foets M., Van der Velden J. **Een Nationale Studie naar Ziekten en Verrichtingen in de Huisartspraktijk. Basisrapport: Meetinstrumenten en procedures**. Utrecht: NIVEL 1990.
- Van Dijk L., Foets M., Doornbos G. **Voorschrijven van geneesmiddelen in de huisartspraktijk: de ontwikkeling van een kwaliteitsmaat**. Utrecht: NIVEL 1991.
- Foets M., Sixma H. **Een Nationale Studie naar Ziekten en Verrichtingen in de Huisartspraktijk. Basisrapport: Gezondheid en gezondheidsgedrag in de praktijkpopulatie**. Utrecht: NIVEL 1991.
- Tijhuis MAR., Verhaak PFM., Wennink J. **Psychosociale problematiek in de huisartspraktijk**. Utrecht: NIVEL 1991.
- Uunk W., Dekker J., Groenewegen PP. **Verwijzingen van huisartsen naar fysiotherapeuten: morbiditeitsspecifieke verwijskansen**. Utrecht: NIVEL 1991.
- Van der Velden J., De Bakker DH., Claessens A., Schellevis FG. **Een Nationale Studie naar Ziekten en Verrichtingen in de Huisartspraktijk. Basisrapport: Morbiditeit in de huisartspraktijk**. Utrecht: NIVEL 1991.
- Verhaak PFM., Bosman J., Van der Velden J. **Preventie**. Utrecht: NIVEL, 1991.
- Van der Velden J., De Bakker DH., Claessens AAMC., Schellevis FG. **Dutch National Survey of General Practice. Morbidity in General Practice**. Utrecht: NIVEL, 1992.
- Foets M., Van der Velden J., De Bakker DH. **Dutch National Survey of General Practice. A summary of the survey design**. Utrecht: NIVEL, 1992.
- De Bakker DH., Claessens AAMC., Van der Velden J. **Man-vrouw verschillen in gezondheid en medische consumptie. Nationale Studie van Ziekten en Verrichtingen in de Huisartspraktijk**. Utrecht: NIVEL 1992.
- De Bakker DH., Sixma H, Van der Velden J. **Ziekten en Verrichtingen in de Almeerse Huisartspraktijk. Morbiditeit in de huisartspraktijk**. Utrecht: NIVEL 1992.

APPENDIX

- Donker G., Hutten J., Bijl D., Flierman H., Van der Velden J. **Hoofdbrekens en kopzorgen**. Utrecht: NIVEL 1992.
- Foets M., Stokx L., Hutten J., Sixma H. **Een Nationale Studie naar Ziekten en Verrichtingen in de Huisartspraktijk. Basisrapport: De huisartsenquête: datareductie door schaalconstructie**. Utrecht: NIVEL 1992.
- Gloerich A., Van der Zee J. **Determinanten van verwijzen door de huisarts naar de tweede lijn**. Utrecht: NIVEL 1992.
- Groenewegen PP., De Bakker DH., Van der Velden J. **Een Nationale Studie naar Ziekten en Verrichtingen in de Huisartspraktijk. Basisrapport: Verrichtingen in de huisartspraktijk**. Utrecht: NIVEL 1992.
- Sixma H., Thomassen C., De Bakker DH., Foets M. **Ziekten en Verrichtingen in de Almeerse Huisartspraktijk. Gezondheid en gezondheidsgedrag in de praktijkpopulatie**. Utrecht: NIVEL 1992.
- Sixma H., De Bakker DH., Groenewegen PP., Van der Velden J. **Ziekten en Verrichtingen in de Almeerse Huisartspraktijk. Verrichtingen in de huisartspraktijk**. Utrecht: NIVEL 1992.
- Stokx L., Foets M., De Bakker DH., Flierman H. **Voorschrijven van geneesmiddelen in de huisartspraktijk I**. Utrecht: NIVEL 1992.
- De Waal M., Donker G, Van der Velden J. **Spijsverteringsziekten onder de bevolking en in de huisartspraktijk**. Utrecht/Breukelen: NIVEL/NLDS 1992.
- Bruijnzeels MA., Van Suijlekom-Smit LWA., Van der Velden J., Van der Wouden JC. **Het kind bij de huisarts**. Rotterdam/Utrecht EUR/NIVEL 1993
- Sixma H., De Bakker DH., Foets M., Van der Velden J. **Ziekten en Verrichtingen in de Almeerse huisartspraktijk. Een samenvattend rapport**. Utrecht NIVEL 1993
- Sixma H., De Bakker DH., Van der Bent J. **Het experiment Gezondheidszorg Almere**. Utrecht: NIVEL/NZI 1993
- Van der Velden J., Hutten JBF., Bijl D., Brandenburg BJ., Hofstra M., Grol RPTM. **NHG-standaarden en kwaliteit van zorg in de huisartspraktijk**. Eindrapport. Utrecht/Nijmegen: NHG/NIVEL/WOK 1993.
- Van der Velden J., Abrahamse HPhH., De Bakker DH. **Een Nationale Studie naar Ziekten en Verrichtingen in de huisartspraktijk. Arbeid en gezondheid**. Utrecht: NIVEL 1993.
- Verheij R., De Bakker DH., Van der Velden J. **De huisarts in de grote stad**. Utrecht: NIVEL 1993.
- De Bakker DH., Van der Velden J., Abrahamse HPhH. **Huisartsgeneeskundige zorg buiten kantooruren**. Utrecht: NIVEL 1994.
- Van den Bos GAM., Van der Velden J., Van Sonsbeek JLA., Nusselder W., Lenior M. **Gezonde Levensverwachting en Chronische ziekten. Vergelijking van databronnen en berekeningen**. Utrecht/Amsterdam/Voorburg: NIVEL/ISG/CBS 1994.
- Kooiker SE. **Leefgewoonten en alledaagse gezondheidsproblemen; een onderzoek met behulp van gezondheidsdagboeken**. Eindrapport. Utrecht: NIVEL 1994.
- Nusselder W., Van den Bos GAM., Lenior M, Van Sonsbeek JLA., Van der Velden J. **Gezonde levensverwachting en het effect van chronische ziekten**. Amsterdam/ Utrecht/Voorburg: ISG/NIVEL/CBS 1994.
- Stokx L., Foets M. **Voorschrijven van geneesmiddelen in de huisartspraktijk II**. Utrecht: NIVEL 1994.
- Verhaak PFM., Van der Pasch M. **Het effect van de communicatiestijl van huisartsen op het beloop van de psychische problematiek van hun patiënten**. Utrecht: NIVEL 1994.

- De Bakker DH., Groenewegen PP. **Kostencijfer huisartsenzorg. Verslag van een onderzoek op verzoek van de Paritaire Werkgroep Huisartsenzorg**. Utrecht: NIVEL 1995.
- Foets M., Stokx LJ. **De ziekte van Parkinson in de huisartspraktijk**. Utrecht: NIVEL 1995.
- Kruidenier H., Buijs P., Van der Velden J. **Sociaal medische begeleiding door de huisarts**. Leiden/Amsterdam/Utrecht: AS/tri/NIA/NIVEL 1995.
- Verhaak P.F.M. **Mental disorder in the community and in general practice**. Aldershot: Averbury 1995.

3. ARTICLES

- Van der Velden J. **Ziekten en Verrichtingen in de Huisartspraktijk: een nationale studie**. *Medisch Contact 1986; 19: 603-6.*
- Van der Velden J. **Een preventiemodel voor de jaren negentig**. *Medisch Contact 1989; 3: 78-82.*
- Van der Zee J., Verhaak PFM. **Preventie: een kwestie van routine?** *Medisch Contact 1989; 3: 75-8.*
- Bensing JM., Berendonk Ph. **Psychosociale problemen in de huisartspraktijk: weten en meten**. *Maandblad Geestelijke Volksgezondheid 1990; 6: 595-618.*
- Groenewegen PP. **Verrichtingen in de huisartspraktijk. Een presentatie van gegevens uit de Nationale Studie**. *Medisch Contact 1990; 45: 608-11.*
- Kerkstra A., Verheij R., Foets M., Van der Velden J. **Behoefte aan nazorg na ontslag uit het ziekenhuis**. *Tijdschrift voor Sociale Gezondheidszorg 1990; 8: 346-54.*
- Kersten TJJMT., Foets M. **De betrokkenheid van huisartsen bij de ziekenhuisbehandeling van hun patiënten**. *Tijdschrift voor Sociale Gezondheidszorg 1990; 1: 36-41.*
- Kluyt I. **Onzekerheid en aanvullende diagnostiek bij enkele huisartsgeneeskundige problemen**. In: D. Kriegsman (red.). *Continuïteit, een leidraad voor de huisartsgeneeskunde. Amsterdam, 1990, 55-62.*
- Nijland A., De Haan J., Van der Velden J., Meyboom-de Jong B. **Sociale en professionele kenmerken van de doktersassistente**. *Huisarts & Wetenschap 1990 (33); 9:350-4.*
- Schellevis F. **Vervolgafspraken van huisartsen met chronische patiënten**. In: D. Kriegsman (red.). *Continuïteit, een leidraad voor de huisarts(geneeskunde), Amsterdam 1990, 80-7.*
- Van der Velden J. **De rol van de huisartspraktijk in Nederland**. *Medisch Contact 1990; 45: 605-8.*
  Van der Zee J. **De Nationale Studie en het overheidsbeleid ten aanzien van de huisarts**. *Medisch Contact 1990; 45: 611-14.*
- Bensing JM., Foets M., Van der Velden J., Van der Zee J. **De Nationale Studie van Ziekten en Verrichtingen in de huisartspraktijk: achtergronden en methoden**. *Huisarts & Wetenschap 1991; 2: 51-61.*
- Bensing JM., De Bakker DH., Van der Velden J. **Health for all by the year 2000. Hoe ziek is de WAO?** *Medisch Contact 1991; 46: 1075-80.*
- Brandenburg BJ., Hutten JBF., Van der Velden J., Grol RPTM. **Acne vulgaris in de huisartspraktijk: een vergelijking met de NHG-standaard**. *Huisarts & Wetenschap 1991(34); 12:535-9.*
- Grol R., Claessens A., Van der Velden J., Heerdink H. **Kwaliteit van zorg bij enkeldistorsie: invoering van een standaard**. *Huisarts & Wetenschap 1991(34); 1: 30-4.*

APPENDIX

- Kluyt I., Zaat JOM., Van der Velden J., Van Eijk JThM., Schellevis FG. **Voor een prikje? Het gebruik van klinisch-chemische, hematologische en serologische bepalingen door de huisarts**. *Huisarts & Wetenschap 1991(34); 2: 67-71.*
- Nijland A., Groenier K., Meyboom-de Jong B., De Haan J., Van der Velden J. **Determinanten van het delegeren van (medisch-technische) taken aan de praktijkassistente**. *Huisarts & Wetenschap 1991(34); 11: 484-7.*
- Nijland A., Meyboom-de Jong B., Groenier KH., Van der Velden J., De Haan J. **Zelfstandige verrichtingen van de assistente in de huisartspraktijk**. *Huisarts & Wetenschap 1991(34); 13: 587-92.*
- Van der Velden J. **Sociaal-economische gezondheidsverschillen en de huisartspraktijk. Wordt de ongelijkheid gelijk getrokken?** *Tijdschrift voor Sociale Gezondheidszorg 1991(69); 1:22-3.*
- Van der Velden J. **Organisatie en bewaking van zorg voor chronisch zieken**. *Tijdschrift voor Sociale Gezondheidszorg 1991(69); 7:252-6.*
- Verhaak PFM., Bosman JM., Foets M., Van der Velden J. **Psychosocial complaints in general practice. A national survey in the Netherlands**. In: Cooper, B et al. *Primary Health Care and psychiatric epidemiology. London: Routledge, 1991.*
- Bensing JM., Van de Brink-Muinen A., De Bakker DH. **Huisarts (m/v): het kleine verschil met de grote gevolgen**. *Medisch Contact 1992(47); 30:879-83.*
- Bohnen A., Bruijnzeels MA., Van der Velden J., Van der Wouden JC. **Otitis media acuta: incidentie en beleid**. *Huisarts & Wetenschap 1992(35); 5:134-6.*
- Foets M. **Gezondheidsmeting in de nationale studie van ziekten en verrichtingen in de huisartspraktijk**. In: Gunning-Schepers LJ. en Mootz M. (red.) *Gezondheidsmeting. Reeks gezondheidsbeleid. Houten: Bohn Stafleu Van Loghum 1992; 87-98.*
- Hart HE., Bruijnzeels MA., Van der Wouden JC., Van Suijlekom-Smit LWA., Van der Does E., Van der Velden J. **Het kind met koorts in de huisartspraktijk**. *Huisarts & Wetenschap 1992 35: 246-8.*
- Kersten TJJMT. **De invloed van de huisarts op de specialist**. *Tijdschrift voor Sociale Gezondheidszorg 1992(70); 3:171-9.*
- Kersten TJJMT, Kreveld D., Van der Zee J. **Contacten tussen huisarts en patiënt in verband met een ziekenhuisopname**. *Huisarts & Wetenschap 1992; 35: 109-12.*
- Kersten TJJMT. **Overleg door huisartsen met specialisten over de ziekenhuisopname van patiënten**. *Nederlands Tijdschrift voor Geneeskunde 1992(136); 23: 1102-6.*
- Kluyt I., Zaat JOM., Van der Velden J., Van Eijk JThM. **Huisarts en beeldvormende diagnostiek**. *Huisarts & Wetenschap 1992(35) 5: 188-91.*
- Peters L., Van der Kooij T. **Huisarts en placebo's. Gegevens uit de Nationale Studie**. *Medisch Contact 1992; 27/28: 850-3.*
- Tijhuis MAR., Foets M., Groenewegen P., Flap H. **Netwerken in Nederland. Een onderzoek naar persoonlijke netwerken van Nederlanders**. *Mens & Maatschappij 1992; 1: 5-22.*
- Tijhuis MAR., Foets M., Groenewegen P., Flap H. **Netwerken in Nederland. Een onderzoek naar persoonlijke netwerken van Nederlanders**. In: Jansen W., Wittenboer GLH. van de (red). *Sociale netwerken en hun invloed. Meppel: Boom 1992 17-38.*
- Uunk W., Dekker J., Groenewegen PP. **Verwijzingen van huisartsen naar fysiotherapeuten: een verklaring en analyse van verschillen tussen huisartsen**. *Mens & Maatschappij 1992(67):389-411.*

APPENDIX

- Van der Ven-Daane I., Van der Ven M., Bruijnzeels MA., Van der Wouden JC., Van Suijlekom-Smit LWA., Van der Does E. **Het voorschrijven van antibiotica aan kinderen in de huisartspraktijk**. *Huisarts & Wetenschap 1992 (35): 8: 272-4.*
- Verburg M.E., Bruijnzeels MA., Van der Wouden JC., Van Suijlekom-Smit LWA, Van der Velden J., Hoes AW., Offringa M. **Incidence of febrile seizures in the Netherlands**. *Neuroepidemiology 1992; 11: 169-172.*
- Zaat J., Schellevis FG., Kluyt I., Van Eijk JThM., Van der Velden J. **Laboratoriumonderzoek bij de klacht moeheid in de huisartspraktijk**. *Huisarts & Wetenschap 1992 (35); 5: 183-7.*
- Bransen E., Van der Velden J. **Het premenstrueel syndroom in de Nederlandse huisartspraktijk**. *Huisarts & Wetenschap 1993(36); 10:322-7.*
- Van Campen L., Bruijnzeels MA., Uijen H., Van der Velden J., Van der Wouden JC. **Astma bij kinderen: diagnostiek en beleid in de huisartspraktijk**. In: Velden J. van der et al. *NHG-standaarden en de kwaliteit van zorg in de huisartspraktijk. Utrecht/Nijmegen: NHG/NIVEL/WOK 1993.*
- Claessens AAMC., Van der Velden J. **Mammografie aanvragen in de huisartspraktijk**. In: Velden J van der et al. *NHG-standaarden en de kwaliteit van zorg in de huisartspraktijk. Utrecht/Nijmegen NHG/NIVEL/WOK 1993.*
- Foets M., Stokx L. **Het voorschrijven van geneesmiddelen in de huisartspraktijk**. *Medisch Contact 1993(48);16:489-92.*
- Foets M., Elie I., Van der Rijt G. **Medische berichtgeving in de media. De invloed op het consultatiegedrag van huisartsen**. *Medisch Contact 1993(48); 20 617-9.*
- Foets M., Visser J. **Het voorschrijven van homeopathische middelen in de Nederlandse huisartspraktijk**. *Medisch Contact 1993(48); 22:683-7.*
- Van Hall EV., Kerdel M., Van der Velden J. **Medicalisering van de overgang**. *Medisch Contact 1993(48); 6:179-81.*
- Hart HE., Bruijnzeels MA., Van der Wouden JC., Van Suijlekom-Smit LWA., Van der Does E., Van der Velden J. **Het kind met koorts in de huisartspraktijk**. *Medisch Journaal Delft: 1993;5:30-36.*
- Hingstman L., Foets M., Riteco JA. **Thuis of in het ziekenhuis bevallen**. *Tijdschrift voor Verloskundigen 1993; 2:66-73.*
- Hofstra M., Hutten JBF., Kaiser V., Stoffers HJEH., Grol R., Van der Velden J. **Perifeer arterieel vaatlijden in de huisartspraktijk** In: Velden J. van der et al. *NHG-standaarden en kwaliteit van zorg in de huisartspraktijk. Utrecht/Nijmegen: NHG/NIVEL/WOK 1993.*
- Hutten JBF., Van der Velden J., Bijl D., Winkens R., Grol RPTM. **Diagnostiek en behandeling van ongecompliceerde urineweginfecties in de huisartspraktijk**. In: Velden J. van der et al. *NHG-standaarden en de kwaliteit van zorg in de huisartspraktijk. Utrecht/Nijmegen NHG/NIVEL/WOK 1993.*
- Konings GPJM., Van der Velden J., Hutten JBF., Grol RPTM. **Het controle- en therapiebeleid van de huisarts bij type II diabetes voor het verschijnen van de NHG-standaard: de nationale studie als voormeting**. In: Velden J. van der et al. *NHG-standaarden en de kwaliteit van zorg in de huisartspraktijk Utrecht/Nijmegen: NHG/NIVEL/WOK 1993.*
- De Melker RA., Kuyvenhoven MM., Van der Velden J. **Voorschrijf- en verwijsgedrag van huisartsen bij bovenste luchtweginfecties**. *Huisarts & Wetenschap 1993(36); 1:3-5.*
- Rooijakkers-Lemmens M., Van de Lisdonk E., Van de Hoogen H., Van der Velden J. **Otitis externa: klachten en medicamenteus beleid**. *Huisarts & Wetenschap 1993(36); 8:221-3.*
- Tasche MJA., Bruijnzeels MA., Van der Poel BNM., Van Suijlekom-Smit LWA., Van der Wouden JC. **Zuigelingen die vaak huilen: frequentie en aanpak van dit pro-**

bleem in de huisartspraktijk. *Nederlands Tijdschrift voor Geneeskunde 1993;137 38: 1927-30.*

- Terluin B., De Bakker DH., Verhaak PFM. **Het gebruik van de diagnosen surmenage en depressie in de huisartspraktijk**. *Huisarts & Wetenschap 1993; 10:312-21.*
- Van der Velden J., Hutten JBF., Grol RPTM. **Verhoogde bloeddruk en hypertensie in de huisartspraktijk. Een eerste verkenning uit de Nationale Studie**. In: Velden J. van der et al. *NHG-standaarden en kwaliteit van zorg in de huisartspraktijk. Utrecht/Nijmegen: NHG/NIVEL/WOK 1993.*
- Van der Velden J., Van den Bos GAM., Schellevis FG., Van Ammers E. **Comorbiditeit en chronische ziekten**. In: Ruwaard D., Kramers P. eds. *Volksgezondheid Toekomst Verkenning. Bilthoven: RIVM 1993 pp 146-51.*
- Van der Ven-Daane I., Bruijnzeels MA., Van der Wouden JC., Van Suijlekom-Smit LWA. **Het voorkomen en de behandeling van impetigo vulgaris bij kinderen in de huisartspraktijk**. *Huisarts & Wetenschap 1993 (36): 291-3.*
- Van der Ven M., Van der Ven-Daane I., Bruijnzeels MA., Van der Wouden JC., Van Suijlekom-Smit LWA., Van der Does E. **De antimicrobiele behandeling van tonsillitis acuta bij kinderen in de huisartspraktijk**. *Huisarts & Wetenschap 1993 (36); 4: 134-6.*
- De Veer A., De Bakker DH. **Praktijkvorm en het handelen van de huisarts**. *Tijdschrift voor Sociale Gezondheidszorg 1993 (71); 5: 279-87.*
- Verhaak PFM., Bijl RV., Van der Bergh Jeths A., Harteloh P. **Psychische problematiek**. In: Ruwaard D, Kramers P (eds). *VTV RIVM 1993 pp. 166-175.*
- Verheij R. **Kloof tussen huisartsen in stad en op platteland**. *Geografie 1993(2); 4:19-20.*
- Bensing JM., Verhaak PFM. **Psychische problemen in de huisartspraktijk veelvormiger en diffuser dan in de psychiatrie**. *Nederlands Tijdschrift voor Geneeskunde 1994(138); 5:130-5.*
- Bijl D., Hutten JBF., Van der Velden J., Grol RPTM. **Hoofdpijn, migraine en spanningshoofdpijn in de huisartspraktijk**. *Huisarts & Wetenschap 1994(37); 5:142-148.*
- Bijl D. **Migraine en andere hoofdpijn in de huisartspraktijk**. *Hoofdzaken 1994; 2:10-11 3:8-10.*
- Delnoy D., De Bakker DH., Groenewegen PP. **Loon naar werken. Vergelijking van een gedifferentieerd abonnement en een gemengde honoreringsstructuur voor huisartsen**. *Medisch Contact 1994(49) 36:1109-11.*
- Foets M., Schellevis FG., Van Dijk L. **Competentie van huisartsen inzake het voorschrijven van geneesmiddelen. Een onderzoek op basis van papieren patiënten**. *Tijdschrift voor Sociale Gezondheidszorg 1994; 1:32-7.*
- Groenewegen PP., Hutten JBF. **De werkbelasting van de Nederlandse huisarts. Resultaten van de Nationale Studie naar ziekten en verrichtingen in de huisartspraktijk**. *Medisch Contact 1994(49) 2:55-60.*
- Van Hall EV., Verdel M., Van der Velden J. **'Perimenopausal' complaints in women and men: A comparative study**. *Journal of Women's Health 1994 (3); 1: 45-49.*
- Koning S., Bruijnzeels MA., Van Suijlekom-Smit LWA., Van der Wouden JC. **Wratten: incidentie en beleid in de huisartspraktijk**. *Huisarts & Wetenschap 1994(37); 10:431-5.*
- De Melker R., Kuyvenhoven MM. **Management of upper respiratory tract infections in Dutch family practice**. *The Journal of Family Practice 1994 (38); 4: 353-357.*
- De Melker RA., Van der Velden J., Kuyvenhoven MM. **Huisbezoek bij luchtweginfecties: gezinsgeneeskunde pur sang?** *Huisarts & Wetenschap 1994(37); 12:521-6.*

- Pater BW., Van der Velden J. **De output van het HIS: ondersteuning door het Huisartsen Informatie Systeem van patiëntenvoorlichting ter verbetering van de therapietrouw**. In: Sollet PCGM. De huisarts en het elektronische medische dossier. 'Op voorschrift'. *Symposiumverslag EMD 1994. Vakgroep Medische Informatica EUR Rotterdam.*
- Schellevis FG., Van Eijk JThM., Van de Lisdonk E., Van der Velden J., Van Weel C. **Richtlijnen en de kwaliteit van de huisartsgeneeskundige zorg**. *Kwaliteit & Zorg 1994(2); 1:3-12.*
- Schellevis FG. **Thuis bij patiënten met een chronische ziekte**. *Huisarts & Wetenschap 1994(49); 12:527-31.*
- Sixma HJM., Spreeuwenberg PMM., Van der Pasch MAA. **Tevredenheid met de huisarts; een vergelijking van de verschillende methoden van data-analyse**. *Kwaliteit & Zorg 1994(2); 2:68-77.*
- Tijhuis M., Flap H., Foets M., Groenewegen PP. **Kenmerken van sociale netwerken en ziekte**. *Tijdschrift voor Sociale Gezondheidszorg 1994(72); 7:341-51.*
- Bakker RHC., De Bakker DH., Kerkstra A., Stokx LJ. **De rol van de huisarts bij psychische en sociale problemen in het verzorgingstehuis**. *Tijdschrift voor Gerontologie Geriatrie 1995; 26:71-9.*
- Suijlekom-Smit LWA van, Bruijnzeels MA, Wouden JC van der, Velden J van der, Visser HKA, Dokter HJ. **Voor welke gezondheidsproblemen bij kinderen wordt de huisarts geconsulteerd en hoe vaak?** *Nederlands Tijdschrift voor Geneeskunde 1995(139); 33:1684-9.*
- Van der Velden J., Abrahamse HPhH., De Bakker DH. **Arbeid en gezondheid gezien vanuit de huisartspraktijk**. *Tijdschrift voor Sociale Gezondheidszorg 1995(73); 5:300-7.*
- De Waal MWM., Schellevis FG., Van der Velden J. **Consultatie van de huisarts door patiënten met reumatoide artritis of gonartrose/coxartrose**. *Nederlands Tijdschrift voor Geneeskunde 1995(139); 26:1331-5.*
- Bensing JM, Schreurs K. **Sekseverschillen bij moeheid**. *Huisarts Wet. 1995 (38); 9: 412-21.*

4. OTHER PUBLICATIONS *(including data input from National Survey)*

- Koopmanschap M.A., Van Roijen L., Bonneux L. **Kosten van Ziekten in Nederland**. Rotterdam IMTA/EUR 1991.
- Meeuwesen L., Ruiter M., Westerbeek van Eerten J., Brinkgreve C. **Vrouwen en gezondheid**. 's-Gravenhage: STEO 1991.
- Stuurgroep Toekomst Scenario's Gezondheidszorg (STG). **Chronische Ziekten in het jaar 2005. Deel 3**. Reumatoïde artritis. Houten/Antwerpen: Bohn, Stafleu & Van Loghum 1991.
- Stuurgroep Toekomst Scenario's Gezondheidszorg. **Toekomstscenario's voor eerstelijnszorg en thuiszorg DI 2-3**. Houten/Zaventem: Bohn Stafleu & van Loghum 1991.
- Sociaal Cultureel Planbureau. **Sociaal en Cultureel Rapport 1992**. Rijswijk: SCP 1992.
  Stuurgroep Toekomst Scenario's Gezondheidszorg (STG). **Chronische Ziekten in het jaar 2005. Deel 4**. Scenario's voor beleid. Houten/Antwerpen: Bohn, Stafleu & van Loghum 1992.
- Stokx L., De Bakker DH., Delnoij D., Gloerich A., Groenewegen PP. **Verwijscijfers belicht**. Utrecht: NIVEL 1992.

APPENDIX

- Van de Kwartel AJJ., Van Dijk JK., Van der Meulen LJR., De Bakker DH., Stokx LJ., Delnoy D. **Curatieve zorg in Nederland. Feiten, ontwikkelingen, knelpunten**. In: Commissie modernisering curatieve zorg. '*Gedeelde zorg: betere zorg'. Zoetermeer, Hageman 1994.*
- Ten Doeschate L., Van der Velden J. **Melkveehouderij en Gezondheid**. Doetinchem: Misset 1993.
- Van Roij L., Michel BC., Koopmanschap MA., Wielink G., Rutten FFH. **Kosten van migraine in Nederland. Deel I en II**. Roterdam: IMTA/EUR 1993.
- Ruwaard D., Kramers P. (eds). **Volksgezondheids toekomstverkenning 1990-2005**. Bilthoven: RIVM 1993.
- Schrameijer F., Brunenberg W. **Psychosociale zorg bij kanker**. Utrecht: NcGv 1993.
- De Bakker DH., Delnoy DMJ., Groenewegen PP. In: Commissie modernisering curatieve zorg. **Een uniform gedifferentieerd honorarium voor huisartsenzorg**. *'Gedeelde zorg: betere zorg'. Zoetermeer: Hageman 1994.*
- Sociaal Cultureel Planbureau. **Sociaal en Cultureel Rapport 1994**. Rijswijk: SCP 1994.
- Miedema H.S. **Reuma-onderzoek meerdere echelons (Rome). Basisrapport**. Leiden: NIPG/TNO 1994.
- Nationale Raad voor de Volksgezondheid/Overlegorgaan Gezondheidszorg en Multiculturele Samenleving. **Gezondheidsbeleid voor migranten. Advies over de Nederlandse Gezondheidszorg in multicultureel perspectief**. Zoetermeer: NRV 1995.