Is questionnaire information valid in the study of a chronic disease such as diabetes? The Nord-Trøndelag diabetes study

Kristian Midthjell, Jostein Holmen, Arild Bjørndal, Per G Lund-Larsen

Abstract

Study objective—The aim was to validate information about diabetes mellitus collected by questionnaire in a large epidemiological survey.

Design—Questions on diabetes diagnosis, medical treatment for diabetes, diabetes duration, and hypertension treatment were selected from the Nord-Trøndelag health survey questionnaires. One of the municipalities was selected for the validation study.

Setting—The health survey 1984–86 addressed all inhabitants ≥20 years of age in Nord-Trøndelag county, Norway; 76 885 (90.3%) of the eligible population participated in answering the question on diabetes.

Participants—All inhabitants in the municipality answering “yes” to the question on diabetes (n = 169) and the persons with the same sex born closest before and after each diabetic patient and answering “no” to the diabetes question (n = 338) were included.

Measurements and main results—A very thorough search was made in the medical files of the general practitioners in the municipality for corresponding information. Compared to the files, diabetes was verified in 163 out of the 169. The commonest cause of discrepancy was renal glycosuria. One out of the 338 registered non-diabetic persons was found to have diabetes. Diabetic patients tended to overestimate diabetes duration significantly. Insulin treatment was verified in 1920 (95%) and treatment with oral hypoglycaemic agents in all 44 with an affirmative questionnaire answer. A negative answer on insulin and oral hypoglycaemic agents was verified in 100% and 99% respectively.

Conclusions—The concordance was considerably higher than in a comparable Norwegian study performed 10 years earlier. Patient administered questionnaires may be a very reliable source of information for epidemiological purposes in a well defined chronic disease such as diabetes mellitus.

J Epidemiol Community Health 1992; 46: 537–542

Patient administered questionnaires are often applied to establish the prevalence of chronic conditions like diabetes mellitus. Epidemiological publications based on questionnaire or interview data often apply advanced statistical methods, while the quality of the “raw data” is often not validated. Most often it is taken for granted that a “yes” or “no” to a question on disease is valid.

Several errors may, however, threaten the validity of epidemiological data based on patient administered information. Random errors may be introduced in the data handling process, such as punching errors, or the patient may put the mark in the wrong box or misunderstand the question. Systematic errors may occur due to denial of the diagnosis; alternatively the opposite may occur: the person may want the benefits of care associated with the patient role. Diabetes mellitus is a disease with relatively clear diagnostic criteria and a diagnosis that is easily communicated to the patient. This should make it easy to establish the validity of a questionnaire answer on the diabetes diagnosis, and it may be the reason why very few investigators have addressed the validity of questionnaire information about diabetes. An earlier study has, nevertheless, shown verification from medical records of a positive answer for diabetes in only 66%. The Nord-Trøndelag diabetes study aims at evaluating diabetes health care in the county of Nord-Trøndelag, Norway. Diabetes prevalence in the adult population in the county has been established through questionnaire and screening in the Nord-Trøndelag health survey (5, unpublished observations). The screening also comprised studies on hypertension, lung diseases, and quality of life.

The aim of this part of the Nord-Trøndelag diabetes study was to validate patient administered questionnaire information on diabetes diagnosis, types of diabetes treatment, diabetes duration, and medical treatment of hypertension in diabetic patients.

Methods

THE NORD-TrøNDELAG HEALTH SURVEY

Nord-Trøndelag is a county located in the middle part of Norway. Its population is fairly representative of Norway. The total population by January 1 1984 was 127 051. The survey addressed all 87 259 inhabitants of Nord-Trøndelag 20 years of age and older by January 1 1984, and took place between January 5 1984 and February 15 1986. All subjects were registered by their national personal identification number. This number contains birth date and code for sex and is unique for each Norwegian citizen. Of the subjects identified 2159 had moved out of the county or had died between the inclusion date and the screening, leaving a survey population of 85 100 (fig 1). A detailed description of the Nord-Trøndelag health survey is given elsewhere.

National Institute for Public Health, Community Medicine Research Centre, Box 234, N-7651 Verdal, Norway
K Midthjell
J Holmen
National Institute for Public Health, Department of Health Services Research, Oslo, Norway
A Bjørndal
National Health Screening Service, Oslo, Norway
P G Lund-Larsen
Correspondence to:
Dr Midthjell
Accepted for publication March 1992
QUESTIONNAIRES

Questionnaire 1 was mailed to the participant along with the invitation to the survey. It included the following questions:

- “Do you have or have you had any of the following diseases?”
  - Diabetes (Yes/No)”,
- “Are you taking or have you taken medicine for high blood pressure (Yes/No)”.

Known diabetes was defined by “Yes” to the question on diabetes. The National Health Screening Service was responsible for the clinical examination and the handling of the questionnaires. The participant was asked to bring the completed questionnaire to the examination. A few who could not attend the clinical examination mailed the questionnaire, and if answering the crucial question, were also included in the calculation of the prevalence of known diabetes.

Questionnaire 2 was handed out together with a prepaid envelope when participants left after the clinical examination, and could be mailed to the Health Screening Service free of charge. One reminder was sent if the participant did not answer. Questionnaire 2 for diabetic patients included among others the following questions:

- “When was your diabetes diagnosed?”
- “Do you take insulin injections for your diabetes?” (Yes/No)
- “Do you take tablets for your diabetes?” (Yes/No)

Those who confirmed treatment for high blood pressure in questionnaire 1 were asked to answer the question:

- “Are you taking medicine for blood pressure at the moment?” (Yes/No)

The complete questionnaires, both in original layout and in English translation are presented in a report on the survey methods.5

HANDLING OF QUESTIONNAIRE DATA

Questionnaire 2 was made in two versions: type A included 17 items on hypertension and 41 questions on known diabetes; type B did not include the items on hypertension and diabetes and was handed out to those who confirmed neither of these conditions in questionnaire 1. To ensure delivery of the correct version of the second questionnaire, questionnaire 1 was checked at the screening site by one of the screening team nurses. If the patient had answered “Yes” for diabetes, but “No” for hypertension, the page with items on hypertension was crossed out and vice versa. The nurses were instructed not to query a yes or no to these questions.

In a very few cases the patient had forgotten to answer the diabetes question, and was then asked by the nurse. The answer was then filled in on the first questionnaire. The screening nurses were thoroughly instructed for their job in a written protocol and in courses before start of the screening. Both questionnaires were punched by the data department of the National Health Screening Service, which has a vast experience in the handling of such data, and syntax programs were applied to discover punching errors causing invalid values. The punching operators are constantly exposed to surveillance and quality control. Frequency tables were made as the punching was finished, and illegal codes and extremes in measurement results were controlled and if necessary compared to the original questionnaires and corrected.

ATTENDANCE

Questionnaire 1 was answered by 76 885 persons (90.3%). The attendance differed, however, considerably between age groups: from 95% for women 60–64 years old to 50% in men ≥ 95 years old and 68% in men 20–24 years old. The clinical examination was attended by 74 977 persons. Questionnaire 2 was returned by mail by 85% of those participating in the clinical examination. The attendance in the different parts of the survey by age and sex is illustrated in fig 1. Those not attending are described in previous reports.6
Validity of questionnaire information in chronic diseases

POPULATION FOR THE VALIDATION STUDY

For practical reasons one of the municipalities, Verdal, was chosen for validation of the questionnaire information. By January 1st 1984 Verdal had 13,140 inhabitants. Verdal is fairly representative for Nord-Trøndelag county in demographic variables except for a greater proportion of inhabitants ≤12 years old and 30–39 years old and a lower proportion ≥70 years old (81.1% vs. 10.3%).

The prevalence of known diabetes in persons over 20 years old in Verdal was 2.3% compared to 2.9% for the total county, mainly due to the somewhat different age composition of the populations. Of known diabetic patients in Verdal, 14% were on insulin treatment, 33% on oral hypoglycaemic agents, and 53% on diet only, versus 20%, 40%, and 40% respectively for Nord-Trøndelag. During the screening period the medical profession in the municipality consisted of 11 physicians. Seven of these were attached to a health centre, with common medical files. Four physicians had single practices with their own files. Two of these were in the pensionable age, a third was engaged in general practice to a very restricted extent, attached to a practice in occupational medicine.

In the screening, 169 persons were identified as known diabetic patients according to “Yes” on diabetes in questionnaire 1. Among those attending 338 individuals were selected on the grounds that they answered “No” to the diabetes question, lived in Verdal, were of the same sex, and were registered as born closest before and after the corresponding diabetic patient in the electoral register.

A search was made for these 507 individuals in the files of the general practitioners in Verdal. The participants had given the name of their doctor on the questionnaire. The examination of the files was performed during the autumn of 1989. The files of those deceased after the screening were also examined. The date of examination by the screening team was recorded, and the doctors’ files were searched for the following items registered closest prior to this date:

- known diabetes
- if identified as a diabetic patient: type of diabetes treatment (insulin, oral antidiabetic agents, or diet only), and year of diabetes diagnosis
- medical treatment for hypertension

If the participant was not found in any of the doctors’ records in the municipality, he or she was defined as not having diabetes. One of the participants had given the name of a non-resident doctor, working in an outpatient clinic at the local hospital in the adjacent municipality. In this case the record in the outpatient clinic was also examined. Some of those answering “No” in relation to diabetes were obviously in good health and had never visited a doctor, or if so only for minor acute illness.

In case of a discrepancy between questionnaire and medical record on diabetes diagnosis or type of treatment the file was examined for a probable explanation. If necessary the reason was obtained from the doctor or the patient in person.

STATISTICAL ANALYSIS

Comparison of mean diabetes duration obtained by questionnaire and by medical records was done by t test. Sign test was applied to evaluate whether an overestimation was more frequent than an underestimation. Level of significance was defined as p < 0.05.

CONSENT

Participation in the survey was voluntary. According to Norwegian law consent to register data and include the personal identification number was obtained in advance from the Norwegian Data Inspectorate. Special permission was obtained from the Data Inspectorate and from the National Health Directorate to procure data from the medical records.

Results

DIABETES DIAGNOSIS

Table I illustrates the validation of the question about diabetes. In six cases a diabetes diagnosis in the questionnaire was not confirmed in the medical records. Three of these patients had once been examined because of glycosuria and were found to have renal glycosuria. In two cases (described below) this had happened two and four years earlier, and the individuals thought that they had had diabetes which had passed off. The first (patient No 3) was a man, 38 years of age, who had glycosuria at age 14. A glucose tolerance test was normal, but he was told by the specialist in charge that it would be wise to be careful with sugar. He had wisely not taken much notice, but had never been quite sure that he did not have diabetes. A new tolerance test performed in 1989 was still completely normal. The second (patient No 4) was a man, 63 years of age, who had glycosuria on one occasion four years previously. According to him the doctor had suggested that this might indicate diabetes. Blood glucose was, however, not tested. Renal glycosuria was thus not confirmed by a glucose tolerance test. Three months later the “diabetes” was said to have passed off, again on the basis of urine analysis alone. The patient had since then been careful with sugar. His doctor had made no notes on this incident, and he was not registered as a diabetic patient. In fact a non-fasting blood glucose taken by a new doctor in 1990 was 3.9 mmol/litre.

| Questionnaire | Medical file | Diabetes no diabetes | Total in questionnaires | Diagnosis verified
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>File</td>
<td>163</td>
<td>6</td>
<td>169</td>
</tr>
<tr>
<td></td>
<td>No diabetes</td>
<td>1</td>
<td>337</td>
<td>338</td>
</tr>
<tr>
<td></td>
<td>Total in files</td>
<td>164</td>
<td>343</td>
<td>507</td>
</tr>
</tbody>
</table>

* expressed as percentage of questionnaire answers verified in medical files

Patient No 5 was a man, 69 years of age. At follow up he was found to be in a nursing home suffering from amnesia due to a cerebral haemorrhage, and the reason for the inconsistency in the completion of the questionnaire could not be identified. The reason might be an erroneous mark by the patient or a punching error. Neither of these two persons had filled in questionnaire 2.

Patient No 6 was a woman who moved to the municipality three years before the screening and told her doctor that she had diabetes. She was only on diet, and had always had normal blood glucose
values. Her present doctor thought that the diagnosis was wrong, hence the conclusion from the medical file.

One of the 338 persons who stated that they did not have diabetes was found to have diabetes at follow up. This was obviously a punching error, because she had answered the diabetes version of questionnaire 2, which was only handed out to those answering "Yes" in questionnaire 1. In one case with a "No" answer the record revealed that an ophthalmologist had diagnosed a few small microaneurysms five years previously and had asked the general practitioner to check for diabetes. The record contained no later entries, and this had obviously not been done. The record thus confirmed that no diagnosis was made.

MEDICAL TREATMENT

The validation of the answers on medical treatment is shown in table II. The correspondence here was almost complete. A 59 year old diabetic woman had changed from oral antidiabetics to insulin during a hospital admission one month after the screening date. She probably did not return the questionnaire answering this question until after the change. She thus appears twice in table II.

Table II: Validity of the questions "Do you take insulin injections for your diabetes?" (1), and "Do you take tablets for your diabetes?" (2).

<table>
<thead>
<tr>
<th>Questionnaire</th>
<th>Medical file</th>
<th>Total in questionnaires</th>
<th>Treatment verified</th>
<th>%</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) Insulin</td>
<td>Yes</td>
<td>19</td>
<td>20</td>
<td>95.0</td>
</tr>
<tr>
<td></td>
<td>No</td>
<td>0</td>
<td>119</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>Unanswered</td>
<td>4</td>
<td>26</td>
<td>96.2</td>
</tr>
<tr>
<td></td>
<td>Total in files, insulin</td>
<td>23</td>
<td>146</td>
<td>100.0</td>
</tr>
<tr>
<td>(2) Oral antidiabetic agents</td>
<td>Yes</td>
<td>44</td>
<td>44</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>No oral antidiabetic agents</td>
<td>1</td>
<td>93</td>
<td>98.9</td>
</tr>
<tr>
<td></td>
<td>Unanswered</td>
<td>12</td>
<td>19</td>
<td>100.0</td>
</tr>
<tr>
<td></td>
<td>Total in files, oral antidiabetic agents</td>
<td>57</td>
<td>112</td>
<td>100.0</td>
</tr>
</tbody>
</table>

*Expressed as percentage of questionnaire answers verified in medical files.

Table III illustrates the distribution on treatment groups of those not returning questionnaire 2. On the basis of the figures in table II the distribution between the three treatment groups (insulin, oral antidiabetic agents, and diet only) was compared to the distribution among all known diabetic patients in Nord-Trøndelag (unpublished observations). The percentages from the validation study were calculated on only 31(30) persons, and did not differ significantly from corresponding figures in those answering the questionnaire.

TREATMENT FOR HYPERTENSION

From those confirming having been or being at the moment taking antihypertensive (questionnaire 1), 108 persons returned questionnaire 2 answering the question on present treatment. According to the questionnaire 57 of these had diabetes.

The second answer on antihypertensive treatment (questionnaire 2) was validated. The concordance was almost complete (table IV). Present antihypertensive treatment claimed by the patient could not be confirmed in the medical files in two cases. The reason for discrepancy in these two cases could not be found.

DIABETES DURATION

In 130 of the diabetic patients the year of diagnosis was obtained both from the questionnaire and the medical records. In 71 cases the same year was found. The remaining 59 patients more frequent overestimated (n=41) than underestimated (n=18) diabetes duration (p=0.004). Mean diabetes duration according to medical files was 7.65 years, and according to the questionnaires, 8.05 years. This difference was also significant (p=0.047). Mean age in these patients was 58 (range 26-91). Figure 2 illustrates the distribution of differences between pairs of observations on diabetes duration.

Discussion

KNOWN DIABETES

The prevalence of known diabetes mellitus may be established in many ways. Even in the Nordic countries different approaches are made: questionnaire or interview,7 medical records,8-10 prescriptions registration,11-14 national insurances,15,16 diabetes registries,17 hospital admissions,18-20 or by combinations of these.21-25

Patient administered questionnaires have, for different reasons, obvious advantages; they are easy to administer, convenient for the patient, and require less funding and personnel resources than many of the alternative sources of information.2

A chronic disease such as diabetes most often affects the life of the suffering individual considerably. It therefore seems unreasonable that a patient should answer a question on diabetes wrongly, particularly for type 1 (insulin dependent) diabetes. Great care must nevertheless be taken in the construction of the questionnaire to avoid misunderstandings. Type 2 (non-insulin dependent) diabetic patients often have less pronounced symptoms, have an easier type of treatment, and their mean age is greater. This increases the risk of misunderstandings between patient and doctor concerning the diagnosis of diabetes.

The question often posed is "Do you have or have you had diabetes?" The rationale for this formulation is that some diabetic patients treated only with diet may believe that they no longer have diabetes because they are in no need of medication. This formulation introduces, however, the possibility of erroneously including those with renal glycosuria.

The comparison with the medical records in this study showed very good correlation between the questionnaire information and the medical records. Very few other studies have addressed the validity of positive questionnaire information on chronic diseases in general and on diabetes in particular, and almost none has also validated a negative answer to chronic disease.
The question on diabetes diagnosis applied in Nord-Trøndelag has been used in other screening investigations in Norway performed by the National Health Screening Service.\textsuperscript{26, 27} In a study in the county of Finnmark\textsuperscript{4} the diagnosis was only verified in 29 (66\%) out of 44 with a positive answer to diabetes. Comparison was done with hospital and primary health care records. The most common reason for discrepancy in this study was also renal glycosuria. Only 24 out of 33 (73\%) gave a positive answer to the same question three years later. This rather low reliability might be due to the fact that those with renal glycosuria in the first screening answered “No” three years later, following better explanation by their doctors.

The concordance in Nord-Trøndelag was far better than in Finnmark. The most probable reason for this is a positive change in health education and doctors’ awareness of diabetes related symptoms over the 10 years between these two surveys. Health education in Finnmark was earlier considered to be particularly poor, as illustrated by differences in cardiovascular risk factor levels.\textsuperscript{26} Finnmark is also more sparsely populated and has had a less stable primary health care service than Nord-Trøndelag. Finnmark has a considerable linguistic minority—the Lapps. The Finnmark questionnaires were translated to Lappish, but as this language is very different from Norwegian, different nuances in the meaning of words might increase the risk of misunderstandings. Choosing to perform the validation in a smaller geographical area and therefore being able to make a very thorough search of the files and a personal follow up of the doctors might also have enhanced the concordance in Nord-Trøndelag.

In a nationwide Norwegian interview health survey in 1968 the results for different diseases were validated, giving an agreement of 95\% to a positive diabetes diagnosis.\textsuperscript{28, 29} The doctor’s diagnosis has not been validated in this survey. In most clinical cases the diagnosis of diabetes is obvious, and physician defined diabetes is found to be a useful and reproducible case definition.\textsuperscript{30} The WHO criteria for diagnosis\textsuperscript{31, 32} have gained wide acceptance among Norwegian doctors, and may make medical files a better “gold standard” today than 10 years ago.\textsuperscript{4} Complete blinding of the researcher was not considered necessary as the diabetes diagnosis was nearly always easy to extract from the files. Medical treatment for diabetes was also easy to follow up as this type of treatment is used exclusively for the treatment of diabetes.

**Table IV Validity of the question “Are you taking medicine for blood pressure at the moment?”\textsuperscript{2}**

<table>
<thead>
<tr>
<th>Questionnaire</th>
<th>Medical file</th>
<th>Treatment</th>
<th>No treatment</th>
<th>Total in questionnaires</th>
<th>Treatment verified \textsuperscript{6} (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Treatment</td>
<td>76</td>
<td>2</td>
<td>78</td>
<td>100</td>
<td>97.4</td>
</tr>
<tr>
<td>No treatment</td>
<td>0</td>
<td>30</td>
<td>30</td>
<td>100</td>
<td>100.0</td>
</tr>
<tr>
<td>Total in files</td>
<td>76</td>
<td>32</td>
<td>108</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

\textsuperscript{6} Answered in questionnaire 2 by those answering “Yes” to the question “Are you taking or have you taken medicine for high blood pressure?” in questionnaire 1.

\textsuperscript{2} Expressed as percentage of questionnaire answers verified in medical files.

The types of diabetes treatment

Data on main types of diabetes treatment can be derived by questionnaire/interview, by search of registries/files, or by studying prescriptions, the latter excluding those not on medical treatment. Questionnaire data on diabetes treatment had very good validity in Nord-Trøndelag when compared with medical files. Three out of a total of 10 women in the 20–29 year age group in the entire population of 2242 registered known diabetic patients in Nord-Trøndelag stated that they did not use medication for their diabetes. As nearly all diabetic patients in this age group are type 1 (insulin dependent) this might indicate a frequent misclassification. The reason for this might, however, be gestational diabetes or glycosuria during recent pregnancies registered by the patient as previous diabetes. None of these were residents of Verdal, and were therefore not included in the validation study.

A considerable proportion of the diabetic patients (40\%) in Nord-Trøndelag were on no antidiabetic medication (unpublished observations). This percentage is probably a minimum estimate. It would be expected that there was an overrepresentation of diabetic patients with a “mild diabetes” (ie, without medication) among
the questionnaires. The figures presented in the bottom part of table 11 seem to fit well with this hypothesis, but due to the small numbers the confidence intervals are too wide to allow any definitive conclusions.

**MEDICAL TREATMENT FOR HYPERTENSION**

The sample derived in table IV is not necessarily representative for all hypertensives, as only those answering “Yes” or “No” to the question on present antihypertensive treatment have been validated. The evaluation of the question on treatment for hypertension was also somewhat more difficult because some patients also had angina pectoris and used β blockers suitable for both conditions. A very thorough search of the files including reports from hospital admissions or hospital outpatient visits was performed, and only cases where the diagnosis was confirmed were considered positive. With these limitations, and taking into consideration the small number included, the validity of the question on present antihypertensive treatment also seemed remarkably good.

Validation on hypertension was not included in the study of Tretti et al, and validation of questionnaire information on hypertension does not seem to have been reported earlier. A more extended study on the validity of such a question is recommended.

**DURATION OF DIABETES**

Our study indicated a tendency to overestimate the duration of diabetes. This is a little surprising as important life events often seem more recent than they are. Considering the high mean age of the diabetic patients (68-2 years) they could be inclined to “live in the past” and thus underestimate the duration. This does not, however, seem to be the case. Some had experienced their diabetes as a great inconvenience (in response to the question on coping) and may have felt that they had had the disease for “ages”. Those in error by exactly 10 years may be a misprint by the patient or undiscovered punching errors.

**CONCLUSION**

Our conclusion is that a patient administered, questionnaire is a convenient and valid way of studying a chronic disease such as diabetes. The validity was much better than in a comparable study 10 years earlier. A probable reason for this is better medical awareness by patients, a greater consensus on diagnostic routines, and an increasing willingness of physicians to take responsibility for providing information and education for their patients about chronic diseases.

The study was supported by a grant from the Norwegian Research Council for Science and the Humanities, by The County Council of Nord-Trendelag, The Department for Health and Social Affairs, The National Institute of Public Health, The Norwegian Diabetes Association, and Hoechst Norway Ltd. The National Health Screening Service is acknowledged for supplying their experience and working power in establishing the study population. Our highly skilled secretaries, Mrs. Inger Duvsete Holbe and Mrs Aina Enes, were responsible for the final layout of the manuscript.

2 Burney P, Chinn S. Developing a new questionnaire for measuring the prevalence and distribution of asthma. Chest 1987; 91(suppl): 79-83S.