

# General practice data retrieval: the Northern Ireland project

Leslie Boydell, Hazel Grandidier, Colm Rafferty, Charles McAteer, Philip Reilly

## Abstract

**Objective** – To develop an epidemiological database of morbidity in the community as presented to general practitioners and to support epidemiological research in general practice.

**Design** – The project is a sentinel network of 23 general practices in Northern Ireland that report the incidence of a list of selected diseases to a central unit for analysis.

**Results** – Results are presented for depression, diabetes, and myocardial infarction. The age – sex distributions of diabetes and myocardial infarction are comparable with other published data but the incidence of depression is underestimated.

**Conclusions** – The importance of complete accuracy of data reported within sentinel networks depends on the purposes for which information is to be used. For some diagnoses, such as diabetes and acute myocardial infarction, the accuracy of the reported incidence is high while for other diseases, such as depression, where diagnostic behaviour varies more between doctors, the figures are much less reliable.

(*J Epidemiol Community Health* 1995;49(Suppl 1):22-25)

The general practice data retrieval project (GPDRP) is a sentinel network of general practices distributed across Northern Ireland which continuously record the incidence of selected conditions on their practice computer systems and report these to a central unit. The project was initiated in 1989 by the Department of Public Health Medicine of the Eastern Health and Social Services Board in collaboration with the Department of General Practice at the Queen's University of Belfast with a view to providing data to inform the assessment of health care needs. Between February 1991 and March 1992, a pilot project was undertaken which involved the recruitment and training of 11 practices using a single software system, the *General Practice Administration System for Scotland* (GPASS). Software was written to enable the extraction of anonymous morbidity data from individual practices and its subsequent aggregation and analysis on a central computer located in the Department of General Practice.

The pilot study established the feasibility of this approach to collecting morbidity data from general practice in Northern Ireland, and in April 1992 a decision was made to expand the project to include practices throughout North-

ern Ireland and to achieve a coverage of 10% of the population of Northern Ireland. This paper presents preliminary findings from the project.

## Aims of the project

The GPDRP aimed:

- To develop an epidemiological database for Northern Ireland of morbidity within the community as presented to general practitioners.
- To support epidemiological research in general practice.

## Methods

### RECRUITMENT

Between April and October 1992 additional practices were recruited bringing the total number of practices to 23 and the number of general practitioners involved to 81. The practices were recruited to reflect the age and sex distribution of the Northern Ireland population, to cover both urban and rural communities and different socioeconomic conditions. The total study population during 1993 was 132 975, equivalent to 8.4% of the Northern Ireland population. It closely resembles the Northern Ireland population in terms of age and sex structure with slightly fewer people in the 5 to 20 year age group.

### DATA SET

General practitioners were asked to record the incidence of an agreed list of 33 diseases. These were selected to include both acute and chronic pathology and to reflect the broad range of morbidity in general practice. The conditions were chosen for inclusion with the following criteria in mind:

- The condition should have reasonable frequency;
- The condition should have service or public health implications beyond the surgery consultation;
- The condition should be clearly identifiable clinically;
- General practitioners must consider the condition important and worth recording on their computers;
- Information from other sources should be unavailable or incomplete.

The dataset was agreed with participating doctors during a workshop. In addition, diagnostic guidelines were agreed for some conditions, although for most diseases the doctors'

Department of Public Health Medicine, Eastern Health and Social Services Board, 12/22 Linenhall Street, Belfast BT2 8BS  
L Boydell

Department of General Practice, Queen's University, Belfast  
H Grandidier  
C Rafferty  
C McAteer  
P Reilly

Correspondence to:  
Dr L Boydell.

clinical opinion was considered sufficient. The diagnostic guidelines reflect good practice in relation to the supporting evidence required for certain diagnoses (that is, laboratory and radiological tests, respiratory function tests, etc).

**FACILITATION OF RECORDING**

Several methods are used to promote accurate and complete recording of morbidity data. A brief set of guidelines is provided detailing the diagnostic criteria for certain diseases and advice on when to record a new diagnosis, which may be uncertain for some chronic diseases, such as chronic obstructive airways disease. In addition, strategies are suggested for ensuring that diagnoses made in different settings are recorded, including home visits. Coloured laminated prompt cards listing the conditions of the data set with relevant Read codes (the coding system used within GPASS) are displayed in surgeries, reception, and computer areas.

Support for participating practices is provided by a nurse facilitator whose role is to visit all practices regularly, to provide training and advice to practice staff both on the use of the computer software and the recording of morbidity.

In addition, practices are given regular feedback of their recorded morbidity, compared with the mean of the study and take part in project meetings where results are discussed.

**CONFIDENTIALITY AND SECURITY**

All data are rendered anonymous and encrypted before transfer from the practice on floppy disk. There is also a monitoring committee which has been established to oversee all potential uses of the data.

**VALIDATION**

The main method of validation is the checking of the data for impossible values and internal consistency as well as the comparison of the

data with data from other sources to assess external consistency. A major audit is currently underway within the study to investigate the completeness of transfer of new diagnoses included in the data set from the paper record to the computer file and to examine whether diagnostic criteria have been used where these have been specified and can be identified within the record. A further effort to improve data quality used is to identify and investigate practices which have very low or high recorded incidence rates compared with the mean rate for all the practices.

**Results**

1993 is the first year for which data has been available for all participating practices. The main source of published data used for comparisons is the weekly returns service of the Royal College of General Practitioners Research Unit (1992).<sup>1</sup> For diseases which are not included in this service, the national morbidity survey (1981) data are used for comparison.<sup>2</sup> Results are presented here for depression, diabetes, and acute myocardial infarction (figs 1-3). Because of the limited period of recording and the relatively small numbers of cases recorded, no attempt has been made to standardise the results. This limits the interpretation of comparisons. It is important that these results are seen as preliminary findings and the comparisons made are aimed at consistency checks rather than comparisons of morbidity.

**DEPRESSION**

The age and sex distribution of the recorded incidence of depression shows a higher incidence in females in the 15 to 64 year old age group but a higher incidence in males over 65 years. The incidence ranges from 1.0 per 1000/annum in males aged 15 to 44 years to 2.7 per 1000/annum in males 65 years and over, with the incidence for females showing less variation between the different age groups. There was major variation between practices with five practices recording no cases and the practice with the highest incidence recording 14 cases/1000/annum.

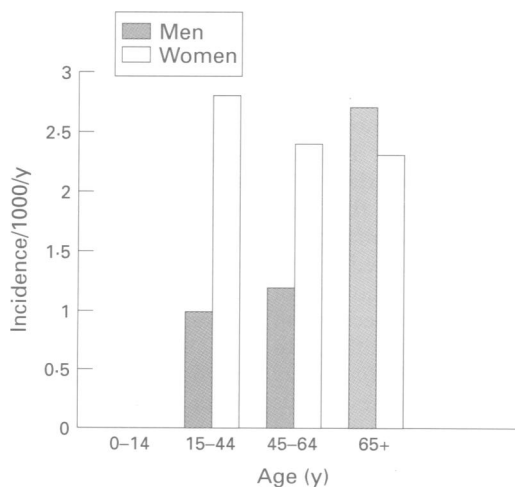


Figure 1 Incidence of depression per 1000 population in relation to age and sex reported by the sentinel network for 1993.

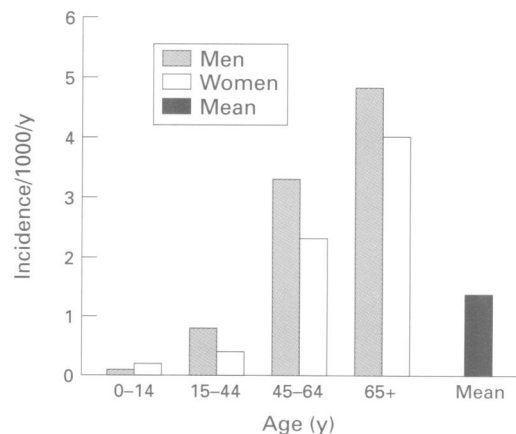


Figure 2 Incidence of newly diagnosed diabetes per 1000 population in relation to age and sex reported by the sentinel network for 1993.

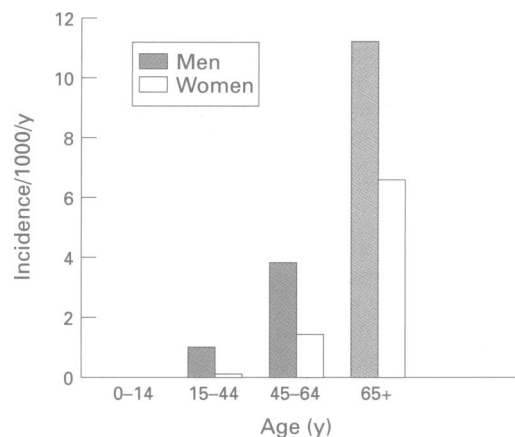


Figure 3 Incidence of acute myocardial infarction per 1000 population in relation to age and sex reported by the sentinel network for 1993.

#### DIABETES (FIG 2)

The incidence of diabetes can be seen to rise with increasing age and be more common in males at all ages, except the youngest age group where the number of cases recorded was very small. The mean incidence is 1.4 per 1000/y.

#### ACUTE MYOCARDIAL INFARCTION (FIG 3)

The incidence of acute myocardial infarction rises steeply with age and reflects the worse experience of males at all ages. The mean incidence rate was 2.00/1000/y.

#### VALIDATION

The results of the audit of the transfer of data from the paper record to the computer and the use of diagnostic criteria is not yet complete but preliminary analysis indicates a wide variation in recording between practices for some diseases.

#### Discussion

The quality of data derived for sentinel networks is particularly difficult to validate since diagnosis of illness in general practice often relies solely on the clinical opinion of the general practitioner. The nature of general practice is that there is often uncertainty around diagnosis and in many cases the general practitioner's task is to rule out significant illness rather than to make a definitive diagnosis.

Most sentinel networks carry out internal consistency checks and compare their data with data published by others. It seems that few networks have undertaken further validation of data.<sup>3</sup> Where differences are observed, these may result from differences in the quality of the data or recording behaviour or because of real differences in morbidity and it is difficult to distinguish between the two.<sup>4</sup> While diagnostic guidelines were used for some diseases in this project, there is doubt whether the existence of these guidelines influences doctors' behaviour.<sup>5</sup> Furthermore, the importance of strict adherence to guidelines depends on the purpose for which the data are being collected. If the main purpose of the sentinel network is to

contribute information for service planning then, as the general practitioner will base decisions regarding intervention on his working diagnosis, it is this which will impact on the service. Similarly, if the network is being used to monitor trends in infectious diseases, it is the trend which is important rather than the absolute value. However, if the aim is to investigate a research hypothesis then diagnostic accuracy is much more critical. Until the results of the validation are available the robustness of the data is uncertain and results must be viewed in this light.

In this paper, results are presented for three diseases. The recording behaviour of doctors with regard to the clinical diagnosis of depression shows great variation, and in some practices no cases were reported. Surveys have shown that one in 10 patients attending the general practitioner can be diagnosed as suffering from depression.<sup>6</sup> The figures presented therefore are clearly an under estimate of the size of the problem in the community. The age and sex distribution of the diagnosis may reflect both the consulting habits of patients and the diagnostic behaviour of doctors rather than any real difference in the incidence of disease. It has been recognised by the Royal College of General Practitioners in England that depression is greatly under diagnosed in general practice and the need to raise awareness of the disease has been addressed in a joint initiative between the Royal College of Psychiatrists and the Royal College of General Practitioners launched in 1993.<sup>7</sup> In addition, the general practitioners within the Northern Ireland network have reported that while they may be willing to diagnose depression and to treat it, they are often reluctant to commit the diagnosis to paper. This has obvious implications for the monitoring of depression through sentinel networks.

The second diagnosis for which results are presented is diabetes.<sup>2</sup> The recorded incidence of diabetes in the Northern Ireland network for 1993 was much lower than that recorded in the weekly return service of the Royal College of General Practitioners for 1992, which reported an incidence rate of 2.5 per 1000/y compared with 1.4 per 1000/y in the Northern Ireland study.<sup>1</sup> One explanation for this is that the figures from the Royal College of General Practitioners include diabetic episodes rather than first ever diagnoses of diabetes (personal communication). This emphasises the importance of the recording of episode type within the general practitioner's record, if data are to be interpretable through sentinel networks. A further explanation of the difference observed between the two studies may be the use of different diagnostic criteria.

The third set of results presented are for acute myocardial infarction. These figures reflect the high incidence of acute myocardial infarction in Northern Ireland.<sup>8</sup> Because of the serious nature of acute myocardial infarction, this diagnosis is normally well recorded within general practice.

There is little published work on evaluation of sentinel networks and the impact which they

have had on health service decision making and outcomes. Sentinel networks have been regarded as the epidemiological bridge between hospital morbidity statistics and population surveys.<sup>4</sup> However, it is important to show that the data produced are of sufficient quality to be able to fulfil that expectation. Within Northern Ireland, the data from the GPDRP have so far been considered in the following projects:

- The planning of a stroke unit;
- An estimate of the numbers of procedures required for osteoarthritis of hip and knee;
- To evaluate a diabetic retinopathy screening service;
- To assist with the setting of targets for diagnosed hypertension by the Regional Strategy Group.

While the limitations of these preliminary data were recognised, the interest in the data came from the fact that there is such a dearth of local information available to assist in planning initiatives. It is difficult to estimate what impact the information provided by the GPDRP had on the decisions made, as many factors were considered in the planning of these issues.

There is an urgent need for the benefits of sentinel networks to be evaluated and for the required validity of the data to be explored in relation to the purpose for which it is to be used. Without this, it can be difficult to get the necessary commitment to support the establishment and maintenance of sentinel networks and to ensure that general practitioners who participate in these networks gain maximum benefit from them in terms of education and development.

- 1 Research Unit of the Royal College of General Practitioners. *The weekly returns service annual report for 1992*. Birmingham: RCGP, 1993.
- 2 Royal College of General Practitioners, Office of Population Census and Surveys, and Department of Health and Social Security. *Morbidity statistics from general practice. 3rd national study (1981/2)*. Series MB5 no. 1. London: HMSO, 1986.
- 3 Boydell L. The epidemiology bridge. In: De Maeseneer J, Beolchi L, Boydell L *et al.* eds. *Telematics in primary health care*. Amsterdam: IOS. (In press).
- 4 Royal College of General Practitioners. *The interface study*. London: RCGP, 1990.
- 5 Fleming DN. The measurement of morbidity in general practice. *J Epidemiol Community Health* 1991;45:180-3.
- 6 Blacker CVR, Clare AW. Depressive disorder in primary care. *Br J Psychiatr* 1987;150:563-9.
- 7 Wright AF. *Depression, recognition and management in general practice*. London: RCGP, 1993.
- 8 Evans A, McCrum EE, Patterson CC. *The change of heart baseline clinical survey Monica project*. Belfast: Queens University of Belfast, 1990.