

A user's guide to data collected in primary care in England

Shamini Gnani and Azeem Majeed, Department of Primary Care
and Social Medicine, Imperial College London



About the APHO

The Association of Public Health Observatories was established in 2000 and has as a main focus facilitating collaborative working between the Public Health Observatories (PHO) in the UK and Ireland. APHO was set up with the following aims:

- To be a learning network for members and participants.
- To be a single point of contact for external partners.
- To be an advocate for users of public health information.
- To coordinate work across public health observatories.

Joint work is facilitated by:

- Each PHO taking the lead in a defined area to avoid duplication at regional and national levels.
- Acting as a major public health resource, raising the public health profile at regional and national levels.
- Developing collaboration through links at regional, national and international levels.

Further information about APHO, the PHOs and their work can be obtained from www.apho.org.uk.

Foreword

As the late Kerr White and others have demonstrated,^{a,b} the majority of health care (both formal and informal) takes place in the community. More NHS contacts occur and more NHS prescriptions are written in primary care than in any other setting. It is no surprise therefore that enormous quantities of data are routinely collected in thousands of separate general practices the length and breadth of the United Kingdom every day. These population-based data on individuals from birth vary, but include information on symptoms, investigations, diagnoses, referrals, treatment and outcomes, and therefore represent an incredible resource for research, education, audit, quality management, service development and planning. In an aggregated form this information could provide the most complete picture about the health of the national population as is conceivably possible. Furthermore, the rapid and widespread uptake of computers in primary care, initially for patient registration and repeat prescribing and increasingly to replace all paper records, means that these data are now more accessible than ever. However, much of the evidence we actually use to inform policy continues to be derived from secondary or tertiary care. Although there have been some notable advances in knowledge that can be attributed, at least in part, to data from primary care (witness the 'fetal origins' hypothesis arising from the meticulous ledgers of midwife Ethel Burnside^c and the use of the network of VAMP computers to identify serious adverse drug reactions^d) this resource has been underused and perhaps occasionally even misused.

Data collected in primary care has strengths and limitations, and it is important for those collecting, analysing and interpreting it to be aware of these. Algorithms have been developed to clean and code computer data, yet every practitioner whose surgery list has run thirty minutes late will know that the quality of data entry can vary. While the enormous size of pooled practice datasets can overcome error when assessing the strength of associations between different variables, it is not so easy to deal with bias. Clearly primary care data is of enormous value in helping practitioners to do simple things well for many people, few of whom actually feel ill (which remains the hallmark of chronic disease care). However, quantifying the population frequency of important health behaviours such as physical activity or the association between such behaviours and disease endpoints will continue to require specialised studies. There is a danger that the availability of a wonderful resource such as a large network of practices downloading medical records daily can start to drive the formation of research questions, which is undesirable in the longer term. In developing the academic discipline of general practice there has been an understandable rush to the randomised trial. The availability of data from large cohorts of patients in primary care should facilitate important and much-needed observational studies of the natural history of a range of symptoms.

The opportunities for using information held in electronic health records in general practice are only likely to increase, for several reasons. The £6.3 billion investment in *Connecting for Health* (formerly the National Programme for Information Technology in the NHS) should start to realise some of its potential before 2010. The UK Biobank, the UK Clinical Research Network and the topic-specific research networks will all be heavily dependent on electronic data held or collected in primary care. General practices now record information on clinical management which is linked to payments as part of the Quality and Outcomes Framework of the new GP contract. This should lead to

improvements in data quality and has renewed interest in the use of general practice data to estimate the frequency, natural history and determinants of different diseases.

This succinct, clearly presented Association of Public Health Observatories report, written by authors with considerable expertise, is very timely. It brings together and summarises information about known sources of primary care data (focusing on general practice teams rather than the professions allied to medicine), the methods and systems used to access primary care data at the moment, and sets out their range, strengths and weaknesses. It should be required reading for anyone who plans to use primary care data in the near future.



Dr Simon Griffin, Programme Leader, Medical Research Council Epidemiology Unit, Cambridge.

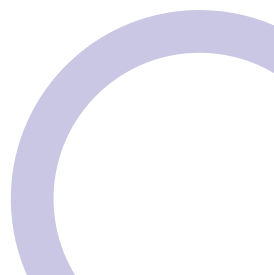
References

- a. White K. L. *NEJM* 1961;265:885–892.
- b. Scambler A. *et al.* *JRCGP* 1981;31:746–750.
- c. Barker D. *BMJ* 2003;327:1428–1430.
- d. Jick H. *et al.* *BMJ* 2000;321:1190–1195.

A user's guide to data collected in primary care in England

Shamini Gnani and Azeem Majeed, Department of Primary Care
and Social Medicine, Imperial College London

Published by the Eastern Region Public Health Observatory
on behalf of the Association of Public Health Observatories



Acknowledgements

We would like to extend our thanks to Dr Douglas Fleming from the Royal College of General Practitioners Weekly Returns Service and Mr Paul Bingham from the Eastern Region Public Health Observatory (erpho) for kindly commenting and reviewing the report, and the team at erpho for their initial comments on the scope of the report.

Title	A user's guide to data collected in primary care in England
Publisher	Eastern Region Public Health Observatory (erpho) on behalf of the Association of Public Health Observatories
Date of Publication	January 2006
ISBN	1 904389 09 0
Further copies from	www.erpho.org.uk Quick Link 12899 email: enquiries@rdd-phru.cam.ac.uk or tel: +44 (0)1223 336101
Copyright information	Please contact the main author, Dr Shamini Gnani (s.gnani@imperial.ac.uk), for permission to reproduce any part of this document.

Contents

.....

Glossary	2
Executive summary with key messages and sources of primary care data	3
Chapter 1 Introduction	13
Chapter 2 Connecting for Health	15
Chapter 3 Using data from primary care to improve health	17
3.1 Needs assessment, health service planning, and commissioning	18
3.2 Regulation, accountability and performance management	19
3.3 Clinical governance and quality improvement	21
3.4 Monitoring health inequalities	21
3.5 Monitoring healthcare use	22
3.6 Monitoring death rates among patients of general practitioners	22
3.7 Prescribing policy and pricing	23
3.8 Pharmaco-epidemiology	23
3.9 Resource allocation, risk adjustment and case-mix	23
3.10 Financial flows and payment by results	25
3.11 Public health research and health services research	25
Chapter 4 QOF data and QMAS	27
Chapter 5 Case studies	32
5.1 Primary prevention and public health surveillance	32
Case A: How do we monitor the prevalence of adult and childhood obesity?	32
Case B: How do we monitor the prevalence of smoking?	34
5.2 Chronic disease management	35
Case A: How do we monitor diabetes mellitus?	35
Chapter 6 A surveillance system in primary care	37
6.1 What is the role of a survey?	39
6.2 What is the role of a register?	39
References	41
Appendices	43
1. Routine NHS primary care activity data	43
2. RCGP Weekly Returns Service	44
3. General Practice Research Database (GPRD)	44
4. National Database for Primary Care Groups and Trusts	45
5. General Household Survey (GHS)	45
6. Fourth Morbidity Survey in General Practice (MSGP4)	46
7. Primary Care Information Services (PRIMIS)	46
8. Morbidity, Information Query and Export Syntax (MIQUEST)	47
9. Primary Care Research Networks; Trent Focus	47
10. Practice based disease registers	48
11. Practice based health promotion data	48
12. Prescribing Analysis and Cost (PACT) Data	49
13. Quality Prevalence and Indicator Database (QPID)	50
14. QResearch	50
15. Secondary Uses Service (SUS)	50
16. MEMO	51
17. The Health Improvement Network (THIN)	51
18. UK Biobank	52
19. UK Clinical Research Collaboration	52
Abbreviations	53

Glossary

Administrative data: information that is a by-product of administering care, mainly for payment or reimbursement for services provided, or to meet regulatory requirements. It typically includes information such as patient demographics, diagnostic codes and procedures performed.

Clinical data: refers to the clinical attributes of patients and represent factors that health care professionals use for patients such as symptoms (e.g. chest pain), vital signs (e.g. blood pressure), and lab test results. They are the types of observations written down by healthcare providers in the medical record and are the data used to diagnose patients and determine treatment plans.

Drugs bill: a term for the allocation of a budget for prescribing.

Indicator: a tool for measuring how well a standard is achieved.

Local Service Providers: those responsible for delivering IT services and support on a local level for five regional clusters of strategic health authorities.

Patient identifiers: aspects of information that can be traced back to the patient, typically date of birth, name and address.

Performance: the extent to which a defined standard is achieved. It is a subjective judgement of a level of performance that should be achieved and may be set nationally or locally.

Primary care: in this report, defined as care provided by general practice teams including health visitors, district nurses and mental health nurses, and excludes care given by dentists, pharmacists, opticians, and midwives. In this report general practice and primary care are used synonymously.

Standard: a statement of an agreed level of care.

Surveillance: a system that involves the scrutiny of all aspects of disease occurrence both communicable and non-communicable including the monitoring of interventions that may reduce disease prevalence.

Executive summary

General practitioners are responsible both for providing primary care services and for organising referrals for specialist care. Consequently, the medical records held by general practitioners include details of all diagnoses and prescribed drugs, in many cases extending from birth. In an increasing number of general practices, these records are held in electronic format and are potentially available for extraction and analysis. Electronic data from general practice therefore offer unique opportunities to plan and monitor health services, measure the quality of care provided by the National Health Service, and undertake population-based research.

The strengths of the data collected from general practice are that they are population-based and not derived from an unrepresentative subset of the population; that most contacts with the NHS take place in primary care; and they contain information on illness, treatments, outcomes and use of healthcare services. Now that the NHS is investing heavily in information technology, through Connecting for Health (formerly the National Programme for Information Technology in the NHS), careful thought needs to be given to how to improve the quality of electronic health records in general practice to derive maximum benefit from the data they contain. For example, there are now opportunities to link information from primary care records with information from other sources, such as hospital episode statistics, to provide a more comprehensive record of the clinical care experienced by patients.

This report describes the main primary care data sources available in England and summarises some recent developments. Comments or suggestions about the content can be sent to the main author, Dr Shamini Gnani, s.gnani@imperial.ac.uk.

Key messages are listed on pages 4 and 5.

Sources of primary care data are tabulated on pages 6 to 12.

Key messages

Chapter 1. Introduction

- Data in general practice are population based and in many cases the patient record extends from birth, and includes details of a patient's diagnoses, management and health outcomes.
- The introduction of a new GP contract with a Quality and Outcomes Framework requires general practices to routinely record detailed information on clinical management.

Chapter 2. Connecting for Health

- *Connecting for Health* aims to centralise the electronic records of 50 million patients.
- The Health and Social Care Information Centre aims to combine information systems for health and social care.

Chapter 3. Using data from primary care to improve health

There are several expected benefits from general practices improving information management and data quality. These include:

- Improving patient care within the consultation: high quality data about patients at the point of care ensures that there is legibility of medical notes, data are complete and comprehensive, there are alerts to clinical errors, and appropriate warnings to support decision making
- Identifying patient groups that may benefit from health promotion and preventive medicine
- Improved communication and patient follow up, especially across health care organisations, for example in the tracking of laboratory investigations
- Better chronic disease management: clinicians, by using templates and protocols can ensure that patients with chronic disease can be monitored effectively
- Practice organisation: practices with high quality data can improve cervical cytology recall, improve letters of referral, ensure claims and payments are made
- Supporting clinical governance: assessment and improvement of the quality of clinical care requires good data.
- Supporting effective commissioning and healthcare planning by PCOs and more recently practices that require operational clinical data on areas such as workload, quality of care, and health care burden
- Providing data to the wider NHS: a great deal of the work of general practice was hidden because either the data were unavailable or there were no requirements for data, in contrast to the hospital sector.

The use of health data, especially primary care data can be classified into a number of broad themes:

- Needs assessment, health service planning and commissioning
- Regulation, accountability and performance management
- Clinical governance and quality improvement
- Monitoring health inequalities
- Monitoring healthcare use
- Monitoring death rates among patients of general practitioners
- Prescribing policy and pricing

- Pharmaco-epidemiology
- Resource allocation, risk adjustment and case-mix
- Financial flows and payment by results
- Public health and health services research

Chapter 4. QOF data and QMAS

- The national QMAS database currently holds indicator data for 8575 practices in England.
- QMAS is a national web-based software tool developed for implementing the new GP contract.
- Data from practices are aggregated to maintain patient confidentiality, and for every practice a set of quality (QOF) scores is calculated.
- Information is collected on the number of patients with a particular chronic disease condition and on the care they receive.
- There is potential to link QOF data with other data collected in primary care and with other data sources. This would allow practices to be compared, adjusting for differences in the underlying population, to help understand differences in the quality of care provided.
- QOF data have limitations and these include not being able to: calculate age-standardised prevalence rates, adjust for socio-economic and ethnicity differences in the population, or adjust for inconsistencies in diagnosing and coding of disease conditions between healthcare professionals.

Chapter 5. Case studies

- Obesity is predicted to have an effect on the health of the population equivalent to tobacco smoking.
- The Health Survey for England is the main source of routine data on the prevalence of obesity in England. Another source is the National Diet and Nutrition Survey.
- The majority of recording of BMI among adults in general practice is ad hoc.
- Obesity levels are increasing among boys and girls.
- Government targets have exposed the lack of information on the extent of childhood obesity.
- Smoking is the single most modifiable risk factor for ill health.
- In primary care the recording of smoking status on general practice computers is variable.
- Smoking data are recorded on separate databases either within hospitals or primary care organisations, and are rarely linked.
- It is important to link the varied sources of data on smoking, improve the extraction of smoking data from GP systems and to take into account, in analysis, the recording of smoking status over time.

Chapter 6. A surveillance system in primary care

- There will be differences in the requirements of potential users of primary care data.
- There is a need to improve the quality of electronic medical records in primary care, in particular the recording discipline of doctors.
- The decisions made as to what data should be collected by a primary care data system should take into account the burden of data collection for individual practices.

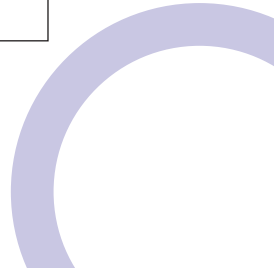
Sources of primary care data

Source and Purpose	Population covered	Strengths	Weaknesses	Areas for development
<p>Routine NHS primary care activity data: General Medical Services (GMS) data</p> <ul style="list-style-type: none"> Collected through administering the system of fees and allowances paid to general practices Used to compare performance among general practices and PCOs Costs of collecting data difficult to measure, as integral to administrative function of health authorities and PCOs Costs of data access are low locally and nationally 	<ul style="list-style-type: none"> Entire population of England (about 49 million people) Information available on all general practices in England (about 9,000) 	<ul style="list-style-type: none"> Commonly used in local primary care indicator packages Straightforward to access new GMS data at local and national level 	<ul style="list-style-type: none"> Largely derived from claims and registration data No clinical information No information on process of clinical care 	<p>See Appendix 1</p> <ul style="list-style-type: none"> Replaced by new GMS contract (see QMAS and OPID database) Publish information on practice's and PCO's Quality and Outcomes Framework performance to allow greater public scrutiny
<p>RCGP Weekly Returns Service</p> <ul style="list-style-type: none"> Publishes the incidence of selected respiratory and communicable diseases twice weekly using recent trend data Produces annual reports on the weekly incidence and prevalence of all diseases Costs about £400,000 per year to run, two thirds of which goes to practices 	<ul style="list-style-type: none"> 600,000 to 700,000 people depending on the number of practices contributing data at any one time Number of practices varies from 70 to 80 	<ul style="list-style-type: none"> Provides timely information on the number of people presenting to GPs with specified conditions Practices that contribute are experienced in collecting morbidity data. Many took part in the Fourth Annual Morbidity Survey in General Practice (see section 6) Provides information on GP workload. 	<ul style="list-style-type: none"> Only tabulated data available; not possible to link data, as no individual patient-level data No information on prescribing or community or hospital care 	<p>See Appendix 2</p> <ul style="list-style-type: none"> Enhance to a daily reporting system Integrate microbiological data Include data on treatment Could be used as the basis of a new morbidity survey

Source and Purpose	Population covered	Strengths	Weaknesses	Areas for development
<p>General Practice Research Database (GPRD)</p> <ul style="list-style-type: none"> Initially set up as a research tool, principally to carry out pharmaco-epidemiological studies (mainly drug side-effects). Increasingly used for health services research, e.g., time trends in disease prevalence and treatment, area and inter-practice variations. Costs of data collection are £2–3 million. This cost has been borne by researchers and other users via a data charge. 	<ul style="list-style-type: none"> 1.4 to 3.5 million people (depending on selection criteria for inclusion of practices) Number of practices included varies, depending on inclusion criteria, from between 200 to 400 	<ul style="list-style-type: none"> Fairly complete recording of morbidity information Includes information on age, sex, and preventive care Good information on prescribing Widely used for research and a new licence paid for by the Medical Research Council will allow 50 projects to use the database free of charge each year for five years.³⁰ Can be used to provide information on current standards of care and the inter-practice variation among contributing practices, for use in 'benchmarking' Can provide time series' data for monitoring trends in healthcare 	<ul style="list-style-type: none"> Initially set up as a research tool, rather than a statistical database Requires high degree of technical expertise to use All projects require approval by a Scientific and Ethical Advisory Group. This slows down access to the database Has a new interface that researchers have little experience of Can not be used to provide information for smaller areas, such as PCOs No direct link between diagnosis and prescribing; indirect links only Projects can take a long time to set up and complete. 	<p>See Appendix 3</p> <ul style="list-style-type: none"> Addition of socio-economic information, initially by the use of practice postcodes, later by patient postcodes. Development of a method for linking mothers' and children's' records Increased use for policy driven research Could be complemented by using information from other primary care databases
<p>National Database for Primary Care Groups and Trusts</p> <ul style="list-style-type: none"> Established by the Department of Health to provide baseline information on PCG/Ts, for use in monitoring and research 	<ul style="list-style-type: none"> All primary care groups and trusts in England; about 49 million people Information is not available at general practice level 	<ul style="list-style-type: none"> Provides population, socio-economic and GMS data on all PCG/Ts in England Updated regularly Free access to the database for registered users 	<ul style="list-style-type: none"> Does not yet contain any information on prescribing, hospital referrals or hospital admissions No information on the supply of secondary care 	<p>See Appendix 4</p> <ul style="list-style-type: none"> Addition of information on prescribing and use of hospital care, as well as on waiting lists and times Addition of supply variables, such as specialists or hospital beds per capita

Source and Purpose	Population covered	Strengths	Weaknesses	Areas for development
<p>General Household Survey</p> <ul style="list-style-type: none"> Main source of information on activity in general practice 	<ul style="list-style-type: none"> Nationally representative sample of private households in Britain 9000 households and about 16,500 adults aged 16 and over 	<ul style="list-style-type: none"> Helpful in examining rates of consultations by GPs and by age groups and in providing a yearly snapshot 	<ul style="list-style-type: none"> Unable to accurately monitor trends in GP consultation rates 	<p>See Appendix 5</p>
<p>Fourth Morbidity Survey in General Practice (MSGP4)</p> <ul style="list-style-type: none"> The survey examined the patterns of disease in general practice, in relation to the age, sex, and socio-economic status of patients Cost about £6 million pounds to complete. Costs of a repeat survey are expected to be considerably lower 	<ul style="list-style-type: none"> 502,000 people from 60 volunteer practices in England and Wales 	<ul style="list-style-type: none"> Provides a complete record of consultations with GPs Good recording of primary care morbidity data Contains individual level demographic, socio-economic and ethnicity data Costs of data access are low 	<ul style="list-style-type: none"> Carried out in 1991-92 No information on prescribing Limited information on the use of secondary care 	<p>See Appendix 6</p> <ul style="list-style-type: none"> More up to date information on GP workload and socio-economic variations is needed. Need to include information on prescribing, use of community services, and hospital referrals in a new survey
<p>Primary Care Information Services (PRIMIS)</p> <ul style="list-style-type: none"> Main aim is to educate practices through local information facilitators in data quality and information management skills Data quality assessment and feedback integral part of the process Regular clinically-focused data extraction done on a variety of topics Costs about £1.2 million per year 	<ul style="list-style-type: none"> Practices spread across 276 PCOs; 90 percent of all PCOs in England. 	<ul style="list-style-type: none"> Uses MIQUEST (see below) 	<ul style="list-style-type: none"> Dependent on PCOs investing in facilitators and their quality and abilities Dependent on willingness of practices to change Data quality improvement is a slow process 	<p>See Appendix 7</p> <ul style="list-style-type: none"> Add socio-economic data on all patients in selected practices by linking postcodes and then use survey data Use of household identification to enable linking of families for genetic studies Already developing web based feedback to practices

Source and Purpose	Population covered	Strengths	Weaknesses	Areas for development
<p>Morbidity, Information Query and Export Syntax (MIQUEST)</p> <p>Primary care system data extraction software tool</p>		<ul style="list-style-type: none"> • Compatible with European laws for confidentiality and consent • Can be used to write flexible queries • Used to monitor NSFs and inequalities, and family studies • Can be used to provide summary information on smaller areas and report the inter-practice variation among contributing practices, for use in 'benchmarking' • Can provide time series data for monitoring purposes • Interview data from staff 	<ul style="list-style-type: none"> • Requires high level of expertise to specify and write queries • Dependent on quality of system interpreters 	<p>See Appendix 8</p> <ul style="list-style-type: none"> • System suppliers to update MIQUEST interpreters and language, and network option to improve response time and make process less people-intensive
<p>Primary Care Research Networks: Trent Focus</p> <ul style="list-style-type: none"> • Network of 60 practices; patients representative of other patients in Trent • Minimum data quality standards • Practices paid £1,000 per year to participate in research studies • Track record of research successfully completed and published 	<ul style="list-style-type: none"> • 800,000 patients from 60 practices spread among 51 PCOs 	<ul style="list-style-type: none"> • Representative sample of practices • National sample • Accessible data, low cost • Validated databases of proven quality • Continuous monitoring of data quality and morbidity • Includes information on age and sex, and on preventive care • Good information on prescribing. Can obtain questionnaires from patients at cost 	<ul style="list-style-type: none"> • Need to validate recording of death data 	<p>See Appendix 9</p> <p>The Clinical Research Network model could be extended nationally via PRIMIS</p>



Source and Purpose	Population covered	Strengths	Weaknesses	Areas for development
<p>Practice based disease registers</p> <ul style="list-style-type: none"> Established to meet requirements of National Service Frameworks Generally no special funding allocated for task by many PCOs Difficult to access data where data have not been collated centrally by PCO or where not accessible using MIQUEST 	<ul style="list-style-type: none"> Typically, one PCO (100,000 – 250,000 people) Number of general practices varies considerably depending on PCO population 	<ul style="list-style-type: none"> Provides local data on morbidity and sometimes on treatment Can provide information to help improve the quality of care and the accuracy of data recording Costs of access low when data accessible by MIQUEST 	<ul style="list-style-type: none"> Only produced for a limited number of conditions (most often coronary heart disease) Wide variations in the accuracy of data between practices 	<p><i>See Appendix 10</i></p> <ul style="list-style-type: none"> Develop guidelines for data recording and an accreditation scheme Demonstrate benefits to GPs of improving the quality of their data Develop capability to compare registers in different areas to allow comparisons of PCOs
<p>Practice based health promotion data</p> <ul style="list-style-type: none"> Traditionally has not taken place in general practice 		<ul style="list-style-type: none"> Potential to give a greater understanding of lifestyle factors (e.g. smoking and obesity) and risk factors for many long term conditions 	<ul style="list-style-type: none"> Not uniformly collected Reliance on national survey data applied at a local level 	<p><i>See Appendix 11</i></p> <ul style="list-style-type: none"> New public health White Paper may influence data collection Further work is required on collection of obesity and smoking data
<p>Prescribing Analysis and Cost (PACT) data</p> <ul style="list-style-type: none"> Initially developed as a by-product of the system for reimbursing pharmaceutical contractors Used extensively for monitoring general practice prescribing and setting prescribing budgets Increasingly used by researchers Prescription Pricing Authority's administrative costs are about £62 million per year 	<ul style="list-style-type: none"> Entire population of England (about 49 million people) Information available on all general practices in England (about 9,000), PCOs and health authorities 	<ul style="list-style-type: none"> Used extensively by different organisations to monitor prescribing and inform policy Provides an accurate and complete record of the cost and volume of GP prescribing Can provide information at general practice, PCO and national levels Costs of accessing data are low, both at national or regional level 	<ul style="list-style-type: none"> Can not be linked to patients. Hence, no information on prescribing by age and sex Can not be linked to diagnostic data. Hence, limited information on prescribing for specific disorders Many of the suggested improvements to PACT may be unfeasible for technical, financial and political reasons 	<p><i>See Appendix 12</i></p> <ul style="list-style-type: none"> Linkage of data to information on patients' age and sex Addition of diagnostic data to information collected on prescriptions Availability of longer time series data at general practice and local levels (three years data available currently). Much longer term time series data already available at national level, through the DH's PCA system

Source and Purpose	Population covered	Strengths	Weaknesses	Areas for development
<p>Quality Management and Analysis System (QMAS)</p> <ul style="list-style-type: none"> National web based software tool, developed by <i>Connecting for Health</i> to extract data from GP practice systems Data extracted for the Quality and Outcomes Framework element of new GP contract 	<ul style="list-style-type: none"> Holds clinical data for approximately 8,500 practices, 50 million people in England 	<ul style="list-style-type: none"> Representative of practices in UK Aggregate data ensures patient confidentiality Covers clinical domains, especially chronic disease 	<ul style="list-style-type: none"> Quality of data Ascertainment of cases Crude prevalence figures for chronic disease only 	<p>See Chapter 4</p> <ul style="list-style-type: none"> Over time will allow for trend analysis May include recording of risk factors Linkage with other data sets such as Hospital Episode Statistics
<p>Quality Prevalence and Indicator Database (QPID)</p> <ul style="list-style-type: none"> Developed by Health and Social Care Information Centre in Leeds 	<ul style="list-style-type: none"> Aim is to widen access to Quality and Outcomes Framework and QMAS data, which covers approximately 8,500 general practices in England 	<ul style="list-style-type: none"> Provide analytical support Undertake national analysis Publication of data as part of Freedom of Information act 	<ul style="list-style-type: none"> Access to data requires prior approval 	<p>See Appendix 13</p> <ul style="list-style-type: none"> Under development
<p>QResearch</p> <ul style="list-style-type: none"> Currently a pilot database of general practice derived data for use in medical research Contains data from nearly 500 English medical practices Recruited 500 general practices throughout England, Wales, Northern Ireland and Scotland, and covers every Strategic Health Authority Area 	<ul style="list-style-type: none"> Contain records of approximately 3.5 million patients currently registered from around 500 practices 	<ul style="list-style-type: none"> Fairly complete recording of morbidity information Includes information on age, sex and preventive care Good information on prescribing Provides some small area socio-economic data Can be used to provide information on current standards of care & the inter-practice variation among contributing practices for use in 'benchmarking' Can provide time series' data for monitoring of trends in care Provides some feedback of data to contributing practices Increasingly used by researchers 	<ul style="list-style-type: none"> Database is still under development, but near completion There are some costs attached to using data All projects require approval by a Scientific Advisory Group Indirect links only between diagnosis and prescribing 	<p>See Appendix 14</p> <ul style="list-style-type: none"> Reduce the costs of data access Increased use for public health and policy research

Source and Purpose	Population covered	Strengths	Weaknesses	Areas for development
<p>Secondary Uses Service</p> <ul style="list-style-type: none"> Part of the NHS Care Records Scheme To be used for purposes other than direct clinical care 		<ul style="list-style-type: none"> Aim to reduce data collection on the NHS Will provide pseudo-anonymised patient data Combine health and social data 	<ul style="list-style-type: none"> Potentially enormous implications for providing information Ambitious task 	See Appendix 15
<p>MEMO</p> <ul style="list-style-type: none"> Originally set up in Tayside to undertake hypothesis testing of pharmaco-vigilance studies Three original datasets: dispensed prescribing, hospitalisation and death certification 	<ul style="list-style-type: none"> Covers about 400,000 people in Tayside 	<ul style="list-style-type: none"> Able to link primary, secondary care and pharmacy datasets Generalisable to populations in UK and Europe 	<ul style="list-style-type: none"> Validity of data recording 	<p>See Appendix 16</p> <ul style="list-style-type: none"> Comparisons of data with other databases especially QOF data
<p>The Health Improvement Network (THIN)</p> <ul style="list-style-type: none"> Developed by two commercial organisations – EPIC and In Practice Systems 	<ul style="list-style-type: none"> Covers approximately 100 practices 	<ul style="list-style-type: none"> Anonymised clinical records Most practices have several years of data and used to contribute data to the previous VAMP systems 	<ul style="list-style-type: none"> Relatively new database started in 2003 	See Appendix 17
<p>UK Biobank</p> <ul style="list-style-type: none"> Project to build largest information resource for researchers 	<ul style="list-style-type: none"> 500,000 volunteers aged between 40-69 years 	<ul style="list-style-type: none"> Will provide information on volunteers' lifestyle, employment and medical and family history Researchers will be able to understand the link between the participant's genes, lifestyles, environment and the disease and conditions they may develop Follow up of 30 years 	<ul style="list-style-type: none"> Involves volunteers, and therefore may not be representative 	See Appendix 18

Chapter 1

Introduction

Key messages

- Data in general practice are population based and in many cases the patient record extends from birth, and includes details of a patient's diagnoses, management and health outcomes.
- The introduction of a new GP contract with a Quality and Outcomes Framework requires general practices to routinely record detailed information on clinical management.

In the UK, most patients' experience of healthcare is within primary care. Although over 80% of general practitioners (GPs) use computers to record patient information, information on the consultations between patients and GPs are not routinely collected, either nationally or locally. Traditionally, the main function of information systems in general practice has been to provide information for general practitioners and other members of the clinical team to use in day-to-day clinical care. The data have also been used for patient registration and, more recently, to help with payments made to practices under the new GP contract.¹

The NHS aims to provide 'cradle-to-grave' care and universal coverage of primary care services by GPs. This means that data collected in general practice is population based, and in many cases the patient record extends from birth, and includes details of a patient's diagnoses, management and health outcomes. By using primary care data, it is possible to extend the understanding of the natural history of illness including access to healthcare by patients. The patterns of care in general practice, such as rates of prescribing and outpatient referrals, vary widely but there is little understanding as to why these variations exist.

There are increasing requirements for GPs to record a wide range of data electronically, following the introduction of a new GP contract with a Quality and Outcomes Framework (QOF). For the first time, this requires general practices to record detailed information on clinical management. Since these data are linked to payments to practices, it is hoped that this will lead to improvements in the accuracy and quality of electronic health records and information in general practice.

The aim of this report is to:

- summarise the main systems of data collection in primary care for practitioners;
- describe the strengths and limitations of existing data sources;
- examine the potential of new sources of primary care data.

In this report, we focus mainly on the care given by general practice teams and acknowledge the care given by community nurses, midwives, dentists, opticians, pharmacists and professionals allied to medicine. We use general practice and primary care synonymously. We list the main databases and data sources in primary care that are accessible to practitioners (see table on pages 6 to 12; with further details in the Appendices).

We do not distinguish between data collected for direct patient care and that collected for indirect or secondary use. We consider the use of primary care data in planning improvements in both individual patient care and population-based care. The new Health Information and Intelligence Task Force has the responsibility to develop a health information and intelligence strategy, as recommended by the Department of Health's public health White Paper *Choosing Health*.² We acknowledge that changes in NHS policy continue to take place, but we hope this report will be of use to public health practitioners who wish to expand their knowledge and use of data collected in primary care.

Chapter 2

Connecting for Health

Key messages

- *Connecting for Health* aims to centralise the electronic records of 50 million patients.
- The Health and Social Care Information Centre aims to combine information systems for health and social care.

In 1998 the Department of Health's strategy *Information for Health*³ committed the NHS to setting up lifelong electronic health records for all. In 2002 the Government set up the National Programme for Information Technology (NPfIT), now called *Connecting for Health*, to overhaul the information technology infrastructure of the NHS by 2010, allocating £6.3 billion to the project.⁴ The aim was to digitise the 50 million patient records creating a central record of patients' care. This would allow information to be shared safely across the NHS, and for individuals to have access to their own electronic records. It would hold a summary of a patient's contact with all care providers and would record the patient's consent for professionals to access their records. By 2010, the NHS Care Records Service (NCRS) or data spine would provide a 'cradle-to-grave' NHS record for each patient holding essential information that is accessible by different local NHS organisations.⁵

Connecting for Health in England awarded the national contract to British Telecommunications to develop software and to support the connection of all NHS organisations including GPs, acute trusts and community and mental health trusts in a single secure system. *Connecting for Health* has been divided into five regional areas, consisting of between five to seven strategic health authority areas: Southern, London, North East, East, West Midlands and North West. Each regional area has contracted with a Local Service Provider (LSP), which is a commercial organisation that has bid for a contract to provide information systems to the local health community.

A new Health and Social Care Information Centre was set up in March 2005, based in Leeds.⁶ The centre is a Special Health Authority incorporating the Prescribing Support Unit. It aims to combine the information systems for health and social care, provide national leadership for data and information, and respond to the information requirements of those in the NHS.

The first Wanless report,⁷ the *NHS Plan*⁸ and *Delivering 21st Century IT support for the NHS*⁹ all recommended that information technology in the NHS should support the development of electronic records. However, there are concerns about the accuracy and sensitivity of information with the NHS Care Records Service. There are three main steps in ensuring patient confidentiality and

appropriate disclosure of data. First, there has to be a legitimate relationship between the patient and the care professional accessing the patient's record. Second, professionals should only have access to information as they need it. Finally, patients may choose to have information placed in 'sealed envelopes' if they do not wish it to be revealed during routine care. The rules governing information held in the service are set out in the NHS Care Record Guarantee.

Patients have a right of access to health information about themselves under the *Freedom of Information Act*,¹⁰ while the *Data Protection Act 1998* governs access to the health records of living patients.¹¹ The *Freedom of Information Act* has implications for health service organisations in how information is made available to the public including patients and carers, consumer organisations, professionals and professional bodies, and local and national government.

There are other components of *Connecting for Health*, which are described in more detail within the Appendices, such as the Secondary Uses Service (SUS) (defined as the use of NHS data for purposes other than direct patient care). SUS will include all NHS activity and other non-patient record based data. It will replace some of the functions of the Office for National Statistics (ONS), including tracing, flagging, migration, cancer registration and the longitudinal study (although the practical aspects of this remain to be determined). SUS will replace everything that exists in the NHS wide clearing system and provide online reporting, analysis and extracts. A key function of SUS will be to support the Department of Health's policy of payment by results, derive dominant Health Related Groups (HRGs), calculate costs,¹² and provide standard reports.

Chapter 3

Using data from primary care to improve health

Many organisations and individuals make use of data from primary care. The key use of data recorded by primary care professionals is in the clinical care of patients. Primary care data have also been used for purposes deemed as indirect care, such as the organisation of health services.

The vast majority of data collected in primary care are by-products of administrative activity, for example General Medical Services (GMS) data and more recently the QOF in the new GP contract. However, there are several specific primary care databases. The most widely used is the General Practice Research Database (GPRD), which has mainly been used for epidemiological and health services research, providing information on trends of disease prevalence, and prescribing patterns.

Prescribing Analysis and Cost (PACT) data provide most of the information on community prescribing; a by-product of reimbursing pharmaceutical contractors. PACT data are also used for analysing prescribing in therapeutic areas, identifying expensive drugs, developing prescribing indicators, setting and monitoring budgets, and health services research. Researchers have used PACT data to investigate variations and trends in prescribing costs for example, between fundholding and non-fundholding practices. Primary care organisations (PCOs) can identify practices that prescribe new and high-cost drugs compared with older drugs of similar efficacy, or drugs that are known to be ineffective, using information from the Prescription Pricing Authority.

However, there are key limitations with data collected by GPs:

- The most heavily used primary care databases have data collected from volunteer practices. Volunteer practices are often large practices that are deemed to provide an above average quality of care, and therefore are often not representative of all practices;
- The quality and completeness of electronic health records is highly variable; and
- Information on social circumstances and ethnic status is often absent.

There are several expected benefits from general practices improving information management and data quality. These include:

- Improving patient care within the consultation: high quality data about patients at the point of care ensures that there is legibility of medical notes, data are complete and comprehensive, there are alerts to clinical errors, and appropriate warnings to support decision making
- Identifying patient groups that may benefit from health promotion and preventive medicine
- Improved communication and patient follow up, especially across health care organisations, for example in the tracking of laboratory investigations

- Better chronic disease management: clinicians, by using templates and protocols can ensure that patients with chronic disease can be monitored effectively
- Practice organisation: practices with high quality data can improve cervical cytology recall, improve letters of referral, ensure claims and payments are made
- Supporting clinical governance: assessment and improvement of the quality of clinical care requires good data
- Supporting effective commissioning and healthcare planning by PCOs and more recently practices that require operational clinical data on areas such as workload, quality of care, and health care burden
- Providing data to the wider NHS: a great deal of the work of general practice was hidden because either the data were unavailable or there were no requirements for data, in contrast to the hospital sector.

The use of health data, especially primary care data can be classified into a number of broad themes, which are discussed in more detail in this section:

- Needs assessment, health service planning and commissioning
- Regulation, accountability and performance management
- Clinical governance and quality improvement
- Monitoring health inequalities
- Monitoring healthcare use
- Monitoring death rates among patients of general practitioners
- Prescribing policy and pricing
- Pharmaco-epidemiology
- Resource allocation, risk adjustment and case-mix
- Financial flows and payment by results
- Public health and health services research

3.1 Needs assessment, health service planning and commissioning

Historically, mortality data were used to estimate the burden of disease in populations. This is due to the comprehensiveness and accuracy of collecting and recording deaths in the UK; every death is reported to the Registrar General. However, it is also important to measure the morbidity associated with diseases (not commonly recorded on death certificates), due to its effect on both the use of health services and the wider costs to society. Data derived from primary care can complement morbidity data derived from other sources such as hospital admissions data to describe the burden of disease in the community.

Unfortunately, the population burden of chronic disease is not well measured and, where it is available, it is based on small studies. There is a low degree of certainty of estimates of disease prevalence associated with single, small studies when applied to the general population. The small numbers of people associated with such studies means it is not possible to accurately calculate the prevalence rates of disease, by age group or sex. Furthermore, single studies only provide information at one point in time, and cannot be used to describe changes in disease prevalence over time.

The limitations associated with small studies can be overcome by using data from large population-based databases of primary care data, such as the General Practice Research Database, or from information derived from projects such as Primary Care Information Services (PRIMIS) or the Weekly Returns Service of the Royal College of General Practitioners (RCGP) (see Appendices). It is possible to calculate age-specific and age-standardised disease prevalence rates from these data sources and thus help to monitor national trends in rates of disease prevalence.

Needs assessment depends on the ability to quantify risk factors, diseases and population subgroups. This requires the use of routine NHS information systems, which general practices may not have access to. Most practice computers are set up to record clinical activity and perform only routine administrative tasks. Practices that plan to commission health services, under the practice-based commissioning scheme will require information from public health departments in PCOs on the prevalence of disease, effectiveness of treatments and utilisation of health services to allow them to plan health services effectively for their practice population.

Recent developments in primary care data such as the QOF, the Quality Management Analysis System (QMAS) and the Quality Prevalence and Indicator Database (QPID) will improve the potential of primary care data to examine the prevalence of disease locally. This in turn should improve the planning of health services and the allocation of resources, according to where there is greatest need.

Data from primary care are useful in implementing and monitoring national service frameworks. For example, a recent study estimated the workload implications of the national service frameworks on coronary heart disease for general practitioners. By combining disease prevalence derived from primary care with population projections it is possible to estimate the number of people with chronic disease locally, and plan interventions to improve the responsiveness of health services. Primary care data can be used to help plan and monitor the provision of services according to need, and identify areas in primary care where services or interventions are effective or missing. PCOs can use this information to help produce 'Local Area Profiles' and secure better services by implementing evidence-based care or best practice, allocating resources appropriately, and in investing in appropriate interventions including education and training. The analysis of estimates of disease prevalence from the QOF may support local surveillance of chronic diseases and the monitoring by public health organisations and integration of epidemiology in service planning and delivery.

3.2 Regulation, accountability and performance management

Previously, the NHS (principally health authorities and PCOs) monitored the performance of general practices and primary care using financial administrative data sources such as general medical

services contract and PACT data. The performance indicators they used were largely based on routine data sources of population health measures applied to general practice. This was partly due to the variability of data collection in primary care and ad hoc information systems among general practices. These 'primary care indicators' were criticised for being crude measures of performance.

Existing performance measures for primary care trusts may mask the significant variations in performance that exist among general practices. However, using indicators to monitor performance at a general practice level requires reliable data collection and validation, and this in turn requires good information technology systems, which has implications on resources.

Data items from the QOF may lead to the development of more robust indicators to monitor the quality of care provided by general practices and PCOs. Indicators based on routine data are easiest to produce, compared with carrying out patient or practice surveys. In contrast to primary care, NHS hospitals have a more developed procedure for collecting information: patient administration systems and Hospital Episode Statistics. The collection of a minimum data set by each hospital allows the production of more clinically focussed performance indicators, which can be aggregated at a regional and national level.

The Department of Health and the Healthcare Commission, which exists to promote improvement in the quality of healthcare in England, are responsible for developing a new performance framework for PCOs. They are all likely to use data from primary care, especially QOF data in the task of performance monitoring NHS and non-NHS organisations. The results from these new performance measures will be vital in assigning star ratings and Foundation Status to organisations.¹³ NHS organisations deemed to have under-performed are likely to face sanctions. Therefore, it is important that performance indicators provide an accurate and valid assessment of quality of care across key areas, and that information is interpreted appropriately, as perverse incentives may operate. National sources of primary care data such as GPRD and PRIMIS and data from GMS QOF can provide information on how practices vary. The 'benchmarking' of PCOs and general practices through the use of national and local standards may help identify areas of clinical practice requiring improvement.

The advantage of using performance indicators is in their potential effect of altering behaviour in a beneficial way. In New York State the publication of indicators of death rates among patients who underwent cardiac surgery was initially followed by a fall in patient death rates.¹⁴ In England results from the QOF have been published, allowing public access to information on the quality of care provided by general practices. Furthermore, following the conviction of Dr Harold Shipman, there is a requirement for improved accountability of doctors including recording of mortality rates among patients of individual GPs.¹⁵ There is debate, as to whether patients will choose to register with practices that perform better, based on the indicator results published.

For most patients, primary health care teams are their first point of contact with the NHS. The care patients receive within primary care has a major knock-on effect on the use of other NHS services, for example, on prescribing, investigations, referrals and hospital admissions. Patients' perceptions of the quality of care that the NHS provides are to some extent also determined by their experience of primary care.

3.3 Clinical governance and quality improvement

Many practices use data from their computerised medical records for clinical audit and clinical governance. Some PCOs have found using data in this way, as a powerful incentive for improving the completeness and accuracy of data recording, and consequently the quality of care. To be useful for quality measurement purposes, clinical data must be computer readable. Before practices became computerised, GPs recorded most, if not all clinical information on paper. This meant that important clinical data in one part of the system was not available or useable across other parts. Although there are now improvements in the recording of clinical information by GPs, there are still many gaps, for example in the recording of results of diagnostic investigations. This is partly due to the opportunity cost and resource implications for practices in recording clinical data on general practice computer systems, for example, in lengthening the consultation time between doctors and patients. In many cases, the audits have used tools and techniques developed by PRIMIS. Some practices are developing the use of clinical data together with guidelines and decision support software. Such links are likely to become more widespread, and may help improve the process of clinical governance.

3.4 Monitoring health inequalities

Reducing inequalities in health status among the UK population is a key priority for the NHS.^{16 17} The NHS has two roles in improving the delivery of healthcare services ensuring equity of access, and in improving population health through prevention. Currently, there are gaps in understanding the extent to which primary care use is related to need among differing socio-economic and ethnic groups, or whether there exists age, sex or ethnic discrimination. Research shows there is an association between deprivation and access to primary healthcare: the inverse healthcare law, which states that patients with the greatest need for healthcare are least likely to access care. Many small studies have reported difficulties in access to primary care for patient groups such as the homeless, asylum seekers, refugees and minority ethnic groups.¹⁸

Data collected in primary care, because of the universal coverage of primary care services, can provide important information on the morbidity of the population and how this varies among different socio-economic and ethnic groups. Collecting this information will allow the NHS to systematically monitor whether there is equitable access to care. Previously the only main sources of patient-level socio-economic data in primary care have been the national morbidity surveys in general practice, the last carried out in 1991-1992. Accurate and systematic allocation of socio-economic information to individual patient records has been undertaken, as part of these surveys. Such data are not available in routinely collected medical records. Subsequent studies by the ONS and the Department of Primary Care, University of Nottingham have used practice postcodes, as a proxy measure of deprivation to examine the patterns of healthcare use and health outcome.¹⁹ Analyses using this approach usually show that the prevalence of chronic disease is highest in the most deprived areas, and is associated with the lowest rates of treatment. A recent study showed the prevalence of coronary heart disease was highest in deprived areas and lowest in affluent areas, whilst the use of lipid lowering therapy showed the opposite pattern. Further work in monitoring health inequalities may be helped by local incentive schemes for GPs to record socio-economic and ethnicity data on their patients.

3.5 Monitoring healthcare use

The monitoring of health service use has focussed mainly on monitoring hospitals and hospital admissions. The use of primary care services was monitored last by the Fourth Survey of Morbidity in General Practice (MSGP4) in 1991-1992.²⁰ The survey provides information on the range of conditions presenting to GPs and their workload. This information has been the basis of contract negotiations about workload between GPs and the Department of Health. The RCGP's Weekly Returns Service provides more up to date information, but person-linked socio-economic analyses are not available, and do not include MSGP4 data.

The main method of monitoring prescribing for specific conditions is from data collected directly from primary care, for example primary care databases, PRIMIS, or local disease registers. PACT data, the other main source of prescribing data allows the monitoring of cost and volume of community prescribing in the NHS. PACT data are accurate but provide limited information on prescribing patterns for specific conditions.

The referral rates for specialist care are an important area for monitoring health service use; the cost of patient care rises substantially on referral to hospital. The NHS plan requires PCOs to have methods of monitoring general practice referral rates. Referral rates can help predict the demand for specialist services and monitor changes over time. For example, the number of referrals is expected to increase, as the population ages, even if the prevalence of disease stays constant. This suggests that some targets set may be difficult to achieve, without either a large increase in the provision of specialist services or radical changes in the balance between primary and secondary care services.

3.6 Monitoring death rates among patients of general practitioners

Before the conviction of Dr Harold Shipman there was no national requirement to monitor death rates among patients of general practitioners. A copy of the death registration of every resident who dies in a district or elsewhere is sent to the Director of Public Health. District mortality data are processed by the ONS and are available as Vital Statistics (VS) returns. VS returns contain some details of the number of people who die in the district, the ICD code relating to the cause of death, and data are presented by age and sex.

Post Shipman, the system of death certification has changed.²¹ The Department of Health is working on the development of a single database containing information on every deceased NHS patient, including the identity of the GP or GP practice with whom the patient was registered, and the cause of death. Monitoring mortality trends requires annual deaths to be examined because of the small numbers involved. The database is to be a surveillance system that identifies GPs with high rates of mortality among their patients and thus allows GPs to be scrutinised more closely. For effective monitoring of mortality rates at individual GP level, records need to be linked between an individual patient registered with a practice and a named GP with that practice.

However, the patient lists that GPs and PCOs hold often have errors. Patients frequently change address without notifying their GP and sometimes patients on the list may no longer be present in the area. This creates 'list variation', a difference between official population estimates and estimates obtained from GP lists. The difference is particularly large in inner city areas and may lead to inaccurate denominators and hence inaccurate rates.

3.7 Prescribing policy and pricing

In England the NHS community prescribing or drugs bill is about £7 billion per year; greater than the costs of all other aspects of general practice combined. Ensuring the NHS obtains value for money in community prescribing is an important task for the Department of Health and PCOs.

The main source of prescribing data is PACT data. The collection of PACT data was initially a by product of reimbursing pharmaceutical contractors. Now it is widely used to monitor the cost and volume of community prescribing in the NHS, the implementation of prescribing policies and to ensure cost-effective prescribing in primary care. Health authorities and PCOs have mainly used this data to help set and monitor the prescribing budgets of practices. Researchers have used PACT data to investigate variations and trends in prescribing costs, for example between fundholding and non-fundholding practices.

But PACT data provide very limited information on prescribing patterns for specific conditions. At present, the main method of monitoring prescribing for specific conditions is to use data collected from primary care databases, such as the General Practice Research Database. In the future, data collected as part of QOF may help to link prescribing data with disease data.

3.8 Pharmaco-epidemiology

Research in pharmaco-epidemiology has important implications for public health and patient safety both in the United Kingdom and elsewhere. The use of primary care data is powerful in examining the side-effects and complications associated with both new and established drugs. Data collected in primary care over time allow researchers to examine the long-term complications linked with drug prescribing. For example, research using primary care data shows an increased risk of thrombo-embolic disease among women taking third-generation combined contraceptive pills. However, there are limitations to these studies due to bias and confounding with the potential for researchers to draw erroneous conclusions about the safety and effectiveness of drugs. These observational studies can be carried out more readily compared to prospective studies and the results disseminated and action taken to protect the public health. This can take place long before any findings would be available from prospective studies, because they use data that are already collected. Large primary care based databases can provide information on the long-term follow up of new and established drugs and help to answer questions about the safety of drugs.

Primary care databases, such as GPRD have mainly been used for pharmaco-epidemiological studies. This research has been driven by the pharmaceutical industry, which has a major interest in this area. They have funded many of these studies and determined the research agenda. It is important that public health priorities should also drive this type of pharmaco-epidemiological research.

3.9 Resource allocation, risk-adjustment and case-mix

Governments of all health care systems are looking to provide services that are cost effective, of high quality and are responsive to people. In the UK the quality of primary care has varied enormously, and much of this variation has been largely unexplained.

Health care systems around the world are all trying to ensure that resources are used appropriately and as part of this process, they are trying to reduce unacceptable variations between providers.

For example, the NHS Plan stated that 'it promises better performance and accountability systems to reduce variations in services across England'. In both primary and secondary care, there are large differences in the way in which doctors practice medicine. However, without taking into account differences in case-mix, we do not know whether such differences are justified clinically. Risk adjustment can help correct such variations for underlying differences in population case-mix and thus could lead to performance measures for providers that are fairer and more accurate than the unadjusted measures available now.

In the USA risk adjustment methods were developed to understand variations in populations. A similar development of methods of risk adjustment in the UK, using accurate primary care data may contribute to understanding and reducing the unacceptable variations in quality of care.

Primary care data currently have a limited role in resource allocation in the NHS. Prescribing budgets to PCOs are determined using PACT data. The current allocation formulas do not use morbidity data from primary care to help allocate NHS resources, either to PCOs or to general practices. This may be an area of future development.

Several diagnosis-based population risk adjustment models have been developed. Diagnostic Related Groups (DRGs) are a classification initially developed at Yale University. DRGs have been used by the Medicare system in the US since 1983 for the reimbursement of health service charges. They enabled a prospective payment system to develop with the objective of controlling charges more strictly.

The most widely used primary care based risk adjustment method is the Adjusted Clinical Group (ACG) system developed at John Hopkins University in Baltimore, USA.^{22 23} The ACG system clusters diagnoses into clinically meaningful categories and for individual patients gives a composite measure of health status, to help predict the patient's future use of health services. Its objective is to help ensure that the morbidity of individual patients is accounted for in allocating budgets. A patient with heart disease and diabetes would be placed in a higher category than a patient with ischaemic heart disease.

In the USA, risk adjustment is now starting to be used to adjust capitation or other types of payments to health providers. For large populations, adjustments for age and sex may be adequate for this purpose, but this is not the case for smaller populations such as those managed by small healthcare providers. Risk adjustment helps ensure that providers of healthcare who manage patients with more complex medical problems have their budgets adjusted to take this into account. The use of risk adjustment systems has also given doctors and providers a powerful incentive to provide more accurate and complete diagnostic data.

Another use of risk adjustment is to measure the health status of a population. The traditional way of doing this in most countries has been to use mortality rates or self-reported measures of chronic illness derived from censuses or surveys of the population. Because many chronic illnesses do not result in death, mortality rates are not always a good measure of the health of a population. Similarly a generic measure of chronic illness based on self-reports may also be an unreliable measure of health status. By contrast, risk adjustment models based on all conditions treated by primary and secondary care physicians take into account the full spectrum of illnesses in a population. Hence they incorporate chronic illnesses, such as arthritis or epilepsy that lead to considerable population morbidity but that are often not recorded on death certificates.

PCOs and primary health care teams need local data to help them plan local health services and monitor the quality of care they provide. They can make use of national data; for example, to provide baseline estimates of disease prevalence or to compare local treatment patterns against national patterns. However, their principal requirement is always likely to be for local rather than national data.

3.10 Financial flows and payment by results

Primary care data are likely to be used to develop accurate pricing of healthcare interventions by different providers within primary and hospital care. As part of the Department of Health's policy of introducing payment by results within the NHS, national tariffs for primary care interventions and procedures are likely to be set.

The aim of the new financial system is to provide a transparent, rules-based system for paying trusts. It also aims to support patient choice and diversity and encourage shorter hospital waiting times. Payment will be linked to activity and adjusted for casemix. Traditionally funding for hospitals relied on historic budgets and the negotiating skills of hospital managers. Under the reforms to NHS Financial Flows, instead of being commissioned through block agreements as previously, hospitals (and other providers) will be paid for the activity that they undertake so PCOs will commission the volume of activity required to deliver service priorities, adjusted for casemix (i.e. the mix of types of patients and/or treatment episodes), from a plurality of providers, on the basis of a standard national price tariff, adjusted for regional variation in wages and other costs of service delivery.

3.11 Public health research and health services research

In 'Securing Good Health for the Whole Population' Derek Wanless was critical of the lack of evidence of interventions that would lead to improved health. Primary care data have been used in health services research; typically studies have examined disease prevalence and treatment effectiveness, time trends, and area and socio-economic variations to help inform public health priorities.

The gold standard for studies of clinical effectiveness is the randomised controlled trial. However, these trials are carried out on carefully selected subset of patients, who usually have more rigorous monitoring and follow up than takes place in routine clinical practice. Therefore they may be atypical of patients and clinical management in the general population. But policy makers and clinicians need to know whether the benefits and risks of treatment found in trials are likely to be observed in routine clinical practice. Using data from primary care may help in determining whether benefits will translate to the wider population.

However, before researchers can use primary care data there are a number of requirements:

- to improve access to primary care data
- to improve the range of data that is collected
- to link primary care data with other sources of data
- to develop the recording of socio-economic status at a patient level in primary care
- to improve the quality of data.

Furthermore, new methods need to be developed to rapidly evaluate health policy and clinical effectiveness of interventions using primary care data. These observational studies are common in the USA, but are carried out less frequently in the UK.

New research initiatives such as Biobank aim to look more closely at linking datasets to the aetiology of conditions and propose treatments or interventions that would be possible at a system level. Biobank will aim to use data from clinical records to record follow-up information on patients.

Chapter 4

QOF data and QMAS

Key messages

- The national QMAS database currently holds indicator data for 8575 practices in England.
- QMAS is a national web-based software tool developed for implementing the new GP contract.
- Data from practices are aggregated to maintain patient confidentiality, and for every practice a set of quality (QOF) scores is calculated.
- Information is collected on the number of patients with a particular chronic disease condition and on the care they receive.
- There is potential to link QOF data with other data collected in primary care and with other data sources. This would allow practices to be compared, adjusting for differences in the underlying population, to help understand differences in the quality of care provided.
- QOF data have limitations and these include not being able to: calculate age-standardised prevalence rates, adjust for socio-economic and ethnicity differences in the population, or adjust for inconsistencies in diagnosing and coding of disease conditions between healthcare professionals.

The Quality Management and Analysis System (QMAS) is a national web based software tool developed, in response to the introduction of a new GMS contract in April 2004, by *Connecting for Health*. Its aim is to extract data from GP practice systems. The Quality Prevalence and Indicator Database (QPID) held by the Health and Social care Information Centre in Leeds aims to improve access to QMAS data, by users in the Department of Health and NHS. The Information Centre will undertake national analyses, for example of disease prevalence using QMAS data and other data sources.

The new GMS contract has for the first time a quality and outcomes framework (QOF), which has 76 performance indicators covering 10 chronic disease areas: coronary heart disease and left ventricular dysfunction, stroke or transient ischaemic attacks, hypertension, diabetes mellitus, asthma, chronic obstructive pulmonary disease, epilepsy, hypothyroidism, cancer, and mental health. The framework also covers areas such as patient records and information, patient communication, education and training, practice and medicines management, patient experience and services such as child health surveillance, and maternity and contraceptive services.

The national QMAS database currently holds indicator data for 8,575 practices; not all practices in England have contributed data, as involvement in the framework is voluntary. Data are organised by practice, and are aggregated to maintain patient confidentiality. For every practice a set of QOF indicator scores are calculated. The maximum score a practice can achieve is 1050 points. The QOF scores achieved are then converted into a payment, in accordance with the GMS contract. Payments are then adjusted for practice list size and practice disease prevalence. The National Health Applications and Infrastructure Services (NHAIS) or 'Exeter' system administers the cancer screening call-and-recall system and patient registration. NHAIS provides QMAS with information on list sizes (information is collected as at 1 January 2005 from GP systems), and a snapshot of all QOF disease registers (collected as at 14 February 2005, National Prevalence Day) to calculate disease prevalence.

QOF results are fed back to practices, PCOs and strategic health authorities; PCOs receive practice level data and strategic health authorities receive PCO level data. PCOs and practices can access information at any time about their QOF achievement against their aspiration (what practices thought that they would achieve), the estimated relative prevalence and their current achievement payment. From April 2005, GP practices have received financial payments that are based on their practice list size, achievement data and aspiration data held on QMAS.

Although data are not interrogated by QMAS, GP software systems had to pass a data quality check. The accuracy of QOF data is important, as GP practices receive financial payments according to the quality of care they provide. QOF data aims to give GP practices and PCOs objective evidence and feedback on the quality of care delivered to patients, measured against national targets set out in the general medical services contract. QMAS allows GP practices to analyse the data they collect about the range of services and the quality of care they deliver, such as maternity services or chronic disease management clinics. It is intended that GPs have a financial incentive to treat most patients in the community rather than referring them to hospital for treatment.

Information is collected, as to the number of patients within a practice that have a particular chronic disease condition such as coronary heart disease or diabetes mellitus. The key information collected by QMAS is on the disease register size and the numerator and denominator for each clinical indicator based on a set of defined Read codes. No demographic information is collected by QMAS. The clinical indicators that relate to each chronic condition are measures of process rather than outcome for patients with chronic conditions. Some of these process indicators, but not all, are linked to better health outcome; good blood pressure and blood glucose monitoring among people with diabetes mellitus is associated with lower mortality rates.

Exception codes

Practices are allowed to exclude patients from the denominator for an individual clinical indicator through exception coding. There are two types of Read exception codes. A high level code that excepts from the whole clinical domain, for example, if a patient has diabetes mellitus and does not wish to attend for review. There are also individual exception codes for indicators within the clinical domain, for example, those declining to have a flu injection. But when data are entered for individual clinical indicators, this overrides the high level exception code for that clinical domain. The exception codes cover the following areas:

- Patients who refuse to attend reviews, and have been invited on at least three occasions in the preceding 12 months

- Inappropriate review e.g. because of frailty or terminal illness
- Patients are on the maximum tolerated dose of a drug but are not within the requirement for a specific clinical indicator e.g. blood pressure control
- Patients newly diagnosed within the practice or who have recently registered, who should have measurements made within three months and delivery of clinical standards within nine months e.g. blood pressure or cholesterol measurements within target levels
- Patients who decline or refuse drugs or treatment
- Patients who are allergic, or experience an adverse reaction to a particular drug, or the drug is not indicated or is contraindicated
- Where a patient has not tolerated medicine
- Where a patient does not agree to investigation or treatment
- Where a patient has a supervening condition which makes treatment of their chronic condition inappropriate
- Where an investigative service or secondary care service is unavailable

Potential uses of QOF data

Using QOF data it is possible to describe differences in the quality of care across the clinical disease areas. Interpreting this information may prove difficult. It is possible to compare QOF scores by practice and by PCO. QOF data could be used to calculate the crude prevalence of disease and thus to estimate population prevalence of disease across England. This allows for differences between practice prevalence (observed) and population prevalence (expected) to be examined.

Along with the collection of information on the organisational aspects of the QOF framework, this may help us understand the factors that are necessary to provide high quality general practice. QOF data signal new possibilities to investigate variations in primary care and test hypotheses, e.g. do practices with larger list sizes achieve higher QOF scores, and thus better care?

To understand the potential of QMAS data to provide information on the prevalence of disease at a local level, it is important to understand the accuracy of the practice list size and the population characteristics of those registered with the practice. Inaccuracies in the list size will alter the denominator for calculating the prevalence, and thus this may lead to underestimates or overestimates of disease prevalence.

Exception reporting is also important to understanding QMAS data. Exception reporting was introduced to prevent practices being penalised under the quality and outcomes framework for factors for which they had no control. The exception code is used so as to not affect the practice's quality point score. There are a number of reasons for exception coding and therefore the clinical indicators need careful interpretation. A denominator for a specific clinical indicator may not be the same as the disease register size to which the clinical indicator relates because GPs can exclude patients with the use of exception codes. Practices may have an incentive to use exception reporting inappropriately to maximise their income by selectively identifying patients. High levels of exception coding may be associated with practices delivering poor quality care, or may be associated with areas of deprivation.

QMAS data are extracted from individual practices and aggregated. Therefore there is potential for linking these data to other data sources at practice or PCO level. This would confer significant benefits. For example, it would provide a much better indication of the effects of practice resources or population factors such as deprivation or ethnicity on practice performance. It would allow benchmarking and inter-practice comparisons, and validation of community profiles; an opportunity to develop a predictive model of expected prevalence by age, sex and ethnicity at a practice level; identify outlying practices. This information would help contribute to the planning and commissioning of services and to improving the management of chronic disease by the primary care team and thus reducing hospital admission rates.

Population based data on lifestyle and risk factors such as prevalence of smoking, high blood pressure and obesity could also be recorded. Analysing anonymised patient-specific data would allow much more powerful analyses of patients at high risk of coronary heart disease, e.g. absolute 10 year risk >30%, because this function is not yet available in GP systems or QMAS. Data that are postcoded would enable linkage with census data. Other identifiers (such as NHS number) may allow linkage with hospital admissions and mortality data. This could lead to identifying those primary factors that lead to hospital admissions. Furthermore, PCOs would be able to produce additional practice and patient level reports for practices such as funnel plots. Currently few PCOs are providing these additional reporting mechanisms, but there are now several software solutions which aim to do this such as CHART from PRIMIS and A3 from Apollo systems. Hence, the use of QOF data by PCOs for such purposes is likely to increase.

Although the QOF attempts to measure the quality of health care delivered in general practice, it may skew the focus of clinical care, by its failure to cover all aspects of general practice especially those which are less readily measurable.

Limitations of QOF data

There are a number of key limitations:

- It is not possible to calculate age-standardised prevalence rates for each of the chronic diseases in the contract framework. This could be modified by linking information that practices hold on their patients' age and sex. Until recent incentives, general practices did not systematically code the ethnicity of their patients. This information is important in determining whether access is equitable and in contributing evidence as to whether certain diseases are more prevalent among different ethnic groups due to genetic differences.
- As the population ages, more people will have more than one chronic condition. This will have an effect on how people use services and how services should be configured to address population need. Information on co-morbidity is not available from QOF data.
- There are limitations in the recording of risk factors within QOF, for example, the prevalence of smoking and obesity within practice populations. This would provide a proxy figure for the general population, and help decision makers anticipate future needs and develop services accordingly.
- There are limitations with QOF data due to the diagnosis and coding of disease and the completeness of practice disease registers. There are no standard methods for reporting the quality of data recorded, or the completion and accuracy of data, or the recording of morbidity

data in primary care. This may be overcome by requiring practices to be involved in a data accreditation scheme. The development of guidelines for using Read codes for diagnoses may reduce the wide variations in coding among practices computerised for a long time. The Department of Health decides the codes that are to be used for defining the disease entries.

It could be argued that under the new contract, GPs have a perverse incentive to use incorrect codes to improve the financial payments that they receive. This could be identified through developing a process of quality assurance for data entered onto clinical systems by GPs. PCOs are required to visit each practice annually, part of the purpose being to audit the data collection processes.

The research value of using QMAS and QPID data would be increased significantly if they were available at patient level. Long term gains may also be achieved by training practices in coding and classification of conditions. The monitoring of data by PCOs and feedback to practices may help in this. Finally, the identification of exception codes used by practices is not possible with the 2004-2005 data. The ability to do this with 2005-2006 data may help to explain variations.

Hospital Episode Statistics (HES) data contains all records of inpatient care provided by NHS hospitals in England, and when it was first introduced was initially thought to be of low quality. Now it is being used increasingly for financial management and clinical audit. Similarly it is likely that the quality and accuracy of QOF data will be poor at the start, but will improve over time. Thus, it will be more useful for undertaking health services research, especially in investigating primary care services. Furthermore, the linkage of QOF data with HES data may improve the analytical power of studies examining the quality of health service care. For example, whether there is an association between the quality of care of patients with asthma and the number of asthma hospital admissions.

Note:

INphoRM 7: Introduction to 2004–2005 QOF data will be published by the Eastern Region Public Health Observatory in Spring 2006.

Chapter 5

Case studies

5.1 Primary prevention and public health surveillance

Case A: How do we monitor the prevalence of adult and childhood obesity?

Key messages

- Obesity is predicted to have an effect on the health of the population equivalent to tobacco smoking.
- The Health Survey for England is the main source of routine data on the prevalence of obesity in England. Another source is the National Diet and Nutrition Survey.
- The majority of recording of BMI among adults in general practice is ad hoc.
- Obesity levels are increasing among boys and girls.
- Government targets have exposed the lack of information on the extent of childhood obesity.

Obesity is predicted to have an effect on the health of the population equivalent to tobacco smoking. In the UK the numbers of adults and children who are overweight and obese are rising markedly. In 2002, 22% of men and 23% of women were defined as obese, with older people and the less wealthy more likely to be obese.²⁴ Obesity is usually defined in terms of the Body Mass Index (BMI), and a BMI of over 30kg/m² is classed as obese. However, a more refined measure of obesity is a high waist to hip ratio, which describes the distribution of fat in the body, and is shown to be associated with high rates of cardiovascular disease.

The Health Survey for England (HSE) is the main source of routine data on the prevalence of obesity in England.²⁵ The survey is carried out each year. Each participant has their height, weight, and waist and hip measurements recorded. However, information on obesity at a local level is not available because of the sample size used in the HSE. The National Diet and Nutrition Survey is another source of data on obesity that includes BMI and waist to hip ratios.²⁶ This was carried out in 1986/87 and in 2000/2001. But, similar to HSE, it is not possible to undertake analyses at a local level. Some health authorities undertake local surveys of obesity, but these tend to be ad hoc and limited to subsets of the population, and the results cannot be assumed to be representative of other areas of the UK.

In England the levels of obesity and overweight among boys and girls shows a steady increase from 1995 to 2002. Obesity in boys almost doubled, rising from 2.9% in 1995 to 5.7% in 2002. The trends observed are more marked in children from households in manual social classes compared with non-manual classes. It is difficult to define obesity and overweight among children. Currently children are defined as being overweight at the 85th percentile, and obese at the 95th percentile, according to UK 1990 reference curves. This approach assumes that the prevalence of obesity is 5% and that of being overweight is 10%. But there is little evidence that these cut-off points relate to morbidity or health outcomes.

The government has set a national target by 2010 to stop the annual increase in obesity among children under 11 years. However, in setting this target, the government has exposed the lack of information on the extent of childhood obesity, as children are not weighed routinely. Undertaking physical assessments of children routinely is viewed as a poor screening procedure, due to the balance of benefit and costs of undertaking an assessment. But in view of the rise in obesity levels among younger people it may be justified to measure height and weight at school entry. The benefit may also extend to measuring obesity at age 10-11 years and 15-16 years, as policy makers may need to develop appropriate interventions at different points in time.

How should obesity be monitored for adults? The majority of BMI recording in general practice is ad hoc. Practices are most likely to record body weight when a patient first registers with a practice, as part of a health check. Weight may also be recorded by general practitioners or nurses when a patient expresses a wish to reduce their weight, or a patient is taking a medication that requires weight monitoring (such as the combined oral contraceptive or hormone replacement therapy), or they appear overweight and are considered to be at risk of for example, heart disease. One method is to collect obesity data routinely, as called for in 'Choosing Health'. An option here may be to require general practices to record the BMI of all their registered patients, and identify those with a BMI that is over 25 and 30, or measure patients' waist-hip ratio in the last 15 months. Another option is to carry out local surveys each year. The Scottish Intercollegiate Guidelines Network recommends that doctors should opportunistically take BMI and waist measurements at least every three years.

There is little comprehensive evidence on the effectiveness of strategies, especially in primary care, that reduce levels of obesity. GPs may be influential in getting patients to change their diets, but most advice given to patients tends to be disease-specific. The extent to which GPs provide nutrition education to patients is reported to depend on their perception of their own ability to influence lifestyle, and confidence in their ability to advise patients about their diet. The British Nutrition Foundation Task Force on obesity identified at-risk groups as: obese children and children with obese parents, rapid weight gainers (greater than 5kg in 5 years), post-obese, pregnant women, smoking quitters, physically inactive and certain ethnic groups. This may help target interventions either practice or hospital-based in the management of obesity.

At present, information on obesity is mainly available at a national level due to present systems of data collection. To monitor levels of obesity more effectively, and its effect on health, it is important to compare national survey data with information collected at a local level, for example the registered population of PCOs who are recorded as obese.

Case B: How do we monitor the prevalence of smoking?**Key messages**

- Smoking is the single most modifiable risk factor for ill health.
- In primary care the recording of smoking status on general practice computers is variable.
- Smoking data are recorded on separate databases either within hospitals or PCOs, and are rarely linked.
- It is important to link the varied sources of data on smoking, improve the extraction of smoking data from GP systems and to take into account, in analysis, the recording of smoking status over time.

Smoking is the single most important modifiable risk factor for ill health. To date information collected on the prevalence of smoking is taken mainly from the Health Survey for England. These data are then often applied to local populations to determine the local prevalence of smoking. Some areas have undertaken their own surveys among groups such as pregnant women, or among ethnic minority groups. Current smoking status was obtained routinely as part of the specific socio-economic enquiry used by MSGP4 and was used as a standardisation criterion for several of the analyses of disease prevalence.

All PCOs are required to collect smoking data. In England, each PCO has a stop smoking service, and its performance is monitored by the Department of Health; PCOs report on a quarterly basis the number of people seen by the stop smoking service. All hospitals have to collect information on the proportion of pregnant women who smoke. Smoking data are also recorded, as part of Child Health Surveillance systems. The smoking status of parents for every newborn is recorded on the discharge sheet. But often this information is frequently not used at a local level. Unfortunately, these data are recorded on separate databases within hospitals and PCOs.

In primary care, the recording of smoking status on general practice computer systems is variable. The recording of smoking status tends to be high for new patients, as this question is part of the registration health check. The recording of smoking status among patients with chronic conditions is also part of the new GP contract quality and outcomes framework.

Data collected within primary care and the community setting may help evaluate interventions that will be effective in reducing smoking prevalence. The smoking prevalence among children of school age is not known. Usually, this information is collected by surveys along with information on drug and alcohol use among school children.

It is proposed that GPs record the percentage of people aged 15 to 75 years who smoke. These data may then be analysed by five year age-sex groups and compared with existing population and survey data. Furthermore, details on the recording of the provision of stop smoking advice and referral to specialist stop smoking service may help develop appropriate services to the need of the population.

However, there are key issues that should be solved before developing a surveillance system for smoking. First, it is important to link the varied sources of data collection on smoking, thus avoiding

the numerous stand-alone databases. Second, to improve the extraction of smoking data from GP clinical systems. Third, to take into account in analysing the data methodological issues associated with recording smoking status over time, and the problem of repeated measures.

5.2 Chronic disease management

Case A: How do we monitor diabetes mellitus?

The prevalence of diabetes mellitus will increase significantly by 2010. Diabetes mellitus is a considerable health problem for the individual and society. The prevalence of disease is higher in areas of socio-economic deprivation and among certain ethnic groups such as South Asians. There is strong evidence that good quality diabetic care improves outcomes, such as the prevention of blindness and heart disease. Before 1970 most patients with diabetes were cared for in hospital, but since then there has been a gradual shift in care of patients to primary care.

In this case history we look at the monitoring of the key factors that influence the health of patients with diabetes. Evidence centres on good blood pressure control and blood glucose monitoring. Many patients are not diagnosed for between five and ten years.

In the QOF there are 18 indicators for diabetes care, shown below:

QOF clinical indicators relating to diabetes mellitus:

- DM 1 The practice can produce a register of all patients with diabetes mellitus
- DM 2 The percentage of patients with diabetes whose notes record Body Mass Index (BMI) in the previous 15 months
- DM 3 The percentage of patients with diabetes for whom there is a record of smoking status in the previous 15 months, except those who have never smoked where smoking status need be recorded only once since diagnosis
- DM 4 The percentage of patients with diabetes who smoke and whose notes contain a record that smoking cessation advice or referral to a specialist service, where available, has been offered in the last 15 months
- DM 5 The percentage of diabetic patients who have a record of glycosylated haemoglobin (HbA1c) or equivalent in the previous 15 months
- DM 6 The percentage of patients with diabetes in whom the last HbA1c is 7.4 or less (or equivalent test/reference range depending on local laboratory) in last 15 months
- DM 7 The percentage of patients with diabetes in whom the last HbA1c is 10 or less (or equivalent test/reference range depending on local laboratory) in last 15 months
- DM 8 The percentage of patients with diabetes who have a record of retinal screening in the previous 15 months
- DM 9 The percentage of patients with a record of the presence or absence of peripheral pulses in the previous 15 months
- DM 10 The percentage of patients with diabetes with a record of neuropathy testing in the previous 15 months
- DM 11 The percentage of patients with diabetes who have a record of the blood pressure in the past 15 months
- DM 12 The percentage of patients with diabetes in whom the last blood pressure is 145/85 or less

- DM 13 The percentage of patients with diabetes who have a record of micro-albuminuria testing in the previous 15 months (exception reporting for patients with proteinuria)
- DM 14 The percentage of patients with diabetes who have a record of serum creatinine testing in the previous 15 months
- DM 15 The percentage of patients with diabetes with a diagnosis of proteinuria or micro-albuminuria who are treated with Angiotensin Converting Enzyme inhibitors (or Alpha 2 antagonists)
- DM 16 The percentage of patients with diabetes who have a record of total cholesterol in the previous 15 months
- DM 17 The percentage of patients with diabetes whose last measured total cholesterol within the previous 15 months is 5mmols/or less
- DM 18 The percentage of patients with diabetes who have had influenza immunisation in the preceding 1 September to 31 March

The indicators predominantly focus on measures of risk factors such as smoking, which are modifiable by changes to lifestyle, or in recording BMI levels, thus identifying patients that are overweight. Other indicators, such as cholesterol measurements may be partly modifiable by lifestyle changes but may also require drug therapy. Many of the indicators are measures of the process of diabetic care, such as recording whether screening for eye or circulatory disease problems has taken place. Other indicators have been linked to poorer health outcome such as identifying patients with protein in their urine, who are more at risk of heart and renal disease.

There is a trend towards patients with diabetes being managed outside hospital in primary care settings, especially with the development of GPs with special interest in diabetes. Data from QOF can help decision makers when commissioning services for patients with diabetes. Now, there are diabetes prevalence data, by general practice and by PCOs. At present QOF data only provide crude prevalence figures, which cannot be analysed further by age, sex or ethnicity. However, if at a local level practice population data can be linked to QOF data, this would provide more detailed information on the burden of diabetes and population need. Further analysis on the type of diabetes and method of treatment such as diet controlled, oral hypoglycaemics or insulin therapy would help determine where diabetic care would be most appropriate to take place such as hospital or within practice. Disease severity among patients is important, and can be determined by examining indicators such as glycosylated haemoglobin (HbA1c), or examining prescribing data to identify patient with co-morbidities such as heart and kidney disease. By monitoring the effect of preventative measures such as exercise, diet, smoking prevention and alcohol, it may be possible to develop more effective interventions.

PCOs are likely to examine cost data to determine current levels of expenditure on diabetic care. Potentially larger practices may have scope to make savings, and primary care data especially QOF, may contribute to determining the cost-effective of primary care based diabetic services, as opposed to services delivered by hospitals or other providers.

Chapter 6

A surveillance system in primary care

Key messages

- There will be differences in the requirements of potential users of primary care data.
- There is a need to improve the quality of electronic medical records in primary care, in particular the recording discipline of doctors.
- The decisions made as to what data should be collected by a primary care data system should take into account the burden of data collection for individual practices.

Publications such as Key Health Statistics from General Practice used primary care data to report national estimates of chronic disease prevalence and management, and thus helped to raise the importance of data collected in primary care.²⁷ More recently, it has been the introduction of a new GP contract and the associated requirements of practices to demonstrate their performance that has raised further interest in using primary care data to help plan, monitor and improve the quality of services.

The requirements of potential users of primary care data will differ. Researchers may request person-based datasets for independent analysis, while clinicians and managers may request information already analysed. Meanwhile the public may demand information that allows them to learn more about the range of services offered, and to compare the quality of care among practices.

The potential of data collected as a by product of the new general medical services contract may help establish a system of surveillance in primary care. However, this is a complex area. There is first a need to understand the limitations of GP contract data and to develop expertise, especially when promoting its use in appraising health policy, service developments and supporting public health activities. In the United States, there are more widespread analyses of administrative databases and surveys, but with QMAS data, which covers approximately 8500 practices, there is great potential to improve the understanding of disease and illness.

Before data held by general practitioners on their computer systems can be used for surveillance, two key issues need to be resolved. First, there is a need to improve the quality of electronic medical records in primary care, particularly in the recording discipline of doctors. Using databases such as GPRD, it is now possible to get good information on chronic disease management at a national level. However, studies that have verified that GPRD data are of high quality are supplied by volunteer general practices and are unlikely to be representative of all UK practices. The recording of data is likely to be far more variable. A study of 47 general practices, all in one PCO in

south-west London showed a wide variation in the coding of patients with ischaemic heart disease.²⁸ No practice had all cases coded (although some achieved coding levels of over 90%) and in some practices, no cases were coded.

One problem that GPs face when improving the quality of their computerised data is the lack of standards in recording morbidity data. Even among practices that have been computerised for a long time, there are wide variations in the coding of clinical information by different practitioners. For example, there are many codes that are used for heart disease. Some researchers have proposed a data accreditation scheme, with standards for completion and accuracy of data, and guidelines for using codes for diagnoses. Standardisation would allow for comparisons of general practices and PCOs, and allow data to be aggregated to produce national estimates. For example, tools such as Morbidity, Information Query and Export Syntax (MIQUEST), a computer program designed to extract information from GPs' computer systems are used to help obtain practice data in a standardised way. MIQUEST has been adopted as a standard within the NHS for data extraction, and practices are now required to be MIQUEST compatible before receiving accreditation. However, there remains in many areas a lack of expertise in writing MIQUEST queries.

The Department of Health introduced several initiatives to overcome some of the problems with data quality in primary care. These include PRIMIS, a no charge support service designed to help primary care organisations improve patient care through the use of their computer systems, and Health Information Project for Coronary Heart Disease, a project to help practices and PCOs measure their progress in meeting the national service framework for coronary heart disease standards. There are also independently funded data quality projects such as Primary Care Data Quality programme. This is an educational intervention that aims to improve data quality by building on existing skills and knowledge within primary care of data recording and the use of technology.

A significant development in the improvement of coding is likely as a result of the introduction of the QOF.

The second issue that requires resolution in setting up a system of primary care surveillance is to reduce the burden of data collected by practices. Decisions need to be made, as to what data should be collected in a primary care data system. This would minimise the development of adhoc databases. A further key approach is to link existing data through data linkage systems. Attempts have been made to do this by the Department of Health with the National Clinical Audit Support Programme for Diabetes project.³¹ However, this would require computer hardware and software systems to be re-designed.

The analyses of primary care data are mainly limited because of issues of data completeness, the difficulty in linking data, and follow up over time (longitudinal study). Statistical analysis of data needs to consider issues such as inter- and intra-practice variation. Although the errors associated with these are limited when using larger databases compared with smaller datasets, expertise is required in working with large primary care databases.

6.1 What is the role of a survey?

Data provided by surveys can help identify specific problems in the delivery of healthcare services or the health status of individuals. Surveys can also be useful for determining patients' view about the care that they receive, the quality of communications between patients and professionals, and in assessing patients' physical and psychosocial function as a result of an intervention. Surveys are cross sectional, and provide information at one moment in time.

Surveys may be used as part of surveillance. However, there are concerns with the accuracy of surveys, the resources that are required to carry them out and the timescales of data collection. Additionally, there are methodological issues with how patients are sampled, and the significance of people who do not respond to the survey and, thus the generalisability of results. These areas need to be considered when using survey data for carrying out surveillance.

6.2 What is the role of a register?

Registration systems can be time intensive and expensive; hence it is important to consider the need for a register. A key component of the QOF and national service frameworks are that practices are required to maintain registers across many disease areas.

Typically a register in primary care contains personal details, such as the patient's address, date of birth and sex, and disease status, details of treatment and outcome, and can be viewed by disease or risk factor status, or by treatment. Data collected by registries in primary care often omit socio-economic and ethnic data information.

Variations in how diagnosis and classification of disease occurs can give rise to problems when comparing data between different countries, between different areas in a country, or between the same populations over time. It is important to validate the information recorded and continually evaluate its quality, as poor quality registers are of little use and may be misleading. But comparing data derived from practice registers, even within the same PCO has been extremely difficult. In many cases, the methods used to produce such registers, including case definitions, were left to individual practices and varied widely. Interpreting the prevalence of chronic disease and treatment rates locally has been complex due to the variability in the completeness and accuracy of coding.

Disease registers are most useful in situations where disease or risk factor status does not tend to change over time; the diagnosis of disease needs to be consistent, and based on a robust diagnostic test. Furthermore, a register is of use when there is a requirement for ongoing health care, for example, retinal screening among patients with diabetes.

Registers can be used for many purposes; to monitor temporal trends, follow up patients, help compare treatment outcomes, help undertake evaluation of services, and study the causes of diseases and help organise services for patients. If case ascertainment is high, prevalence and incidence rates can be calculated and analyses of risks and aetiology can be explored using individual and area characteristics. With follow up data, outcomes such as survival rates for cancer can be measured. If registries are maintained over time, they can produce evidence of change as in epidemics or in the effectiveness of interventions, and be of use in surveillance.

Registers can be used to assist in the management of chronic disease management in clinical settings, help trigger the follow-up care for people with for example, diabetes or asthma within general practice. Registers can form the basis for clinical audit and quality improvement, and for providing services such as child protection registers.

There are several limitations of primary care disease registers. They will not tell you about which patients on the register have died over time, those patients that are newly diagnosed, or whether they were admitted to hospital. The usefulness comes with linking data across more than one registry.

Medical record linkage is the process whereby health records from two or more different sources and containing different types of information are brought together to provide a single file for the individual. For example, linking cancer registration statistics with mortality data enables survival rates to be compared between different groups of people with different types of cancer.

A registry must establish systems to maintain the reliable notification or identification of cases within the studied population, ensure comparability and strict rules in diagnosis. A revision must minimise under-coverage (cases not being included when they should be) and ensure that duplication of cases within the register does not exist, by keeping the register updated, removing those who have recovered, died or moved out of the area.

To date, creating primary care disease registers by searching on coded patients with a disease and drugs was a relatively simple strategy, but it is resource intensive. It is likely that registers need to be coded prospectively using hospital discharge summaries and outpatient letters. There will be a need to factor in the cost of coders at a practice level.

References

- 1 NHS Executive. *General Medical Services (GMS) contract 2004*. London: NHS Executive, 2004.
- 2 Department of Health. *Choosing Health – making healthier choices easier*. London: Department of Health, 2004.
- 3 NHS Executive. *Information for health. An information strategy for the modern NHS 1998-2005*. Leeds: NHS Executive, 1998.
- 4 NHS National Programme for Information Technology. *Making IT happen: Information about the national programme for IT*. Birmingham: NHS Information Authority, 2004.
- 5 Department of Health. *Delivering 21st Century IT support for the NHS: National specification for Integrated Care Records Service- consultation document*. London: Department of Health, 2002.
- 6 Connecting for Health (previously NHS National Programme for Information Technology)
<http://www.connectingforhealth.nhs.uk/>.
- 7 Wanless D. *Securing our future health: taking a long-term view. Final report*. London: HM Treasury, 2002.
- 8 Department of Health. *The NHS Plan. A plan for investment. A plan for reform*. London: Department of Health, 2000.
- 9 Department of Health. *Delivering 21st Century IT support for the NHS. National strategic programme*. London: Department of Health, 2002.
- 10 Home Office. *Freedom of Information Act 2000*. London: Cabinet Office, 1999.
- 11 Data Protection Act 1998. London: Stationery Office, 1998.
- 12 Department of Health. *Reforming NHS Financial Flows: introducing payment by results*. London: Department of Health, 2002.
- 13 Department of Health. *A guide to foundation trusts*. London: Department of Health, 2002.
- 14 Marshall MN, Shekelle PG, Leatherman S et al. Public disclosure of performance data: learning from the US experience. *Quality in Health Care* 2000; 9:53-57.
- 15 Aylin P, Best N, Bottle A, Marshall C. *Monitoring of mortality rates in primary care*. London: Imperial College of Science, Technology & Medicine, 2003.
- 16 Department of Health. *Tackling health inequalities: a programme of action*. London: Department of Health, 2003.
- 17 Hunter DJ, Killoran A. *Tackling health inequalities: turning policy into practice?* London: Health Development Agency, 2003.
- 18 Majeed FA, Chaturvedi N, Reading R, Ben-Shlomo Y. Equity in the NHS Monitoring and promoting equity in primary and secondary care. *BMJ* 1994; 308:1426-29.

- 19 Hippisley-Cox J, Pringle M. Inequalities in access to coronary angiography and revascularisation: the association of deprivation and location of primary care services. *Br J Gen Pract* 2000; 50:449-454.
- 20 McCormick A, Fleming D, Charlton J. *Morbidity statistics from general practice. Fourth national study, 1991-1992*. London: HMSO, 1995.
- 21 Home Office. *Death certification and investigation in England, Wales and Northern Ireland. The report of a fundamental review 2003*. London: HMSO, 2003.
- 22 Majeed A, Bindman AB, Weiner JP. Use of risk adjustment in budget setting and performance measurement in primary care: how it works. *BMJ* 2001; 323:604-607.
- 23 Majeed A, Bindman AB, Weiner JP. Use of risk adjustment in budget setting and performance measurement in primary care: Paper II. Advantages, disadvantages, practicalities. *BMJ* 2001; 323:604-607.
- 24 Controller and Auditor General. *Tackling Obesity in England*. London: National Audit Office, 2001.
- 25 Department of Health. *Health Survey for England*. London: Department of Health, 2000.
- 26 National Diet and Nutrition Survey. (accessed by <http://www.food.gov.uk/science/101717/ndnsdocuments/>).
- 27 Office for National Statistics. *Key Health Statistics from General Practice 1998*. Office for National Statistics: London, 2000.
- 28 Gray J, Orr D, Majeed A. Use of Read codes in diabetes management in a south London primary care group: implications for establishing disease registers. *BMJ* 2003; 326: 1130-1134
- 29 Hammersley V, Hippisley-Cox J, Wilson A, Pringle M. A comparison of research general practices and their patients with other practices – a cross sectional survey in Trent. *Br J Gen Pract* 2002; 52:463-468.
- 30 Kmietowicz Z. UK academics gain free access to database of patients. *BMJ* 2005 331:924.
- 31 The National Clinical Audit Support Programme's National Diabetes Audit is described at URL http://www.icservices.nhs.uk/ncasp/pages/audit_topics/diabetes/default-new.asp

Appendices

1. Routine NHS primary care activity data
2. Royal College of General Practitioners Weekly Returns Service
3. General Practice Research Database (GPRD)
4. National Database for Primary Care Groups and Trusts
5. General Household Survey (GHS)
6. Fourth Morbidity Survey in General Practice (MSGP4)
7. Primary Care Information Services (PRIMIS)
8. Morbidity, Information Query and Export Syntax (MIQUEST)
9. Primary Care Research Networks; Trent Focus
10. Practice based disease registers
11. Practice based health promotion data
12. Prescribing Analysis and Cost (PACT) data
13. Quality Prevalence and Indicator Database (QPID)
14. QResearch
15. Secondary Uses Service (SUS)
16. MEMO
17. The Health Improvement Network (THIN)
18. UK Biobank
19. UK Clinical Research Collaboration

1. Routine NHS primary care activity data

Health authorities and PCOs collect some information on NHS activity in primary care. This mainly concerns details of patients registered with GPs, preventative health screening such as cervical screening and childhood immunisation, and item of service target payments for GPs (for example, to undertake minor surgery). Information is also collected on staff employed in general practice, such as GPs and practice nurses, as part of General Medical Service statistics.

In the 1990s many health authorities began to use routine NHS activity information to produce performance indicators on general practices in their area. These indicators were generally used internally, and when they were supplied to general practices, did not reveal the identity of individual practices. Indicators were rarely made available to the public. In many PCOs this information was combined with other information from age-sex registers and published as primary care indicator packages. Some of these data are now available to primary care organisations, in the national database stored at the National Primary Care Research and Development Centre in Manchester.

Most recently, the NHS now collects and publishes information on general practices collected as part of the new GP contract (described in Chapter 4 of this report)

2. Royal College of General Practitioners Weekly Returns Service

The Birmingham Research Unit of the Royal College of General Practitioners (RCGP), funded by the Department of Health is responsible for the Weekly Returns Service, which has operated continuously since 1964 and is now a fully automated system. The Weekly Returns Service can provide weekly information on the total rates of consultation with GPs by age and sex, as well as weekly consultation rates for any disease. The system facilitates the estimation of annual consultation rates, and since 2001 has also provided annual prevalence data. For this purpose prevalence is determined by the need to consult for a particular problem or group of diseases, classified according to the structure of the International Classification of Diseases, during the course of the year.

The research unit manages a network of about 80 to 90 general practices in England and Wales. Each week data are sent electronically from practices to the research unit. The data sent provides the number of episodes of care and information on practice populations. An episode of care represents one consultation between a patient and their general practitioner, and is recorded as first, new or ongoing. The mean weekly consultation rates per 100,000 population are calculated based on first and new episodes. The research unit also publishes mean weekly incidence for selected respiratory and communicable diseases and several symptom complexes (including flu-like illness) compared to a background over the previous 10 years. Other published statistics include information on chronic diseases such as eczema, depression and heart failure. The Weekly Returns Service could contribute to public health surveillance, by providing information on illnesses and symptom complexes that may mark bio-terrorist activity.

3. General Practice Research Database

During the late 1980s VAMP Health started to install computer systems and practice management software in GP surgeries throughout the UK. The aim was that data collected on these computers could be used for practice administration and research. The General Practice Research Database (GPRD), owned by VAMP Health was taken over by Reuters Plc, and subsequently given to the Department of Health. The Office for National Statistics (ONS) operated the database on behalf of the DH between 1994 and 1999. From 1999, the responsibility for the database was with the Medicines Control Agency (MCA).

The GPRD contains information entered by GPs onto their practice computers; the number of practices varies but is typically around 300. Data collection is available from 1987 for a limited number of practices. Most of the general practices that participate in the GPRD provide data of sufficient quality for its use in research; an analysis by the ONS showed that the 211 general practices passed all data quality checks during the period 1994 to 1998 inclusive.

There are several key strengths of GPRD data. First, it is population based, the population is large enough to study rare diseases, information on preventive care and secondary care are recorded, and finally there is access to original records. The GPRD has been mainly used for disease epidemiology, pharmaco-epidemiology and health services research. A limitation to using the GPRD has been the cost in accessing data but a new licence paid for by the Medical Research Council will fund 50 projects to use the database free of charge each year for five years, and will also pay for extra staff and a support service.³⁰

4. National Database for Primary Care Groups and Trusts

This database was developed by the National Primary Care Research and Development Centre in Manchester and the University of Manchester's Department of Geography, as a resource for primary care researchers, practitioners and managers. The database links information on population characteristics, health service provision, and health status for all the PCOs in England. Socio-economic and demographic data derived from the 1991 and 2001 Censuses are linked to information on the characteristics and activities of general practice from the GMS statistics. It provides a national tool for monitoring and evaluating the performance of PCOs over time. Data are available to registered users from NHS organisations and the academic community.

At present the database does not include information on prescribing, hospital referrals or hospital admissions data. Once these are added the database will become a powerful tool for monitoring PCOs, and in examining the links between population and patient characteristics, and health outcomes. Additional datasets are proposed and include: the composition and organisation of PCO boards and their budgets, information about the quality of care and prescribing, information on measurable health outcomes such as morbidity, mortality and health related behaviours such as smoking, local authority data on community and social services, health authority derived secondary health care information.

5. General Household Survey

The General Household Survey (GHS) is the main source of information on activity in general practice. Respondents of the survey are asked if they have made contact with their GP in the previous two weeks, the type and frequency of contact, such as a home visit or a telephone consultation, and with whom, either a GP or practice nurse. GHS is widely acknowledged, as useful in examining the rates of consultations with GPs by patient age, and in providing a snapshot of annual workload. But data from GHS does not accurately monitor trends in GP consultation rates.

The University of Kent calculates for the Department of Health the cost of a consultation. It takes into account the average consultation time and travel time, using 1992 data. Databases such as QResearch (see Appendix 14) may be able to provide more information on the types of consultation than the current GHS, once data validity is verified. The QResearch database covers 6% of all patients registered with GPs, and nearly 500 practices spread throughout the UK. The sample data can be re-weighted to make it more representative of the UK population.

6. Fourth Morbidity Survey in General Practice

The Fourth Morbidity Survey in General Practice (MSGP4) was a prospective cohort study of around 500,000 patients (1% of the general population) registered with 60 volunteer general practices in England and Wales, undertaken in 1991-92. Earlier surveys were carried out in 1952, 1971-72 and 1981-82. The aim of the survey was to examine the workload and pattern of disease in general practice in relation to patient age, sex and socio-economic status. Many of the practices contributing to MSGP4 continue to provide data for the RCGP's Weekly Returns Service, which observes the same recording discipline.

Before the survey started, doctors and staff from each practice attended three two day training sessions on how to record morbidity data. Practices then collected data for two to four weeks before the start of the survey. These data were analysed and any errors or inconsistencies reported back to the practices. Once the morbidity survey started, general practitioners and nurses recorded information on all face to face contacts with patients. Each reason for consulting and the place of contact was directly entered into patient records on the practice computer. Every consultation was assigned an ICD-9 (International Classification of Disease Ninth Revision) code. When patients presented with more than one problem, doctors were asked to record a separate ICD-9 code for each problem. Although the number of diagnoses recorded was greater than the total number of contacts with general practitioners, the vast majority of contacts were for one problem only. Data supplied by the practices were subject to regular checks to ensure its validity.

The practices that took part in the survey compared with the average general practice were bigger, were more likely to be computerised and to show a greater interest in the collection of morbidity data. However, the sample of patients was representative of the population of England and Wales for characteristics such as age, sex and social class.

The key strength of MSGP4 is that it provides information on consultation rates and disease patterns by socio-economic and ethnic group. For this reason, despite the length of time that has elapsed since the 1991-1992 survey was carried out, interest in its findings remains high. But its relevance is waning.

7. Primary Care Information Services

Primary Care Information Services (PRIMIS) originated from a pilot project called 'Collection of Health Data from General Practice'. By April 2000 PRIMIS was more widespread. PRIMIS was funded by the NHS Information Authority. It is now managed by *Connecting for Health*. The service is led by the Division of General Practice at the University of Nottingham. The aim of PRIMIS is to provide education, training, and analysis to local facilitators, to help them assess the data quality of practice systems and to assist practices in using clinical computer systems. PRIMIS works with over 200 local information facilitators in England, who cascade their knowledge and skills to around 3000 practices.

Data recorded, as part of routine clinical activity, are extracted from the computer systems in general practice using MIQUEST queries. The queries are designed to examine data quality including completeness, accuracy and timeliness. The analyses are fed back by facilitators to each

practice, to help practices examine and improve their data quality and clinical practice. PRIMIS offers a Comparative Analysis Service that provides clinically focussed analysis and feedback on key areas, such as heart disease, asthma, epilepsy and severe mental illness.

8. Morbidity, Information Query and Export Syntax

Morbidity, Information Query and Export Syntax (MIQUEST) was developed between 1992 and 1994; a project jointly funded by the UK NHS Executive Information Group and the former Northern Regional Health Authority. The aim of the project was to develop a Windows software system to extract and aggregate data from different GP computer systems. Currently MIQUEST is the standard within the NHS for data extraction accreditation.

MIQUEST has a very strong security mechanism for protecting patient confidentiality. An external enquirer (that is anyone outside the practice) using MIQUEST is unable to extract data with strong patient identifiers from a practice system. PRIMIS facilitators are able to extract data authorised by practices, following the signing of a confidentiality agreement between practices and their PCO. The practice controls the access and release of data extracted using MIQUEST; query sets can only be run with a practice's approval and response files can only be transmitted externally, with their express approval.

PRIMIS uses MIQUEST, as the preferred approach in supporting analyses of data quality in practices and extracting data for comparative analyses on clinical topics, such as coronary heart disease and diabetes mellitus. It is also used by Trent Focus and by some other research groups and networks. The use of MIQUEST is part of the free training that PRIMIS provides to information facilitators, who are employed by PCOs to help practices make the best use of their clinical computer systems. The facilitators will usually run the first MIQUEST data extraction in each practice and then train a practice staff member to run subsequent extractions themselves. However, many practices have been reluctant to allocate staff time to MIQUEST training and data extraction.

9. Primary Care Research Networks

Primary care research networks consist of professionals interested in undertaking research in primary care. Most networks aim to support new researchers, help with project management, assist researchers in writing articles for publication, and fund more established researchers.

The UK Federation of Primary Care Research Organisations was established in 1998 to bring together all primary care research networks in the UK. The aim of the federation is to promote learning across networks and to secure the long-term future of networks. Any primary care research network within the UK that supports the aim of the federation may join. An example of a primary care research network is Trent Focus.

Trent Focus

Trent Focus is a Collaborative Research Network set up in 1994, funded by Trent Regional Health Authority, with the aim to improve research knowledge and skills in primary care. A key objective of Trent Focus was to establish a network of general practices to take part in practice based research.

The Network Board approves all research studies. Initially 72 practices were recruited, but some practices dropped out due to problems with committing time to research. Practices are paid £1,000 to collaborate in a research project at least once a year. All practices had to reach a specified standard in their recording of diagnostic, lifestyle, prescribing and referrals data. The accuracy and completeness of the clinical databases in these studies has enabled cross sectional and case-control studies to be undertaken. Although practices involved in Trent Focus are representative of other practices in Trent in terms of population and morbidity, this may not apply to the rest of England.²⁹

10. Practice based disease registers

Before 1999 and the introduction of national service frameworks across clinical areas (the first for coronary heart disease), the use by practices of disease registers was sporadic. The introduction of national service frameworks led to a widespread adoption of disease registers in primary care. Practices had to identify the number of patients registered with the practice with a particular condition, for example, coronary heart disease. This was not as simple as it was first made out. Patients had to be coded with a diagnosis, but many patients were either not diagnosed or inaccurately diagnosed.

The data for such registers tends to be available most readily from large, well-organised general practices that are highly computerised. Therefore, they may not be representative of all practices in the locality. The introduction of disease registers was one of the first steps towards chronic disease surveillance in primary care, by providing information on the prevalence of disease in the population. But setting and maintaining a disease register is also very resource intensive.

For the present, however, major obstacles remain in trying to use this data obtained from general practice computer systems. The principal limitation is that not all GPs are currently recording details of their consultations, or other encounters between patients and the NHS on their practice computers. Even where GPs are working in paperless practices (recording all clinical and administrative information about patients), there is no standard method for reporting on the quality of the data recorded. Hence, the accuracy and completeness of the data may vary widely between practices. Problems also remain in trying to extract comparable data from the different computer systems currently on the market. Furthermore, practices do not uniformly collect socio-economic and ethnicity data, or risk factor data.

Practices need to meet specific standards for recording information on morbidity and healthcare use. In the longer term, the number of practices that can supply high quality information could be increased by better integration of information systems between primary and secondary care (reducing the need to record data more than once), clear guidelines on recording data, regular feedback on data quality and incentives for practices to become fully computerised.

11. Practice based health promotion data

The development of NSFs and the new GMS contract has placed a greater emphasis on recording data on risk factors of disease, such as smoking and obesity. 'Choosing Health' documented an epidemic rise in obesity levels and the implications for public health. Determining the extent of

particular lifestyles or risk factors within a population is an important aspect of assessing population health need. As with morbidity data, the sources of data on lifestyle are very disparate in primary care. The main source of data is from national surveys, which are then applied to local populations.

12. Prescribing Analysis and Cost (PACT) data

Prescribing Analysis and Cost (PACT) data are a by-product of reimbursing pharmaceutical contractors. Information on all dispensed NHS prescriptions is collected by the Prescribing Pricing Authority (PPA). The information collected includes the name and cost of the drug, and the number of items dispensed. The drugs dispensed are then used to calculate the cost of each item. Drugs are categorised by the section of the British National Formula that they fall in. Information is available for individual drugs, for categories of drugs such as bronchodilators and for therapeutic areas such as respiratory drugs.

Analysed information is fed back at practice, PCO, SHA and national level. But users do not have access to local data. As prescribing is heavily influenced by practice demography, data are adjusted for age and sex to compare the rates and costs of prescribing in different practices or health authorities.

GPs generally receive a standard PACT report that contains information on the practice's rates and costs of prescribing along with comparative information. GPs can ask for more detailed information but this report is often very unwieldy. PCOs receive PACT reports on GPs' prescribing and can access computerised PACT data through PACTLINE; 'epact.net'. PCO pharmaceutical advisers can obtain information on the prescribing of specific drugs online directly from the PPA.

PACT data have been used mainly for two purposes. First, as a financial tool to help health authorities set and monitor prescribing budgets in general practice. Second, in allocation formulas for prescribing budgets by the Department of Health. Before the new GMS contract and its allocation formula, health authorities had to use their judgement in allocating the prescribing budget for practices. This may in part have led to the development of practice formularies. PACT data have also been used in audit and research.

But there are limitations to PACT data:

- They provide information mainly on what drugs are prescribed and how much the prescribed drug costs.
- They cannot be linked to demographic or clinical data on patients.
- They cannot be used to calculate age and sex specific prescribing rates.
- They cannot be used to look at prescribing rates for specific conditions.
- PACT does not include private prescriptions, or prescriptions, that the patient does not have dispensed.
- The number of items prescribed is not always an accurate measure of the amount of a drug actually prescribed.

It may be possible to expand the use of PACT data by including diagnostic data on the prescription issued.

13. Quality Prevalence and Indicator Database

The Quality Prevalence and Indicator Database (QPID) is being developed by the new Health and Social Care Information Centre in Leeds. The aim of QPID is to widen the access to QMAS data by users in the Department of Health and NHS. QPID (and the Information Centre) aims to provide analytical support and undertake national analysis on prevalence using QMAS data and other data sources. They are to publish data in readiness of the freedom of information legislation.

To access the QPID data users will have to inform the QPID team of the proposed use of data or information from QPID, which is then considered for approval by the QPID Gateway Committee. Approved requests lead to detailed user specifications. The QPID team and the user will sign an agreement on the provision and use of data.

14. QResearch

The aim of QResearch is to develop and maintain a high quality database of general practice derived data for use in medical research. QResearch now contains data from 468 general practices in the UK, with records for 3.3 million current patients and 4 million past patients. The database includes socio-economic details such as patient's postcode. But the data extracted will contain no strong patient identifiers. Patients can opt out. When it is fully established the QResearch database will be one of the largest aggregated databases containing records for nearly 8 million people.

The database will be open to researchers with ethical committee approval, and information will be provided to answer their research question only (not the whole dataset) either at patient or practice level. The costs of using the data will be carefully controlled, to allow the scheme to be self-funding, but allow good access to academic researchers. Analyses will be undertaken to demonstrate the accuracy and completeness of the data and will be made available for morbidity analyses.

QResearch also has links with EMIS, who supply clinical systems to just over half of all general practices. There has been a recent joint initiative to provide information through development of a database on the number of people with flu. By examining regional differences, general practices and other parts of the health service will be aware of epidemics in time to introduce preventative measures.

15. Secondary Uses Service

The Secondary Uses Service (SUS) is part of the NHS Care Records Service (NCRS). The service aims to provide pseudo-anonymised patient data, for purposes defined as other than direct clinical care. This includes planning, commissioning, public health, clinical audit, benchmarking, performance improvement and clinical governance. The aims of SUS are to reduce the burden on the NHS of collecting, abstracting and submitting data. Furthermore, to provide national comparators, increase timeliness and quality of data available to the NHS and its partners and access to national, transparent data from the NHS and other organisations e.g. ONS. As a result of better more timely public health information the quality of care for patients can be improved.

Initially SUS will take on the data flowing through the current NHS-Wide Clearing Service (NWCS) and support the implementation of Payment by Results. Over time other data sources will be included, including cancer waiting times, clinical audit information, and central returns. It will also include data from non-patient based sources (e.g. ONS, workforce data) and collect data from the NCRS patient records right across the care pathway. The service will also generate some central returns and on-line access to analytical tools and services will be also be available for research.

The information within SUS will have rigorous access controls, providing online access to the NHS and to other related agencies and health organisations. SUS will provide facilities for quality reporting, standard analyses, user analysis, and extraction.

16. MEMO

MEMO was originally set up in Tayside in Scotland to undertake pharmaco-vigilance studies using datasets on dispensed prescribing, hospitalisation, and death certification. MEMO is able to record link a wide range of primary care, secondary care and pharmacy datasets. This information, although collected in Scotland, may be applicable to improving health services and public health in England, as the demographics of the Tayside population are representative of the UK and Europe.

17. The Health Improvement Network

EPIC and In Practice Systems, two commercial organisations, have developed 'The Health Improvement Network' (THIN) primary care database to facilitate the use of NHS electronic databases of primary care records for medical and pharmaceutical research. It is a new database of anonymised clinical records. EPIC has a long history of using electronic medical records in research and In Practice Systems in developing and supplying the Vision general practice computer system. Data collection started in January 2003. Over a hundred practices using Vision software have joined THIN. Most practices have recorded several years of data on their system and where contributing practices had previously used VAMP systems, data entries extend to 1987.

An independent research panel will review all research proposals and THIN data will be supplied as flat text files, including data quality indicators to inform researchers of the completeness of data recording in practices. Additional information will also include the geographical location of practices by health administrative area, and socio-economic indicators at individual patient level.

Practices are trained to record information and data collected from practices are subjected to a continuous programme of data quality control. Recording is assessed against various quality indicators and national statistics are used for comparison. THIN quality standards cover a wide area of clinical recording including: asthma, coronary heart disease, diabetes mellitus, epilepsy, menopause, hypertension, hypothyroidism, leg ulcers, heart failure, warfarin use, lithium use, oral contraceptive use, pernicious anaemia, rheumatoid arthritis, secondary stroke prevention, lower back pain, mental health, and smoking. Practices will receive quarterly reports measuring performance in specific clinical areas.

18. UK Biobank

UK Biobank is a project currently under way, which will follow the health of 500,000 volunteers aged 40-69 in the UK for up to 30 years. Its aim is to build the world's largest information resource for researchers to develop new and better ways of preventing, diagnosing and treating common illnesses such as cancer, heart disease, diabetes and Alzheimer's disease. It is funded by the Medical Research Council, The Wellcome Trust, the Department of Health and the Scottish Executive.

The project will provide information about volunteer's lifestyle such as exercise patterns and diet, employment history, and their medical and family history along with samples of their blood and urine. Simple measures such as weight and blood pressure will also be recorded. People are randomly selected via health registers and asked if they are willing to participate. This information will be linked with the participants' medical records so that researchers can study the links between the participants' genes, lifestyle, environment and the diseases and conditions they may develop.

19. UK Clinical Research Collaboration

The UK Clinical Research Collaboration (UKCRC) is a partnership of organisations with the aim to establish the UK as world leader in clinical research. The partnership is represented by the main funding bodies for clinical research in the UK, academic medicine, the NHS, regulatory bodies, industry and patients (See *Igniting our potential – An introduction to the UK clinical research network*, available at URL <http://www.ntrac.org.uk/News/Summary/UKCRC%20Introduction%20leaflet%20Dec-04.pdf>). The NHS Confederation is setting up a health services research network, which has been supported by the Department of Health, the NHS Service Delivery and Organisation Research and Development Programme, the Health Foundation and the Nuffield Trust. All these clinical research collaborations will require data from primary care to develop sampling frames for clinical trials, to provide follow up data on patients in trials, and data for other types of research.

Abbreviations

ACG	Adjusted Clinical Groups
BMI	Body Mass Index
COPD	Chronic Obstructive Pulmonary Disease
DH	Department of Health
DRG	Diagnostic Related Group
GHS	General Household Survey
GMS	General Medical Services
GPRD	General Practice Research Database
HES	Hospital Episode Statistics
HRG	Health Related Group
HSE	Health Survey for England
ICD	International Classification of Diseases
LSP	Local Service Provider
MCA	Medicines Control Agency
MIQUEST	Morbidity, Information Query and Export Syntax
MSGP4	Fourth Morbidity Survey in General Practice
NCRS	National Health Service Care Records Service
NHAIS	National Health Applications and Infrastructure Services
NPC	National Prescribing Centre
NPfIT	National Programme for Information Technology (now <i>Connecting for Health</i>)
NSF	National Service Framework
ONS	Office for National Statistics
PACT	Prescribing Analysis and Cost
PCO	Primary Care Organisation
PPA	Prescription Pricing Authority
PRIMIS	Primary Care Information Services
PSU	Prescribing Support Unit
QMAS	Quality Management and Analysis System
QOF	Quality and Outcomes Framework
QPID	Quality Prevalence and Indicator Database
RCGP	Royal College of General Practitioners
SHA	Strategic Health Authority
SUS	Secondary Uses Service
THIN	The Health Improvement Network
UKCRC	United Kingdom Clinical Research Collaborations
VS	Vital Statistics



Eastern Region Public Health Observatory

Institute of Public Health
University Forvie Site
Robinson Way
Cambridge
CB2 2SR

t: 01223 330353

f: 01223 330345

e: enquiries@rdd-phru.cam.ac.uk

www.erpho.org.uk

January 2006

ISBN 1 904389 09 0

Typeset and printed by Piggott Black Bear, Cambridge, UK.