Informatics in Primary Care 2009;17:137–9

Editorial

Using patient records from general practice for research

Nigel Mathers BSc MD PhD MRCGP
Professor of Primary Medical Care, University of Sheffield, UK

Nicola Perrin MPhil MSc BA (Hons)
Senior Policy Adviser, Strategic Planning and Policy Unit, The Wellcome Trust, UK

Graham Watt MD FMedSci FRCGP FFPH FRCP (Glas)
Professor of General Practice, University of Glasgow, UK

General practice records are a unique source of information which can help us as medical researchers to improve our understanding of disease, develop potential new treatments and improve the care of our patients. The use of general practice electronic records is increasing and until recently there was little consensus on how such records could be accessed and used for research. The recently published document by the Wellcome Trust[1] provides guidelines for best practice in the use of electronic patient records for research and is the result of a national consensus meeting held in 2008 with general practitioners (GPs), researchers and patient groups. This report is very timely and proposes three overarching principles as a basis for best practice. One of these key principles is that personal information held within patient records may be both sensitive and private so security and confidentiality must be safeguarded at all times if the general public, patients and healthcare professionals are to have confidence in the processes used by researchers to access the records.

Patient **confidentiality** can be maintained at a technical level by the use of the best available electronic technologies which ensure security and confidentiality. The introduction of **safe havens** and **honest brokers** provide further mechanisms for maintaining the confidentiality of data. **Safe havens** are defined as designated physical or electronic areas that provide the most appropriate level of security for the use of data. Researchers working in safe havens should be 'bound by a strict code, preventing disclosure of any personally identifying information' (p.14). The UK Government in its response to the Data Sharing Review[2] has accepted a recommendation for the development of safe havens to minimise the risk of individuals being identified and has committed itself to developing a system to ensure that only 'accredited researchers' work within safe havens.

An **honest broker** is a trusted custodian of data, with the dual role of 'ensuring patient data confidentiality and security and ensuring scientific integrity of data' (p.14), i.e. he/she is responsible for ensuring that the coding and anonymisation processes are correctly implemented and for carrying out data quality checks that, for reasons of confidentiality, it is not possible for researchers to do themselves.

The consensus document distinguishes between three levels of identifiable information: the first level is **anonymised** (other terms which may be used are: irreversibly de-identified; unlinked anonymised information and unidentifiable). At this level, it is not possible to identify an individual because, although data are provided at an individual level, there is no way of establishing a link with the original, identifiable clinical record. Data at this level do not include identifiers such as name, address, full postcode, full date of birth or NHS number. (These are so-called strong identifiers.)

The second level of identifiable information is **coded data** where it is not possible to directly identify an individual but where a key is available which enables the identity of the patient to be linked to the data by the person who holds the key. This coded identifier should be globally unique and the key needs to be held under strict conditions. Alternatively, the data may become identifiable when used together with other data sources. The likelihood of such identification is increased when data relates to individuals with rare illnesses or exposures, or to small or unusual populations. (Other terms used for coded data include: pseudonymous; key-coded; reversibly de-identified; linked anonymised; masked and encrypted.)

The third level of identifiable information is, of course, any **personal data** that directly identifies individuals. (The other terms which may be used at

this data level include: identified; personal and nominative.) In these circumstances, individual informed consent is normally required before identifiable data can be used for research. However, in England special permission may need to be obtained where it is not possible or practical to seek consent, in which case the Ethics and Confidentiality Committee of the National Information Governance Board for Health and Social Care would consider such applications.[3]

In addition to these technical safeguards, patient confidentiality also needs to be controlled at the researcher level by ensuring that only accredited, approved researchers have access to identifiable patient information. The report recommends that researchers are placed under the same duty of confidentiality as health professionals and that appropriate and substantive, possibly criminal, sanctions should be applied for any breaches of confidentiality.

There are considerable implications for us as researchers arising from this guidance. It is clear that academic researchers need to work closely with GPs and healthcare professionals who retain ultimate responsibility for ensuring that data are accessed appropriately. The Wellcome Best Practice Guidance provides a framework for a greater mutual understanding of the different roles and responsibilities of clinicians and researchers. However, although the public are generally supportive of research and two-thirds of people are likely or certain to allow personal health information to be used for research, there is little public understanding of what this actually means in practice. A national awareness raising programme is, therefore, recommended by the report to highlight the importance of using patient records for research, describing the difference between identifiable and non-identifiable data and explaining the safeguards which are in place to protect privacy. Some controversy remains, however – particularly in the areas of 'opting in/opting out' and 'consent for consent'. The report recommends that patients should be informed on a number of occasions, such as when registering with a general practice, that they can if they wish opt out of the use of their identifiable information in research. Other opportunities for informing patients about opt-out could include prominently displayed notices in waiting rooms, and there is a need to develop a process whereby dissent to research can be recorded within individual medical records. Informed consent is, of course, required for participation in individual research projects when identifiable data are to be used – a detailed account for best practice in these circumstances is provided by the Wellcome report.

When patients are to be invited to take part in research, greater clarity is also needed about the mechanism for contacting potential study recruits. GPs are sometimes required to contact patients in the first instance to ask whether they are happy to be contacted at a later time with information about a study. Only after this initial contact can researchers contact patients to invite them to participate in the study. The Data Sharing Review report[4] described this need for consent to gain consent as a 'problem that requires a solution' (p.27). It is clear that response rates may be higher when patients are invited by their GPs to participate in a study: possible reasons for this include the high levels of trust in GPs consistently expressed by the public and the high status given to a GP's endorsement of a particular study. However, GP involvement in contacting patients may require significant time and resources which can be a substantial barrier to conducting research.

It is clear that the Wellcome report is a big step in the right direction for researchers wishing to use data from general practice patient records for research and the report demonstrates clearly that there is a growing consensus on best practice between GPs, patients, the public and researchers, although a number of outstanding issues remain. The full report can be accessed on www.wellcome.ac.uk/gprecords.

REFERENCES

1 Wellcome Trust. *Towards Consensus for Best Practice: use of patient records from general practice for research.* London: Wellcome Trust, 2009. www.wellcome.ac.uk/gprecords
2 Ministry of Justice. *Response to the Data Sharing Review Report.* London: Ministry of Justice; 2008. www.justice.gov.uk/publications/docs/response-data-sharing-review.pdf (accessed 1 June 2009).
3 The Legal Framework: the Data Protection Act (1988), the Human Rights Act (1998), the Common Law of Confidentiality, Section 60 of the Health and Social Care Act (2001) and Section 251 of the NHS Act (2006).
4 Thomas R and Walport M. *Data Sharing Review Report.* London: Ministry of Justice, 2008. www.justice.gov.uk/reveiwsdocs/data-sharing-review-reportdoc.pdf (accessed 1 June 2009).

ADDRESS FOR CORRESPONDENCE

Nigel Mathers
Academic Unit of Primary Medical Care
University of Sheffield
Samuel Fox House
Northern General Hospital
Herries Road
Sheffield S5 7AU
UK

Accepted October 2009