REPORT

TOWARDS CONSENSUS FOR BEST PRACTICE

Use of patient records from general practice for research

June 2009

wellcome trust

TOWARDS CONSENSUS FOR BEST PRACTICE:

Use of patient records from general practice for research

Published by the Wellcome Trust, June 2009

Main contributing authors
Professor Nigel Mathers, Professor of Primary Medical Care, University of Sheffield
Professor Graham Watt, Professor of General Practice, University of Glasgow
Nicola Perrin, Senior Policy Adviser, Wellcome Trust

Citing this report
We suggest that you cite this report as follows:

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 [accessed 1 June 2009].

If you have any queries about this report, please contact:

Strategic Planning and Policy Unit
Wellcome Trust
Gibbs Building
215 Euston Road
London NW1 2BE, UK
T +44 (0)20 7611 8646
F +44 (0)20 7611 8742
E sppu@wellcome.ac.uk

www.wellcome.ac.uk/GPrecords

TOWARDS CONSENSUS FOR BEST PRACTICE:
Use of patient records from general practice for research

OVERVIEW

Research is a core part of the NHS, as highlighted both in the 60th anniversary celebrations and the new NHS Constitution (Department of Health, 2009). Such research, "to improve the current and future health of the population", relies heavily on the information contained in patient records held in general practice. But patient information is both sensitive and private. The general public, patient groups and individuals must have confidence that the security and confidentiality of personal information are protected, and that appropriate procedures are in place to safeguard data.

There is considerable uncertainty, and a lack of consistency, about the processes that should be used when information from patient records is required for research. The purpose of this document is to provide guidance for best practice when general practice patient records are used for research. It is primarily intended for GPs and responsible healthcare professionals, but the guidance will also be useful to researchers, patient groups, health informatics professionals and advisory bodies. It is intended as the first step in a process to ensure that patients and GPs have confidence in the processes used to access patient information.

The guidance was developed during discussions with GPs, researchers and patient groups, hosted at the Wellcome Trust in May 2008. The document has been endorsed by the Royal College of General Practitioners, the British Medical Association, the HODS group (heads of departments of general practice and primary care in the UK and Ireland), the Health Improvement Network (THIN), QResearch, the Research Capability Programme in NHS Connecting for Health/National Institute for Health Research, the Society for Academic Primary Care and the Wellcome Trust.

The guidance is based on three overarching principles:

- **The overriding importance of safeguarding patient confidentiality and privacy** and the need to clearly define the processes and procedures for the use of data. Wherever possible, the best available technologies should be used to improve security and enhance confidentiality. Where researchers are to have access to identifiable patient information, there must be mechanisms of accreditation and accountability. A formal process of accreditation should be introduced that places researchers under the same duty of confidentiality as a health professional. This should also define appropriate and substantive sanctions so that researchers may be held to account should there be any breaches of confidence.
- **GPs and healthcare professionals should play the role of patient's advocate.** The first priority of GPs must always be to deliver high-quality healthcare but the GP must also protect patients if patient records are to be used in research. The GP and the practice must retain ultimate responsibility for ensuring appropriate access to data. The GP should also be in a position to provide advice to patients about taking part in a trial, and the research process. It is recognised that where a GP endorses a study, this increases patient trust in that study.
- **The need to improve public awareness and understanding** about the use of patient records in research. We recommend that information should be provided through two routes: a national programme to raise awareness about both the importance of research and the need to use patient records to inform that research; and the provision of generic

information locally through individual general practices. Transparency is essential: the information should make it clear that patients can opt out of the use of their data in research if they wish. Care should never be conditional on the willingness to take part in research.

This document provides guidance for best practice, focusing on the two main ways that information from patient records may be used in research. First, research may depend on examining the information contained in existing patient records, without needing direct patient participation. This may use data that are anonymised or coded. Different levels of anonymity offer different risks and benefits, and this document sets out guidance for the uses of different types of data. It is important to note that researchers are usually not interested in the identity of an individual person but may want to have data at an individual person level (rather than aggregated, anonymised data for a large number of patients).

Alternatively, patient records may be used as a starting-point to identify potential participants for clinical trials or interventional research who are then contacted to seek consent to participate. This document provides guidance as to how records should be reviewed to identify potential participants, who should contact patients to invite them into a study, and how best to seek consent.

Most general practices now keep patient records electronically. The use of electronic records opens up the possibility of conducting studies that require access to very large quantities of patient data, to address research questions that would previously have been unanswerable. The best practice guidance described in this document should enable patients and GPs to have confidence in the processes used for accessing information, and enable everyone to benefit from the huge research potential of medical records.

INTRODUCTION

1. General practice is the primary route through which the National Health Service (NHS) provides comprehensive healthcare to the UK population, 'from the cradle to the grave'. Nearly all people in the UK are registered with a general practice, and there are high levels of contact, continuity and trust between patients and general practitioners (GPs). Patient records in general practice are therefore a unique and valuable resource.

2. In addition to being a vital tool for the delivery of healthcare to individual patients, information in patient records holds significant potential for research. Research evidence is needed to inform decisions in general practice and primary care, improve understanding of disease, and evaluate existing and new interventions. Because data held throughout the NHS are representative of the whole population, and generalisable samples are available, research findings can be relevant to a wide number of people. Alternatively, it is possible to use the data to identify sub-populations with specific conditions.

3. However, patient records, at the point of care, contain personal information; some medical details may be considered particularly sensitive. The general public, patient groups and individuals must have confidence that the security and confidentiality of information are protected, and that appropriate procedures are in place to safeguard data.

4. The purpose of this document is to provide guidance for best practice when general practice patient records are used for research. It is primarily intended for GPs and responsible primary care professionals, who will expect researchers to commit to the processes outlined in the document. The guidance will also be useful to researchers, patient groups, health informatics professionals and advisory bodies. It is the first step in a process to ensure that patients and GPs have confidence in the processes used to access patient information.

5. The guidance was developed during discussions with GPs, researchers and patient groups, hosted at the Wellcome Trust in May 2008 (see Annexe A for a list of participants), when delegates reached consensus about a number of key issues. The document has been endorsed by the Royal College of General Practitioners, the British Medical Association, the HODS group (heads of departments of general practice and primary care in the UK and Ireland), the Health Improvement Network (THIN), QResearch, the Research Capability Programme in NHS Connecting for Health/National Institute for Health Research, the Society for Academic Primary Care and the Wellcome Trust.

6. The document begins by providing background information about the different ways in which patient records may be used in research, and the type of information that might be used. It then highlights a number of overarching principles that were agreed during the meeting:
   - the overriding importance of safeguarding patient confidentiality and privacy, and the need to clearly define the processes and procedures for the use of data
   - the role of the GP as the patient's advocate
   - the need to ensure transparency, and to improve public awareness and understanding about the use of patient records in research.

7. The main focus of the document is to provide guidance for best practice, including discussion of several specific issues that were identified as needing particular clarification:
   - the need to distinguish between the use of identifiable and non-identifiable information
   - how patient records may be used to identify potential study participants, including when researchers outside the care team may be involved in accessing patient records

- who should contact patients to invite them into a study
- when and how patient consent should be obtained
- appropriate safeguards to maintain confidentiality of patient records throughout the research process.

1. BACKGROUND INFORMATION

1.1 The use of patient records in research

8. Research underpins modern medicine, treatment and healthcare. Examples of research studies carried out using information in medical records include:

   - audit, to inform healthcare management[1]
   - surveillance, for example detecting and monitoring trends in infectious diseases (see Box 1 below)
   - monitoring safety and efficacy of prescribed drugs (pharmacovigilance), vaccines and devices used in healthcare
   - epidemiological research, for example longitudinal studies to assess the health consequences of exposure to occupational, environmental or other risks
   - clinical trials or research to evaluate new drugs, treatments and interventions.

9. Different types of research use patient records in different ways. The first four types of study listed above depend on examining the information contained in existing patient records, without needing direct patient participation (discussed further in section 3.2). Patient data may be anonymised or coded but may sometimes need to be directly identifiable (see section 1.2). The final type of research – clinical trials or different types of interventional research – uses patient records as a starting-point for participatory research. In this case, patient records are used at the preliminary stages of a study to identify potential participants who are then contacted to seek consent (see section 3.3). This document sets out guidance for the different uses of patient data.

10. Most general practices now keep patient records electronically. The use of electronic records opens up the possibility of conducting studies that require access to very large quantities of patient data, to address research questions that would previously have been unanswerable. Connecting for Health (CfH), which has been set up to deliver the NHS National Programme for IT (NPfIT), includes a Research Capability Programme that offers significant additional potential for such secondary uses of patient information. Egton Medical Information Systems (EMIS) and other providers of GP electronic systems are also starting to set up shared systems locally. Secure systems to facilitate access to, and linkage of, electronic records will be highly beneficial – both to improve patient care and public health, and for research to advance medical knowledge.

11. Researchers with an interest in carrying out studies using information from patient records might be part of the health service, but may also be employed by, or affiliated with, academic institutions or public research organisations. Alternatively, they may be part of commercial companies, for example the pharmaceutical industry or companies, such as Dr Foster, which access large volumes of NHS data for research purposes. It is suggested that the guidance in this document should apply equally to any users of such data, whether academic or commercial researchers.

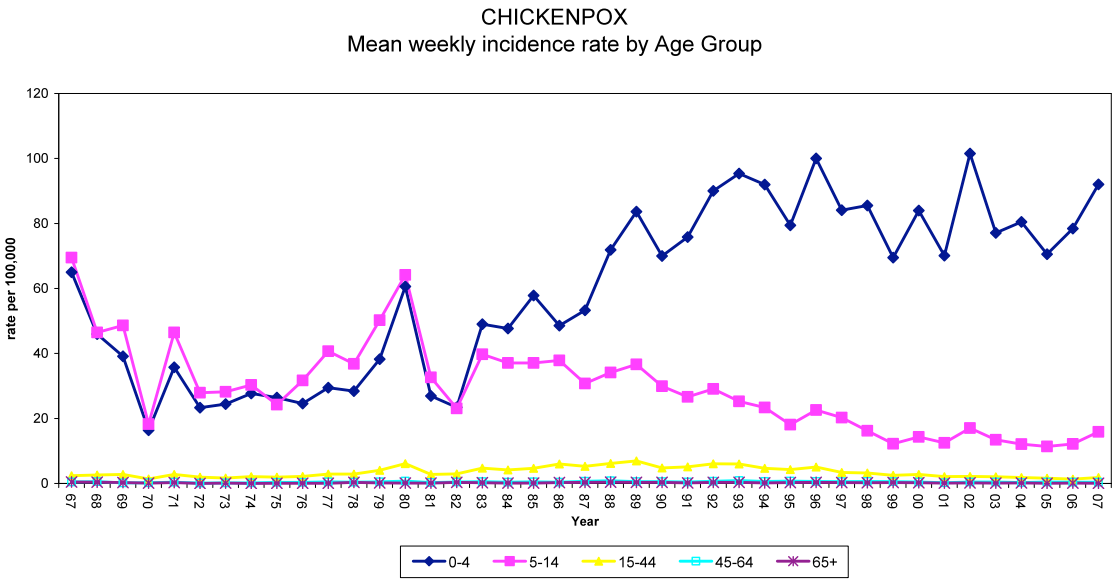
[1] Audit, which is conducted as a routine part of healthcare, for example to check quality and standards of care, is often considered to be separate from research. However, the issues raised by the use of patient data are the same as those discussed throughout this document, and so it is suggested that the guidance should apply equally to audit, particularly when conducted by those outside the healthcare team.

Box 1: Surveillance of chickenpox

The Weekly Returns Service of the Royal College of General Practitioners provides information on a weekly basis about the incidence of common diseases. Since 1999, this information system has been fully electronic and all data have been collected without any patient identification. Figure 1 demonstrates the variation from year to year in the incidence of chickenpox and shows a very interesting shift in the early 1980s, after which the incidence increased in pre-school children and fell in older children.

This surveillance information is currently proving very useful to guide the decision as to whether we should introduce a national policy of childhood vaccination against chickenpox.

Figure 1: Chickenpox mean weekly incidence rate by age group, 1967–2007



1.2 Identifiability of data used during research

12. Some research can be conducted with anonymous information, from which it is not possible to identify an individual. In other cases, researchers need access to information from which it may be possible directly, or indirectly, to identify a patient by some means.[2] Different levels of anonymity offer different risks and benefits and it is important to distinguish between them. At present, a confusing range of terms are used to define the level and identifiability of data – for example, anonymous, anonymised, de-identified, coded, pseudonymous and identifiable. It is important to ensure there is a common understanding of the terms used, and to raise awareness of what each entails. The terms used throughout this document are described in Box 2. It should be noted that very few data are truly 'anonymous' and all clinical data may potentially be considered sensitive to a patient.

13. It should be recognised that although researchers may want to link to contents of records in order to have data at a person level, they usually do not need to know the identity of an

[2] Even where researchers are not interested in the identity of the individual, the nature of the data means it rarely can be anonymous.

individual (i.e. who the person is). However, the controls used will need to distinguish between using person-level data that is not identifiable and using personally identifiable data.

14. The method used to display data can also be used to reduce the risk of identifiability, for example when using aggregated data, no personal information is provided and the results are presented as summary tables.

Box 2: Identifiability of information[3]

Anonymised information: Data must not include any identifiers – for example, name, address, full postcode, full date of birth or NHS number. Although data are provided at an individual level, there is no way of establishing a link with the original, identifiable clinical record – in other words it is not possible to identify an individual. The Committee on Standards of Data Extraction (COSODE) has defined a number of 'strong identifiers'. (Other terms sometimes used include: irreversibly de-identified, unlinked anonymised information, unidentifiable.)

Coded information: This type of information cannot directly identify an individual, but a 'key' is available that enables the identity of the patient to be re-linked to the data (by the person who holds the key). For example, personally identifiable data may be removed from a record but an arbitrary code maintained, from which it is possible to re-link the record with identifying data held in a different source. The coded identifier should be globally unique, and the key held in one place under strict conditions. Alternatively, the data may become identifiable when used together with other data sources. In cases where data relate to individuals with rare illnesses or exposures, or small or unusual populations, the likelihood of potential identification is increased. (Other terms used include: pseudonymous, linked anonymised, key-coded, reversibly de-identified, masked, encrypted.)

Identifiable information: This type of information contains personal data that directly identifies individuals. (Other terms used include: identified, identifying, personal, nominative.)

1.3 Context

15. The workshop took place against the background of considerable uncertainty about the processes that should be used when information from patient records is required for research. There is currently significant variation in the approaches adopted across the UK, and delegates commented about the lack of consistency in decisions taken by different bodies, at both national and local levels. Researchers have also expressed frustration at the time taken to get approval for studies, which can be up to 15 months. It is therefore clear that GPs, researchers and patients would benefit from agreed guidance and examples of best practice.

16. A number of reports have considered the use of personal information in research, including from the Academy of Medical Sciences (AMS, 2006) and the Care Record Development Board (2007). There is also some guidance relating to the use of patient records in research: for example, a Medical Research Council (MRC) Ethics Series report (2000); and the General Medical Council (GMC) is currently updating its guidance on research and consent, and confidentiality.[4]

17. The *Data Sharing Review Report* – published by the Information Commissioner, Richard Thomas, and Mark Walport, Director of the Wellcome Trust (Thomas and Walport, 2008) – also includes a number of recommendations relating to the use and sharing of personal information for research and statistical analysis. The Review recommendations are entirely

[3] For further background, see Lowrance (2002) and Care Record Development Board (2007).
[4] Current version available at www.gmc-uk.org/guidance/current/library/research.asp [accessed 6 May 2009].

consistent with the guidance for best practice given in this consensus document. The government responded to the Review in November 2008 (see Box 3). These statements should be a first step towards clarifying best practice for the use of patient records for research.

Box 3: Government response to the *Data Sharing Review Report*

Recommendation 17: The NHS should develop a system to allow approved researchers to work with healthcare providers to identify potential patients, who may then be approached to take part in clinical studies for which consent is needed.

The Government has announced plans to ensure that patients, from every part of the country, with any illness or disease, are made aware of research that is of particular relevance to them; and to enable them to choose whether to take part in appropriate clinical trials.

The Department of Health will develop a system to allow approved researchers to work with healthcare providers for this purpose, under a duty of confidentiality equivalent to the duty owed by health professionals. The Department will develop mechanisms to help healthcare providers operate the system consistently, and will ensure they work with the employers of the approved staff to deal effectively with any breaches of confidentiality. The independent National Information Governance Board will monitor the operation of the system.

The Research Capability Programme in NHS Connecting for Health will develop secure ways to speed up the operation of the system and reduce the need for approved staff to process identifiable patient information. (Ministry of Justice, 2008)

18. As part of the celebrations to mark the 60th anniversary of the NHS, the Secretary of State for Health held a seminar to highlight the importance of research for the NHS. The NHS Constitution, and the accompanying handbook, also enshrine research as a core part of the NHS (see Box 4, overleaf) and include pledges about access to information and informed choice. The NHS Constitution and the Care Record Guarantee both emphasise the importance of the confidentiality and security of patient information.

Box 4: Commitments to research in the NHS

From *The Handbook to the NHS Constitution for England*:

Research is a core part of the NHS. Research enables the NHS to improve the current and future health of the people it serves. The NHS will do all it can to ensure that patients, from every part of England, are made aware of research that is of particular relevance to them. The NHS is therefore putting in place procedures to ensure that patients are notified of opportunities to join in relevant ethically approved research and will be free to choose whether they wish to do so. (Department of Health, 2009, p. 51)

19. Issues relating to the use of information from children's records have not yet been considered fully and so are not included in this guidance. We recognise that this is an important topic that should be addressed in future guidance.

1.4 The legal framework

20. The Academy of Medical Sciences report (2006) described the "complicated patchwork of statutory and common law that operates in this area". Laws governing medical research using personal data include the Data Protection Act (1998), 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).

21. Section 251 of the NHS Act (2006) – originally enacted under Section 60 of the Health and Social Care Act (2001) – permits the common law duty of confidentiality "to be set aside in specific circumstances for medical purposes", where it is not possible to use anonymised information and where seeking individual consent is not practicable (see Box 5).[5]

22. From 2001, the Patient Information Advisory Group (PIAG) was responsible for assessing applications for the sharing of identifiable patient information in accordance with these provisions. The Health and Social Care Act (2008) established the National Information Governance Board for Health and Social Care (NIGB) as a statutory body, to provide leadership and promote consistent standards for information governance across health and social care.[6] The role of PIAG was transferred to the Ethics and Confidentiality Committee (ECC) of the NIGB in January 2009. The ECC is therefore now responsible for applications under Section 251, and provides advice to the NIGB on ethical issues relating to the processing of health or social care information.

23. The guidance notes of the Health and Social Care Act (2008) state that the new arrangements will "enable identifiable patient information to be disclosed and used for essential NHS activity and medical research without patient consent where activity is sufficiently in the public interest". However, it should be noted that although Section 251 permits the common law of confidentiality to be set aside, it does not override obligations under the Data Protection Act.

24. The situation in the devolved administrations is different. While the PIAG had a remit to cover Wales, the NIGB currently does not. In Scotland, the Confidentiality and Security Advisory Group for Scotland concluded in 2002 that a formal legal arrangement was not necessary. However, the Group recommended improvements in the way NHSScotland protects the privacy of patient data, and suggested that there should be a central service to anonymise national data. The Scottish Executive Health Department therefore asked NHS National Services Scotland to establish an Information Governance Programme for NHSScotland. Patient information is brought together and managed at a national level by the Information Services Division (ISD), part of NHS National Services Scotland.[7] The ISD works in partnership with a wide range of organisations and individuals, including health boards, GPs, hospitals and local authorities, to ensure the information is used "securely and efficiently". A Privacy Advisory Committee provides advice on the use of personal health information for research.[8]

25. The situation in Northern Ireland is different again, and is currently rather fragmented. Each Health and Social Services Board has a Patient Council that acts as a patient advocate and provides advice regarding policy related issues to the Boards. There is also a Research Ethics Committee to scrutinise protocols, and an Information Commissioner to provide

[5] The 2006 AMS report describes how Section 60 of the Health and Social Care Act (2001) came about in response to concerns that the reporting of cancer incidence to cancer registries by doctors would be jeopardised by the GMC guidance on confidentiality (see note 4 above) – since updated. It is interesting to note that the public are supportive of cancer registries. In 2005, Cancer Research UK surveyed nearly 3000 people and found that, when asked to consider whether, if they had cancer, they would mind their information being stored on a cancer registry and used in research, people did not consider this to be an invasion of their privacy – even with information such as name or address still attached (Barrett *et al.*, 2006).

[6] For further information, see www.nigb.nhs.uk.

[7] See www.isdscotland.org/isd/4489.html.

[8] The Privacy Advisory Committee was set up to advise the ISD and the General Register Office for Scotland on the processing and release of patient identifiable data. The Committee also advises the ISD on linkage of datasets that have not been linked previously, whether or not identifiable data are to be released.

overarching advice. Discussions are ongoing about the need to introduce a new framework for information governance. We are aware that different issues may also be raised by increasing interoperability of health records throughout Europe, which may need further consideration in future.

Box 5: Section 251 of the NHS Act (2006)

From the NIGB website:[9]

Section 251 of the NHS Act 2006 re-enacted Section 60 of the Health and Social Care Act 2001. The terms Section 60 and Section 251, when used in relation to use of patient information therefore refer to the same powers. These powers allow the Secretary of State for Health to make regulations to set aside the common law duty of confidentiality for medical purposes where it is not possible to use anonymised information and where seeking individual consent is not practicable. Under the Health and Social Care Act 2008, responsibility for administering these powers, was transferred from the Patient Information Advisory Group to the National Information Governance Board.

Why was Section 251 created?

Section 251 came about because it was recognised that there were essential activities of the NHS, and important medical research, that required use of identifiable patient information but because patient consent had not been obtained to use people's personal and confidential information for these other purposes, there was no secure basis in law for these uses. [NB. There are a few exceptions where there is a legal basis for disclosure e.g. reporting of notifiable diseases]. Section 251 was established to provide a secure legal basis for disclosure of confidential patient information for medical purposes, where it was not possible to use anonymised information and where seeking consent was not practicable, having regard to the cost and technology available.

It was anticipated when Section 251 powers were originally established that the NHS would develop mechanisms to seek, record and implement consent. Also that the NHS would endeavour to improve data quality and develop processes to link data in pseudonymised form, reducing the need for identifiable data to be used. These mechanisms are still being developed. It has since been acknowledged that there will continue to be a need for Section 251 powers, for some uses, on a more long-term basis.

What does Section 251 do and what are its limitations?

Section 251 powers lift the common law duty of confidentiality and therefore provide protection to disclosing bodies from being pursued for breach of confidence. Section 251 only applies to data generated in England and Wales. Both disclosing bodies and organisations receiving data are still required to comply with other legal requirements, in particular, the Data Protection Act and Human Rights Act. Organisations receiving data are themselves under a duty of confidentiality and cannot disclose information received under Section 251 to third parties without this being part of the approval given or, more commonly, the third party seeking separate approval.

How is Section 251 administered?

The Health Service (Control of Patient Information) Regulations 2002 (SI 1438) were made under Section 60 of the Health and Social Care Act 2001 and continue to have effect under Section 251 of the NHS Act 2006. These regulations established a class support mechanism whereby Section 251 powers could be used by the Secretary of State without needing to lay regulations before Parliament for each use of the powers. The classes of support are:

1. To reduce the identifiability of data.
2. For the past or present geographical location of patients for medical research.
3. To identify and contact patients with a view to inviting them to participate in medical research or to allow their data or tissue to be used for medical research or to allow their tissue to be used for other medical purposes.
4. To link information from more than one source or to validate the quality and completeness of confidential patient information or information derived from patient data.

[9] www.nigb.nhs.uk/ecc/about-the-ecc [accessed 6 May 2009].

5. For the audit, monitoring and analysing of the provision made by the health service for patient care and treatment.

6. To allow access to confidential patient information for one or more of the above purposes.

[Paraphrased from Schedule 1 of the 2002 Regulations]

Additionally, these regulations provided specific support under Section 251 for the Health Protection Agency and other public health staff to collect data relating to communicable disease surveillance and for surveillance of other risks to public health. They also provided specific support for Cancer Registries to collect data relating to cancer.

The National Information Governance Board and its Ethics and Confidentiality Committee, do not have statutorily delegated authority from the Secretary of State but rather advise the Secretary of State on the use of Section 251 powers and make recommendations about applications for approval under Section 251. The NIGB has agreed to delegate its responsibility for administering Section 251 powers to its Ethics and Confidentiality Committee.

As and when the above regulations need revision, Section 251 requires that the Secretary of State consults the NIGB prior to revised regulations being laid before Parliament.

The application process established by PIAG will continue to be used for the foreseeable future to facilitate continuity, although changes to the application process will be made as the Ethics and Confidentiality Committee agree are appropriate.

2. OVERARCHING PRINCIPLES

2.1 Safeguarding patient confidentiality

26. A particular feature of general practice records, given their comprehensive and continuing nature, is that they may contain sensitive information concerning a person's medical, personal or family history. The privacy of such information must be protected; GPs and patients must have confidence that data will remain secure. It is therefore crucial to have a clearly defined process for the use of different types of data – anonymised, coded or identifiable – with appropriate checks and balances, and procedural controls to minimise any potential risks.

27. It has been suggested that patient information that is particularly sensitive, for example information about mental health, should be treated with enhanced security measures. However, delegates at the workshop agreed that *all* clinical data should be treated as potentially sensitive. It is not possible to predict how different patients will rate the sensitivity of specific information, and labelling some conditions as particularly 'sensitive' may lead to increased stigmatisation. There was therefore a strong consensus that all data should be handled in the same way.

28. Electronic technologies offer potential advantages for maintaining the confidentiality of medical records. Privacy-enhancing technologies should be applied at all times. Wherever possible, systems should be designed to minimise the unauthorised release of data, and to process data so that only the minimum amount of identifiable information is made available. The inclusion of an audit trail in any electronic system would also significantly improve accountability.

Box 6: Maintaining confidentiality of data

Secure systems are essential to facilitate access to electronic records while maintaining data confidentiality. The Care Record Development Board report (2007) defined honest brokers and safe havens as follows:

Honest brokers: a trusted custodian of data who has responsibility to implement systems of access according to the Care Record Guarantee, the web of complex legislation and in accordance with the decisions and guidance of relevant Information Governance Boards. The honest broker would have responsibility for ensuring that the coding and anonymisation processes are correctly specified and implemented. An honest broker would also be responsible for:

- carrying out any permitted statistical linkage of different sources of identifiable health and social care data; and
- carrying out data quality checks which are not possible for researchers and other data users to do themselves for reasons of confidentiality.

The UK Clinical Research Collaboration R&D Advisory Group to Connecting for Health (2007) also described the importance of a dual role of the honest broker "in ensuring patient data confidentiality and security as well as scientific integrity of data delivered to the research community" in order to ensure public and professional trust.

Safe haven: a designated physical or electronic area that provides the most appropriate level of security for the use of data, including – and importantly – the most sensitive and confidential information. The Data Sharing Review recommended that "Government departments and other wishing to develop, share and hold datasets for research and statistical purposes should work with academic and other partners to set up safe havens" and that researchers working in safe havens should be "bound by a strict code, preventing disclosure of any personally identifying information" (Thomas and Walport, 2008).

29. The introduction of honest brokers and safe havens for the storage of data are also likely to provide significant benefits (see Box 6). The scope and delivery mechanism for a nationwide system is currently being defined by the Research Capability Programme of Connecting for Health. It will be important to ensure training for all involved, to improve understanding and confidence when new electronic systems are used, and to ensure appropriate controls and accreditation are in place to ensure standards are maintained.

Box 7: Government response to the *Data Sharing Review Report* recommendations relating to safe havens

In relation to the Department of Health in England, we accept the recommendation that:

- 'safe havens' are developed, as an environment to assist with appropriate processing for the purpose of population-based medical research and statistical analysis for medical purposes, to minimise the risk of identifying individuals; and
- a system is devised to ensure that only accredited people do work within safe havens.

Through the Research Capability Programme, established via the NHS Connecting for Health in 2007 programme, the Department of Health is working with the Information Centre for Health and Social Care to develop safe havens. They will be designed to enable appropriate processing for health research purposes of patient information and other data derived from patient information.

The aim is to provide a secure environment in which suitable investigators and research professionals can work under conditions of confidentiality, with expert support from health professionals and staff who owe a duty of confidentiality equivalent to that of a health professional.

In this context, the Government will commission a code or codes for the use of safe havens, and a scheme for accrediting researchers. The Government will continue to consider the appropriate legal structures for the different types of processing that might in future be carried out using 'safe havens' and in relation to the use for non-medical purposes of data derived from patient information. (Ministry of Justice, 2008)

30. But technology alone will not be enough. There must also be mechanisms of accreditation and accountability to reinforce trust where a researcher does have access to identifiable patient information.

31. The Data Sharing Review also developed the idea of an 'approved researcher', who would be "bound by the same duty of confidentiality as the clinical team providing care, and face similar penalties in the case of any breach of confidentiality". The report cited the example of the Statistics Board, which can grant 'approved researchers' access to personal information held by the Statistics Board for statistical research, with criminal sanctions for any breach of data security.[10] The Data Sharing Review recommended that "a system of approving or accrediting researchers who meet the relevant criteria to work within a safe haven is established…preventing disclosure of any personally identifying information, and providing criminal sanctions in case of breach of confidentiality".

32. In its response to the Data Sharing Review (see Box 7), the government stated that it would commission a code or codes for the use of safe havens, and a scheme for accrediting researchers. The response also stated that the Department of Health "will develop a system to allow approved researchers to work with healthcare providers for this purpose, under a

[10] "Researchers are then bound by a strict code, which prevents disclosure of any personal identifying information. Any deliberate or negligent breach of data security by the approved researcher would entail criminal liability and the prospect of a custodial sentence up to a maximum of two years." (Thomas and Walport, 2008, p. 71)

duty of confidentiality equivalent to the duty owed by health professionals". A formal process of accrediting researchers should be introduced to achieve this. The process will need to define appropriate and substantive sanctions so that researchers may be held to account should there be any breaches of confidence.

33. Public confidence in the use of patient records will depend not only on statements of good practice but also on evidence that such statements are adhered to and that there are appropriate and visible procedures for identifying, reporting and responding to lapses in good practice. Independent audit of the governance, management and use of such data will be important.[11] The aim should be to work towards a partnership between practice, researcher and patient.

2.2 The role of the GP and healthcare professionals

34. Delivering high-quality care and protecting patients must always remain the first priority of GPs. In relation to research, the primary role of the GP, or the responsible primary care professional, should be to act as an advocate for the patient. As set out in this guidance for best practice, the GP and the practice must retain ultimate responsibility for ensuring confidentiality and appropriate access to data, and that the associated processes are adhered to. The GP needs to be given sufficient information about the research process in order for them to be in a position to assure patients that research is conducted according to best practice. They may also be asked to provide advice to patients about whether to take part in a trial, or need to provide support where any feedback is provided following research. It is recognised that where a GP endorses a study, it is often seen to increase patient participation in that study.

35. There will need to be consideration of the burden that these responsibilities will place upon GPs. Support and resources may be needed to enable GPs to fulfil this advocacy role and to help to facilitate a culture that is supportive of research. Resource implications should be taken properly into account for all research projects and the burden should be with researchers rather than GPs. For example, training may be needed to help GPs and practice teams become familiar with key elements of the research process, such as consent. GPs should also understand the methods of a study and the research design. It was suggested that it might be helpful to provide a GP with a checklist to help them to satisfy themselves both that the preconditions for research are met, and that appropriate data security measures are in place.

2.3 Improving public awareness and engagement

36. It is essential to establish public confidence in the use of patient records for research. The overriding message is the need for openness and transparency, to improve patients' understanding of how their records might be used. Currently, there is little understanding of how patient records are used beyond the direct provision of care. It is important to improve public awareness about all potential uses of records, but particularly to improve understanding of potential research uses.

37. Members of the public are generally supportive of medical research and appear to recognise the value of using personal information to advance medical science. In a recent survey conducted by Ipsos MORI (2007), commissioned by the Medical Research Council, 60 per cent of people in the UK reported feeling that they have a responsibility to allow their personal health information to be used in medical research, and 69 per cent reported that

[11] For example, ISO 27001 ensures best practice with the amount and nature of data extraction.

they would be 'likely' or 'certain' to allow their 'personal health information' to be used for research purposes.[12]

38. However, there appears to be little understanding of what this might mean in practice. It is very important that all NHS patients understand how their medical records may be used, and the safeguards that have been put in place to protect their privacy during the research process. Patients should also be aware of the reasons that medical records may sometimes be accessed for research without the individual being contacted, to ensure there are 'no surprises'.

39. To improve transparency, information should also be provided about the difference between the use of identifiable and non-identifiable data, and the different implications for individuals. A key message to convey is that researchers are usually not interested in the identity of an individual, but may need to access personally identifiable data in order to conduct research (rather than aggregated, anonymised data for a large number of patients).

40. There are two potential routes to improve public knowledge and understanding. First, it is recommended that there should be a national programme to raise awareness of the importance of research, and the need to use patient records (whether anonymised, coded or identifiable) to inform that research. This should provide simple, clear messages both about different types of research, and how information from patient records may be used. It should be balanced and non-coercive. It would be helpful to provide specific examples (such as Sir Richard Doll's work identifying the link between smoking and lung cancer), and there may be lessons to learn from successful campaigns for blood donation, which raise many similar issues.

41. Second, information should be provided locally. Individual general practices may use a variety of mechanisms to let patients know how their records might be used, including to support research. Patients should receive information when registering with a practice, and information leaflets and posters should be displayed in a practice. For example, there are systems in place already to inform patients about some existing research datasets. Above all, this information should make it clear that patients can opt out of the use of their identifiable information in research, if they wish.

42. The NHS Constitution recognises that research is a core part of the NHS, and commits to informing patients about relevant trials, wherever possible. There may be opportunities to develop an information campaign about the importance of research in order to help to deliver this commitment. Such a campaign could also improve awareness about the processes of research and the use of data within the NHS. The UK Clinical Research Collaboration (UKCRC) has already established a subgroup to consider issues relating to public awareness. There may be a role for an organisation such as the National Institute for Health Research Clinical Research Network or the General Practice Research Database to provide appropriate information that can be used at a practice level, and adopted nationally.

43. The government response to the *Data Sharing Review Report* also made a number of commitments to improve awareness of research uses of patient information, through the National Institute for Health Research and the NHS.

[12] The survey considered 'personal health information' – in other words, identifiable data. The use of anonymised and coded data was not considered.

Box 8: The electronic Primary Care Research Network

The electronic Primary Care Research Network (ePCRN)[13] is an electronic tool that facilitates research in primary care, jointly funded by the National Institutes for Health (in the USA) and the National Institute for health Research National School for Primary Care Research (in the UK). ePCRN aims to enhance the potential for using electronic health records in primary care by:

1. Improving patient recruitment to clinical trials: ePCRN has developed generic methods for accessing information from electronic health records to identify and recruit patients for randomised controlled trials and cohorts. Authorised researchers can access the system via a secure web portal, and provide information about the trial and subject eligibility criteria. ePCRN then uses a secure federated database to run a query to identify and flag patients who meet the eligibility criteria. The researcher is given an anonymous 'count' of potential subjects, while all patient information remains confidential in the clinic. The clinical team is then able to identify and recruit the eligible subjects because of local flagging in the electronic health record. This protects the data confidentiality of patients, while allowing the use of the minimal information necessary for research.

2. Enabling data standardisation: Researchers in primary care have faced difficulties because there is no standardised language or standard method to collect and compare data. ePCRN allows researchers to define 'common data elements', giving full data definitions for a data item such as blood pressure, including measurements standards, permissible values and the way data should be entered on a form. This allows faster, more effective design and organisation of randomised clinical trials, with rapid creation and re-use of data collection forms and data elements. Elements and forms can be archived and re-used, and an underlying XML-structured study database created.

[13] www.ePCRN.org.

3. GUIDANCE FOR BEST PRACTICE

3.1 Preconditions for research studies in general practice

44. When a general practice is approached by researchers and asked to take part in a research study, the practice should make sure that the proposed study satisfies a number of conditions before agreeing to take part. The research should have been through a process of independent and expert review and approval by a range of other bodies, including the research ethics committee, an NHS or university R&D office, and the research funder. These review processes will generally ensure that the study design is appropriate to the research question, and that the proposed setting, management and conduct of the research comply with current standards of research governance, as set out by the NHS R&D programme.

45. Before agreeing to take part in a study, GPs should additionally assure themselves that:
    - The practice provides an appropriate setting for the study to be carried out.
    - The practice has the resources – which may be provided by researchers – to deliver what is proposed in the study protocol.
    - Any potential conflicts of interest related to the practice or the researchers have been identified and addressed. Different approaches may need to be taken depending upon the type or research and sources of funding involved.

46. As the NHS now contracts with general practices and other primary care providers rather than individual practitioners, the responsibility for ensuring that studies meet these conditions, and have been approved by an independent research ethics committee, is held at practice level.

47. Where a practice is engaging directly with the research process (rather than simply providing access to information about a patient population), the GP and the practice will need to be more involved with the process of research review and approval. The Royal College of General Practitioners (RCGP) has developed an online accreditation programme called 'Research Ready', to help practices ensure that they are up to date on the governance and resource requirements needed to take part in research studies.[14] The General Medical Council (GMC) also provides guidance on the role and responsibilities of doctors in research.[15]

3.2 Research using existing patient records alone (i.e. without patient participation)

3.2.1 Using anonymised data

48. The use of anonymised records, where all details that could identify an individual are removed, enables research to be conducted without compromising the confidentiality of individual patients. Electronic patient records can be readily anonymised, and offer the potential for large volumes of data to be used with all identifiable information removed.

49. Several large research datasets have been established, for example the General Practice Research Database (GPRD), QResearch and The Health Improvement Network (THIN), collating and sharing anonymised patient information from large numbers of practices with researchers. In these services, no identifiable data leave the general practice and all

[14] Details of the Research Ready Self-accreditation are available at www.rcgp.org.uk/clinical_and_research/circ/research/primary_care_research_team_ass/research_ready.aspx [accessed 6 May 2009].
[15] Available at www.gmc-uk.org/guidance/current/library/research.asp [accessed 6 May 2009].

research is undertaken within very strict guidelines. These are exemplar services, and are welcomed by GPs. The model where no identifiable data leave the GP practice could be used when setting up other services for the use of anonymised data (see case studies in Boxes 9–11).

50. **Consent:** Where anonymised, non-identifiable records are used, it is not necessary to seek consent from individuals. If, however, a patient has indicated that they do not wish their anonymised data to be used in research, this should be honoured. However, as discussed above, it is important to raise awareness among the general public, and patients within a specific practice, that records may be used in this way.

51. **Process of anonymisation of data:** The process of anonymisation should make use of available technologies and be as automated as possible, to ensure there is no need for researchers to view identifiable records at any stage of the process. It is important to ensure that data are genuinely anonymised – in other words that it is impossible to identify a specific individual.[16] It is best practice for the anonymisation to take place within the GP domain (i.e. in the practice or on the practice database).[17]

52. If the anonymisation is not done within the practice itself, the GP needs to have confidence that the process used conforms to the highest standards of best practice. It may be useful to develop a process of accreditation for individuals responsible for the anonymisation process based on the concept of safe havens and honest brokers. The Data Sharing Review recommended that government departments should work with academic and other partners to set up safe havens.

[16] Where data relate to individuals with rare illnesses or exposures, or small or unusual populations, the likelihood of potential identification is increased. In such instances, the data should be treated as pseudonymised, or coded data.
[17] As with the GPRD, a key system would be needed to allow linkage (see Box 11).

Box 9: QResearch

QResearch[18] is an established not-for-profit joint venture between the University of Nottingham and Egton Medical Information Systems (EMIS), supplier of IT systems for 56 per cent of UK general practices covering a combined population in excess of 39 million NHS patients nationwide. The core aim of QResearch is to enable the highest-quality hypothesis-driven original research aimed at improving health and healthcare. Established in 2002, QResearch specialises in the access to and analysis of primary care data for research.

The organisation hosts two major databases – the QResearch database (anonymised patient-level data on 10m patients for research) and QSurveillance (the UK rapid response surveillance system containing age–sex aggregated data with 140 indicators, on 23m patients). The main QResearch database is one of the largest anonymised general practice patient-level consolidated databases for research worldwide, with records from over 550 general practices in the UK. It contains over 40m person years of observation, with data extending back to the early 1990s.

Safeguarding patient confidentiality: QResearch is regulated by a National Advisory Board comprising representatives of national professional (including the British Medical Association, the RCGP, the General Practitioners' Committee and the EMIS National User Group) and patient organisations. It has an established governance framework with scientific peer review of all proposals to use the database. Practices contribute data to QResearch for free in exchange for feedback.

In addition to containing coded clinical data for demographics, prescriptions, referrals, consultations, diagnoses, clinical values, laboratory tests and screening, QResearch also contains socioeconomic and geographical data linked to the patient's postcode (the postcode, however, is not extracted to preserve anonymity). QResearch also contains linked Office for National Statistics cause of death data back to 1993 linked at individual level.

QResearch has been validated by comparing birth rates, death rates, consultation rates, prevalence and mortality rates with other data sources including the General Household Survey, the GPRD and THIN. Detailed analyses have shown good levels of completeness and consistency, and comparison with non-research practices suggests results are likely to be generalisable.

Research using QResearch: The QResearch database has been used for a wide range of studies from academics and health service organisations.[19] Many studies have been published in high-ranking journals such as the *BMJ* and the *Lancet*. Most recently, the QResearch database was used to develop QRISK2, a new cardiovascular disease risk algorithm,[20] which is being implemented in clinical practice.

[18] www.qresearch.org.
[19] For a list of published outputs see www.qresearch.org/public/publications.aspx [accessed 6 May 2009].
[20] www.qrisk.org.

Box 10: The Health Improvement Network

A new data collection scheme, The Health Improvement Network (THIN), was started in 2003 as a collaboration between the providers of the Vision practice software (INPS) and EPIC,[21] both companies now being part of Cegedim. One of the aims of THIN, as well as creating a data source for research, was to improve the quality and completeness of recording of clinical data in GP practices. This is achieved by providing participating practices with support and training as part of their agreement to supply data as well as supplying feedback on the recording of a range of clinical conditions. EPIC compiles the data into a research dataset and provides access to THIN data for researchers.

THIN data contain anonymised computerised information entered by GPs throughout the UK. Data are collected from over 410 practices with the most recent dataset including over 6 million patients. This amounts to approximately 37m patient years of research-quality data. The computerised information, extending as far back as the early 1990s, includes demographics, GP recorded prescriptions, signs and symptoms, diagnoses from specialist referrals and hospital admissions, and the results of laboratory tests. The data also contain links at patient postcode level to Townsend score and other environmental and social variables (the link to the postcode is done at the practice; to maintain anonymity, postcodes are not extracted).

The data have been validated against external sources such as the Office for National Statistics, QResearch and Quality and Outcomes Framework sources in order to demonstrate that the data are generalisable to the UK.

The database is accessible for research by academic, government, regulatory, public health or pharmaceutical company entities so long as the research fulfils ethical and scientific criteria as well as obeying data confidentiality rules. Researchers sign a confidentiality agreement and agree to security provision for extracted data before they are supplied.

Safeguarding patient confidentiality
Before initiation of the collection scheme, prior approval of the NHS South-east Multi-centre Research Ethics Committee (MREC) was gained to conduct such research. The data are collected using secure automated collection software. No name, address or postcode fields are collected; neither is full date of birth nor NHS number. The practice identifier is encrypted as well as the patient identifier, such that the practice can be contacted only by THIN and not by researchers, and that a patient can be identified only by the practice. Patients are made aware of the data collection scheme through practice posters and information in practice leaflets. It is made clear to patients through these documents and practice staff that they may opt out of the collection scheme. If additional information is to be collected as part of the research this needs to be approved by an MREC.

Research using THIN data
All research studies to be undertaken require prior ethical and scientific approval. The database's use for research is relatively new but already over 100 publications have been based on the THIN data, including a validation of QRisk scores in collaboration with QResearch.[22]

[21] www.epic-uk.org.
[22] www.epic-uk.org/thin_publications.htm.

3.2.2 Using coded data

53. In many cases, researchers will need to use data that are coded, allowing a link back to the original data. This may be important in studies where the validity of the data is being checked or additional supporting information is required. The use of coded data means it might be possible directly, or indirectly, to identify patients. The chances of identification might be small – depending upon the type of data, the rarity of certain data values, other data sources that could be used in conjunction with it (particularly if it is linked with identifiable data via a code) and how the data have been processed.

54. It is important to ensure that coded data are managed appropriately with safeguards that are proportionate to the risks. The main issues to consider are the likelihood of individuals being identified, the risks that would entail and how the chances of identification can be minimised. This will depend how and where any linking keys are held. It should also be noted that, in cases where data relate to individuals with rare illnesses or exposures, the likelihood of potential identification is increased.[23]

55. **Consent:** As described above (section 2.3), patients should be kept informed, including when registering at a practice, about the potential for records to be used in research. As part of this awareness-raising exercise, patients should be given the opportunity to opt out of research using their records if they wish. It should therefore not be necessary for patients to give individual consent for their records to be used in a coded manner for research. Experience from the GPRD and Tayside shows the rate of such opt-out is very small, and of a level that would not affect the validity of the research.

56. **Process of coding data:** New technologies and methods of encryption should be used wherever possible to minimise the risk of identification of patients. Where a code is used to link coded data with identifying data, multiple keys should be used, with one held at practice level and another at the level of the research repository. The process should be handled by honest brokers, in a safe haven (see Box 6). In this way, GPs should be able to satisfy themselves that a researcher will only be given access to data once it has been appropriately coded by an accredited system.

[23] Where small numbers of patients are involved, the possibility of 'inference attacks' – where a user is able to infer information about a patient by putting together different pieces of information without directly accessing the patient record – may be more likely. There may be a role for honest brokers in minimising this risk.

Box 11: General Practice Research Database

The General Practice Research Database (GPRD)[24] is the world's largest database of anonymised longitudinal medical records from primary care, established over 20 years ago. It is owned by the Secretary of State for Health, and managed by the GPRD Group within the Medicines and Healthcare products Regulatory Agency (MHRA), the UK's medicines and devices regulator, on a non-profit-making basis.

Currently data are collected on over 3.6 million patients from more than 450 primary care practices throughout the UK (5.5 per cent of all practices). With comprehensive observational data collected from GPs, the GPRD is used for medical and health research purposes in a range of areas including clinical research planning, clinical epidemiology, disease management, outcomes research and drug safety. The data are updated on a regular basis, and are used worldwide for research by academic researchers, the pharmaceutical industry, clinical research organisations, regulators and government departments. There are currently about 120 ongoing studies, of which 50 are under an MRC–GPRD arrangement

Safeguarding patient confidentiality: The data in GPRD are heavily quality-controlled. Data collection software is built into the Vision system and extracts the required data from a practice computer. Personal identifiers, including names and addresses, are not collected. Patient/prescriber identities are encoded so that it is impossible for anyone other than the practice to identify them. The GPRD uses a two-key approach. The first key is held by the GP practice within its IT system; as data leaves the practice en route to the GPRD, the key creates a new ID. The second key is held at the GPRD where the research-level ID is created. The only way back is via both keys.

GPRD data is made available to researchers via a secure online service, or as project-specific datasets or aggregate tables on CD-ROM. All uses of data are covered by:

- a legal agreement between the GPRD and the researchers organisation, covering how the data must be stored and managed, the uses to which data can be put, and the way data use must be documented
- scientific approval, by the Independent Scientific Advisory Committee
- ethics approval
- NIGB approval where required: for linkage studies, for instance where an additional file is downloaded to a Trusted Third Party who undertakes linkage using the NHS number; and to process the text to remove any identifying text before research use (this is done mainly by in-house built software).

Patient engagement: Practices are supplied with posters and leaflets that explain to patients that their anonymised data may be used for research and that they can, if they so wish, ask the doctor not to allow their data to be used. The exact number of patients who have used the opt-out option is not known, but discussions with GPs suggest that it is extremely small (one report suggests the opt-out rate is less than 1 per 1000 patients).

All activities with patients are undertaken by GPs or their practice staff. There have not been any complaints from GPs, healthcare professionals or patients about the system, and there have been no breaches of patient confidentiality.

[24] www.gprd.com.

3.2.3 Using identifiable data

57. The use of identifiable records, which may be needed to answer certain research questions, requires patients to be informed about the research and asked to consent to the use of their information. Consent should follow the process described in section 3.4. The default position should always be that consent should be sought to access identifiable patient records. There may, however, be circumstances where obtaining consent is impractical or even, rarely, inappropriate. In situations where it is not possible to seek informed consent to use identifiable records (see Box 12), researchers must apply for special permission to the NIGB Ethics and Confidentiality Committee (see section 1.4). To ensure that properly accredited researchers are not unduly impeded there has to be a clear and agreed legally enforceable mechanism in place to reassure the public that their data are safe and protected. Currently this mechanism is through NIGB.

58. **Consent:** As described above (section 2.3), patients should be kept informed about the potential for records to be used in research, using a variety of approaches including information leaflets and posters on display in a practice. There was consensus at the workshop that patients should also be given the opportunity to opt out of the use of their identifiable information in research, if they wish.

59. Further thought needs to be given as to the best mechanisms for achieving this. One way would be to include it during the electronic registration system. The General Practitioners' Committee of the British Medical Association, for example, has suggested that a new check box, with a default position of 'not consenting', should be added to the GMS1 registration form to allow patients to consent to be approached to participate in research (similar to the organ donation tick box). Other delegates at the workshop suggested that it would be more appropriate for patients to be able to opt out of the use of their identifiable information in research, using a flag system on the electronic record. This would allow patients to review their position and update their consent as and when necessary.

Box 12: Difficulties obtaining consent

The Academy of Medical Sciences (2006) identified a number of circumstances where it may not be practicable to seek consent for the use of identifiable patient records in research. The report stated that "an insistence on explicit and specific consent often comes at a considerable, and sometimes prohibitive, cost to research. Such an insistence can also give rise to selection, recruitment and participation biases, leading to potentially misleading research results and the exclusion of disadvantaged social groups from research findings."

The report cited examples where:

- seeking consent may be impracticable, for example if a large proportion of the study population are untraceable
- seeking consent may compromise effective population coverage
- seeking consent may cause distress or harm (for example, the MRC gives the example of contacting people about a study examining correlations between parents' mental health and unexplained child deaths, which might cause serious distress)[25]
- seeking consent may lead to bias
- seeking consent may prevent appropriately large studies, leading to a reduction in the reliability and generalisability of the research findings.

[25] See MRC (2000), p. 19.

3.3 Research using patient records as the starting-point for participatory research

60. Patient records may be used as a starting-point to identify groups of patients with a particular disease or condition to invite them to participate in a study, or to identify a representative sample of the general population, for example as a control group. General practice records are particularly useful as the starting-point for studies of common conditions that are unlikely to be referred to hospital or for which hospital patients are likely to represent a biased sample of patients with the condition.

61. Issues that need to be addressed include: how patient records are used to identify participants; who has access to the records; what level of anonymisation is possible at different stages of the process; and, where appropriate, who approaches patients to invite them to take part in a study.

Box 13: MIQUEST

MIQUEST (Morbidity Information Query and Export Syntax)[26] is a Department of Health-sponsored data extraction application that can query and extract data from the different brands of GP computer system. In its remote enquirer mode it will code data using a method carefully evaluated by the University of Nottingham. Names are removed and replaced with a number that can only be decoded in the practice. The system has restrictions such as only allowing the first part of the postcode and the year (rather than full date of birth), to ensure that data are de-identified. MIQUEST is a robust tool that can be used to identify cases that meet inclusion and exclusion criteria set out in study protocols. Once a search has identified the number of patients, the local enquirer mode of MIQUEST can be used to print off the patients details at a practice level.

3.3.1 Reviewing records to identify potential participants

62. In order to identify potential participants for a research study it is often necessary to review medical records to determine whether patients meet the eligibility criteria for the study (such as diagnoses, age or gender). Following the initial screening exercise, those patients that meet the eligibility criteria are contacted to invite them to take part in the study.

63. The conditions and procedures by which medical records may be accessed at the start of the research process need careful clarification. In particular, questions arise as to who should initially screen the records – the GP, the care team or the researcher. In some cases, it may need to be a member of the research team, for example where there is a particular need for a uniform selection process across multiple practices or if the practice does not have the resources to undertake a search.

64. The GP must retain ultimate responsibility for the process of accessing records, but can make a judgement, on a case-by-case basis, as to whether it is appropriate for a researcher to view records.

65. Wherever possible, the best available technology should be used to reduce access to identifiable data. Increasingly, there will be technical solutions that allow records to be screened automatically, eliminating the need for researchers to access identifiable personal information in a record. This is the ideal solution. However, in the interim it may still be necessary for an individual to access records to identify potential participants, in which case appropriate checks must be in place.

[26] www.connectingforhealth.nhs.uk/systemsandservices/data/miquest.

66. Where approval has been given for a researcher to access records, there must be a formal process of accreditation as described above (section 2.1). This should place researchers under the same duty of confidentiality as a health professional, and must define substantive sanctions for any instances where a researcher is found to have breached the agreed conditions for access to data.

67. When researchers do have access to records, they should only have access (by using the best available technologies) to the minimum amount of information necessary to undertake the process of identifying potential participants. This should be defined and justified in discussion with ethics committees, research networks, the practice and Caldicott Guardians, and be clearly documented in the approved research protocol. GPs should make the final decision as to what is acceptable and appropriate in their practice.

68. As part of the public awareness-raising exercise described above (section 2.3), patients should be informed that records may be used in this way to identify potential participants for research. The information provided should make clear that only those patients that meet a study's eligibility criteria would be contacted and invited to take part in the research, following an initial screening process. Patients should be given the opportunity to opt out of the use of their records for research if they wish.

3.3.2 Inviting patients to take part in a study

69. When patients are to be invited to take part in research, greater clarity is 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* (Thomas and Walport, 2008) described this need for "consent to gain consent" as a "problem that requires a solution".

70. 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 of GPs' 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.

71. Patients need to be assured that they are free to choose whether or not to participate in a study without affecting the quality of their care or their relationship with the GP. Patients must never feel coerced into participating in research because of an overriding wish not to upset their doctor. Patients should also be contacted in a way that avoids possible worry or embarrassment. The initial letter of invitation will usually offer the opportunity for discussion between the patient and GP or researcher as part of the consent process.

72. Ideally, once a list of potential participants has been drawn up, the GP should screen the list before patients are contacted in order to identify those whom it would be inappropriate to contact, for example as a result of a recent bereavement or serious illness. As the practice retains overall responsibility for the process, this decision is up to the individual practice. There may be ways to enhance the efficiency of the process, for example by using a template for the procedure. It is appropriate for researchers to make resources available to a practice to facilitate this where necessary. This should be included in the project budget, and research funders will need to recognise it as a legitimate cost on research grants.

73. In order to maintain patient trust, best practice is for the invitation to participate in a study to come from the GP, signed and on practice headed paper. It is both appropriate and

acceptable for researchers to provide clerical and administrative support to facilitate this process – for example by preparing the letters on the basis of the identified list of potential participants, and providing postage. Again, research funders will need to recognise this as a legitimate cost on research grants.

74. Wherever possible the best available technologies should be used to improve security and maintain confidentiality. For example, in the future, privacy-enhancing technologies might allow letters to be processed and addressed without revealing identifiable information to clerical staff, or the process might be fully automated (as, for example, with the generation of monthly payslips).

75. The practice retains ultimate responsibility and accountability for the process of contacting participants to invite them to participate in research.

76. This guidance for best practice is in line with the recommendations made by the Data Sharing Review. The report concluded that "the NHS should develop a system to allow approved researchers to work with healthcare providers to identify potential patients, who may then be approached to take part in clinical studies for which consent is needed. These approved researchers would be bound by the same duty of confidentiality as the clinical team providing care, and face similar penalties in the case of any breach of confidentiality" (Thomas and Walport, 2008). The government accepted this recommendation in its response (see Box 3).

3.4 Informed consent

77. The default position should always be that patients should be asked to consent to the use of their identifiable information for research, as described above (paragraph 57). Consent is also required for patients to participate in a clinical trial or other research study. The process for seeking consent, once the researcher or GP is working directly with the patient, is well established and has been described by the GMC as "a process involving open and helpful dialogue".[27] The GMC guidelines suggest that "participants' consent is legally valid and professionally acceptable only where participants are competent to give consent, have been properly informed, and have agreed without coercion".

78. Potential participants must be provided with information about the purpose and nature of the research, their role as a participant, what will happen to information and/or samples collected during the study, and other relevant information, including the right to withdraw from the study at any time (see Box 14). This information is usually provided via a letter, with an accompanying information leaflet, and should be approved by the ethics committee. Potential participants should be given the opportunity to ask questions or discuss any issues and be given time in which to reach a decision.

79. It should be clear that participation is voluntary and that whether or not a patient participates in a particular research study will have no bearing on the care that the patient receives from the NHS or their relationship with healthcare staff. Participants must be informed and understand that they have the right to withdraw from a study at any time.

80. On some occasions, research may involve adults who lack capacity, in which case it is important to give particular consideration to the consent process. The Mental Capacity Act (2005) and accompanying codes set out when research with adults who lack capacity can be conducted, and clarifies how to seek consent from those who lack capacity or their legal representative. The MRC (2007a) has published related guidance for medical research involving adults who cannot consent. Issues relating to the use of information from children's

[27] See www.gmc-uk.org/guidance/current/library/research.asp [accessed 6 May 2009].

records have not been considered in this guidance document, although both the MRC and the GMC have also provided guidance for consent processes when undertaking research with children, which should involve someone with parental responsibility.[28]

81. **Mechanism for recording consent:** Consent (or dissent) should be recorded in the patient record, unless a patient specifies otherwise. It may be helpful to develop a template, with a defined code, to improve consistency when recording this information.[29] A patient's request that their decision not be recorded in their medical records should be respected and followed. However, the patient should be informed that this may result in their being re-contacted to participate in future research.

Box 14: Inviting people to take part in research

The GMC provides guidelines on the roles and responsibilities for doctors when conducting research. This includes a section on consent,[30] which suggests that "obtaining consent is a process involving open and helpful dialogue, and is essential in clarifying objectives and understanding between doctors and research participants". Ethics committees will also review the process by which it is intended to seek consent. The GMC guidelines suggest that the following information should be provided to potential participants of research:

- what the research aims to achieve, an outline of the research method, and confirmation that a research ethics committee has approved the project
- the legal rights and safeguards provided for participants
- the reasons that the patient or volunteer has been asked to participate
- if the project involves randomisation, the nature of the process and reasons for it, and the fact that in double-blind research trials neither the patient nor the treatment team will know whether the patient is receiving the treatment being tested or is in the control group
- information about possible benefits and risks
- an explanation of which parts of the treatment are experimental or not fully tested
- advice that they can withdraw at any time and, where relevant, an assurance that this will not adversely affect their relationship with those providing care
- an explanation of how personal information will be stored, transmitted and published
- what information will be available to the participant about the outcome of the research, and how that information will be presented
- arrangements for responding to adverse events
- details of compensation available should participants suffer harm as a result of their participation in the research.

3.5 Feedback of results to study participants

82. Before participants consent to take part in a research study they should understand how their personal results and the overall findings of the study will be used. Participants should be informed what, if any, findings they will receive about themselves and should have the opportunity to decline to receive this feedback if they choose.

[28] See MRC (2007b) and www.gmc-uk.org/guidance/current/library/research.asp#special_consideration [accessed 6 May 2009] respectively.
[29] For example: when a patient was approached, if they agreed or refused to participate, whether they are currently participating in a trial.
[30] See www.gmc-uk.org/guidance/current/library/research.asp#research_consent [accessed 6 May 2009].

83. The researcher has a responsibility for reporting the overall results of the study, and should ensure that GPs are informed of the results of the research. This is particularly important if research is likely to be widely reported in the media, so that GPs (of practices that have participated in the trial) have the appropriate information to be ready to respond to patient concerns. It is also important if there are adverse results affecting either individuals or groups of patients.

84. Feedback to study participants should be provided as initially agreed. It should be presented in a way that participants can understand, and be sensitive to the needs of the participants as well as the communities in which they live. The GP, acting as the patient's advocate, will have a particular responsibility for discussing a patient's personal results with them, and any associated issues.

4. CONCLUSION

85. Patient records held in general practices offer significant potential for research. In order to maximise the potential of this research, it is vital to agree and follow the highest standards of research practice. The first task must be to ensure clearly defined processes for uses of data, with appropriate procedural controls – including accreditation and sanctions – for those who have access to data. Wherever possible, available technologies should be used to improve security and enhance confidentiality.

86. It is also crucial to improve public awareness, and to ensure public understanding and support for the various research uses of patient records. Transparency is essential. Once these objectives are achieved, it will then be possible for the GP, or the responsible primary care professional, to focus on their main role: to act as an advocate for the patient. The GP and the practice must retain ultimate responsibility for ensuring appropriate access to data, and they must be in a position to provide advice to patients about the research.

87. The best practice guidance described in this document is intended as the first step in a process to ensure that patients and GPs have confidence in the processes used to access patient information, and to enable everyone to benefit from the significant research potential of medical records.

ANNEXE A

Workshop delegates

Towards Consensus in the Use of Patient Records for Research in General Practice 22–23 May 2008

Ms Amanda Brewster, Strategic Planning and Policy Unit, Wellcome Trust

Dr Tony Calland, Chairman, Medical Ethics Committee, British Medical Association

Mr Antony Chuter, Patient Partnership Group**,** Royal College of General Practitioners

Dr Jimmy Courtney, Chair, British Medical Association Northern Ireland subgroup on IT

Dr Paul Cundy, Chair, British Medical Association General Practitioners' Joint IT committee

Dr Thomas Davis, Chairman, National Vision Users Group

Professor Brendan Delaney, Professor of Primary Care**,** University of Birmingham

Professor Ian Diamond, Chief Executive, Economic and Social Research Council and Chair, External Reference Group, Connecting for Health Research Capability Programme

Dr Richard Fitton, Tameside and Glossop Primary Care Trust

Dr Douglas Fleming, Director, Royal College of General Practitioners Birmingham Research Unit

Dr Pat Goodwin, Head of Pathogens, Immunology and Population Health**,** Wellcome Trust

Professor Julia Hippisley-Cox, Professor of Clinical Epidemiology and General Practice, University of Nottingham

Dr David Jewell, Editor, *British Journal of General Practice*

Mrs Loretta Jones, Patient Partnership Group**,** Royal College of General Practitioners – Northern Ireland

Mrs Susan Kinsey, Patient Partnership Group, Royal College of General Practitioners – Scotland

Dr Simon de Lusignan, Reader, Head of General Practice and Primary Care, St George's, University of London

Dr David Lynn, Head, Strategic Planning and Policy Unit**,** Wellcome Trust

Professor Sir Alex Markham, Chair, Research Capability Programme, Connecting for Health

Professor Nigel Mathers, Professor of Primary Medical Care, Unit Director, Academic Unit of Primary Medical Care**,** University of Sheffield, and Chair of Research, Royal College of General Practitioners

Dr Beth McCarron-Nash, British Medical Association General Practitioners' Committee

Ms June McGill, General Practice and Primary Care**,** University of Glasgow

Miss Rachel Merrett, Senior Policy Research Manager**,** British Medical Association

Dr Bob Milne, Senior Lecturer (Clinical), University of Aberdeen, and Chair, Scottish Clinical Information Management in Practice

Dr Helen Munn, Medical Science Policy, Academy of Medical Sciences

Professor Irwin Nazareth, Research Department of Primary Care and Population Health, University College London

Dr John Parkinson, Director, General Practice Research Database, Medicines and Healthcare products Regulatory Agency

Ms Nicola Perrin, Senior Policy Adviser, Strategic Planning and Policy Unit, Wellcome Trust

Mr Dave Roberts, NHS Information Centre for Health and Social Care

Professor Frank Sullivan, Community Health Sciences Division**,** University of Dundee

Mr Michael Walker, Patient Partnership Group, Royal College of General Practitioners

Professor Paul Wallace, UK Primary Care Research Network

Dr Mark Walport, Director, Wellcome Trust

Professor Graham Watt, Professor of General Practice**,** University of Glasgow

Mrs Carol Webley-Brown, Patient Liaison Group**,** British Medical Association

Ms Marlene Winfield, NHS Connecting for Health

BIBLIOGRAPHY

Academy of Medical Sciences. Personal Data for Public Good: Using health information in medical research. London: AMS; 2006. www.acmedsci.ac.uk/download.php?file=/images/project/Personal.pdf [accessed 6 May 2009].

Armstrong V, Barnett J, Cooper H, Monkman M, Moran-Ellis J, Shepherd R. Public Attitudes to Research Governance: A qualitative study in a deliberative context. London: Wellcome Trust; 2006. www.wellcome.ac.uk/About-us/Publications/Reports/Public-engagement/WTX038446.htm [accessed 21 April 2009].

Barrett G, Cassell JA, Peacock JL, Coleman MP; National Cancer Registry. National survey of British public's views on use of identifiable medical data by the National Cancer Registry. BMJ 2006;332(7549):1068–72.

Care Record Development Board. Report of the Care Record Development Board Working Group on the Secondary Uses of Patient Information. London: National Health Service; 2007. www.connectingforhealth.nhs.uk/crdb/workstreams/secusesreport.pdf [accessed 6 May 2009].

Department of Health. The NHS Constitution for England. London: Department of Health; 2009. www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_093419 [accessed 6 May 2009].

Department of Health. The Handbook to the NHS Constitution for England. London: Department of Health; 2009. www.dh.gov.uk/en/Publicationsandstatistics/Publications/PublicationsPolicyAndGuidance/DH_093421 [accessed 6 May 2009].

General Medical Council. Research: The role and responsibilities of doctors. Manchester: GMC; 2002. www.gmc-uk.org/guidance/current/library/research.asp [accessed 21 April 2009].

General Medical Council. Consent: Patients and doctors making decision together. London: GMC; 2008. www.gmc-uk.org/news/articles/Consent_guidance.pdf [accessed 21 April 2009].

Ipsos MORI. The Use of Personal Health Information in Medical Research. London: Medical Research Council; 2007. www.mrc.ac.uk/Utilities/Documentrecord/index.htm?d=MRC003810 [accessed 6 May 2009].

Lowrance WW. Learning from Experience: Privacy and the secondary use of data in health research. London: Nuffield Trust; 2002. www.nuffieldtrust.org.uk/ecomm/files/161202learning.pdf [accessed 6 May 2009].

Medical Research Council. Personal Information in Medical Research. London: MRC; 2000. www.mrc.ac.uk/Utilities/Documentrecord/index.htm?d=MRC002452 [accessed 6 May 2009].

Medical Research Council [2007a]. MRC Ethics Guide: Medical research involving adults who cannot consent. London: MRC; 2007. www.mrc.ac.uk/Utilities/Documentrecord/index.htm?d=MRC002409 [accessed 6 May 2009].

Medical Research Council [2007b]. MRC Ethics Guide: Medical research involving children. London: MRC; 2007. www.mrc.ac.uk/Utilities/Documentrecord/index.htm?d=MRC002430 [accessed 6 May 2009].

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 21 April 2009].

Royal College of General Practitioners. Research Ready Self-accreditation. www.rcgp.org.uk/clinical_and_research/circ/research/primary_care_research_team_ass/research_ready.aspx [accessed 21 April 2009].

Thomas R, Walport M. Data Sharing Review Report. London: Ministry of Justice; 2008. www.justice.gov.uk/reviews/docs/data-sharing-review-report.pdf [accessed 6 May 2009].

UK Clinical Research Collaboration R&D Advisory Group to Connecting for Health. Report of Research Simulations. London: UKCRC: 2007. www.ukcrc.org/pdf/CfH%20report%20June%2007%20full.pdf [accessed 6 May 2009].

The Wellcome Trust is the largest charity in the UK. It funds innovative biomedical research, in the UK and internationally, spending over £600 million each year to support the brightest scientists with the best ideas. The Wellcome Trust supports public debate about biomedical research and its impact on health and wellbeing.

www.wellcome.ac.uk

Strategic Planning and Policy Unit
Wellcome Trust
Gibbs Building
215 Euston Road
London NW1 2BE, UK
T +44 (0)20 7611 8646
F +44 (0)20 7611 8742
E sppu@wellcome.ac.uk

SP-4356.2/2k/05-2009/JS