

February 2008

Summary

- UKCRC Partners all agree on the potential benefit for patients of high quality research using patient data provided it can be done within the context of an appropriate regulatory framework
- UKCRC Partners collectively recognise the legitimate public concerns over the security and confidentiality of personal data of all kinds
- Some vital medical research cannot be conducted without access to identifiable information about patients
- Researchers are concerned about the disproportionate time, effort and cost currently required to obtain access to personal medical data. It has an impact on major research advances which in turn have an impact on economic development and quality of life in the UK
- The current law on data sharing is extremely complex and is open to differing interpretations with respect to use of identifiable medical data in circumstances where consent is not available
- There is confusion about the relationship between section 60 of the Health and Social Care Act 2002, the Data Protection Act, the common law of confidentiality and the Human Tissue Act 2004

Question 1 - Background

The UK Clinical Research Collaboration (UKCRC) welcomes the opportunity to respond to this consultation on the review of data sharing.

The UKCRC brings together organisations that shape the clinical research environment and includes the main UK funding bodies, academia, the NHS, regulators, industry, patient and public representatives. It was established in 2004 by the then chancellor of the Exchequer to establish the UK as a world leader in clinical research and in so doing to improve national health and increase national wealth.

As a broad-based partnership, the UKCRC encompasses a range of views on the use of personal information in the public and private sectors. However, UKCRC Partners all agree on the potential benefit for patients of high quality research using patient data provided it can be done within the context of an appropriate regulatory framework.

In 2005, the then Chancellor of the Exchequer and the then Health Secretary, stated a new commitment to develop the capability within the National Programme for IT to facilitate 'the gathering of data to support groundbreaking work on the health of the population and the effectiveness of health interventions'. This was reflected in Sir David Cooksey's Review of the UK Health Research, which identified an essential need 'to ensure that research is fully embedded in and integral to the NHS IT programme, and prioritised on par with other service uses for the system'. The unique opportunity to access data from the National Health Service, providing healthcare to a population of more than 50 million patients, is one that the research community is keen to develop.

Personal medical data that are of interest for research purposes such as safety monitoring and clinical trial research are often also of interest to those conducting public health studies, NHS clinical treatment and NHS management activity. Such data can be collected directly from patients, extracted from health service records and disease registers or other data sources.

The UKCRC has contributed to this agenda through the work of the UKCRC R&D Advisory Group to Connecting for Health and individual Partner organisations have produced a range of reports on the subject including: the Academy of Medical Sciences report, *Personal data for public good: Using Health Information in Medical Research*, published in January 2006; the Ipsos MORI/ MRC report on the use of personal health information in medical research; and the Wellcome Trust study on the public perspectives on the governance of biomedical research, both published in June 2007.

In this response, our focus is on the issues raised in sharing personal information for the purposes of clinical, public health and health services research. Our submission concentrates on issues raised by access to, sharing and use of personal medical information for these purposes. Since this research is varied in its aims and methodology, the data required may be in a form that can be fully anonymised or pseudonymised or for some purposes may need to be identifiable.

Question 2

What in your view are the key benefits of sharing personal information to a) individuals and b) society? Please provide examples.

Data are collected for many different types of research study such as clinical trials, safety monitoring and epidemiology. In the case of epidemiological research where researchers require information about groups of patients for their studies rather than individual patients, data still need to be linkable and comprehensive at the individual patient level for reasons of research methodology. Sometimes, to ensure maximum value can be obtained from data and to allow quality and completeness of data to be validated, some identifiers need to be available to researchers.

The benefits for public health are in establishing connections between data sources such as primary care patient records and disease registry records. This enables exploration of trends and associations which may have a potential impact on public health and patient care. Examples of such studies include the health of miners¹ and the social distribution of cancer². Importantly, society benefits from improvements to population health resulting from research outcomes achieved through linkage of currently available data rather than research outcomes determined from exposing members of the public to the risks of experimental or long-term treatments.

Medical research is performed primarily for public rather than individual personal benefit however there can also be benefit at an individual patient level. An example of personal benefit is often seen in oncology where patients have the opportunity to access new therapies as part of enrolment in clinical trials. Access to these clinical trials relies on sharing personal medical information about patients with clinical researchers who are recruiting to studies.

Question 3

What in your view are the key risks of sharing personal information to a) individuals and b) society? Please provide examples.

UKCRC Partners and stakeholders acknowledge that the opportunity to increase access to personal health data brings new challenges and responsibilities for safeguarding patient privacy. The risks mainly arise from inappropriate sharing or loss of data where there is a lack of well-trained staff or effective systems of data security.

However modern health services cannot advance or be delivered for the benefit of patients without clinical, public health and health services research. Personal information held as e-health records can be extremely sensitive and inappropriate use or disclosure has the potential to cause harm or distress. The

¹ Fox AJ, Goldblatt P & Klein LJ (1981) A study of mortality of Cornish tin miners. *British Journal of Industrial Medicine* 38, 378 – 80

² Kilne LJ (1988) The longitudinal study and the social distribution of cancer. *British Medical Journal* 297, 1070

research community is overwhelmingly aware that experience of inappropriate disclosure might induce patients to withhold information from a health professional or even avoid treatment altogether.

The research community is keen that systems to regulate access to personal data strike a balance between the need to protect patient privacy and the potential for great individual and public health benefit in using this data for effective research. However it should be noted that the great majority of the research uses of personal data are of low risk, focusing on patient or population groups rather than individuals. It should also be noted that there is an established system of compulsory research ethics review which oversees research operating procedure in the UK and researchers in the NHS understand and accept that they are also bound by a duty of confidentiality.

Question 4

What scope and what methods of personal information sharing, in your view, pose the greatest opportunities or risks? Please explain the reasoning behind your response.

The UKCRC R&D Advisory Group to Connecting for Health was established in 2006 under sponsorship from the Department for Health and NHS Connecting for Health in England to consider the feasibility of using electronic personal data from the NHS Care Records Service to conduct high quality research for patient benefit. The report³ of the group was published online in summer 2007 and made a number of recommendations including some on data access, availability, data quality and mandating use of the NHS number. There was a close working relationship with the Care Record Development Boards (CRDB) working group on secondary uses of patient data⁴. Both reports discuss the need for demonstrably effective information governance that is also responsive to the needs of research. The reports also propose the development of an enhanced infrastructure. This could link and integrate patient information in a secure format, to provide a comprehensive service for research, ultimately with direct benefits for patients.

These two reports recommend development of honest brokers as a way of engendering trust amongst patients, clinical professionals and the research community. The honest broker would have a dual role, ensuring patient confidentiality and security as well as the scientific integrity of the data delivered to the research community.

In addition, the CRDB report recommended the development and use of safe havens to facilitate access to identifying data for research where this is essential. Safe Havens are designated physical or electronic areas which provide appropriate levels of security for the use of the most sensitive and confidential

³ UKCRC R&D Advisory Group to Connecting for Health: Report of the research simulations. June 2007
<http://www.ukcrc.org/publications/reports.aspx>

⁴ Report of Care Record Development Board (CRDB) Working Group on Secondary Uses of Patient Information. August 2007 <http://www.connectingforhealth.nhs.uk/crdb/workstreams/secusesreport.pdf>

information. Examples of this model are found in Canada where research data centres⁵ are used as a means of providing researchers with access to confidential micro data in a controlled environment.

The reports demonstrated that development of an enhanced infrastructure presents opportunities for research in the UK. It will enhance our research communities' ability to link together information about the same individual in order to assess patterns in the outcome or causes of disease, from which general inferences can be drawn. The greatest risks are likely to occur where secure procedures for data-handling are absent or are not understood or monitored. Data in which personal identifiers are minimal and encrypted or held separately from the main clinical database reduce significantly the risk of disclosure. The focusing of data handling and linkage into centres where procedures are well established will go a long way towards reducing this risk.

Question 5

Please provide examples of where, in your view, the public authorities hold too much data or not enough personal information, and the reasoning behind your response.

UKCRC Partners call for an improvement in the completeness of disease registries, such as cancer registration, where missing information introduces biases in the data, leading to potentially misleading claims or even costing lives. Some UKCRC stakeholders have also indicated that birth and death data held by ONS (GRO⁶ in Scotland), NHS outpatient and inpatient records and GP data could be linked more effectively. For example, one use of such data might be to trace outcomes after the antenatal anomaly scan to determine whether anomalies are detected accurately or missed. Often negative scans are not followed up in research studies, especially if women attend a regional centre, and so 'missed' diagnoses are not effectively recorded. Anomaly registers are incomplete as they require anomalies to be reported at birth. In cases of termination/miscarriage, linking of records from the maternal antenatal period and child after birth would provide a more complete picture of congenital anomalies in the UK.

The Human Fertilisation and Embryology Authority (HFEA) holds information on all women undergoing IVF treatment. However, researchers have experienced difficulties in accessing this data and it is not held in a way that facilitates research use. There is therefore a marked lack of research information as to the medium- and long-term health impacts of IVF treatment on both women and children. Also, it is not presently possible for this information to be effectively linked to other registers or personal health records to allow long-term effects to be discovered.

Question 7

Please provide examples of cases where you believe the sharing of personal information between two or more bodies would be beneficial, but where it is not currently taking place.

⁵ Statistics Canada. The Research Data Centre www.statcan.ca/english/rdc/index.htm

⁶ General Register Office for Scotland <http://www.gro-scotland.gov.uk/>

Please explain as fully as possible why information is not being shared, detailing what the barriers to the sharing of personal information are – e.g. legal, cultural, financial, institutional – and how these barriers can be overcome.

In the report of the UKCRC research simulations there is an emphasis on the importance of improving the ability to access and link data sources at the individual patient level from primary health care, NHS pathology services, private sector treatment and disease registers.

In large national cohort studies of individuals followed up long-term, researchers require access to data from a wide range of records, both within and without the health services. The UKCRC research simulation report outlines the importance of taking steps to ensure that, in future, NHS data structures and systems include the ability to combine data from different sources comprehensively at the individual level and family level using a unique patient identifier. This could be achieved relatively simply for data sources within the current NHS Care Record Service by building on the achievements of NHS Numbers for Babies. The use of the NHS number could be mandated as a unique identifier in all health care settings and encounters, starting with key maternity and birth datasets.

Question 9

In your view, how well does the DPA work? Please outline the DPA's main strengths and weaknesses and any proposals for changes you would like to see made, including suggestions for their implementation.

Researchers consider that the main weaknesses of the data protection act as the law governing data sharing is the confusing and overly complex language. The resulting variation in interpretation of the Act by those who locally govern access to personal data appears to contrast unfavourably with the approach taken by the Information Commissioner's office to similar questions.

UKCRC Partner concerns about data sharing stem directly from the fact that there is a lack of agreement about whether researchers can access identifiable patient data without explicit consent, if they are not part of the direct care team. This introduces two important issues for researchers:

- The need to obtain consent to access medical records in order to identify eligible research participants to invite them to participate in a research study
- Confusion about whether general consent or opt-out consent qualifies as valid consent for information processing under the Data Protection Act 1998

The use of personal data from patient health records in research was the subject of a study by the Academy of Medical Sciences in 2006. The study culminated in the publication of a report⁷, *Personal Data*

⁷ Personal Data for Public Good: Using Health Information in Medical Research. Academy of Medical Sciences January 2006 <http://www.acmedsci.ac.uk/p48prid5.html>

for Public Good: Using Health Information in Medical Research. Publication of this report was followed in June 2006 by a symposium⁸ involving senior members of the legal profession and academics.

In its report, the Academy of Medical Sciences discuss the definition of fair processing as defined in the Data Protection Act and its difficulties for research using personal data. Research is often a secondary use of the personal data that has been collected. As a result it is not always possible to predict the research purposes for which data might be used in the future. Research using personal data often involves very large numbers of individuals' data who cannot be readily contacted with fair processing information.

Question 13

Are there any other aspects of UK or EU law (such as EU Directive 95/46/EC) that impact positively or negatively on data sharing or data protection? Please provide examples.

The impact of Section 60 of the Health and Social Care Act is the subject of ongoing debate amongst UKCRC Partners. There is also widespread confusion about the relationship between section 60, the Data Protection Act and the common law of confidentiality. The complex overlap between the Data Protection Act and the Human Tissue Act 2004 also introduces another element of confusion for the medical research community.

Members of the medical research community seek to obtain consent for access to records wherever possible. They however remain unconvinced that it is practical to obtain consent to perform the 'identification sweep'.

Question 16

Is it clear whether and when you need individuals' consent to share information about them? Are you clear about the form that consent should take? Please provide examples.

Please provide details of any initiative you have been involved in that has been based on consent.

As outlined in question 9 above, there is considerable confusion about 'consent to consent': using local clinicians to obtain consent for a member of the research team to obtain their consent for enrolment in a clinical study. Researchers indicate that this requirement fails to recognise the reality of relying on clinical teams. In particular, GP practices have many other competing priorities.

Explicit consent is often given by patients to an authorised member of the research team for the particular research. In practice, studies are sometimes extended or varied and so re-consent would be required. As a result it is important to clarify whether general consent or opt-out consent qualifies as valid consent for these purposes. Whilst the research community accepts that explicit consent for access to personal data

⁸ Personal Data for Public Good: Report of proceedings at the legal symposium. Academy of Medical Sciences June 2006 <http://www.acmedsci.ac.uk/download.php?file=/images/project/1170326729.pdf>

for research purposes is ideal, many researchers consider that it may not always be practical to obtain this on every occasion. An example of this is where blood or tissue samples have been obtained as part of a clinical trial and consent obtained for specific measurements or uses. If at a later date researchers have a need to look at those samples again or share information about, for instance, a new safety issue or marker further consent would need to be obtained. Tracing patients to obtain further consent is time consuming, difficult and expensive.

If explicit consent is demanded then this will have a negative impact on the ability to conduct large scale studies. There is a risk of research bias: individuals who move often may be difficult to trace for consent, it may be inappropriate to approach some individuals for consent (for example to ask parents if they agree to their child's death registration being examined) or certain groups may withhold consent for reasons not linked to the research.

Question 17

What, if any, barriers would a requirement for gaining consent create to the sharing of personal information? Please explain your reasoning.

Researchers are concerned about the disproportionate time, effort and cost currently required to obtain access to data. In particular, some UKCRC partners have expressed concern about the impact of section 60 of the Health and Social Care Act 2002. They point to the resulting impact on their ability to recruit to studies and make progress with major research advances.

Whilst people have reservations about the implications of enabling access to their medical records for research, for their privacy, they are generally sympathetic to medical research. If people feel in control of their own information and its potential uses, they are likely to support research.

Available evidence suggests that the current level of public awareness in relation to the use of medical records in research is low. The UKCRC is keen to see urgent work to increase public engagement around the value of research using health care records and the circumstances and procedures by which such records may be accessed for research purposes.

In two recent reports by UKCRC Partners - the Ipsos Mori/ MRC consultation on public attitudes toward the secondary use of personal health information for medical research (June 2007) ⁹ and the Wellcome Trust Public attitudes to research governance (June 2007) ¹⁰ – researchers investigated whether patients are content to have people other than those providing direct care (in this case researchers) accessing their medical records. The findings revealed that when the public is informed about what medical research

⁹ Ipsos MORI/ MRC report on the use of personal health information in medical research – General public consultation (June 2007) www.mrc.ac.uk/utilities/documentrecord/index.htm?d=MRC003810

¹⁰ Wellcome Trust report on the public perspectives on the governance of biomedical research: A qualitative study in a deliberative context (June 2007) www.wellcome.ac.uk/doc_wtx038446.html

entails they are generally positive towards it. In order to build public trust in the activities of medical researchers there needs to be better communication. There is a tension in the public mind between individual privacy and the greater good, and anonymity and consent are central to this. As a result it is crucial that communication should include a discussion about anonymity and consent as part of building trust.

Question 18

Do you have any suggestions on how to make the sharing of information more transparent?

For example, should individuals be given strengthened access rights? And if so, how? Should organisations be expected to do more to explain their use and sharing of personal information to the public? And if so, how?

There is a real need to re-focus this question and to address a wider issue than simple transparency about what data are held and where. It is now important to start a public discussion about why data are shared and the benefits of data held and accessed for the public good. There is an assumption that individuals want organisations to merely describe what is happening during data-sharing, whereas it may be more important for organisations to describe why data sharing is important and the benefits it may bring.

It is already good practice for researchers to disseminate the outcome of their research to a wider stakeholder community including research participants. This is distinct from enabling stronger access rights, for research participants, to information collected or generated about them in the course of research. For example, it could weaken confidentiality of research information (because data and analyses would need to be easily identified with individuals). While strengthened access rights might work in other sectors, this could pose serious problems for information collected in the course of medical research.

Question 21

Should the law mandate specific technical safeguards for protecting personal information?

For example, should there be an explicit requirement that all personal information held on portable devices be encrypted to a particular standard?

Technical requirements can provide guidance and a basic level of security that should be adhered to. However, technical standards are in a continual state of development and change and consideration should be given to the fact that very specific requirements in law may become rapidly outdated with technical advances.

Question 22

How, in your view, could 'privacy enhancing techniques', such as the anonymisation or pseudonymisation of personal information, help safeguard personal privacy, whilst facilitating activities such as performing medical research?

Is sufficient advice about the deployment of such techniques available? Are you confident about using them? What are the barriers to using them?

Pseudonymisation is very poorly defined - in some cases, the NHS number appears to be considered as a pseudonymised identifier but at other times it is not. Any recommendation about pseudonymisation of data needs to clearly define how the terms are being used and how it is implemented in practice.

In addition to personal identifiers collected for identifying individuals are those personal identifiers which are also used for clinical purposes, e.g date of birth and death in survival analyses, postcodes in geographical analyses of survival, ethnicity and sex. It is difficult to see how clinical data can be anonymised, pseudonymised or removed without severely limiting the outcomes from a study.

UKCRC Partners consider that it is important to undertake anonymisation as far as possible within a study database. The Partners also consider that identifiers that are required for clinical analysis and those that are not should be separated. However, further security measures are then more likely to involve password-protection, encryption of data for transfer, etc.

Question 24

Do you have any international examples of good practice in the sharing of personal information that could or should be adopted by the UK?

Examples of international good practice are listed in the Wellcome Trust/ UKCRC report of the frontiers meeting on the use of electronic patient records for research and health benefit. They include the US Veterans Health Administration data system which demonstrates worthwhile care benefits for patients in allowing access by medical researchers.

In the report of the UKCRC research simulations there are a number of examples of International linked systems which benefit population health. These include Denmark, where a powerful set of health data sources have been developed as an integrated system. These data sources are linked in a framework with appropriate governance and have allowed a range of trials, patient and public safety studies and epidemiological research to be carried out. An impressive example of this was their ability to link MMR-vaccination status obtained from the Danish National Board of Health to information on the children's autism status from the Danish Psychiatric Central Register, as well as to the Danish Medical Birth Registry, the National Hospital Registry, and Statistics Denmark, which were sources of relevant variables and potential confounding factors. Using these data, 537,303 children were identified, of whom 440,655 (82%) had received the MMR vaccine. Among these children, 316 had received a diagnosis of autistic disorder and 422 a diagnosis of other autistic-spectrum disorders. The authors found no association between the age at the time of vaccination, the time since vaccination, or the date of vaccination and the development of autistic disorder, providing strong evidence against the hypothesis that MMR vaccination causes autism.

Question 27

Are there any additional issues on the sharing of personal information and protection of personal information that this review should be considering?

Do any of these issues apply specifically to your sector?

There have, in recent months, been examples in many settings of the way in which data sharing can lead to major research advances which have impact on the economic development and quality of life of the UK. One of the features of these examples has been the multiplicity of data types and sources being used. For example, the power of high speed computers is increasingly leading to the ability to use a combination of social survey, textual, visual and administrative data to address important social phenomena.

A lot of investment is being made by the public and charity sectors to take forward this methodology and, as a result, it is highly likely that the trajectory of new opportunities over the next few years will be steep. Therefore it is important that a climate exists where the UK has the opportunity to be at the forefront of world developments in these areas. This requires a) that any regulation enables research, while being firm on any misuse of responsibility; b) that specification of data covered by the review is flexible so as to allow for future advances and opportunities; c) that it is acknowledged that a programme of proper public engagement is urgent and necessary; and d) that encouragement is given to further research, both substantive and methodological.