The Academy of Medical Sciences

Personal data for public good: using health information in medical research

January 2006

The independent Academy of Medical Sciences promotes advances in medical science and campaigns to ensure these are translated as quickly as possible into benefits for patients. The Academy's Fellows are the United Kingdom's leading medical scientists from academia, hospitals, industry and the public service.

The aims of the Academy are to:

- Give national and international leadership in the medical sciences
- Promote the application of research to the practice of medicine and to the advancement of human health and welfare
- Promote the aims and ethos of medical sciences with particular emphasis on excellence in research and training
- Enhance public understanding of the medical sciences and their impact on society
- Assess and advise on issues of medical science of public concern

The Academy of Medical Sciences was established in 1998 following the recommendations of a working group chaired by Sir Michael Atiyah OM FRS HonFMedSci, Past President of the Royal Society. The Academy currently has a Fellowship of over 800.

There is an elected Council of 23 Fellows that includes the five Honorary Officers of the Academy:

President	*Sir Keith Peters FRS PMedSci*
Vice-President	*Sir John Skehel FRS FMedSci*
Vice-President	*Sir Michael Rutter CBE FRS FBA FMedSci*
Treasurer	*Professor Ian Lauder FMedSci*
Registrar	*Professor Patrick Vallance FMedSci*

The Academy's Executive Director is *Mrs Mary Manning*.

For more information about the work of the Academy see www.acmedsci.ac.uk

The Academy of Medical Sciences is a company limited by guarantee.

Registered Charity No. 1070618
Registered Company No. 3520281
Registered in England
ISBN No. 1-903401-11-9

The **Academy** *of*
Medical Sciences

Personal data for public good: using health information in medical research

A Report from the Academy of Medical Sciences

January 2006

Acknowledgements
The Academy of Medical Sciences is most grateful to Professor Robert Souhami CBE FMedSci and to the members of the Working Group for undertaking this report. It thanks the Review Group, the Academy's Officers and all respondents to the consultation for their instructive comments and support. The Academy is grateful to Cancer Research UK for its generous support of this project.

Cover photograph: Science Photo Library

Contents

Abbreviations

AMRC	Association of Medical Research Charities
BHF	British Heart Foundation
BMA	British Medical Association
CCTV	Closed Circuit Television
CHI	Community Health Index
COREC	Central Office for Research Ethics Committees
CRDB	Care Record Development Board
CRG	Care Record Guarantee
CRS	Care Record Service
CRUK	Cancer Research UK
CSAGS	Confidentiality and Security Advisory Group Scotland
DPA	Data Protection Act
DPP	Data Protection Principle
EASTR	Epidemiology and Survival of Transfusion Recipients
EU	European Union
GMC	General Medical Council
GPRD	General Practice Research Database
HCR	Honorary Contract for Researchers
HIPAA	Health Insurance Portability and Accountability Act
HIV	Human Immunodeficiency Virus
HPA	Health Protection Agency
HPS	Heart Protection Study
HRA	Human Rights Act
HSCA	Health & Social Care Act 2001
ICO	Information Commissioner's Office
IT	Information Technology
LREC	Local Research Ethics Committee
MEMO	Medicines Monitoring Unit
MRC	Medical Research Council
MREC	Multi-centre Research Ethics Committee
NHS	National Health Service
NPfIT	National Programme for Information Technology
NWCS	NHS-Wide Clearing Service
OST	Office of Science and Technology
PCT	Primary Care Trust
PIAG	Patient Information Advisory Group
R&D	Research and Development
RCT	Randomised Controlled Trials
REC	Research Ethics Committee
RGF	Research Governance Framework
SARS	Severe Acute Respiratory Syndrome
SUS	Secondary Uses Service
UC	Ulcerative Colitis
UKCRC	UK Clinical Research Collaboration

Summary

Countless lives have been saved or improved because of medical research using health information. This kind of research has identified important causes of disease, led to effective measures for control of epidemics, demonstrated the long-term effects of treatment, and shown how the health of the population can be improved by the better provision of services. The United Kingdom already has an outstanding record in this area of research. We now have the potential to become a world leader through the opportunities afforded by the National Health Service (NHS) and new initiatives to develop national electronic care records.

However, evidence submitted to the Academy shows that advances in this field are increasingly inhibited by inappropriate constraints on the use of personal health data. These constraints arise through confusing legislation and professional guidance, bureaucracy of process and an undue emphasis on privacy and autonomy. It is essential that data about the health of individuals are only used for research under conditions of confidentiality that enjoy public support. However, evidence of public attitudes towards the use of health information in research is largely absent, forcing regulatory and advisory bodies to make assumptions about what the public might find acceptable. These factors have created a conservative culture of governance, where disproportionate constraints are imposed on research that can compromise its quality and validity. The difficulties of the current situation are a significant disincentive for researchers to undertake work in this field and are detrimental to research aimed at improving public health.

The public, patients and researchers have a common interest in ensuring that research using personal data is conducted efficiently and to the highest standards. Implementing solutions to alleviate the current situation will require coordinated and concerted effort by all concerned with this research. We have been encouraged by the strong desire expressed by those we consulted to see the position improved and hope this report will provide the stimulus for effective action.

In this summary we present the major conclusions on which our recommendations are based. Further discussion can be found in the relevant sections of the main report. Our conclusions and recommendations are presented in the following areas:

1. **Interpreting the legal framework**
2. **Improving regulatory processes**
3. **Developing good practice in research using personal data, including issues related to anonymisation and consent**
4. **Harnessing the opportunities of the NHS National IT programme**
5. **Engaging the public**

1. Interpreting the legal framework

The legal framework around the use of personal data in research is a complicated patchwork involving UK legislation, case decisions and European directives, augmented by various guidance documents. There are many areas of imprecision, and the courts have not tested the legislation as it applies to medical research. Those responsible for research approval decisions have made their judgements within this uncertain legal framework. The resulting variable legal interpretations have been a source of great difficulty, delay and disillusionment for researchers.

Legal uncertainty and an undue emphasis on privacy and autonomy have created a conservative culture of research governance, in which regulatory and professional bodies promote a policy of 'consent or anonymise'. The Academy firmly believes that researchers should employ adequate data security policies, which may involve anonymisation or pseudonymisation techniques where appropriate,

and should seek consent where it is feasible and proportionate. However, the 'consent or anonymise' policy advocated by some authorities is **not** a strict legal requirement. The rigid application of this policy has been detrimental to research in terms of financial and time resources, as well as scientific opportunity and value. Measures designed to protect autonomy and privacy must be considered against the societal costs of diminishing the quality of the research, or of not doing the research at all.

The key point is one of necessity and proportionality: **the law will allow the use of identifiable data for medical research without consent, provided that such use is necessary and is proportionate with respect to privacy and public interest benefits.**

Identifiable data can be used for medical research without consent, provided that such use is necessary and is proportionate with respect to privacy and public interest benefits. Research governance bodies, including the Patient Information Advisory Group, Information Commissioner's Office, research ethics committees, NHS research governance offices and General Medical Council should accept this interpretation in their guidance and approval decisions.

2. Improving regulatory processes

Research involving personal data has been damaged by the complexity, inconsistency and length of time involved in the assessment of research proposals. There is an urgent need for a simplified scheme for assessing a research proposal involving personal data that maintains standards but also reduces the number of steps a proposal must take. The Academy considers that, in the short term, consistent decision making would be facilitated through improved and more formalised communication channels between regulatory bodies, in collaboration with research funders and researchers, and greater transparency of the reasoning behind decisions on individual projects. The Academy considers that the development of joint electronic application forms (including PIAG, RECs and NHS R&D) and expansion of the 'Research Passport' scheme for honorary NHS contracts should also be accelerated.

The UK Clinical Research Collaboration should lead the bodies involved in governance of research using personal data in developing a simple scheme of assessment for proposals and issue clear guidance on the approval process.

The Patient Information Advisory Group (PIAG) is a temporary statutory body that decides whether research projects should be able to use identifiable data without consent and so be granted exemption from the common law of confidentiality. It was established as a temporary body, pending a time when all research using patient data would be conducted with 'consent or anonymisation'.

The Academy considers that 'consent or anonymisation' will never be feasible for a great deal of research using personal data, regardless of potential technical developments. The role of PIAG in a statutory system that can provide immunity from liability is an important way of providing data controllers with reassurance that they may legitimately release data to researchers, often a key component in progressing a research project. The Academy considers that there will be a continuing need for a body with special authority in this area.

This report describes several areas of concern regarding PIAG's current approach, processes and membership. In its communications, PIAG currently stresses its role in protecting privacy and confidentiality, without equal emphasis on

the public benefits derived from well-conducted research. The Academy considers that PIAG should more actively promote its role as a facilitator of research. Relations with the research community have not been aided by the lack of a mechanism for independent appeal of PIAG's decisions (in contrast with the research ethics committee system). In its operations, PIAG should develop an extended and explicit system of class support, whereby applications meeting specific criteria are fast-tracked through the system without detailed review by the committee. The Academy also considers that the current membership of PIAG should include greater representation of active researchers and the inclusion of lay members from medical research charities.

There is a continuing need for a body such as the Patient Information Advisory Group (PIAG) with statutory authority in this area. However, PIAG should address the difficulties of approach, process and membership identified in this report and develop an extended and explicit system of class support, whereby its involvement in research proposals becomes the exception, rather than the norm.

3. Developing good practice in research using personal data, including issues related to anonymisation and consent

To obtain and deserve public support, the research community must demonstrate that research using personal data is always performed to a high standard and within appropriate safeguards. It is essential that researchers working with personal data are fully aware of the relevant legislation and underlying ethical principles, as well as of research governance policy and processes. The Academy considers that the development of Good Practice Guidance would encourage high standards of research, as well as facilitating consistency in approval decisions. The guidance should be used as a set of practical exemplars around which researchers can develop research proposals and not as a checklist for assessment. It should also take account of developments in research methodologies through regular review and involve newly established bodies with special responsibilities in this area (e.g. Human Tissue Authority, Connecting for Health).

The UK Clinical Research Collaboration should lead an initiative involving the regulatory and professional bodies, the medical research community and the public to develop Good Practice Guidance in research using personal data. Such guidance should encompass issues related to data security, anonymisation, consent and the use of health records to identify research participants.

Areas to be addressed in the Good Practice Guidance are outlined on page 7.

4. Harnessing the opportunities of the NHS National IT programme

The NHS National Programme for IT (delivered through Connecting for Health) offers an exceptional opportunity to allow research to inform all aspects of health care. However, the Academy is concerned that research needs are not being integrated into its development. This may undermine the research capability of the NHS and weaken the additional opportunities arising from UK Treasury commitments to large-scale financial support for NHS research[1].

The Academy is concerned with the current wording of the Care Record Guarantee (CRG), which sets out for the public the rules that will govern information held in the NHS Care Records Service. The Guarantee makes commitments that, if strictly interpreted, would prevent many research projects from using Connecting for Health data. The Academy has

1 http://www.hm-treasury.gov.uk/newsroom_and_speeches/press/2005/press_100_05.cfm.

held a constructive consultation with the Care Record Development Board and has suggested changes to the CRG to further clarify the position around the use of information in research. A revision of the CRG is underway at the time of going to press. We strongly urge the development of effective methods of research support within Connecting for Health and the promotion of the benefits of research during the associated public engagement campaign.

'Connecting for Health' should take urgent steps to address the needs of research through the establishment of a Research Advisory Committee. The Care Record Guarantee should be further revised to include support for research as an important and legitimate secondary use of Connecting for Health data, while emphasising the appropriate safeguards.

5. Engaging the public

The Working Group's consultation with patients and patient representatives revealed strong support for research using personal data and confidence in the integrity of research practices. However, evidence of public attitudes and opinions on the specific issue of research using personal data is largely lacking. The absence of such knowledge, and the lack of public debate, forces regulatory and advisory bodies to make assumptions about what the public might find acceptable. Development of good practice should be informed, as far as possible, by empirical evidence on public and patients' awareness and attitudes.

Research funders should encourage and fund research into public awareness and attitudes towards medical research using personal data.

The ethical basis for accessing and using patient records for a research study, with or without consent, depends greatly upon public expectations about how routine health records are used. Urgent work is needed to increase public engagement about the value of research using health care records and the arrangements under which records are held and accessed.

Researchers, research funders, regulatory bodies and universities could do much to engage the public around the benefits of research involving personal data and to demonstrate that high standards are consistently applied. Charities with strong patient/user input could play a particularly important role in more actively advocating the value of research using personal data. Collaborative activity between the members of the UK Clinical Research Collaboration (UKCRC) would be beneficial. Ultimately, there is a need for the UK Departments of Health to undertake a programme of public engagement around these issues.

The UK Departments of Health, working with the UK Clinical Research Collaboration, should develop public engagement programmes around the purpose and value of using personal data in medical research.

Areas to be included in Good Practice Guidance

1. Data security and anonymisation

Anonymisation of data is never an absolute process; there are different degrees of anonymisation that depend on the particular context. For reasons outlined in the report, most important research using personal data requires access to identifiable data at some point for some purpose. Reversible anonymisation (involving key-coded data) can provide a solution. We consider that the additional level of security gained from pseudonymisation (where researchers do not have access to the key) is extremely small compared with the use of coded identifiable data sets under a strict data security policy. Destruction of the key should almost never be necessary. However, we emphasise the responsibilities of researchers in implementing adequate data security policies and consider this an area where improvements could be made.

Research organisations should take steps to review the adequacy of their data security policies. Similarly, funding agencies should be satisfied that researchers and their host institutions have appropriate data security arrangements in place.

Good Practice Guidance should address:

- **methods of data security, including physical, technical and procedural security**
- **who can carry out anonymisation and under what circumstances**
- **'strong' and 'weak' identifiers and the hierarchical removal of identifiers to leave 'more' or 'less' identifiable data**
- **the holder of the encryption key and management of access.**

2. Consent

Researchers experience variable consent requirements for research using personal data from different regulatory bodies. Insistence on explicit consent can impose insupportable time and resource costs. It can also lead to bias in population coverage thus diminishing the value of the research, to the detriment of sections of society.

Good Practice Guidance should be developed around consent requirements for research using personal data with reference to the following criteria:

- **the risk of introducing bias that will endanger the validity of the results**
- **the size of the study population and the proportion likely to be untraceable**
- **the overall financial and time burdens imposed**
- **the risk of inflicting harm or distress by contacting people.**

3. The use of health records to identify research participants

Unresolved questions remain over whether, how, and by whom identifiable patient records may be accessed in order to identify and subsequently contact potential research participants. An insistence that only the medical practitioner responsible for an individual's care can access records imposes significant time and financial costs that can exclude a large research population and cause bias in the research results.

Good Practice Guidance should include:

- **the conditions and procedures by which health records may be accessed at the start of the research process**
- **the mechanism for contacting potential study recruits**
- **the mechanism for registering agreement or refusal to participate.**

Recommendations

1. Interpreting the legal framework
Identifiable data can be used for medical research without consent, provided that such use is necessary and is proportionate with respect to privacy and public interest benefits. Research governance bodies, including the Patient Information Advisory Group, Information Commissioner's Office, research ethics committees, NHS research governance offices and General Medical Council should accept this interpretation in their guidance and approval decisions.

2. Improving regulatory processes
The UK Clinical Research Collaboration should lead the bodies involved in governance of research using personal data in developing a simple scheme of assessment for proposals and issue clear guidance on the approval process.

There is a continuing need for a body such as the Patient Information Advisory Group with statutory authority in this area. However, PIAG should address the difficulties of approach, process and membership identified in this report and develop an extended and explicit system of class support, whereby its involvement in research proposals becomes the exception, rather than the norm.

3. Developing good practice in research using personal data
The UK Clinical Research Collaboration should lead an initiative involving the regulatory and professional bodies, the medical research community and the public to develop Good Practice Guidance in research using personal data. Such guidance should encompass issues related to data security, anonymisation, consent and the use of health records to identify research participants.

4. Harnessing the opportunities of the NHS National IT programme
'Connecting for Health' should take urgent steps to address the needs of research through the establishment of a Research Advisory Committee. The Care Record Guarantee should be revised to include support for research as an important and legitimate secondary use of Connecting for Health data, while emphasising the appropriate safeguards.

5. Engaging the public
Research funders should encourage and fund research into public awareness and attitudes towards medical research using personal data.

The UK Departments of Health, working with the UK Clinical Research Collaboration, should develop public engagement programmes around the purpose and value of using personal data in medical research.

Scope of the Report

The Academy of Medical Sciences 2003 report 'Strengthening Clinical Research' highlighted both the opportunities for population-based research in the UK and the concern among the medical research community that advances in this field were being inhibited by unnecessary constraints on the use of personal data.

In 2004, the Academy was increasingly aware of anecdotal evidence that medical research using personal data was being impeded by legal and administrative complexity and confusion. In response to these concerns, the Academy established a Working Group with a remit to '*examine the current and likely future UK position regarding the use of personal data in medical research*.'

In so doing, to:

- analyse the development of the present position regarding the use of personal data and the necessity and requirements for this type of research
- analyse the advantages and problems of the national and international regulatory environment in this area
- make recommendations for dealing with key issues of consent, security of data, confidentiality and public engagement.

The remit excluded the Working Group from:

- providing practical guidance for researchers on how to negotiate the current environment governing the use of patient data in research
- considering the use of human tissue as governed by the Human Tissue Act
- providing practical guidance to clinicians on obtaining patient consent.

Details of the Working Group and preparation of this report are given in Appendix I.

The Academy issued a call for evidence in October 2004, to which 70 written submissions were received from a wide range of individuals and organisations. Selected quotes from these submissions can be found throughout the report. Oral evidence from key organisations and individuals, extensive legal advice and many published papers have also been considered in producing this report. The Academy held a consultation meeting in May 2005, which was well attended by those representing the public as patients or participants in research. Organisations and individuals who were consulted, and who have forwarded evidence, are listed in Appendix II.

This report is designed for policy makers in Government, research funders, universities, NHS Trusts, patient groups and relevant professional and regulatory bodies, as well as all other interested parties.

Personal data

The Data Protection Act 1998 defines personal data as:
'Data which relate to a living individual who can be identified from those data, or from those data and other information which is in the possession of, or is likely to come into the possession of, the data controller.'

Throughout this report we use the term personal data to refer to information about individuals that may be used in medical research. This information can include health data (e.g. cholesterol level or hospital visit dates) and non-health data (e.g. postcode or occupation).

1 History, opportunities and challenges

Summary

- *Research using personal data has benefited public health by identifying the causes and changing patterns of disease, improving therapeutic practice and the use of health care services, and by indicating promising areas of research. The UK has an outstanding scientific record in this area.*
- *New opportunities for research using personal data are now available in the UK, including the development of the NHS electronic care record. Data derived from patient care within the NHS will provide one of the largest sources of research information in the world.*
- *These exceptional research opportunities are accompanied by important challenges concerning the right to privacy, the sensitive nature of some health data and the importance of patient's trust in the confidentiality of their care. Research must be undertaken within the framework of the law and in accordance with public expectations.*
- *The legal framework is complex and there are numerous UK regulatory agencies whose decisions impact on research programmes using personal data. Evidence indicates that current UK regulatory mechanisms are presenting barriers to medical research that are disproportionate to the risks involved.*

1.1 Introduction

Research using personal data has benefited the health of the public and greatly reduced the burden of disease. The public, patients and researchers have a common interest in ensuring that research using personal data is conducted efficiently and to the highest standards.

Exceptional failures in medical practices (such as the storage of organ samples without consent at Alder Hey) emphasise the general need to protect public trust and ensure that research and other medical practices demonstrably conform to high standards. These standards must apply even when the research relies on medical records, and there is no contact with patients or relatives. Increasingly sophisticated methods of data collection, storage and analysis have generated powerful new research opportunities, but have also led to calls for greater controls on the use and transfer of data. In all aspects of public and commercial life, the legal framework concerning data protection and the right to privacy has been changing to meet these concerns.

The medical record has also evolved in recent years and is no longer simply a summary of patient consultations, but an essential method of sharing information between healthcare professionals. A wide range of personal data are now included in medical records, such as information on lifestyle and family history, clinical and social factors, as well as diagnostic and other test results.

Accurate and timely sharing of personal data is essential for functions not directly connected to individual treatment, but which help to ensure that the delivery of health care is high quality, cost-effective, efficient and evidence-based. Such **secondary uses** of data include: medical research; clinical and financial audits; health service planning; resource management; teaching and training; national statistics; public health surveillance and drug safety monitoring.[2,3]

To fulfil its intention to provide the UK with universal effective health care, the NHS requires information and evidence based on the whole population. Medical research at the population level requires access to large, representative samples of accurate patient and population data. Although researchers can often generate new information using

2 Tranberg H & Rashbass J (2004) *Medical records: use and abuse.* Radcliffe Medical Press, Oxford.
3 Lowrance W (2002) *Learning from experience: privacy and the secondary use of data in health research*. The Nuffield Trust, London.

questionnaires and surveys, a great deal of relevant information will already exist in routine medical records and patient databases. Re-use and linkage of this existing information has a great many advantages:

- very large numbers of patients can be studied, with complete coverage of particular populations, producing more reliable results
- greater accuracy; with increasing time patients may have poor recall of their health history or treatment
- the information is derived from day to day clinical practice in a variety of settings
- the duration and costs of the research programme are reduced, facilitating more, rapid and efficient translation of research findings into improved patient care.

1.2 Secondary data research

Secondary data research encompasses a range of activities, which differ in the type and extent of data required and the manner in which the data are used. The main purposes of secondary data research are outlined below.

Identifying the causes of disease

Secondary data research can identify a variety of disease risk factors, whether biological, physical or socio-economic. An example is the analysis of changing patterns of cancer incidence, which is the most effective way of detecting unexpected cancer risks as early as possible. In fact, it has been argued that all causes of cancer have been detected by unexpected increases in cancer incidence.[4]

> *As two thirds of cancers are potentially preventable, patient data have been especially important to help identify some of the factors that put people at risk of cancer in the first place. Evidence from* ***Cancer Research UK***

Detecting cancer risks can only be achieved through the continual monitoring of all known malignant diseases, which is undertaken in the UK through a programme of cancer registries. In addition to identifying risk factors, patterns of disease incidence can be compared between different populations, allowing those community sectors that are at particular risk to be identified and targeted for preventative treatment.

High voltage power lines and childhood leukaemia

A recent UK study investigated whether proximity of home address at birth to high voltage power lines is associated with increased risks of childhood cancer.[5] Cancer registries were used to identify 33,000 children with cancer, aged between 0 and 14 years. Birth information was subsequently obtained on 31,000. For each case, a control was selected from birth registers matched for sex, approximate date of birth and birth registration district. The final data set comprised 29,081 matched case-control pairs (9700 for leukaemia) that could be mapped with respect to power lines. No active participation from data subjects was required. The study showed that, compared with children who lived greater than 600m from a line at birth, those who lived within 200m had a relative risk of leukaemia of 1.69 (95% confidence interval 1.13–2.53). Children born between 200 and 600m had a relative risk of 1.23 (1.02–1.49).

The authors of this study stressed that there is no accepted biological mechanism to explain their findings and emphasised that the results may be due to chance or some other confounding factor. The debate over whether there is a causal link between overhead power lines and childhood leukaemia will continue. What is clear is that, given the small numbers involved (annual incidence of childhood leukaemia in England and Wales is 42 cases per million), further studies will require access to data on a similar, or even larger, scale.

4 Dudeck J (2001) *Informed consent for cancer registration*. The Lancet Oncology **2**, 8–9

5 Draper G, Vincent T, Kroll ME, Swanson J (2005) *Childhood cancer in relation to distance from high-voltage power lines in England and Wales: a case-control study*. British Medical Journal **333**, 1290–4.

The late Sir Richard Doll CH OBE FRS FMedSci made outstanding contributions to our understanding of a great many disease risk factors, including the effects of diet, radon gas and contraception. However, he is most widely known for proving that smoking causes lung cancer.[6] Sir Richard was insistent that access to medical records was an essential part of his research.

> *Much of my research on the effects of ionising radiation and the use of oral contraceptives, leave alone smoking, would have been impossible without the facility of obtaining unbiased access to medical records. Evidence from*
> ***Sir Richard Doll CH OBE FRS FMedSci***

Evaluating and improving preventive and therapeutic practices

Linkage between prescription data and routine health care records is crucial to the investigation of drug usage in different categories of patients and the identification of possible side effects. Such observations cannot be made on the basis of the experience of individual clinicians and require the collation and analysis of large volumes of data.

Prescription tranquillisers and road traffic accidents

In a UK study of over 40,000 people, linkage of prescriptions issued by General Practitioners (GPs) with data on hospital admissions and deaths indicated a highly significant association between the use of minor tranquillisers (e.g. diazepam) and the risk of serious road traffic accidents.[8] Patients were not contacted during this study and records were accessed without consent. This study had considerable implications for the safety of patients prescribed such treatment, as well as for other road users.

Understanding the utilisation of health care services

As society changes and medical care becomes more complex and expensive, it becomes increasingly important to understand how the provision and utilisation of health care

Smoking and lung cancer

In 1947, Sir Richard Doll began a series of investigations into the link between smoking and lung cancer that would continue for over 50 years.[7] Mortality data collected by the Registrar-General showed a phenomenal increase in deaths attributable to lung cancer in the first half of the 20th century. At the time, two main causes for this increase had been put forward: firstly, general atmospheric pollution from car exhaust fumes, from the surface dust of tarred roads and from industrial activities; and secondly, the smoking of tobacco. Sir Richard and his team were the first to undertake a study on a sufficiently large scale to determine whether lung cancer patients differed materially in terms of their smoking habits, or some other way that might be related to the pollution theory.

Their study involved 20 London hospitals in which lung cancer patients were identified by clinicians who then forwarded the records to the research team. The team conducted extensive interviews with the identified patients around their lifestyle and smoking habits. Interviews were also conducted with sex and age matched non-cancer 'control' patients, who were also identified from medical records. In demonstrating the real association between lung cancer and smoking, the findings paved the way for further large-scale prospective studies carried out by Doll and others, including the Survey of British Doctors.

6 Doll R & Hill A B (2004) *The mortality of doctors in relation to their smoking habits: a preliminary report. 1954*. British Medical Journal **328**, 1529–33.

7 Doll R & Bradford Hill A (1950) *Smoking and carcinoma of the lung*. British Medical Journal **2**, 739–48.

8 Skegg D C, Richards S M & Doll R (1979) *Minor tranquillisers and road accidents*. British Medical Journal **7**, 917–9.

services affects the health of communities. Health care records are an essential resource for this research.

Social factors and breast cancer survival
Several studies have shown that affluent women have a higher incidence of breast cancer than socially deprived women. However, research has also shown that socially deprived women have significantly poorer survival from breast cancer. Several studies have attempted to explore the reasons underlying this important disparity.

One study examined whether differences in outcome were related to differences in the management of patients by their hospitals and GPs. The study involved the detailed analysis of hospital and GP records, investigating the type of treatment received, waiting times experienced, length of hospital stays, and number and nature of outpatients' appointments. A series of factors, including home address, were used to determine social status. Patients were not contacted during this study and records were accessed without consent. The study showed that access to health care and quality of treatment were similar for women from affluent and socially deprived areas. Poorer survival of women from deprived areas was instead associated with health problems unrelated to breast cancer (known as co-morbidities), which were significantly higher in this group.[9]

As a prelude to randomised controlled trials
Secondary data research can provide the necessary preliminary evidence to determine whether a randomised controlled trial (RCT) is ethically and clinically justified. Such studies can identify areas where there is doubt about the best clinical treatment and help to establish the appropriate comparisons that should be tested in the RCT. Similarly, research studies using personal data are the only recourse where RCTs would be unethical, e.g. in studies of smoking or radiation exposure.

> *Routine clinical data can be used to get evidence to decide whether a full RCT is justified. Evidence from* ***Professor Mike Pringle CBE FMedSci***

1.3 Further secondary uses of personal data

Re-use of data for a new research purpose
In addition to the secondary use of routine clinical data for research, it is often possible to test new hypotheses using data that have previously been collected for a different study. Such re-use of research data provides many of the benefits described previously, particularly in reducing research time and cost. Furthermore, existing data can be quickly re-analysed in the light of new methods of analysis, often unforeseen at the time of the original study. For instance, researchers may wish to revisit previous research data in the light of new information about genetic influences on disease incidence, as well as the efficacy and safety of medication.

> *It is not sufficiently well appreciated that a preliminary study that can be carried out quickly, using biological or documentary data, can show whether or not there is a case for a larger, definitive study. This can be done ethically with the necessary safeguards for maintaining confidentiality and without the delay caused by full review. Evidence from* ***Professor Tom Meade FRS FMedSci***

Monitoring communicable diseases
Personal data are used to monitor trends and patterns of communicable diseases, for example to address an emergent threat, such as a measles or influenza outbreak, and to detect any novel infections such as severe acute respiratory syndrome (SARS).[10] Such monitoring provides the essential

9 Macleod U, Ross S, Twelves C, George W D, Gillis C & Watt G C M (2000) *Primary and secondary care management of women with early breast cancer from affluent and deprived areas: retrospective review of hospital and general practice records*. British Medical Journal **320**, 1442–5.
10 Turnberg L (2003) *Common sense and common consent in communicable disease surveillance*. Journal of Medical Ethics **29**, 27–9.

information upon which effective public health programmes, such as vaccination or disease screening, are based.

The Health Protection Agency (HPA) is a UK body with a remit to reduce the impact of infectious disease and other health hazards. The HPA may urgently need to find the source of a disease outbreak or to examine trends and possible links to help prevent infection. Access to accurate and complete personal data is a vital part of this work. Certain communicable diseases are notifiable, and doctors who diagnose such cases are required by law to report them to the appropriate health protection officer. However, many serious diseases, such as Legionnaires' disease, influenza and antibiotic resistant infections, are not notified in this way and effective surveillance relies on voluntary reporting by health professionals.

Identifying potential research participants

Patient data sets and medical records provide a valuable resource for identifying potential medical research participants. While the electoral roll might be used to recruit random population samples, GP or other NHS registers have the advantage of being more reliable and up to date. They also allow identification of people in particular age groups. In addition, GP and hospital records can be used to identify patients with particular conditions, in order to invite them to participate in a research study on that condition (see box below).

1.4 Opportunities

UK researchers have used research methods involving personal data to make outstanding contributions to health improvement. The importance of this work will grow over the coming years, when understanding and

Medical Research Council (MRC)/British Heart Foundation (BHF) Heart Protection Study

The MRC/BHF Heart Protection Study (HPS) is the largest trial in the world of cholesterol-lowering therapy for people at increased risk of heart disease. Before initiation of this study, there had been substantial uncertainty about the long-term benefits of cholesterol-lowering drug therapy for particular types of patient, and it was used to only a limited extent.

From a coordinating centre in Oxford, 130,000 suitable patients were identified without prior consent from local hospital and health authority records. Sixty thousand of these patients attended local assessment clinics in response to a written invitation produced by the coordinating centre on behalf of local investigators. After fully informed written consent had been obtained, more than 20,000 patients were randomly allocated to receive simvastatin (to lower blood cholesterol levels) or a placebo.

HPS showed unequivocally that statins cost-effectively reduce the risk of heart attacks and strokes in a very much wider range of high-risk people than had been previously been thought to benefit.[11]

The efficient recruitment strategy undoubtedly increased the number of suitable participants and so significantly enhanced the certainty of the study results. These findings rapidly led to changes in guidelines and practice around the world. In the UK alone, it was estimated that the study results were directly relevant to about 3 million people who were not being given cholesterol-lowering treatment, with about 5,000 lives now being saved annually for every extra million who have taken up the treatment.

11 Heart Protection Study Collaborative Group (2002) *MRC/BHF Heart Protection Study of cholesterol lowering with simvastatin in 20,536 high-risk individuals: a randomised placebo-controlled trial*. Lancet **360**, 7–22.

tackling health priorities will require access to a wide variety of population data. Foreseeable uses include: detection and monitoring of emerging infectious diseases, particularly in tracking changes in distribution caused by greater population mobility; determination of the health and social needs of an ageing population; investigating the complex interplay between genetic and environmental factors in causing disease; and understanding the long-term outcomes of complex treatment interactions.

The Government has identified public health as a priority. Full utilisation of personal data resources will be required to develop an evidence-based national public health programme. This was recognised in the recent Public Health White Paper, which called for the establishment of a Health Information and Intelligence Task Force to develop a comprehensive strategy for gathering and utilising data from various sources, including the NHS National Programme for IT (NPfIT). Recent proposals from HM Treasury give a commitment to finance an increase in the medical research capability within the NHS. Included in these proposals is an assurance that the Department of Health will play its part by: '*ensuring the capability will exist within the NHS National IT System to facilitate, strictly within the bounds of patient confidentiality, the recruitment of patients to clinical trials and the gathering of data to support work on the health of the population and the effectiveness of health interventions.*'[12]

A recent report by the Council for Science & Technology, 'Better use of personal information: opportunities and risks', stated that Government ambitions to deliver more effective public services are: '*dependent on the intelligent use of information about individual people*.' It concluded that public health constituted an '*under-used opportunity for better linkages between, and access to, personal datasets*.'[13]

Many commentators have highlighted the opportunities for research using personal data presented by the unique features of UK health care, in which the population size captured by the NHS is greater than any other health system in the world. The recent Royal Society report 'Personalised medicines: hopes and realities' stated that: '*The newly created NHS Connecting for Health agency is establishing IT systems in the NHS to store a comprehensive record of the patient's history. As part of the programme, the Department of Health should consider carefully the research implications of these data, including pharmacogenetics research.*'[14]

The development of national electronic patient databases (such as NPfIT in England and comparable systems in Wales and Northern Ireland) could provide researchers with access to comprehensive, standardised, accurate and up to date health information, which can be rapidly analysed on a potentially enormous scale. In short, the UK is extremely well positioned to take advantage of the exceptional opportunities to understand, prevent and treat disease at a population level.

1.5 Challenges

The exceptional opportunities for research using personal data face several challenges.

Sensitivity of personal data

Information held in health records can be extremely sensitive. The examples given in the previous section include research on patients prescribed oral contraceptives or tranquillisers; information which the data subjects might reasonably wish to keep private. Data about sexual or mental health, alcohol or substance abuse, violence or termination of pregnancy are also particularly sensitive.

Inappropriate use or disclosure of personal health information, whether accidental or deliberate, has the potential to cause embarrassment or distress. It may have other serious consequences, for instance if health

12 http://www.hm-treasury.gov.uk/newsroom_and_speeches/press/2005/press_100_05.cfm.

13 Council for Science & Technology (November 2005) *Better use of personal information: opportunities and risks* http://www.cst.gov.uk/cst/reports/files/personal-information/report.pdf.

14 Royal Society (2005) *Personalised medicines: hopes and realities.* The Royal Society, London.

data were passed to insurance companies, banks or employers. Patients' trust in health care professionals relies on the assurance of patient confidentiality. Experience or fear of inappropriate disclosure might induce patients to withhold information from a health professional or even avoid medical treatment altogether.[15] Protecting confidentiality has become increasingly complex as records are computerised and shared between large health care teams, sometimes also stored at remote sites.

Privacy and autonomy

The opportunities for greater use of personal data in medical research come at a time of growing public concern about the prevalence of Governmental and corporate surveillance. The introduction of more pervasive Closed Circuit Television (CCTV), aggressive use of data for commercial marketing purposes and the national debate over identity cards have influenced the climate in which issues related to research using personal data are discussed. The potential concerns of patients and the public over research uses must therefore be viewed within this wider context.

Policies that emphasise choice within health care, as within other aspects of modern life, focus on the value of individual autonomy. That is, patients should be afforded the opportunity to make decisions based on their own values. An emphasis on individual autonomy presents challenges for activities such as medical research, which are performed for public, rather than individual, benefit. It could be maintained that a patient has the right to say *'use my data to treat me, but not to improve care for others'.*[16] Or more starkly, *'use evidence from other people's data to treat me, but don't use my data to help them.'* Whereas some commentators believe that individuals have an absolute right to determine how their medical data are used, there will be a range of views among any group (ethicists, politicians, scientists, members of the public) on the relative importance of individual autonomy compared with the need to undertake research for the public good.

Public engagement

Although the importance of privacy and autonomy has been much discussed by those concerned with medical research, it is unclear how closely the conclusions drawn reflect the views of patients and the public. Evidence of public attitudes and opinions on the specific issue of research using personal data is largely lacking. The absence of such knowledge, and the lack of public debate, leads regulatory and advisory bodies to make assumptions about what the public might find acceptable. Researchers and others have argued that these assumptions often give greater weight to the importance of privacy and autonomy than would be expected or desired by the public.

For this reason a priority for the Working Group was to discuss the concerns related to research using personal data with a wide range of patient representatives.[17] Patients' support for research using personal data was much in evidence. In particular, the consultation confirmed indications that patients place greater weight on the public benefits of research, and less on individual rights to privacy, than is sometimes assumed by regulatory bodies. Further issues related to public engagement are discussed in section 5.

Legal and regulatory complexity

Over recent years legal and regulatory changes have had an important impact on how research using personal data is carried out. These changes have often produced improvements in the handling of data and reduced the potential risk of patient harm.

However, as shown in this report, the legislative and regulatory environment has also become increasingly inhibitory to research using personal data. Changing legal and

15 Although this may be the case, the Working Group notes the absence of evidence to support the assertion that the use of medical records for approved research would have such an effect.

16 Detmer D (2000) *Your privacy or your health – will medical privacy legislation stop quality health care?* International Journal for Quality in Healthcare **12**, 1–3.

17 A summary of this consultation meeting can be downloaded from http://www.acmedsci.ac.uk.

ethical standards have led to confusion and uncertainty among researchers, regulators, professional organisations, advisory bodies, hospital Trusts and the public.[18] This is reflected in the plethora of guidance documents from various statutory and professional bodies, of which variations in jurisdiction, interpretation and emphasis have exacerbated the confusion.

Legal uncertainty and an increasing emphasis on autonomy and individual rights have created a conservative culture of research governance in this field, in which the constraints imposed do not always appear proportionate to the potential risk of harm. This conservative approach is compounded by complex regulatory mechanisms that, despite (or perhaps because of) the many sources offering guidance, have become increasingly difficult for researchers to negotiate successfully.

Current difficulties have been recognised in several recent reports and in evidence submitted to the Working Group:

'While the technical capacity to gather information that could be used in public health research has increased immeasurably, the regulatory environment concerning access to personal information… has become increasingly adverse.' The Wellcome Trust Report 'Public Health Sciences: Challenges and Opportunities' (2004).

'The Department of Health and the NHS should consult with the scientific community as to how the data generated by the NHS could be improved, the regulatory framework simplified, and the bureaucracy removed.' House of Lords Science and Technology Committee Report 'Ageing: Scientific Aspects' (2005).

'Poor regulatory frameworks relating to personal identifiable data may constrain population-based health research… It is the Academy's view that the UK should attempt to avoid an overly bureaucratic system where privacy concerns represent a growing barrier to participation in research.' Academy of Medical Sciences report 'Strengthening Clinical Research' (2003).

> *The ambiguities in the current legislation and the inconsistencies between legislation and professional guidance cause anxiety to researchers and will deter some projects, particularly for smaller studies. Evidence from* ***Royal College of Physicians of Edinburgh***

> *It is clear that some projects are having to be abandoned because of delays in obtaining necessary approvals and a frequent comment has been that people will be thinking carefully before embarking on new projects in the future. Evidence from* ***Royal College of General Practitioners Research Group***

> *Lives could be threatened, far less protected, by excessive data protection and bureaucracy so complex as to discourage researchers. Evidence from* ***Royal College of Obstetricians and Gynaecologists***

1.6 Meeting the challenges: a proportional approach

The research community emphasises that there has never been a legal action in the UK for abuse of a patient's right to confidentiality against a researcher conducting research in an approved programme. Conversely, there have been several instances where increased risks of cancer or other diseases were not detected early, leading to unnecessary disease and morbidity. This was shown in 2000 in Japan, where a decision by the Hyogo prefecture to halt cancer registration on the basis of privacy concerns was widely criticised for delaying the detection of a significant cluster of asbestos-related mesothelioma cases.[19]

Constraints on research using personal data carry a real and tangible risk to health.

18 Ward H J T, Cousens S N, Smith-Bathgate B, Leitch M, Everington D, Will R G & Smith P G (2004) *Obstacles to conducting epidemiological research in the UK general population*. British Medical Journal **329**, 277–9.

19 Submission from GlaxoSmithKline.

To quote Sir Richard Doll: '*Confidential sharing of information about patients between doctors and* bona fide *medical researchers has done no harm and has achieved much good. Why destroy it?'*[20]

The law accounts for the balance between the interests of individual privacy and those of the wider public through the concept of *proportionality*, a theme to which we return throughout this report. In law, the right to privacy is given much weight, but is not absolute. Proportional interferences in privacy are permitted, if a number of criteria are fulfilled. Measures to protect patients' interests may have real and substantial costs for research in terms of financial and time resources and may compromise the reliability and generalisability of research results, so delaying or preventing the acquisition of knowledge necessary to understand, prevent and treat disease. For most research projects using personal data, the risk of inadvertent or damaging disclosure of sensitive information is extremely low. Measures taken to protect patients' interests and their right to privacy must therefore be proportionate to the risks involved and the value of the research in question.

Equally the research community must adhere to demonstrably high standards in using personal data and engage with the public to ensure that its aims and methods are supported. These issues require urgent attention if the UK is to take advantage of its opportunity to make major contributions to population-based research.

20 Doll R & Peto R (2001) *Rights involve responsibilities for patients*. British Medical Journal **322,** 730.

2 The Legal and governance framework

Summary

- *The legal framework governing personal data is complex and confusing, particularly around the use of identifiable data without consent. Regulators and other agencies have adopted a cautious approach to legal interpretation (particularly the Data Protection Act 1998 and common law of confidentiality), resulting in unnecessary restrictions on the use of personal data in research.*
- *The regulatory framework consists of multiple bodies, both statutory and advisory, that assess and advise on research programmes involving personal data. There has been little attempt to harmonise legal interpretation, guidelines or procedures among these bodies.*
- *Despite the complexity of the law, we consider that it is mainly the current regulatory framework, rather than the legislation, that is damaging research using personal data.*
- *In addition to calling for more streamlined and effective procedures for research governance in this area, we also make recommendations designed to promote a more stable and consistent interpretive framework, in which regulators, researchers and the public are jointly engaged. We conclude that there is nothing in the law itself that prevents the use of identifiable data for research without consent, provided that such use is necessary and proportionate with respect to privacy and public interest benefits.*

2.1 Introduction

In this section we outline the legal framework within which research using personal data takes place and the functions of the agencies that regulate the field. It is not our intention to provide researchers with advice on how to negotiate the various legislative and regulatory requirements, nor to suggest practical solutions to facilitate their work. Rather, our aims are: to describe the main features of the legislative and regulatory framework in this area; to explain the origins of current difficulties; and to propose realistic solutions that are compatible with the law, including the provisions of the Human Rights Act. Although the law in this area is undoubtedly complex, we endorse the view submitted to us that the difficulties experienced by the research community arise from current interpretations of the law that do not reflect the original intentions of the legislation.

Ethicists and those involved in health care draw an important distinction between confidentiality and privacy. Confidentiality is perceived to protect information imparted within a relationship of trust, ensuring that it does not exit that relationship without authorisation. Privacy is regarded as having a greater connection with an individual's right to control their personal matters and identity; a right connected with autonomy and dignity. Informational privacy, unlike confidentiality, is seen to protect the information from unauthorised *use* as well as *disclosure*, and to protect the information whether or not it has been imparted in a relationship of confidence.

> *There is considerable feeling within the cancer research community that the current legislation is, in fact, not the key barrier to the conduct of medical research. …it has become clear to us that the legislation is widely misinterpreted and misapplied in a manner that inhibits responsible medical research. Evidence from* ***Cancer Research UK***

In the past, this distinction had practical significance because English law recognised a legal right of confidentiality, but not of privacy. Although the conceptual distinction can still be made, its practical significance has been diluted by the enactment of the Human Rights Act 1998, the Data Protection

Act 1998 and recent developments in the common law of confidentiality. As a result of these legal developments, the medical profession is obliged to observe rights of privacy as well as confidentiality.

This shift towards privacy and autonomy is one of the main drivers behind the 'consent or anonymise' policy promoted by various regulatory and professional bodies. Such a policy seeks to restrict the use of identifiable data without consent and insists that data are anonymised where consent is not possible. The difficulties imposed on research using personal data by this policy are explained further in sections 3 and 4. In this section, we explain that 'consent or anonymise' is **not** a legal requirement and that identifiable data may be used for medical research without the consent of data subjects, provided that such use is necessary and proportionate with respect to privacy and public interest benefits.

2.2 The legal framework

The most important laws governing medical research using personal data include:

- Data Protection Act 1998
- Common law of confidentiality
- Human Rights Act 1998
- Section 60 of the Health & Social Care Act 2001

One of the over-riding problems faced by researchers wishing to use personal data, and by the controllers of that data, is the complicated patchwork of statutory and common law that operates in this area. In most instances, compliance with several different schedules and conditions within a single act is required, in addition to cross-compliance with other acts and common law. This increases uncertainty around whether a particular research practice complies with all the interwoven clauses and conditions, both within and between pieces of legislation. Of course, such legal minefields are not exclusive to research using personal data. However, that does not diminish the significant challenges presented by the legal framework in this area.

2.2.1 Data Protection Act 1998

The Data Protection Act 1998 (DPA) is the principal statute relevant to the use of medical information in research.[21] It is a notoriously unwieldy and confusing piece of legislation. Here we outline the aspects of the Act that bear on research using personal data.

The following terms apply within the DPA:

- '*Data'* include information recorded, or intended to be recorded, as part of a relevant filing system or part of an accessible record.
- '*Personal data'* are data relating to an individual who can be '*identified from those data or from a combination of those data and other information which is in the possession of, or is likely to come into the possession of, the data controller'.*
- '*Processing'* includes the acts of obtaining, recording or holding information as well as its retrieval, disclosure and destruction.
- A '*data controller'* is defined as a person '*who determines the purposes for which and the manner in which any personal data are, or are to be, processed'.*[22]
- A '*data processor'* is '*any other person (other than an employee of the data controller) who processes data on behalf of the data controller'*, i.e. someone who is contracted to perform specific operations on the data, as defined by the data controller.

The definition of personal data given in the Act means that its conditions do not apply to data from which an individual cannot be identified (whether from the data itself or in combination with other data in possession of the data

21 The Act repealed an earlier Data Protection Act 1984 and implemented the provisions of the EU Data Protection Directive (agreed in 1995) in the UK.

22 The duty to comply with the DPA rests with the data controller, who must register with the Information Commissioner and ensure that their employees, and other persons under their control, process data in accordance with the Act. The Information Commissioner's Office (ICO) confirms that a data controller can be an organisation, such as a GP or hospital Trust. Since researchers are generally employees of larger organisations, the legal obligations of the Act are therefore directed at their employer, except where the researcher has signed a contract to process data on behalf of another organisation (see section 2.2.4).

controller), i.e. anonymised data. Further discussion on anonymised data is given in section 3.

Data Protection Principle 1 (DPP1)

Of the eight DPPs, the first is highly relevant to medical research and where most interpretive confusion had arisen. It stipulates that the processing of personal data must comply with several criteria, namely that it must be fair and lawful.

Fair processing

The DPA describes several requirements for data to be processed fairly. These include the obligation to provide information stating:

- the identity of the data controller or his nominated representative
- the purposes for which the data are intended to be processed
- '*any further information which is necessary, having regard to the specific circumstances in which*

The DPA: Data Protection Principles and Schedules 1,2 and 3

Schedule 1 of the DPA sets out eight data protection principles (DPPs) that apply when '*processing personal data'*. These have been helpfully summarised by William Lowrance:[23]

Personal data must be:

1. Fairly and lawfully processed
2. Processed for specified, lawful, limited purposes
3. Adequate, relevant, and not excessive in relation to the purposes
4. Kept accurate, and where necessary, kept up to date
5. Not kept longer than necessary
6. Processed in accordance with data subjects' rights
7. Kept secure against unauthorised or unlawful processing
8. Not transferred to countries not ensuring adequate protection

Schedules 2 and 3 of the DPA set out conditions relevant to the first data protection principle by which data may be legitimately processed. Each schedule has several alternative conditions, the most relevant of which are reproduced below.[24] Processing of personal data must comply with at least one condition in Schedule 2, whereas processing of sensitive personal data (which include health records) must comply with at least one condition of Schedule 3.

Schedule 2

Condition 1: The data subject has given his consent to the processing.

Condition 5(d): The processing is necessary for the exercise of any other functions of a public nature exercised in the public interest by any person.

Schedule 3

Condition 1: The data subject has given his explicit consent to the processing of the personal data.

Condition 8: (1) The processing is necessary for medical purposes and is undertaken by-

(a) a health professional,or

(b) a person who in the circumstances owes a duty of confidentiality which is equivalent to that which would arise if that person were a health professional.

(2) In this paragraph 'medical purposes' includes the purposes of preventative medicine, medical diagnosis, medical research, the provision of care and treatment and the management of health care services.

23 Lowrance W (2002) *Learning from experience: privacy and the secondary use of data in health research.* The Nuffield Trust, London.
24 Full schedules can be downloaded in full from http://www.opsi.gov.uk/acts/acts1998/19980029.htm.

the data are, or are to be, processed to enable processing in respect of the data to be fair.'

As a shorthand, such information is hereafter referred to as *fair processing information*.

The provision of fair processing information can often cause difficulties for research using personal data. This is because research is usually a *secondary* use of data, whereby researchers have no involvement in the collection of data and hence no opportunity to provide fair processing information. Similarly, it is nearly always impossible to predict the research purposes for which data might be used in the future.

In meeting the fair processing requirements, the legislation distinguishes two situations:

(a) those where data are obtained from the data subject; and
(b) those where data about the data subject are obtained from some other source.

Both of these situations require that data subjects are provided with, or have access to, fair processing information. In situation (a), the researcher or health service is required to take steps *'so far as practicable'* to provide the data subject with fair processing information. In situation (b) the information need not be provided if it would involve *'disproportionate effort.'* When data are used for the secondary purpose of research, however, both situations may pertain. This is because the researcher often obtains the data from a source other than the data subject, situation (b), but the person from whom the researcher obtains the data (e.g. a hospital or doctor) often obtains the data directly from the data subjects, situation (a).

If the legislation uses two different terms, it should be generally assumed that there is a distinction between them: *'so far as practicable'* implies that the action should be taken if it is reasonably possible; *'not disproportionate effort'* implies that, although such action may be possible, it is not feasible (i.e. it is unduly onerous, difficult or expensive in comparison with the importance of the processing). Clarity about this distinction is important for secondary data research because it often involves very large numbers of individuals who could not be contacted with fair processing information without considerable financial and time costs (i.e. it may be possible in principle but not reasonably possible or even feasible).

Lawful processing

The prevailing view is that for data processing to be lawful, it must not contravene laws external to the DPA, such as the common law of confidentiality, the Human Rights Act or administrative law. An alternative view submitted to the Working Group has argued that the statutory phrase *'lawful processing'* means no more than compliance with a condition in Schedule 2 and 3[25], for instance, the public interest condition of Schedule 2 or the medical purposes condition of Schedule 3.

2.2.2 Processing personal data under the Data Protection Act

As mentioned previously, in addition to complying with DPP1 (i.e. processing data fairly and lawfully), processing of personal data must also satisfy at least one condition in Schedule 2 of the Act, and, in the case of *'sensitive'* personal data, a condition from Schedule 3. Generally speaking, research activities that satisfy any of the conditions of Schedule 3 will usually meet a condition of Schedule 2. Consent is one of the conditions most relevant to medical research. However, there are alternatives to consent, which are described below.

Consent

Both Schedules 2 and 3 recognise consent as a basis for legitimate data processing. However, where Schedule 2 refers to *'consent'*, Schedule 3 refers to *'explicit consent'*. The Act does not define these terms; it was intended that they should be interpreted in the light of relevant judicial decisions. Unfortunately, there is no specific judicial doctrine of consent and the

25 Liddell K (2005) *The Mythical Connection Between Data Protection Law and Confidentiality: Processing Data "Lawfully"*. Bio-science Law Review **6**, 215–22.

courts have yet to address consent in the context of medical research using personal data.

The validity and limits of various types of consent relevant to research using personal data are discussed in section 4. For the purposes of the Act, the implication is that where consent is the *only* justification a researcher has for health data processing, it must be explicit (to satisfy the demands of Schedule 3). However, consent is not the only basis on which data can be used for research using personal data. Indeed, as stated by the Information Commissioner's guidance: *'It is a common misconception that the Act always requires the consent of data subjects to the processing of their data.'*[26]

Alternatives to consent

Schedules 2 and 3 list a range of alternative justifications to the consent requirement (sometimes called 'exceptions'). The most relevant is Condition 8 of Schedule 3, which allows the processing of data if it is necessary for '*medical purposes'* and is carried out by a '*health professional'* or '*a person who owes an equivalent duty of confidentiality'*. Importantly, the term '*medical purposes'* explicitly includes medical research.

For the purposes of the Act, the term '*health professional'* includes doctors and other registered health professionals, such as nurses and therapists, and scientists who are heads of departments in health service bodies. Other researchers owe a duty of confidentiality because of the legal principles that govern the common law of confidentiality. Assuming the duties are equivalent, all researchers are therefore able to conduct research under this exemption.[27]

Although the Act therefore provides for the use of data without consent for medical research carried out by *bone fide* researchers, other conditions remain. Importantly this exception applies only when the processing for medical purposes is '*necessary'*; a term that is not defined in the Act.[28] (It can be argued that medical research that has been approved by a properly constituted ethics committee is, by definition, *necessary* since it would not otherwise be ethically appropriate to conduct it).

This exception relieves the researcher from the requirement to obtain explicit consent (under Schedule 3) but does not exempt the researcher from complying with:

- requirements for fair processing in the first data processing principle; and
- a Schedule 2 condition, for example by showing that the data subject has given *implied* consent or by claiming that processing is justified in the public interest and of a public nature (again, it can be argued that ethically approved medical research is, by definition, in the public interest).

2.2.3 Exemption for historical and statistical research (section 33)

Section 33 of the DPA exempts the data processor from a number of obligations, but only for the purposes of '*statistical or historical research'*. The most relevant exemptions of Section 33 are from the requirements:

- to provide information to the subject on request;
- for timely destruction of the data (DPP5);
- to obtain personal data '*only for one or more specified and lawful purposes'* and not to further process the data '*in any manner incompatible with that purpose or those purposes'* (DPP2).

These exemptions are extremely important for most research activities, given the large number of data subjects involved, the value of retaining data sets to address multiple research questions and the difficulty of predicting future research questions for which a data set may be used.

26 Information Commissioner's Office (2002). *Use and Disclosure of Health Data: Guidance on the Application of the Data Protection Act 1998*.

27 To establish equivalence, researchers should check that their employment contracts include a clause that misuse of confidential information may give rise to disciplinary proceedings. If not, it is prudent to seek an honorary contract with a health service body including such a clause (see section 2.3.5).

28 Given the relationship between the right of privacy and the DPA, the Courts are likely to interpret it in accordance with the doctrine of necessity developed under the Human Rights Act 1998 (UK) and the European Convention of Human Rights.

Importantly, Section 33 also states that data must not be processed to support decision-making with respect to particular individuals (e.g. clinical decisions) and that processing must not be likely to cause substantial damage or distress. The nature of most medical research involving personal data is highly unlikely to contravene either of these conditions.

Despite the apparent usefulness of Section 33 in supporting research using personal data, it is little used. Several respondents to the Academy's call for evidence queried whether Section 33 could resolve some of the difficulties experienced by researchers and provide a legal basis for medical research on non-consented data.

There are several ambiguities relating to Section 33 that may be responsible for its limited use. The most obvious relates to the term '*statistical or historical research'*, which is not defined in any detail. However, consultation with the Information Commissioner's Office (ICO) indicated that most medical research involving health records would be considered as historical or statistical research.

Importantly, with regard to the limits of the exemption regarding DPP2, sub-section 2 of Section 33 states that: '*For the purposes of the second data protection principle [DPP2], the further processing of personal data only for research purposes in compliance with the relevant conditions is not to be regarded as incompatible with the purposes for which they were obtained.'*

Based on this sub-section, the Working Group supports the argument that Section 33 permits data to be processed for research without a requirement to notify the individual of this new purpose (the fair processing requirements of DPP1: section 2.2.2), provided that the data were *originally* collected in compliance with DPP1.[29] With that proviso, a researcher would not need to provide data subjects with further fair processing information when re-using an existing data set for any historical or statistical research purpose.

This view has not yet been tested in the Courts. Some, including the ICO, argue that, despite the provisions of Section 33, researchers are *not* released from further fair processing requirements of DPP1 and are therefore exposed to the associated difficulties discussed in section 2.2.2. The Working Group consider this position to be overly demanding, to render Section 33(2) otiose, and to overlook the endorsement of scientific and statistical research in the European Data Protection Directive.

With regard to fair processing requirements, the ICO does distinguish between research using current or old health records (the latter being of patients who are no longer being treated for their condition). For old records, *'those patients who cannot be contacted without disproportionate effort need not be given the fair processing information although the researcher should record this fact.'*[30] Although this would appear to provide some help to researchers, in practice it is extremely difficult to make a clear distinction between current and old records and this statement narrows the categories of research that can be supported under Section 33 considerably.

Consequently, the Working Group does not support the interpretation of the ICO. Instead, we strongly endorse the view that the Section 33 exemption was clearly designed to allow further data processing for research purposes to be carried out without revisiting fair processing requirements, providing that the processing is unlikely to cause substantial damage or distress and is not used to support decisions taken concerning the individual. We believe this should apply to both current and old records. We consider that the ICO should enable greater use of the Section 33 research exemption by clarifying the definition of research and

29 If researchers were required to provide fair processing information, section 33 subsection 2 would be redundant, which we do not believe could have been the intention of the legislators.

30 Information Commissioner's Office (2002) *Use and Disclosure of Health Data: Guidance on the Application of the Data Protection Act 1998*.

31 Available online at http://www.ctsu.ox.ac.uk/projects/search.shtml.

32 Liddell K (2005) *The Mythical Connection Between Data Protection Law and Confidentiality: Processing Data "Lawfully"*. Bio-science Law Review **6**, 215–22.

reviewing the circumstances in which this exemption applies.

2.2.4 Data controllers and data processors

In some situations there is another route for alleviating difficulties associated with fair processing obligations. For example where a researcher requesting identifying records from a hospital Trust plans to take steps to de-identify the records at an early stage or to use the information to contact patients for their consent. Researchers have successfully used an arrangement in which the data controller (e.g. the GP practice or NHS Trust) engages the researcher as a '*data processor'*, as defined by the DPA (see box below). Under this arrangement, researchers are permitted to undertake all data processing procedures that may be performed by the data controller and its employees. In effect, this approach brings the researcher within the same legal entity as the GP or health service, which absolves the need for a disclosure to a third party. For this arrangement to be lawful, the processing must be: '*carried out under a contract which is made or evidenced in writing and under which the data processor* [the researcher] *is to act only on instructions from the data controller* [e.g. the GP or health service].'

Such arrangements are already used widely as part of the routine functions of the UK health service. For example, the National Tracing Service is run by a commercial organisation on behalf of the NHS to maintain up-to-date contact details for all registered patients. Similarly, NHS agencies and Primary Care Trusts often use commercial mailing companies to process invitations for screening services.

2.2.5 Common law of confidentiality

The DPA supplements, but does not replace, the common law of confidentiality.[32] The NHS Code on Confidentiality summarises the common law in the following terms: '*The key principle of the duty of confidence is that information confided should not be used or disclosed further in an identifiable form, except as originally understood by the confider, or with his or her subsequent permission.'*[33]

The general rule is that those who disclose information in situations where they know, or ought to know, that the information is confidential will be liable under the common law, unless the disclosure is justified. Most information provided to doctors (and other health professionals) by patients is confidential, and will therefore be subject to this duty. It includes all information that has the necessary quality of confidence about it,[34] including information that one would expect to be considered private.[35]

The SEARCH trial

The SEARCH trial[31] run from Oxford University is a successful example of the data controller/data processor arrangement. Medical collaborators who were employees of UK hospital Trusts took responsibility for helping to identify potential trial participants from local Trust computerised records using diagnostic codes of hospital admissions. In this instance the inclusion criterion was a history of heart attack and discharge diagnoses were used to identify potential participants. The data were sent securely to researchers at Oxford University, where the information was processed confidentially on behalf of the Trust. This facilitated the sending of 85,000 invitation letters from the local medical investigators in over 90 UK hospitals resulting in the recruitment of 12,000 people.

33 NHS Code of Practice (2003) *Confidentiality*. Available online at http://www.dh.gov.uk/assetRoot/04/06/92/54/04069254.pdf.

34 Coco v A.N. Clark (Engineers) Ltd. [1969] R.P.C. 41.

35 The concept of private information has been tested in the courts, most notably in the recent case *Campbell v MGN Ltd* (2004). Personal information is considered confidential when the person publishing the information knows or ought to have known that the other person had a reasonable expectation that the information in question would be kept confidential. Private information in the public domain can be considered confidential if the proposed disclosure 'would be highly offensive to a reasonable person of ordinary sensibilities'. *The Source Informatics* case further limited the obligations of confidence. The Court of Appeal held there to be no breach if information is anonymised. Significantly, the *Source Informatics* judgement was premised on the fact (agreed by the parties) that, in this case, the data were completely and irreversibly anonymised. This case does not therefore define all the circumstances in which much secondary research is carried out. It is possible that the courts would consider reversibly anonymised data to be confidential information. Quite likely, the judges would adopt a test similar to the DPA definition of *personal data*, which takes account of the opportunity (currently and in future) for the data controller to re-identify the individual.

Importantly, confidentiality is not an absolute obligation. There are several valid and lawful justifications for the use and disclosure of confidential information including:

- the consent of the confider; or
- statutory duties to disclose; or
- the public interest.

The latter, sometimes called the 'public interest defence', is now judged, at least insofar as it applies to confidential personal data, according to the concepts of necessity and proportionality developed in human rights law.[36] An interference in privacy is justified if it is directed at a legitimate public purpose as set out in the Human Rights Act 1998 and if the benefits are *proportionate* to the interference. Legitimate public interests relevant to medical research include the protection of health.

Although it is clear from this brief account that the use of confidential data in research will not necessarily be considered a breach of confidentiality, the precise limits of the public interest defence have not been articulated. In the face of this uncertainty, regulatory bodies (aside from the Courts) have taken a cautious approach in interpreting the boundaries of a public interest defence. The result is that the situation relating to processing data without consent in the public interest is unclear. Although many argue that medical research *per se*, carried out with the appropriate ethics committee approval, is in the public interest,[37] there is no consensus or legal authority around this view.

The nature of the common law means the use of confidential data in research should be judged with respect to each individual concerned.[38] For practical reasons, the degree of interference is typically evaluated without looking in depth at each individual's circumstances. The possibility of special circumstances must be borne in mind; each individual has a right to challenge the claim that the interference with their right of confidentiality was proportionate. Section 60 of the Health & Social Care Act 2001 (section 2.2.7) provides a mechanism to set aside obligations of confidentiality on the grounds of public interest and so removes the risk of litigation by particular individuals. An application under section 60 to PIAG can therefore provide useful reassurance for a researcher.

Although each research project and its effect on the individuals concerned must be considered on a case-by-case basis, the Academy is confident in stating that most instances of medical research on data without consent, where the appropriate funding and ethics committee approval has been given, would be deemed lawful under judicial scrutiny, provided the requirements of necessity and proportionality are met.

2.2.6 The Human Rights Act 1998

The Human Rights Act (HRA) 1998 protects an individual's '*right to respect for his private and family life, his home and correspondence.*' Article 8(2) qualifies this right, requiring a balance to be struck between individual rights and public interests so that the right to respect for private and family life may be breached if it is in accordance with the law and is necessary in a democratic society in the interests of national security, public safety, the economic well being of the country, the prevention of disorder or crime, for the *protection of health* or morals, or for the protection of the rights and freedoms of others.

European and UK courts have explained that an interference in privacy can be considered '*necessary in a democratic society'* if it pursues one of the legitimate aims set out in Article 8(2), i.e. it meets a '*pressing social need'* and '*is no greater than is proportionate to the legitimate aim pursued'* and is justified by reasons that are relevant and sufficient for this purpose. Relevant factors include the type and amount of personal information and the number of recipients to whom it

36 Campbell v MGN Ltd (2004) 2 A.C. 457 (HL)
37 Peto J, Fletcher O & Gilham C (2004) *Data protection, informed consent, and research*. British Medical Journal **328**, 1029–30.
38 The Human Rights Act 1998 gives each individual a right to respect for their private life.

might be disclosed. To ensure proportionality, non-consented use and disclosure of data must be subject to safeguards to maintain security and backed by legally enforceable remedies.[39]

2.2.7 Section 60 of the Health and Social Care Act 2001

In the late 1990s, the confusion surrounding the DPA, HRA and common law of confidentiality meant that research ethics committees and hospitals adopted a conservative interpretation of the law and were increasingly reluctant to support studies based on the use of identifiable personal data. Matters were brought to a head with the publication of the 2000 General Medical Council Guidance on Confidentiality, which jeopardised the reporting of cancer incidence to cancer registries by doctors (for further discussion see section 2.3.2).

Section 60 of the Health & Social Care Act 2001 was enacted in response to this situation. The Act applies in England and Wales, but not Scotland, and empowers the Secretary of State for Health to support the use of '*patient information'*, where it is impractical to obtain consent and where anonymised data will not suffice, for certain medical purposes in the public interest

In the Act, '*patient information'* means: '*Information (however recorded) which relates to the physical or mental health or condition of an individual, to the diagnosis of his condition or to his care or treatment, and information which is to any extent derived, directly or indirectly, from such information, whether the identity of the individual in question is ascertainable from the information or not.*'[40]

Provisions pertaining to Section 60 are described in the Health Service (Control of Patient Information) Regulations 2002. Of these, paragraph 4 provides relief from the obligations of the common law of confidentiality: '*Anything done by a person that is necessary for the purpose of processing confidential patient information in accordance with these Regulations shall be taken to be lawfully done despite any obligation of confidence owed by that person in respect of it.'*

Significantly, although this clause affords protection against a common law action for breach of confidence, it does not override obligations under the DPA.

Paragraphs 2 and 3 of the Regulations grant specific dispensation for the use of identifiable data without consent to cancer registries and to bodies carrying out communicable disease surveillance respectively. For other types of medical research, confidential patient information may be processed in accordance with the circumstances set out in the Schedule of the Regulations. These circumstances permit identifiable information to be handled without consent in order to make individuals less identifiable, to invite data subjects to participate in research, to conduct medical research according to geographical location (e.g. postcodes), to link data, or to 'clean' data for medical purposes.

It was envisaged that the provisions captured in the Schedule of the Regulations would be used to develop a system of class support, under which proposals falling into standard categories and having research ethics committee approval could proceed. However, researchers are required to submit applications for approval before relying on these provisions. The Patient Information Advisory Group (PIAG) examines section 60 applications and advises the Secretary of State whether permission should be granted (section 2.3.3).

The Government presented Section 60 as a transitional measure, on the assumption that researchers will eventually conduct all research with consent or full anonymisation.

39 Save where it is absolutely essential, it would not be acceptable to publish a journal article which identified individuals or which posed a real risk of identification. A Local Authority v Health Authority [2004] 2 W.L.R. 926.

40 Under the Act patient information is 'confidential' where: (a) the identity of the individual in question is ascertainable (i) from that information, or (ii) from that information and other information which is in the possession of, or is likely to come into the possession of, the person processing that information, and (b) that information was obtained or generated by a person who, in the circumstances, owed an obligation of confidence to that individual.

PIAG therefore requires all section 60 applications to demonstrate an exit strategy. Researchers express considerable concern over the vulnerability of activities supported by this temporary arrangement. This is discussed further in section 2.3.3.

Section 60 provides researchers with a pathway for certifying that research without consent is compatible with the common law, but does not appear to make the process mandatory. In removing the need for judgements regarding the public interest to be made with respect to each individual (see section 2.2.5), section 60 does provide useful reassurance for researchers and data controllers. However, it appears that Section 60 operates in parallel with the common law public interest defence, rather than replaces it. If this is correct, it is a matter of discretion if researchers make an application to PIAG and, if appropriate, the public interest defence of the common law may be relied upon as a justification.

2.2.8 Research using personal data in Scotland

In Scotland and Northern Ireland, where health is a devolved matter, the evolution of this legislative area has differed from England and Wales. In 2000, the Confidentiality and Security Advisory Group for Scotland (CSAGS) was set up '*to provide advice on the confidentiality and security of health related information to the Scottish Executive, the public and health care professionals*.'[41] Following a public consultation, CSAGS published a report in 2002 that stopped short of recommending a legislative solution in Scotland.[42] Instead, the report set minimum standards, including the provision of better information to patients and the universal adoption of a working practice of always questioning the need for any data collected or shared to be patient-identifiable. Although requiring a public interest defence for the use of identifiable information without explicit consent, CSAGS acknowledged that the arguments in favour of permitting implied consent for some uses (e.g. disease registers) were '*persuasive*'.

The Working Group note the different arrangements concerning the governance of research on non-consented personal data in Scotland, i.e. without the legal arrangements and statutory body provided by Section 60 of the Health & Social Care Act. These less formal arrangements may be more practicable in the Scottish situation. The Working Group considers it unlikely that the system could be successfully transposed to the larger research and health care setting in England and Wales.

2.2.9 Considerations relating to deceased persons

Evidence submitted to the Working Group indicates that researchers experience considerable difficulty in obtaining data relating to deceased persons. These problems are pertinent to research using relatively old data sets, where it is likely that a proportion of the data subjects may have died. There are also difficulties relating to follow up studies of recently deceased individuals, where approaches to relatives for consent might cause distress. A lack of consistency between European and other countries in the laws and perceptions relating to the data of deceased persons presents further problems for international studies.

> *It may be extremely unkind and insensitive to approach next of kin who may find it distressing or be elderly and in poor health themselves, and if death occurred some time ago, it may even be impossible to contact them in the first place. Evidence from* ***Professor Tom Meade FRS FMedSci***

Unlike the EC Data Protection Directive 95/46/EC, the UK Data Protection Act 1998 is expressly limited to living persons. Hence it is generally understood that the requirements of the DPA do not apply to

41 Fraser A (2003) *Privacy and the secondary use of data in health research in Scotland*. Journal of Health Services Research & Policy (Suppl. 1), 12–6.
42 The Confidentiality & Security Advisory Group for Scotland (2003) *Protecting Patient Confidentiality. Final Report*.

deceased persons. Access to records of deceased patients is governed by the Access to Health Records Act (1990). The common law duty of confidence can survive death, but other obligations may compete; each case would have to be considered on its own merits.

Guidance from regulatory and professional bodies indicates they consider a duty of confidentiality to apply after death:

- For deceased persons, the General Medical Council quotes an obligation to keep personal information confidential. Where a particular study falls outside the terms described in the original consent form, it proposes respect for any known views and to treat each case individually.[43]
- The British Medical Association believes that *'all patients are entitled to expect a duty of confidentiality from all their carers and that duty extends up to and beyond death*.' Where consent for a study was not included in the original consent form, posthumous disclosure needs the consent of the executor of the estate or a close relative.[44]
- Guidance from the Medical Research Council points out that if data disclosures about a deceased person intrude into the privacy of their relatives, the relatives may be able to take action under Human Rights legislation.[45]

Although there is no statutory basis for records to be treated as confidential beyond death, issues of access to such data are still under consideration by various regulatory bodies, including the Departments of Health and Constitutional Affairs. In the meantime, varying, and sometimes erroneous, interpretations of the relevant legal and regulatory requirements are causing difficulties for medical research.

It is important that these issues are resolved to ensure that researchers have appropriate access to mortality data across countries. Similarly, establishing whether former study participants have died, and from what cause, provides extremely valuable research information. The Working Group supports the view that, unless research relates to particularly sensitive issues (where there may be significant ethical concerns), a default position should be adopted whereby researchers are legitimately permitted to access data relating to deceased persons.

> *A… more worrying occurrence was when we asked recently for cause of death data on about 800 people contained in a colorectal cancer cohort. The Office for National Statistics expressed their view that data protection approval was necessary, even though it was put to them that these data were not 'personal data' because all subjects were dead. Evidence from* ***Dr Steve George***

2.2.10 Data Transfers

The situation relating to processing personal data is equally complex outside the UK. Within the EU, despite the national data protection legislations being derived from the same Directive (EC Data Protection Directive 95/46/EC), each country has transposed the Directive into its own national interpretation. Hence there are marked differences in some aspects of the legislations around Europe.

Going beyond Europe, legislation, if it exists, is even more varied, with some countries using legislation that is considered inadequate by the European Commission. This leads to issues over how data may be transferred lawfully from one country to another.

The Eighth Principle of the EC Data Protection Directive imposes a prohibition on the transfer of personal data to countries outside Europe unless the country of destination provides an adequate level of protection. However, the Directive sets out several exemptions from

43 General Medical Council (2004) *Confidentiality: Protecting and Providing Information*. Available online at: www.gmc-uk.org/guidance/library/confidentiality.asp.

44 British Medical Association (2004) Appendix 2. *In: Medical Ethics Today: The BMA's Handbook of Ethics and Law*, 2nd edition. BMJ Books, London.

45 Medical Research Council (2000) *Personal Information in Medical Research*.

the application of this Principle, one of which is the granting of consent to the transfer by the data subject.

> *An important consideration is the transfer of patient databases from one institution and /or country to another. No one institution is able to generate sufficiently large patient groups even of common diseases to identify the influence of multiple gene/environment interactions. Thus multi-centre merging of data has become the mantra for EU Biomed applications, amongst others. Evidence from* ***Professor John Warner FMedSci***

In cases where no exemption applies, some means of showing adequacy should be established before transfers can take place. Several options are available including model contract clauses, binding corporate rules and, for EU/US transfers, joining the Safe Harbor Agreement.[46] All of these options are somewhat complicated and a detailed discussion is outwith the scope of this report.

It should be stressed that the Eighth Principle applies only to identifiable data. If the data are anonymised appropriately, the above transfer rules will not apply and data can be transferred relatively freely. When the data are coded (pseudonymised) the picture is much less clear (anonymisation and coded data sets are discussed further in section 3). Certain EU countries take a strict view, considering coded data sets to be identifiable and therefore subject to EC data protection laws. Other countries regard coded data as sufficiently secure and consider it to fall outside the legislation. The Working Group supports the principle that the transfer of coded data sets within the EU should be permitted, provided that the key identifying the subjects of the data sets is held in the country of origin.

2.3 The Governance Framework

In addition to the legal standards found in statute and common law, there are numerous bodies that interpret, implement, supplement and monitor compliance with the legal requirements. These bodies include the Information Commissioner's Office (ICO), General Medical Council (GMC), Medical Research Council (MRC), Department of Health and British Medical Association (BMA), all of which have issued guidance in this area. The focus, context and jurisdiction of each governance body differ slightly. For most, particularly the ICO, Department of Health and GMC, issues of medical research form only a small proportion of their work.

> It is clear that many practitioners are confused between the requirements of the DPA and those of the various regulatory and representative bodies within the sector including the GMC, MRC and BMA. To some extent the advice issued by these different bodies may reflect their different roles. At the same time, as private litigation increases throughout society, many health service bodies have adopted a more cautious approach towards the use and disclosure of patient data, fearing that uses and disclosures of data which previously seemed unexceptionable might attract action for breach of confidence.
> **Elizabeth France, former Information Commissioner** [47]

In the following sections, we show how current governance arrangements have led to several areas of difficulty for researchers. Foremost among these are:

- bureaucracy of process in applying and gaining approval for research projects involving personal data
- variations in emphasis and expression between guidance documents creating confusion over researchers' obligations
- differing interpretations between governance bodies
- a more cautious and conservative approach to research governance than the law requires.

46 For more information see http://www.export.gov/safeharbor.
47 Information Commissioner's Office (2002). *Use and Disclosure of Health Data: Guidance on the Application of the Data Protection Act 1998.*

It would appear that the interpretation of the meaning of the words of the Data Protection Act is much more cautionary than the original Act may have intended. Evidence from ***UK Faculty of Public Health***

2.3.1 Information Commissioner's Office (ICO)

Evidence submitted to the Working Group shows that members of the medical research community have been frustrated by the performance of the ICO in this area. Respondents have complained that guidance from the ICO relating to particular projects can appear idiosyncratic and contradictory. Correspondence shared with the Working Group shows that controversial decisions are sometimes overturned, but only following strenuous advocacy. Initial steps by the ICO to address these problems for medical research are welcome and should be progressed as rapidly as possible. The Working Group believes that researchers would particularly benefit from a more standardised ICO response procedure to their enquiries, including a minimum/maximum response time. Similarly, it would be helpful if letters or case studies illustrating particular ICO advice in this area were more freely available, such as through publication on the ICO website.

2.3.2 General Medical Council (GMC)

The GMC has issued guidance describing the principles of good medical practice and standards of competence, care and conduct expected of doctors in all aspects of their professional work. Serious or persistent failures to meet these standards may put a doctor's registration at risk. Guidance on confidentiality forms one of the core documents issued by the GMC.

The GMC's guidance on confidentiality issued in 2000 stated that doctors should not send patient information to cancer registries without the patient's consent: '*The automatic transfer of personal information to a registry …before informing the patient that information will be passed on, is unacceptable save in the most exceptional circumstances. These would be where a court has already decided that there is such an overwhelming public interest in the disclosure of information to a registry that patients' rights to confidentiality are overridden; or where you are willing and able to justify the disclosure, potentially before a court or the GMC, on the same grounds.*'[48]

This GMC guidance prompted an outcry from the medical research community, who saw the pronouncement as a serious threat to routine disclosures to cancer registries and to cancer research.[49,50,51] Cancer registries were not perceived to be the only activity in jeopardy, and other commentators drew attention to the negative impact on medical research in general[52], health services research[53] and public health surveillance.[54]

Despite criticism by researchers and the House of Commons Science and Technology Committee[55] it was not until 2004 that the GMC published further guidance, intended to be a core statement of principles, accompanied by a continually updated booklet of frequently asked questions. This updated guidance advises that personal information should not be disclosed in the public interest unless the doctor is '*satisfied that identifiable data are necessary for the purpose, or that it is not practicable to anonymise the data*.'[56] The Working Group believes that, unless the circumstances are exceptional, individual

48 General Medical Council (2000) *Confidentiality: Protecting and Providing Information*.
49 Ballantyne A (2000) *Silence could be a death sentence*. The Times, October 27.
50 Brewster D H, Coleman M P, Forman D & Roche M (2001). *Cancer Information Under Threat: the case for legislation*. Annals of Oncology **12**, 145–7.
51 Paterson I C (2001) *Consent to cancer registration – an unnecessary burden*. British Medical Journal **322**, 1130.
52 Warlow C (2001) *Ethical barriers to research into diseases of the human brain*. Advances in Clinical Neuroscience & Rehabilitation **1**, 10–3.
53 Cassell J & Young A (2002) *Why we should not seek individual information consent for participation in health services research*. Journal of Medical Ethics **28**, 313–7.
54 Verity C & Nicoll A (2002) *Consent, confidentiality and the threat to public health surveillance*. British Medical Journal **324**, 1210–13.
55 House of Commons Science & Technology Select Committee (March 2002) 'Cancer Research: A Follow Up', http://www.publications.parliament.uk/pa/cm200102/cmselect/cmsctech/444/44402.htm.
56 General Medical Council (2004) *Confidentiality: Protecting and Providing Information.*

doctors should be guided by the decision of the research ethics committee that the medical research requiring disclosure is in the public interest.

Consultation with the GMC revealed that it is seldom contacted by individuals over problems concerning confidentiality in the context of research. This makes it important that the GMC's advice is aligned with that given by other regulatory bodies, particularly PIAG and ICO, and the possible consequences of such guidance discussed with those expert in this form of research.

2.3.3 Patient Information Advisory Group (PIAG)

PIAG is a supervisory authority established to advise the Secretary of State whether particular research projects, recurrent processing for medical research or disease registries should be granted immunity from the common law of confidentiality. It applies only in England and Wales and was established as a temporary body, on the assumption that researchers would eventually conduct all research with consent or full anonymisation (see section 2.2.7).

PIAG has essentially two functions:

- to advise the Secretary of State on the use of powers provided under Section 60 of the Health & Social Care Act 2001; and
- to provide advice to the Secretary of State in any matter relating to patient information where this advice is sought.

PIAG has informed us that, each year since its establishment, it has approved a higher proportion of applications. This may be due to an improvement in the quality of applications, but it may also reflect a change in the PIAG's attitude to the practical problems of research. Evidence submitted to the Working Group indicates that some research groups have found it useful to apply to PIAG. However, others clearly perceive it to be an unhelpful obstacle in the path of research. Although the number of applications to PIAG has increased, the number is still surprisingly low, with only 26 applications submitted in year 2004/2005.[57] It may be the case that many researchers (and the data controllers providing the data) are bypassing this process and justifying their research on the basis of the common law public interest defence. It might also indicate that, for reasons described in this report, the level of activity in this research area is sub-optimal.

The Working Group has identified several areas of difficulty with PIAG's role.

1. Policy setting roles

From the outset the focus and role of PIAG was intended to include not just review of applications but also the development of general principles. Its Chair, Professor Joan Higgins, has acknowledged that: '*the broad aim of PIAG is to change the culture of the NHS and attitudes to patient information, in order to develop greater focus on the patient. In this way, PIAG is not concerned simply with the narrow application of the law, but focuses more broadly on questions of confidentiality and privacy.'*

Although admirable, this approach creates difficulties for research because PIAG has set a policy direction that appears to ratchet up existing legal standards. Rather than assess whether applications involve proportionate interference in privacy, PIAG applies a stricter standard of absolute and proven necessity.

2. Legal status

As discussed in section 2.2.7, Section 60 of the Health & Social Care Act 2001 appears to operate in parallel with the common law of confidentiality, rather than to replace it. It is clear from the submitted evidence that this has led to confusion about whether it is mandatory for researchers to seek Section 60 support or whether they may rely on the public interest defence at common law and the relevant exemptions of the DPA.

57 Information provided by Department of Health.

3. Temporary nature

In the Government's view PIAG, and the Section 60 exemptions it administers, are transitional measures only. Accordingly, Section 60 exemptions are reviewed on an annual basis. It is envisaged that advances in technological capabilities relating to obtaining patient consent and anonymisation will allow the system to be wound up within a few years: '*Organisations need to recognise that Section 60 [S60] is an interim measure and therefore need to develop an exit strategy from needing S60 support, either through the National Programme for Information Technology's (Connecting for Health) Secondary Uses Service or by obtaining consent.*'[58]

Elsewhere in this report we discuss the likely position with respect to the NHS National Programme for IT delivered by Connecting for Health (see section 3.6). For understandable reasons, Connecting for Health will concentrate on gaining public confidence in the use of the electronic clinical care record, with anonymisation as central to that goal. We see no sign that the work of a PIAG-like body will be made unnecessary because of advances in NHS information technology. Furthermore, as discussed in sections 3 and 4, the nature of a great deal of research using personal data is such that a complete solution based on anonymisation or seeking consent is not feasible. The assumptions conveyed in the phrase 'anonymise or consent' are misleading in this respect.

> *The non-permanent nature of PIAG concerns us a great deal. If PIAG is time limited, pending 'anonymisation or consent', then we will need to seek other means to undertake surveillance legally for infectious disease and other serious risks to health. Evidence from* ***Health Protection Agency***

4. Operations

PIAG, in its evidence to the Working Group, makes it clear that it is concerned to assist researchers and to '*lessen the daunting array of administrative challenges when embarking on their research projects*.'[59] However, the wording

CAP Prostate cancer screening trial

An application to Trent MREC for the CR-UK funded CAP study (a screening trial in prostate cancer) was first submitted in August 2003. In October 2003 provisional approval was given to flag the men for the primary outcome (mortality) but the inspection of the men's records without consent (for validation of the primary outcome and collection of the secondary outcome data) was not approved and we were not given permission to apply to PIAG on this issue. An application to PIAG was made for the flagging of the men's records in November 2003. In December 2003 PIAG gave provisional approval for flagging, subject to some clarification. Final PIAG and MREC approval for flagging was received in April 2004 – eight months after the original submission.

An application to Trent MREC for case-note review without patient consent was again made in June 2004 and to PIAG in July 2004. This time provisional MREC approval, subject to PIAG approval, was given in August 2004. At the PIAG meeting in September 2004 approval was not given and further clarification was requested in time for the December 2004 meeting.

Thus it took 8 months to obtain both MREC and PIAG approval for flagging—something that is not controversial and does not involve any patient contact. The collection of data from case-notes for the crucial secondary outcomes for this study was still not possible 16 months after the first MREC application form.

Evidence from **Cancer Research UK**

58 Patient Information Advisory Group Second Annual Report July 2003 – June 2004 Available online at http://www.dh.gov.uk/assetRoot/04/12/11/86/04121186.pdf.

59 Submitted evidence. See http://www.advisorybodies.doh.gov.uk/piag/ams-patientinfoinresearch.pdf.

of its annual reports does not convey this attitude and instead the emphasis is on the regulatory, rather than facilitative, nature of its work. Evidence submitted to the Working Group suggests that researchers have found the PIAG process to be overly bureaucratic and lengthy.

Applications to PIAG are considered by the committee members every three months (12–15 per meeting). Frequently, PIAG will discuss the applications, or their opinion, with the researchers who have contacted them, a process that researchers generally find helpful. PIAG asks for proof that anonymised data cannot be used, or that consent cannot be obtained. Pilot studies to demonstrate this may be requested. The process may take many months to complete and end in rejection.

> *The PIAG process is lengthy. To collate the information requested took 4 months in total. As the committee only meets every 3 months, delays are inevitable if additional information or clarification is requested, as in our case. We also question the necessity of annual reviews. Evidence from* ***The EASTR Study Group, National Blood Service***

Researchers have also been frustrated at the lack of coordination between the application processes of PIAG and other governance bodies. The Working Group warmly welcomes initiatives to develop joint electronic application processes between PIAG and the REC system. We hope that such work will be accelerated and extended to include NHS R&D procedures. Several respondents argued that the requirement for annual review and the identification of an 'exit strategy' is not appropriate for finite projects where consent or anonymisation is not feasible. The Working Group supports calls for this requirement to be relaxed in these instances. Moreover, unlike the MREC system, there are no independent procedures in place to allow PIAG decisions to be appealed.

5. Expertise and balance

Although PIAG has undoubtedly built up a large body of expertise, particularly with regard to technological issues related to data anonymisation and security, our evidence indicates a concern amongst researchers that PIAG lacks a full understanding of some areas of research using personal data. When deliberating applications, the committee does not seek advice from external experts in the research field in question. Several respondents to the call for evidence emphasised the need for a more balanced membership of PIAG, with greater representation of active researchers and the inclusion of lay members from medical research charities. We acknowledge efforts by PIAG to recruit members with a research background and encourage researchers to actively engage with this, and other, research governance processes.

PIAG, like all organisations concerned with research governance, does not have a reliable body of knowledge on UK public opinion surrounding specific issues arising in research using personal data that could serve as a guide to what might be reasonably acceptable to the public. In the absence of such knowledge, regulatory authorities, including PIAG, often appear to adopt a more conservative approach than patients and the public might themselves favour (as was found to be the case during the Working Group's consultation meeting with patient groups).

> *Standards of confidentiality for this work [research using personal data] are often considered inadequate by regulators, even though they are acceptable to patients. Evidence from* ***Genetic Interest Group***

6. Transparency

The judgements of PIAG should be influential with research ethics committees, Trusts, and other relevant approval bodies. Through its examination of individual applications PIAG is developing substantial experience that should inform future policy in the use of personal data

for medical research. This continued evolution of experience is important because new issues will arise in research using personal data, particularly as applied biomedical technologies open new avenues of research that will impact on treatments and outcomes.

Although PIAG publishes a list of approved projects, the reasoning behind judgements relating to approval or rejection is not presently available. Such information would be valuable to the research community, both in terms of informing them about what PIAG considers to be acceptable and unacceptable research practices and opening up the debate around the judgements made. Such information should more explicitly describe practices falling under specific class support, in addition to illustrating novel issues and research practices as they arise.

7. Relations with researchers and the public
Researchers, REC members and NHS officials are often unclear of the criteria that determine when a research proposal should be submitted to PIAG and how it fits into the wider research governance framework (GMC, research ethics committees, NHS R&D). This is acknowledged by PIAG and we understand that the group intends to undertake a programme of engagement with researchers and research funders, in addition to improving its advice to RECs on when researchers should seek PIAG approval.

PIAG is a useful public mechanism for scrutinising the assumptions of researchers and should provide public assurance that research in this area is subject to proper checks and balances. PIAG has so far not undertaken any kind of public engagement but greater publicity of its role and activities would improve public awareness of research using personal data.

2.3.4 Research ethics committees (RECs)

The processes and decisions of local and multi-centre research ethics committees (LRECs and MRECs) were criticised in submissions to the Working Group. Evidence received suggested that RECs can appear to over-interpret the legal framework. RECs are free to conclude that a research proposal is unethical, even where it is lawful, but in our view they should do so cautiously and only on the basis of clear and definite concerns.

> *We are concerned that research using patient data is being impeded by the excessive demand that Ethics Committees are putting upon researchers. We know from experience, how difficult it has now become to undertake these sorts of studies compared to 10 years ago. Indeed it has become so cumbersome that many people are being put off undertaking such studies. We believe we seriously need a re-evaluation of the way in which Ethics Committees and particularly MREC committees are operating. The excessively large information sheets that are required to obtain written consent are seriously hindering research. Evidence from* ***Royal College of Physicians***

Researchers are also clearly frustrated at the variable and idiosyncratic responses from different RECs in approving or rejecting projects (see box overleaf).

The Working Group acknowledges that, during the production of this report, the REC system has been improved. The greater harmonisation of REC operations and the extended jurisdiction of one MREC decision, where it is accepted by all other LRECs, is particularly welcome.

The Working Group also notes the recent report of the Ad Hoc Advisory Group on the operation of NHS research Ethics Committees.[60] Although the group unfortunately did not comment on REC's remit with regard to the law, we welcome many of the report's recommendations, particularly with regard to the lack of consistency between REC decisions and the need to develop guidelines for when ethical review is *not* necessary. However, we are concerned that proposals to create a

60 Department of Health (2005) *Report of the Ad Hoc Advisory Group on the Operation of NHS Research Ethics Committees.*

smaller number of more 'professional' RECs will increase the commitment required of members to the point where scientists with active and substantial research programmes are unable to participate. The exclusion of researchers, particularly those at a more senior level, will leave RECs at a greater risk of becoming distanced from the challenges of conducting research in the current environment.

Decision tool of urinary tract infections
I received a small amount of money when I was a Senior Lecturer in Bristol and supervised a GP registrar who was working as an academic research fellow for 6 months. We planned to collect structured data on symptoms/signs and diagnostic test (dipstix) on adults presenting to their GP with symptoms of urinary tract infection (UTI). We planned to develop a clinical prediction rule for UTI on this basis. All this would mean is that we would secure agreement with participating GPs to fill out a structured data collection form at each consultation for an adult presenting with UTI. The GP registrar was going to access each patient's record to assess his or her outcome at 1 month (re-consultation, treatment received, etc.). No patient contact was planned and no intervention given. We were asking GPs to follow routine care.

We approached three RECs in the Bristol area. One viewed the proposal as audit and gave immediate approval; the second refused to consider the application unless individual patient consent was obtained; the third asked us to make extensive revisions, re-write the data collection form and wanted us to attend the meeting in Gloucester which wasn't scheduled for about 6 weeks after we received their letter (over half way into the research time of the registrar). We opted to do the study in only the first REC area. Evidence from **Professor Tom Fahey**

With respect to research using identifiable personal data, it will be necessary for RECs to be guided by other bodies concerned with research governance in this domain. It is essential that there is consistency in approach between different RECS and between RECs and other research governance bodies.

2.3.5 NHS Research Governance
Research using identifiable data must comply with the Government's recently published 'Research Governance Framework' (RGF). The RGF is a statement of principles of good practice and has recently been updated to take account of recent statutory developments, in particular those relating to clinical trials and biological samples. One of the specific requirements of the RGF is that researchers sign an Honorary NHS Contract (including provisions about breaches of confidentiality) if they are not employed by an NHS organisation and interact with individuals in a way that has a direct bearing of their care.

There is no doubt that developments in NHS research governance should improve the standards of research undertaken within the NHS. However, attempts to drive up standards need to be handled carefully in order not to unnecessarily obstruct important research. Concerns have been raised with the Academy about the impact of NHS research governance processes across a range of medical research activities, including studies involving personal data.

> *While the increased efficiency [of RECs] is appreciated, this benefit is being eroded by the requirement to obtain approval from the Research Governance Office of all involved Trusts. We have direct experience of this new hurdle producing further delays. Evidence from* ***Health Protection Agency***

For research using personal data, applications are subject to consideration by up to three different entities: NHS Trust R&D Offices, Data Protection Officers and Caldicott Guardians.

NHS Trust Human Resources departments must also be involved where Honorary Contracts are required. The main difficulties lie in the following areas:

1. Inconsistency

> *There is an urgent need for clarification and simplification of the way multi-site studies are dealt with by Trust R&D departments. Serious problems encountered by a recent nationwide population-based cohort study of women recruited through the NHS screening programme (the 'Million Women Study') demonstrate why this is particularly important. The lack of a central mechanism for R&D approval means that investigators running multi-centre trials require R&D approval from every PCT in the UK to pursue the next phase of the study. There are around 70 PCTs in England alone. These organisations ask for different amounts of documentation, use different forms and appear not to have clear guidance on what to do. Evidence from* ***Cancer Research UK***

Variations between Trust R&D Offices in both the requirements of the application process and in decisions to approve projects lead to inevitable delays, confusion and frustration, all of which increase the costs and duration of multi-centre research projects. There is little centralised support or guidance for Caldicott Guardians and Trust Data Protection Officers, a situation that can lead to idiosyncratic and variable decisions.

2. NHS contracts for researchers

Many researchers undertaking studies in NHS settings hold NHS employment contracts, but others do not. Honorary Contracts for Researchers (HCRs) were introduced to ensure that non-NHS researchers are contractually bound to take proper account of NHS duties of care and research governance procedures. In turn, the issuing of an HCR ensures that non-NHS researchers are covered, like NHS staff, by NHS indemnity

The HCR is a means of establishing the division of responsibilities between a researcher and the NHS organisation. This is most important in the context of clinical trials and studies involving direct patient contact, which involve a greater degree of risk. However, there has been confusion about when HCRs are necessary and which members of a research team are required to hold them, particularly in the context of research using personal data. Researchers complain of the bureaucracy and time involved in issuing HCR's, thereby delaying research that has been funded by public money.

> *Some R&D offices are happy to award honorary NHS contracts to researchers employed by other institutions on receipt of a CV and a signed declaration of familiarity with the relevant data protection/confidentiality acts. Others, however, are being advised by their HR departments that police checks need to be carried out before awarding any honorary contract, irrespective of whether or not a previous check has been conducted by the employing non-NHS institutions. This is being applied to all studies, even those which do not require access to patients. Evidence from* ***Dr Bridie Fitzpatrick***

In 2003, the Department of Health issued the following guidance: '*To avoid having to obtain multiple contracts, the researcher may obtain an honorary contract from the main NHS organisation she/he works with. Other NHS organisations where the researcher is working should be given a copy of this contract and asked to send a short letter to the researcher indicating that they also accept the terms of the contract. That letter then constitutes a contract with the second organisation on the same terms as those with the main NHS organisation.'* Despite this guidance, submissions made to the Working

Group indicate that many Trusts still insist on researchers obtaining multiple contracts.

> *A frequent plea has been for uniformity of requirements across PCTs and Hospital Trusts with regard to the documentation they require; a consistency in Trust approval processes; and a common approach to the awarding of Honorary Contracts together with a speedier process, in some cases, for the issue of these. Evidence from* ***Royal College of General Practitioners Research Group***

In response to some of these difficulties, Greater Manchester Strategic Health Authority developed a document for use by Trusts to provide guidance and an outline for issuing HCRs in Greater Manchester that might serve as an example for other Trusts. They have also developed an outline document for a Research Passport. This is a document issued to researchers prior to applying for an HCR, which allows the Trust to indemnify qualified researchers and allows other Trusts involved in multi-centre studies to accept that the necessary checks on the researcher have been carried out.[61] The Working Group considers the introduction of research passports to be a significant step towards streamlining the process for awarding HCRs and recommends that the programme is rolled out to other regions as a matter of urgency.

3. Lack of guidance and support for researchers
Researchers suffer from a lack of guidance on how to manage NHS research governance requirements. Given the multiple NHS offices that can be involved in a project application, the lack of a central point of contact means that a considerable amount of time must be spent in identifying and contacting the appropriate personnel. Pursuing the proposal through the various NHS approval stages compounds the time and cost burden.

Although progress in reducing the impact of NHS governance requirements has been made, most notably the development of a joint application form between NHS R&D and COREC, further rationalisation and simplification of the processes is urgently required. In this context, the recent NHS R&D consultation document lays welcome emphasis on improving the NHS research environment through reducing bureaucracy and assisting researchers.[62] We strongly endorse this endeavour and welcome further initiatives undertaken by the UK Clinical Research Collaboration (UKCRC).

> *Clear guidance from a central Government source was not readily available, and each hospital Trust Data Protection Officer's interpretation of the legislation varied. Evidence from* ***The EASTR Study Group, National Blood Service***

2.4 Discussion, conclusions and recommendations

Interpreting the legal framework
The legal framework is a complicated patchwork involving UK legislation, case decisions and European directives, augmented by various guidance documents. There are many areas of imprecision, and the courts have not tested the provisions of the various acts in so far as they apply to medical research. Those responsible for permitting or refusing requests for research have made their judgements within this uncertain legal framework. This is not a straightforward task, and the multiple interpretations of these laws have been a source of great difficulty, delay and disagreement for researchers. The onus has fallen heavily on research groups to argue their case, often repeatedly, with the various gatekeeper bodies, to the detriment of research for public benefit.

The Academy is concerned by recent calls for further restrictions around the use of personal data. The Privireal project[63], funded by the European Framework, recently published

61 Available online at http://www.gmsha.nhs.uk/workforcematters/may05/seven.pdf.
62 Department of Health (2005) *Best research for best health. A new national health research strategy Consultation document.*
63 PRIVIREAL is a European Commission Framework 5 funded project examining the implementation of the Data Protection Directive 95/46/EC in relation to medical research and the role of ethics committees. It was created to gather information regarding the implementation of this Directive across Europe

a series of recommendations indicating that a number of European member states (including the UK) '*have not gone far enough*' in implementing the European Data Protection Directive.[64] The report calls for: a broader definition of personal data to encompass a wider range of anonymised data; a narrower interpretation of the exemptions that might apply when data are used in medical research; and more extensive prior checking when researchers use anonymised data or rely on research exemptions. It questions the legitimacy of an exception for medical research in Schedule 3 of the DPA, argues for stricter fair processing requirements and rejects a broad interpretation of Section 33. It is the Academy's view that, should this approach be adopted, it would seriously hamper medical research and undermine evidence-based medicine to the detriment of public health.

We considered whether a return to a governance model based on confidentiality might solve the problems currently facing research with health data. Adopting this approach, health data might be used within the trusted NHS community for medical purposes including research, but not disclosed to, or used by, others unless anonymised or authorised by the individual. Coupled with strong enforcement mechanisms and efforts to raise public awareness, the proposal has much to recommend it. However, the legal and social environment appears to have moved beyond the point of return. The concept of informational privacy is now entrenched in the legal and governance system. A return to a regulatory approach based purely on confidentiality, irrespective of its history and ethical merits, is not consistent with concepts of privacy established by the Human Rights Act 1998 and binding European legislation.

The DPA is undoubtedly complex and confusing and, like many other pieces of legislation, would greatly benefit from simplification. Several respondents to the call for evidence suggested that the DPA should be replaced with a new statutory instrument, that would bring together and simplify the rules relating to medical data research that are currently spread over the DPA, common law of confidentiality and Section 60 of the Health & Social Care Act 2001. This could alleviate the lack of clarity around key DPA terms such as '*substantial damage or distress*' (sections 10 and 33), '*necessary*' (Schedules 2 and 3), '*functions of a public nature exercised in the public interest*' (Schedule 2), notification '*as soon as practicable*' (Schedule 1 part II 2(1)(b)) and '*data controllers*'. The Academy does not consider a new statutory instrument to be necessary for the interpretation put forward in this report. However, new legislation should be considered if the interpretation we propose is not adopted by the regulatory and governance bodies. The way forward we believe is through interpretation of the prevailing legal framework, which protects both confidentiality and privacy. At present, it is applied in a strict and unyielding way, emphasising at every turn the need to 'anonymise or seek consent' before using health data for research.

Promotion of a 'consent of anonymise' policy imposes substantial costs for research in terms of financial and time resources, as well as scientific opportunity and value. For most research using personal data, the risk of inadvertent or damaging disclosure of sensitive information is extremely low. It is crucial that the societal costs of diminishing the quality of the research, or not doing the research at all, are considered by governance agencies both in their guidance and approval decisions.

The Academy considers that the research exemption under Section 33 of the DPA was clearly designed to allow further data processing for research purposes to be carried out without additional fair processing requirements, under conditions where the

64 Available online at http://www.privireal.org/content/recommendations/

processing is not likely to cause substantial damage or distress and is not used to support decisions taken against the individual. We consider that the ICO should enable greater use of the Section 33 research exemption by clarifying the definition of research and reviewing the circumstances in which this exemption applies. We also consider it essential that researchers are permitted appropriate access to data relating to deceased persons, unless the research relates to particularly sensitive conditions (as judged by an ethics committee).

Identifiable data can be used for medical research without consent, provided that such use is proportionate with respect to privacy and public interest benefits. Research governance bodies, including the Patient Information Advisory Group, Information Commissioner's Office, research ethics committees, NHS research governance offices and General Medical Council should accept this interpretation in their guidance and approval decisions.

Improving regulatory processes

Research in this area is impeded by the confusing mixture of responsibilities assumed by PIAG, RECs and the different bodies involved in NHS Research Governance processes. The present position requires urgent remedy. Initiatives by the UKCRC and acknowledgements in the NHS R&D Strategy are welcome, but much remains to be done.

The Academy considers that, in the short term, harmonised decision-making would be facilitated through improved and more formalised communication channels between regulatory bodies and greater transparency of the reasoning behind decisions on individual projects, as well as general approaches towards different types of research activity. The Academy considers that the development of joint electronic application processes (including PIAG, RECs and NHS R&D) and expansion of the 'Research Passport' scheme for honorary NHS contracts should also be accelerated.

It is essential that governance personnel have a full understanding of the legal and regulatory framework and underlying ethical issues. In particular, the Academy considers that NHS R&D offices should be provided with more centralised support and guidance in this area.

The UK Clinical Research Collaboration should lead an initiative to harmonise the approval processes for research using personal data of the bodies involved in research governance, including those of the Patient Information Advisory Group, Information Commissioner's Office, research ethics committees and NHS research governance bodies.

The Academy supports the view that application to PIAG for research using non-consented identifiable data is not mandatory. Several commentators and respondents to the call for evidence have questioned the need for the continuing existence of PIAG, instead suggesting that approval could simply be given by a relevant ethics committee.[65]

The Academy considers that a body with statutory authority in this area will continue to be of value to the research community and the public. This function could not be performed by a system of research ethics committees. In many instances, the statutory basis of PIAG approval provides data controllers with reassurance that they may legitimately release data to researchers, often a key component in progressing a research project. Regardless of technical developments related to Connecting for Health, there will always be instances where either anonymisation or consent is not feasible. New problems will emerge and the existence of a body with statutory authority will help to reassure the public that

65 Peto J, Fletcher O, Gilham C (2004) *Data protection, informed consent, and research*. British Medical Journal **328**, 1029–30.

issues of privacy and confidentiality are properly and appropriately considered, as well as to guide researchers, RECs and NHS R&D departments as to the actions they should take.

Although the Academy supports the continuing need for a statutory body, this report highlights several areas of concern regarding PIAG's current approach, operations and membership. In its communications, PIAG currently stresses its role in protecting privacy and confidentiality, without equal emphasis on the public benefits derived from well-conducted research. The Academy considers that PIAG should more actively promote its role as a facilitator of research. Relations with the research community have also not been aided by the lack of a mechanism for independent appeal of PIAG's decisions (in contrast to the research ethics committee system). In its operations, PIAG should develop an extended and explicit system of class support, whereby applications meeting specific criteria are fast-tracked through the system without detailed review by the committee. The Academy also considers that the current membership of PIAG should include greater representation of active researchers and the inclusion of lay members from medical research charities.

There is a continuing need for a body such as the Patient Information Advisory Group with statutory authority in this area. However, PIAG should address the difficulties of approach, process and membership identified in this report and develop an extended and explicit system of class support, whereby its involvement in research proposals becomes the exception, rather than the norm.

Developing good practice in research using personal data

The public, NHS Trusts, regulatory bodies and funding agencies must have confidence that research using personal data is always carried out to the highest standards. The Academy considers that the development of Good Practice Guidance would encourage high standards of research, as well as facilitating greater harmonisation and consistency in approval decisions. The guidance should be used as a set of benchmarks and exemplars around which researchers can develop research proposals, and not as a checklist for assessment. The guidance should be developed through wide consultation with regulatory and professional bodies, the medical research community and the public. It should also take account of developments in research methodologies through regular review and involve any newly established bodies with special responsibilities in this area (e.g. Human Tissue Authority, Connecting for Health). Areas to be addressed in the Good Practice Guidance are explored further in sections 3, 4 and 5.

The UK Clinical Research Collaboration should lead an initiative involving the regulatory and professional bodies, the medical research community and the public to develop Good Practice Guidance in research using personal data. Such guidance should encompass issues related to anonymisation, consent, data security and the use of health records to identify research participants.

3 Confidentiality, security of data and anonymisation

Summary

- *It is lawful to carry out medical research on anonymised data without consent, provided certain conditions are met. However, there are degrees of anonymisation and there is no clear, practical definition in law of what constitutes anonymisation for specific purposes.*
- *Population-based research often requires data to be identifiable to some degree. This is necessary to detect and avoid double counting, to facilitate longitudinal research where follow up data on individuals must be added, for accurate linkage between data sets from different sources, to ensure that data sets cover a valid or representative sample of the population and where identifying data contain useful research information. There are also situations where it is technically very difficult to remove identifying information.*
- *Reversible anonymisation (involving key-coded data) can provide a solution in many instances. Retention of the key by the data provider (often a hospital or PCT) imposes a considerable burden on them, rather than on the researcher, to carry out the necessary data matching and checking. We support the case for allowing researchers to have the key and so take responsibility for handling data in compliance with legal requirements. The additional contribution to data security from denying the researchers access to the key is likely to be small. We stress that researchers have a legal obligation to ensure the security of data, which necessitates a high standard of practice and training, whether they hold the key or not.*
- *Connecting for Health is a programme to establish electronic health care records for the UK population. It offers an exceptional opportunity for research and would allow the NHS to be the first health system in the world in which services are intelligently designed on the basis of research evidence. We are concerned that the associated Secondary Uses Service has not yet fully considered research requirements, and that the Care Record Guarantee currently makes commitments to the population that, if strictly interpreted, would prevent a significant number of research projects from using Connecting for Health data.*

3.1 Introduction

The respect and protection of personal information is one of the foremost responsibilities owed to patients by health professionals. However, in a modern health care setting, the normal processes of care require ever more frequent judgements to be made about when, how and with whom sensitive personal data can be shared. Rather than absolute secrecy, confidentiality involves the intentional sharing of sensitive information in a trusted environment. The manner in which data are shared must reflect the obligations and expectations of confidentiality implicit in a professional relationship, namely:

- effective procedures should be in place to prevent the unintentional disclosure of sensitive data
- data should only be used for properly authorised purposes
- those handling sensitive data must understand and respect patients' interests.

Patients must have confidence that it is safe to disclose sensitive personal information to clinical staff, even though they may expect them to share this information with others as appropriate. Recent developments in electronic health records across Europe have led to a re-assessment of the safeguards that are required to preserve confidentiality and trust.[66]

Research, as a secondary use of personal data, adds another level of complexity to decisions about confidentiality. In these situations, there is often little or no direct benefit to the data subject. It is therefore especially important that the interests of the data subject are protected. Steps must be taken to avoid inappropriate or unintentional disclosure of information by ensuring high standards of data security, using

66 Kluge E H (2004) *Informed consent and the security of the electronic health record: some policy considerations*. International Journal of Medical Information **73**, 229–34.

both technical and procedural means. It should be noted that mechanisms of data security and anonymisation primarily address the risk of inappropriate disclosure, and do not necessarily circumvent issues associated with an individual's right to autonomy in relation to the use of their data.

3.2 Why is identifiable information needed for research?

There are several reasons why constructing a research data set might require access to identifiable information.[67]

1.To assess/avoid double counting

Population-based research will normally involve more than one source to ensure that all cases are accurately ascertained. It is essential that cases involving the same individuals represented in different data sources are recognised to avoid double counting (see box below).

Some epidemiological techniques are based on the identification of the same case across multiple data sources (capture-recapture techniques). These techniques allow information to be derived from incomplete data sources in order to estimate the true prevalence of a condition.[68] The effectiveness of these methods largely depends on the ability to identify the same individuals in the different sources.[69]

2. For longitudinal research

Longitudinal research is essential to assess the health consequences of exposure to risks, be they occupational,[71] environmental,[72] health-care related[73], or social.[74] Without longitudinal research based on large, complete data sets the risks of these everyday exposures would not be known with certainty. The general public and the responsible authorities have benefited for years from research that provides reliable evidence of this kind based on the use of routine records of vital statistics, hospital activity and cancer registration.[75]

If a data set has been irreversibly anonymised it is impossible to add new data about those individuals in the future. We cannot then know how early exposure to a risk factor influences future health. Because the numbers of adverse events are often rare, very large

Congenital anomaly registers

Congenital anomaly registers are an important source of information on the possible teratogenic effects of exposures in pregnancy and were set up in response to the thalidomide tragedy. Many anomalies are not apparent at birth but are diagnosed at a later stage and so it is essential that the registers receive notifications from multiple sources, including paediatricians, midwives, general practices, health visitors, child health services and genetic counselling services. In many instances, notification of the same individual will be received from several sources and matching reliable personal information is the only way to identify duplicates and avoid double counting. Reliable personal information is required in order to identify the many individuals who are notified by more than one agency.[70]

67 Lako C J (1986) *Privacy protection and population-based health research*. Social Science & Medicine **23**, 293–5.

68 Smeeton N C, Rona R J, Sharland G, Botting B J, Barnett A & Dundas R (1999) *Estimating the prevalence of malformation of the heart in the first year of life using capture-recapture methods*. American Journal of Epidemiology **150**, 778–85.

69 Laska E M, Meisner M, Wanderling J & Siegel C (2003) *Estimating population size and duplication rates when records cannot be linked*. Statistical Medicine **22**, 3403–17.

70 Richards I D, Bentley H B & Glenny A M (1999) *A local congenital anomalies register: monitoring preventive interventions*. Journal of Public Health Medicine **21**, 37–40.

71 Fox A J, Goldblatt P & Kinlen I J (1981) *A study of mortality of Cornish tin miners*. British Journal of Industrial Medicine **38**, 378–80.

72 Kinlen L & Doll R (1981) *Fluoridation of water supplies and cancer mortality.III: A re-examination of mortality in cities in the USA*. Journal of Epidemiology and Community Health **35**, 239–44.

73 Wald N J, Terzian E, Vickers P A & Weatherall J A (1983) *Congenital talipes and hip malformation in relation to amniocentesis: a case-control study*. Lancet **2**, 246–9.

74 Kinlen L J (1988) *The longitudinal study and the social distribution of cancer*. British Medical Journal **297**, 1070.

75 Wald N J, Law M, Meade T, Miller G, Alberman E & Dickinson J (1994) *Use of personal medical records for research purposes*. British Medical Journal **309**, 1422–4.

numbers of people must be studied and complete data sets are crucially important to ensure validity.

3. For linkage between data sets

Population-based research often addresses questions where information on risk factors is derived from a different source from the information on outcomes. As with the problem of double counting, this can only work effectively if the same individuals can be reliably identified in the different data sets. The information on risk factors may have been collected in the past and so more recent identifiers (such as the NHS number) cannot be used. There are many examples of this sort of research – in vaccine safety,[76] occupational safety,[77] and health service research.

4. For validation

The value of large-scale routine electronic records lies in their size, scope and population coverage. The disadvantage is that the quality of the data can vary. For this reason it is essential to be able to test the validity of a sample of records. This is generally done by taking a random sample and retrieving original paper records for manual review. Without such validation the reliability of the overall results is uncertain. This process requires access to identifiers in order to retrieve the original records.

5. When the identifiers contain useful information

Removing commonly used identifiers degrades the data for some purposes because identifiers themselves may contain information.[78] Post-code, for example, provides a proxy indicator of several factors (such as social deprivation) that characterise local populations and which can be determined from aggregated census data. However, many other data items are sometimes used in this way: date of birth and death can be important for seasonality studies, occupation is a very important proxy for income[79] or for more specific exposures such as wood-dust[80] or radiation. Even surname can be an important source of information for genealogical studies.

> *In relation to our experience in DNA marker disease association case-control studies, we have become interested in using information on surnames, which can give valuable information as to the origins of where people come from, and so contribute to an understanding of population structure. Evidence from* ***Sir Walter Bodmer FRS FMedSci***

6. When records cannot be anonymised

Although removing identifiers from electronic records can be a relatively straightforward process, doing so for paper records is laborious and may be simply impossible on a large scale. Paper records remain the basis of historical research as well as being essential for the validation of case definitions. Even electronic health records increasingly contain free text, photos and scanned copies of letters. These items are rich in useful information and extremely difficult to anonymise because their data structures are highly variable and unpredictable. Attempts are being made to automate the process using sophisticated informatics techniques but these are not yet available (and may never be in such a form that allows all types of research activity).[81]

3.3 Anonymisation of personal data

Many guidance documents now emphasise anonymisation as the preferred approach for research using personal data. It can be an

76 Miller E, Goldacre M, Pugh S, Colville A, Farrington P, Flower A, Nash J, MacFarlane L & Tettmar R (1993) *Risk of aseptic meningitis after measles, mumps and rubella vaccine in UK children*. Lancet **341**, 979–82.

77 Morris J, Densem J W, Wald N J & Doll R (1995) *Occupational exposure to hydrazine and subsequent risk of cancer*. Occupational & Environmental Medicine **52**, 43–5.

78 Ohno-Machado L, Silveira P S & Vinterbo S (2004) *Protecting patient privacy by quantifiable control of disclosures in disseminated databases*. International Journal of Medical Information **73**, 599–606.

79 Marmot M G, Smith G D, Stansfeld S, Patel C, North F, Head J, White I, Brunner E, Feeney A (1991) *Health inequalities among British civil servants: the Whitehall II study*. Lancet **337**, 1387–93.

80 Acheron E D, Hadfield E H & Macbeth R G (1967) *Carcinoma of the nasal cavity and accessory sinuses in woodworkers*. Lancet **1**, 311–2.

81 For instance the Clinical e-Science Framework (CLEF). Available online at http://www.grid.ucl.ac.uk/research/CLEF.html.

important tool in reducing the risk of harm from inadvertent disclosure. However, what is meant by anonymisation of data is not always clear.

Information in clinical data systems can be broadly divided into two categories:

1. Data that serve to identify individuals for administrative purposes (e.g. name, address, NHS number).
2. Data that provide clinical or other information about those individuals (e.g. diagnoses, drugs prescribed, test results).

The following terms then apply:

- Data are generally referred to as '*anonymised'* or '*de-identified'* if the first category of data items (i.e. identifiers such as name, address, NHS number) are removed and replaced with a meaningless identification number.
- The resulting data set is said to be '*key-coded'.*
- Data are '*reversibly anonymised'* if the original data provider or the holder of the coded data set has access to the key that links the code to the original identifiers.
- Data are '*irreversibly anonymised'* if the key is destroyed.

Neither the DPA nor the common law of confidentiality give a categorical definition of data that can be regarded as anonymised. The most that can be said is that, to be considered anonymised, the information should not fall within the definition of personal data in the DPA, nor confidential or private information under the common law (sections 2.2.1 and 2.2.5 respectively). This means that key-coded data are considered to be personal data unless the key has been destroyed or put beyond the reach of the person holding the coded data set.

The release of anonymised or pseudonymised data sets for research (i.e. where the researcher would not have access to the key) may seem an attractive solution to issues of confidentiality, privacy and data security in policy terms, but there are a number of limitations to this approach.

1. 'Anonymised' and 'identifiable' are not distinct categories of data

There is little consensus guidance on the identifiers or type of identifiers that should be removed from a data set to render it anonymised. The US Health Insurance Portability and Accountability Act of 1996 (HIPAA) gives a list of identifiers that should be removed during anonymisation, although it

The NHS Code on Confidentiality uses the following definitions:[82]

Anonymised information:
This is information which does not identify an individual directly, and which cannot reasonably be used to determine identity. Anonymisation requires the removal of name, address, full postcode and any other detail or combination of details that might support identification.

Pseudonymised information:
This is like anonymised information in that in the possession of the holder it cannot reasonably be used by the holder to identify an individual. However, it differs in that the original provider of the information may retain a means of identifying individuals. This will often be achieved by attaching codes or other unique references to information so that the data will only be identifiable to those who have access to the key or index. Pseudonymisation allows information about the same individual to be linked in a way that true anonymisation does not.

82 NHS Code of Practice (2003) *Confidentiality*. Available online at http://www.dh.gov.uk/assetRoot/04/06/92/54/04069254.pdf.

is notable this list was modified following difficulties related to the reduced value and usability of research data sets.

Practical solutions regarding the identifiers that should be removed from a data set will have to be judged according to the research case in question. Nevertheless, the Academy considers the development of improved guidance in this area to be a priority for researchers, research funders and regulatory bodies. Such guidance should explore good practice around so-called 'stronger' and 'weaker' identifiers and the hierarchical removal of identifiers to leave 'more' or 'less' identifiable data.

We consider that anonymity and identifiability should be regarded as the poles of a continuous spectrum and not as two distinct categories. Judgements as to when a data set becomes 'anonymous' for practical purposes must be somewhat arbitrary and should be continuously revisited.

2. A truly anonymous data set is unlikely to be useful for much research

Incomplete understanding of the impact of removing identifiers has perhaps clouded discussion on this topic more than any other issue.[83] It is often assumed that anonymised data sets can somehow be created that are 'safe' to release to researchers and yet still allow scientific work and surveillance to continue. There is a clear message from researchers that this is simply not the case.[84]

> *Excessive concern can remove so many potential identifiers that the data become of no value for research. Evidence from* ***Faculty of Pharmaceutical Medicine***

3. The burden of managing a research data set is considerable

Providing researchers with only pseudonymised data sets (as envisaged by the NHS and others) presents formidable practical difficulties for the data provider. This approach requires that the data provider undertake all the processes required to construct high quality pseudonymised data sets. This includes many of the processes described in the previous section, including linkage to eliminate double-counting, addition of follow up data on a regular basis, amalgamation of data sets from different sources, as well as validation both internally and against external standards such as paper records. Experience shows that these are not straightforward tasks and the quality with which they are undertaken determines the quality of the subsequent research.

4. Anonymisation is not in itself an adequate data security policy

Complexity in this area is inevitable because anonymity is context dependent and not an intrinsic attribute of the data set itself. The level of anonymity of a given data set depends on what other information is available to the person viewing the data. To illustrate: *Removing the name of the author from a poem by a well-known poet may render it anonymous to the average reader but not to the scholar of English literature. In order to disguise the author's identity from the scholar the essence of the poem itself would have to be changed. It would not then be the same poem.*

A key problem for research is that a data set containing enough data to be useful for research often contains sufficient information for a determined person to identify individuals. Unless a great deal of meaningful data such as postcode, diagnosis, etc. is removed from a data set, it will still be susceptible to an unauthorised user making a deliberate attempt at inferential data mining or identification.[85] Anonymisation is not a sufficient strategy for protection against a deliberate attempt to breach confidentiality and other data security measures must be in place. The Academy stresses that a researcher

83 Regidor E (2004) *The use of personal data from medical records and biological materials: ethical perspectives and the basis for legal restrictions in health research*. Social Science & Medicine **59**, 1975–84.

84 Wylie J E & Mineau G P (2003) *Biomedical databases: protecting privacy and promoting research*. Trends in Biotechnology **21**, 113–6.

85 Malin B A (2005) *An evaluation of the current state of genomic data privacy protection technology and a roadmap for the future*. Journal of the American Medical Information Association **12**, 28–34.

who took such deliberate action would be guilty of a significant act of scientific and professional misconduct, against which strong safeguards already exist.

Accidental breach of confidentiality is reduced if individual data subjects cannot be readily identified from the data provided by health care agencies to academic research groups. However, the Academy believes that the additional level of security gained is very small compared with the use of coded identifiable data sets by academic groups operating under a rigorous data security policy. In practice, the Academy is not aware of any experience of an inadvertent breach of security during a study that would have been prevented by withholding the key from research groups.

5. It is not clear who is permitted to create anonymised data sets

The production of an anonymised data set requires access to original identifiers (in order to remove them). The wording of the DPA suggests that the act of anonymisation is itself a form of data processing and thus requires consent from the data subject, unless it is done by the person or body who originally collected the data (e.g. the treating GP or hospital) or someone acting on their behalf.[86]

As described above, the construction and maintenance of high quality pseudonymised research data sets require complex data management tasks that have previously been undertaken by specialist academic groups. Shifting the onus of this work on to hospitals and GPs would be inconceivable.

Issues surrounding the creation of anonymised and coded data sets are clearly complex. Many research organisations have gone to considerable lengths to ensure that appropriate high quality pseudonymised (but not necessarily conforming to the ideal of 'anonymous') data sets are available without infringing regulations, NHS policy or the law. Examples include the Tayside Project, the General Practice Research Database (GPRD) and the Oxford Record Linkage Study.

MEMO: Example of the creation of anonymised datasets

Since 1988, the MEMO (Medicines Monitoring Unit) project has been using anonymised data on 400,000 people in the Tayside region of Scotland for the study of the safety, efficacy and cost of prescribed medication. Researchers, based in the University of Dundee, use data from general practice, hospitals, pharmacies, laboratories, and other community health sources. Over 150 research studies have been undertaken using this resource.

To comply with recent regulation the organisation of MEMO has changed.
The group has split into four units:

1. A technical group employed by the NHS which improves the flow of NHS data but does not anonymise.
2. An interface unit employed by the NHS which creates electronic records abstracted from NHS paper records.
3. A university-employed informatics centre that undertakes linkage and anonymises the data sets by replacing names and addresses and Community Health Index (CHI) numbers with another number (retaining the original data in look-up files).
4. Researchers who undertake analyses on anonymised data sets.

All the above processes comply with a set of explicit Standard Operating Procedures that are approved by an external Confidentiality Advisory Board, NHS Caldicott Guardians, and audited annually. All research is conducted on the basis that patients living in Tayside are informed by the NHS that their data may be used in an anonymous manner for research with a right of opt-out.

86 The *Source Informatics* case went some way in clarifying that prior consent is not needed for anonymisation, which is considered to safeguard individuals' privacy, rather than invade it. However, uncertainty still surrounds the question of whether a researcher can anonymise data collected from a source other than the data subject.

It is of course essential that the organisation undertaking the file building has access to adequate identifying information. It is also important that the file-building organisation is competent to link data sets from other sources such as occupational data sets, historical records or bespoke registers of various kinds. The ingenuity with which groups have approached this problem is laudable. However, the Working Group considers many of the methods adopted to be examples of complexity driven by regulation, rather than by the need to protect patients' interests.

3.4 Data security

A high level of data security is required whether the key is held by the researcher or by a third party, to minimise the risk of inadvertent disclosure. Access to the key should be limited to those who require it for legitimate purposes and to people who have a clear obligation of confidentiality to the data subjects. For this reason the Working Group considers the question of who holds the key to coded data sets, and how access to it is managed, to be the central issue.

Whatever the route to obtaining personal data, researchers have a legal obligation to ensure data security. Public trust in research using identifiable data can only be gained and maintained if it is demonstrably clear that high levels of security are in place. A detailed description of the technical considerations associated with data security is beyond the scope of this report, but the general approach is summarised in the box below.

Evidence submitted to the Working Group suggests that current standards of data security are variable. Evidence from PIAG indicates that some research applications they receive involve data security systems that are not of an adequate or acceptable technical standard. This is of concern and is an area where researchers and research funders should make improvements.

3.5 The way forward

Researchers can approach regulatory difficulties associated with anonymisation in several ways.

Overview of data security principles[87]

Information security includes the following:

1. Confidentiality: ensuring that information is accessible only to authorised people
2. Integrity: safeguarding the accuracy and completeness of the data
3. Availability: ensuring that authorised users have access to information systems when required

Achieving security is a management and a technical challenge. A comprehensive security policy must cover:
Physical security – secure rooms, buildings and connections
Logical security – encryption and use of key codes
Technical security - password processes and access rules
Procedural security – ensuring that staff are well trained and that procedures are audited

Maintaining security requires:
Identification of likely threats to security (e.g. natural disaster, power failure, hardware failure or theft, or deliberate attempts to gain access to data by hackers); *evaluation* of the their seriousness and likelihood; and *application* of appropriate control measures.

87 Cavalli E, Mattasoglio A, Pinciroli F & Spaggiari P (2004) *Information security concepts and practices: the case of a provincial multi-speciality hospital*. International Journal of Medical Information **73**, 297–303.

The first is to show that the proposed research is within the provisions of the law. We share the view put to us that the law permits researchers to anonymise identifiable records without prior consent where access to identifiable data can be shown to comply with Condition 8 of Schedule 3 of the DPA (i.e. for medical purposes, which include medical research)[88] and the common law public interest exception. That is, researchers should be permitted to access identifiable data for the purposes of anonymisation when:

- the research is in the public interest;
- alternative routes to anonymised information are not practical; and
- sufficient data security safeguards apply.

In other words, disclosure of data prior to anonymisation is necessary and involves a proportionate interference in privacy. This approach is based on the following premises:

- the research is unlikely to cause substantial distress;
- the data will not be processed to support measures or decisions with respect to particular individuals.

Another route for researchers to anonymise data collected by other organisations is through application to PIAG for section 60 approval (section 2.3.3). To gain approval, researchers need to show that the public interest favours the release of data, and that the interference in privacy is proportionate. It should be noted that, since DPA obligations continue even after section 60 approval, the researcher would also need to show that the disclosure was lawful under this Act.

Researchers have also successfully used another arrangement in which the relevant data controller (e.g. the GP practice or NHS Trust) engages the researcher as a *data processor*, as defined by the DPA (section 2.2.4). Under this arrangement, researchers are permitted to undertake all data processing procedures that may be carried out by the data controller and its employees. In effect, this approach brings the researcher within the same legal entity as the GP or health service, which absolves the need for disclosure to a third party.

A different solution submitted to the Working Group involved the creation of a centralised NHS 'anonymising service', which would perform all the data management tasks associated with creating anonymised data sets for researchers. The development of NPfIT provides an opportunity to create such a service and is explored in more detail below.

3.6 Connecting for Health; the National Programme for Information Technology (NPfIT)

NPfIT is a major initiative to create the best possible IT infrastructure and systems to support the work of the NHS. Programmes in Wales and Northern Ireland are being developed alongside that in England. This includes the creation of a new high-speed network, systems for booking outpatient appointments, archiving images, and transmission of prescriptions to pharmacies, as well as a central core electronic health record that would be available to health care professionals and patients wherever, and whenever, it was needed. This latter component, the national Care Record Service (CRS), is perhaps the most ambitious of the whole programme.

Electronic health records are already the norm in general practice and in large parts of the hospital service. The challenge is to link these disparate systems together to allow an efficient transfer of data when required, while maintaining confidentiality and security. The current paper record systems are known to be inefficient and insecure, but electronic records are perceived to have greater potential for exploitation, error, fraudulent or criminal use. To ensure success, public and professional confidence has to be won and maintained.

88 And a condition of Schedule 2, for example the research is in the public interest and of a public nature.

The Secondary Uses Service

The Secondary Uses Service (SUS) is an important part of the work associated with developing the CRS. It will provide data and information for purposes other than direct clinical care, including: research; planning services; commissioning health care; public health; clinical audit; benchmarking; performance improvement; and clinical governance.[89]

Initially SUS will only manage the data currently flowing through the NHS-Wide Clearing Service (NWCS). Over time, SUS will include other data sources, including cancer waiting times, clinical audit information, central returns, and non-NHS data such as vital statistics. When the CRS is fully developed and integrated into SUS, the service will be able to supply linked person-based data from the whole clinical pathway from GP to hospital.

The plan is for information within SUS to be available in pseudonymised form to users, although some NHS staff will have access to identifiable data when necessary through a strict role-based access control system. SUS will also provide results of standard analyses, bespoke analyses for users, and extraction of anonymised data sets for users. Online access to analytical tools and services will be available for research.

One aim of SUS is to reduce the burden of data collection, abstraction and submission on local NHS services by centralising and automating these processes. At the same time, access to standard national analyses of patient-based activity will improve. However, researchers are concerned whether SUS will be able to provide the flexible access needed to allow existing methods of analysis to be applied to the new data sets. The plan relies on SUS undertaking much of the data management work currently done by research groups, such as linkage, validation and additional data collection from patients. Although this is theoretically possible, it is unlikely to be high priority among the other calls on the resources of NPfIT. Strong concern was expressed by respondents to the call for evidence that SUS has not yet fully considered the data requirements for good quality research and surveillance.

As discussed in section 1, the development of the NHS IT programme offers unparalleled opportunities for research using personal data that could have a real and significant impact on future health in the UK. The recent commitment from HM Treasury to develop the capability of the IT system to support work on population health increases this potential. The Working Group shares the growing concern in the medical research community that such opportunities will not be realised, to the detriment of public health.

> *There is concern that data consistency issues and public health outputs are not being given sufficient attention during the design stages of the clinical electronic patient record systems that will generate the data for those outputs, including the Secondary Uses Service. Evidence from* ***Health Protection Agency***

The Care Record Guarantee

Within Connecting for Health, the CRS works closely with the Care Record Development Board (CRDB). The Board aims to bring together *'patients and service users, the public, and social and health care professionals'* to help ensure a successful outcome for the CRS. The CRDB has recently drawn up and published the NHS Care Records Guarantee, which it will review and revise every six months.[90] The Guarantee sets out for the public the rules that will govern information held in the NHS Care Records Service when it is implemented in 2006. It will shortly be the basis of a public information campaign about the confidentiality of the CRS.

The Academy is concerned that the current version of the Care Record Guarantee seems

89 Available online at http://www.isb.nhs.uk/pages/information/docs/2005/isb05_137.pdf.
90 Available online at http://www.e-health-insider.com/tc_domainsBin/Document_Library0282/nhscr_guaranteev1.pdf.

to be based on the assumption that all work with identifiable data will be accomplished within SUS and that research and public health users will be supplied with anonymised output from SUS. It includes statements that seem to preclude any use of Connecting for Health data outside the NHS for research purposes. The Academy welcomes the mention in the document that data might be used to *'help with research'*. However, we are concerned about the explicit pledge that the new IT system will *'allow only those involved in your care to have access to records about you from which you can be identified.'* A public statement of this kind invalidates the legal basis on which public health professionals and clinical researchers currently access identifiable data for research. In the face of such a statement, any release of data would seem to be a breach of the DPA and could not be supported, even with PIAG approval.

3.7 Discussion, conclusions and recommendations

Data security and anonymisation

There are already strong existing legal and regulatory safeguards on the use of personal data for research. They specifically recognise that medical research in the public interest is a legitimate purpose and that access to identifiable data may be required, sometimes without consent.

Most types of research using personal data require access to identifiable data at some point for some purpose. If researchers are not allowed access to the key to coded data, those that do hold the key (i.e. GP practices and hospital Trusts) will need to undertake many tasks on behalf of the research teams. This includes many of the processes described in the previous section, including linkage to eliminate double-counting, addition of follow up data on a regular basis, amalgamation of data sets from different sources, as well as validation both internally and against external standards such as paper records. Experience shows that these are not straightforward tasks and the quality with which they are undertaken determines the quality of the subsequent research.

The additional level of security gained from pseudonymisation (where researchers do not have access to the key codifying the data set) is extremely small compared with the use of coded identifiable data sets by academic research groups operating under a strict security policy. The Academy considers that deliberate destruction of the key should almost never be necessary. We support the view submitted to us that destruction of the key removes forever the research potential of a data set and could therefore be considered an act of vandalism.

The Academy considers that allowing researchers with adequate data security policies to access identifiable data does not present a significantly increased risk to data security. In contrast, denying researchers access to identifiable data will greatly restrict opportunities to use personal data for research. However, we emphasise the responsibilities of researchers in implementing data security policies. Evidence suggests that data security practices across research groups are variable. The Academy believes this is an area where standards should be particularly improved. We consider that all research organisations using personal data should take steps to review the adequacy of their data security policies. Similarly, research funders should require an assurance of adequate data security policies as part of the decision to grant funding.

The Academy considers that the Good Practice Guidance should address:

- **methods of data security (physical, technical and procedural security)**
- **who can carry out anonymisation and under what circumstances**

- **'strong' and 'weak' identifiers and the hierarchical removal of identifiers to leave 'more' or 'less' identifiable data**
- **the holder of the encryption key and management of access**.

Harnessing the opportunities of the NHS National IT programme

Connecting for Health offers an exceptional opportunity to allow research to inform all aspects of health care. Although Connecting for Health is operating in England, each part of the UK is developing its own IT system. It is essential that these systems are compatible and integrated if health data are to be available for UK-wide research. The Academy questions the untested assumption that the pseudonymisation process that will support SUS will be able to provide for the needs of medical research. We consider it highly unlikely that SUS will be able to provide a service that removes the need for access to identifiable data. We strongly urge the development of effective methods of research support within Connecting for Health and the promotion of the benefits of research during the associated public engagement campaign. To this end, the Academy strongly supports the establishment of a Research Advisory Committee to consider the practical ways in which research needs can be met within the IT programme.

The Academy also considers that the current wording of the Care Record Guarantee presents severe problems for medical research. The Guarantee should make a clear distinction between *bone fide* researchers acting in the public interest and 'third parties' who should not access records (such as employers, insurers, the press and other members of the public).

In delivering the National Programme for IT, Connecting for Health should take urgent steps to address the needs of research through the establishment of a Research Advisory Committee. The Care Record Guarantee should be revised to include support for research as an important and legitimate secondary use of Connecting for Health data, while emphasising the appropriate safeguards.

4 Consent

Summary

- *In this section, we consider the nature of consent and the need for consent in research practices using data derived from health records. The laws applicable to research using personal data allow the use of identifiable data without consent, provided that such use is necessary and proportionate with respect to benefits, likely harms and practicability. Researchers experience variable interpretation of the requirement for consent by regulatory bodies (such as PIAG and RECs) when approval is sought.*
- *The difficulties in obtaining consent are greatest in large studies, especially when several different data sets are involved. An insistence on consent can lead to bias in population coverage, consequently diminishing the value of the data obtained. It may also reduce the size of studies that can be conducted with the available resources, which can significantly reduce the reliability of the findings.*
- *Particular care is required in the initial use of GP records as a basis for recruitment into studies.*

4.1 Introduction

In Section 2.2.2 we outlined the legal requirements concerning consent in the use of patient information. Consent is extremely important as a method for providing assurance that patients and others are neither deceived nor coerced.[91] There is no 'right to consent' or 'right to withhold consent' in the same way that there are moral and legal rights to autonomy, confidentiality and privacy. Consent is a signal from an individual that he or she is willing to accept an action that would otherwise interfere with their moral or legal rights.[92] In other words consent is a justification or authorisation for action, rather than an absolute right or requirement.

This understanding of consent means that not all actions require consent; only those actions where justification is needed. In the absence of a legal or moral right being at stake,[93] consent is unnecessary. This is why the law does not require consent when anonymised data are used in research: there is negligible interference in the legal right of privacy or confidentiality when data are anonymised. Considered in this way, consent is one of several possible justifications for interference in an individual's right to privacy. Another justification is necessity in the public interest. When consent is understood as a signal for authorisation, debate about the validity of explicit, implicit, specific and general consent becomes clearer.

4.2 Explicit, implicit and specific consent

Two variables will determine whether consent is likely to be considered valid. The first is the specificity of the information provided to the subject and the second is the explicitness of the subject's response.

Explicit consent, as the name implies, involves data subjects clearly and unambiguously expressing their consent, either orally or in writing. By contrast, implied consent is assumed when subjects take some action (such as participating in or complying with an activity) in the knowledge that doing so will indicate their consent to another occurrence. It is useful to consider the following example: '*Rugby players do not commit assault every time they tackle an opposition player, because, in electing to play, the opponent is taken to have implied consent to the normal rough and tumble of the game*.'[94]

In general the DPA requires explicit consent for the use of records containing information on mental or physical health, because such

91 O'Neill O (2003) *Some limits of informed consent.* Journal of Medical Ethics **29**, 4–7.
92 Beyleveld D & Brownsword R. *Consent in the Law* (In preparation).
93 Some people prefer to use the terminology moral 'interests' because rights-theory is contentious.
94 Tranberg H & Rashbass J (2004) *Medical records: use and abuse*. Radcliffe Medical Press, Oxford.

information is regarded as sensitive data. However, this does not mean that explicit consent is *always* required, nor that implied consent is irrelevant. Rather, in the absence of explicit consent, researchers must ensure that the research is authorised on some other ground (section 2.2.2).

The specificity of the information supplied to the subject raises a complex set of issues. Subjects must be given sufficient information about the activities to which they are being asked to consent and the consequences of their decision. However, it is not always practicable to provide subjects with highly specific information, for instance in open-ended studies (as with much observational epidemiology). In addition, such information, especially if voluminous and complex, may not be desired by subjects.

In judging the validity of consent, it is likely that the Courts would consider the specificity of the information provided and the explicitness of the subject's response; as the specificity of the information provided and the explicitness of the subject's response decreased, so the validity of the consent would progressively be called into question. This relationship will depend on the particular circumstances of the research, such as the sensitivity of the data, the goals of the research and the reasons why specific information could be not supplied nor explicit consent be obtained.

For most forms of medical research explicit consent is given by the patient to an authorised member of the research team and is usually defined for the particular research programme (e.g. a therapeutic trial). In contrast, many other medical activities already rely heavily on implied consent, for instance patients give implied consent for the sharing of identifiable health data between members of a health care team involved in their treatment.

However, the degree to which implied consent can be assumed when identifiable data are used for research is more contentious. Data from clinical care that are routinely collected in cancer registries, for example, are a powerful source of research information, but the assumption of implicit consent to such use has rested on uncertain ground. One approach has been to ask patients for explicit and general (i.e. not related to a specific research programme) consent to the use of their data for research purposes. This method has been used in some academic treatment centres but is not part of standard NHS practice.

In the following sections we discuss some of the problems of seeking consent for the secondary use of identifiable data. We then address issues related to access to identifiable records in order to approach potential study participants.

4.3 Problems of consent in research using personal data

Seeking and obtaining consent from data subjects can incur significant costs for research. These costs are best explained through examples taken from the evidence submitted to the Academy. The examples given opposite illustrate 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. Evidence suggests that demands for consent have in some instances prevented studies being undertaken and in other cases, have reduced the size of the project. These costs to research reduce the evidence available to improve public health.

1. Seeking consent may be impracticable
Unlike other types of research, studies using personal data can take place at a distance, both temporally and geographically, from the

patients themselves. This frequently has an impact on the practicability of seeking consent. Although either generalised or specific consent can be sought at the time of the patient's interaction with the health service, at that stage the health care provider usually has little idea of the nature of research likely to be undertaken in the future. The advantages of using personal data in research, namely the sample size and timescale that can be encompassed, mean that seeking consent from each individual can often incur insupportable time and expense. Costs are further increased if a proportion of the individuals have moved or died since the data were collected.

The use of data years after collection: the Barker Hypothesis

In the 1980s, Professor David Barker FRS FMedSci of Southampton University developed a hypothesis that adverse conditions during pregnancy and infancy may increase the risk of cardiovascular disease in later adult life.[95] Testing this hypothesis required linking information on birth weight and living conditions during infancy for people born at least 60 years ago with their current cardiovascular health. After searching for several years, Professor Barker's team finally identified a large and detailed collection of birth records in Hertfordshire dating back to 1911.

Fifteen thousand records from this collection were analysed and linked with data from other sources, including death records. Patients were not contacted for consent to use their records for this research. Indeed, 3000 data-subjects had died, making consent impossible. Results from the analysis have linked low birth weight with adult high blood pressure, increased risk of type II diabetes, reduced bone density and different hormonal profiles. The identification of foetal development as a potential risk factor for several conditions in later life has allowed preventative measures for these common diseases to be investigated.

2. Seeking consent may compromise effective population coverage

Population coverage is most important for disease registries, for which inclusivity and universality are key requirements.[96] Evidence shows that a requirement for informed consent can lead to a significant diminution in the quality of registry data. For example, a belief that consent was needed to include personal data in a UK diabetes register contributed to the register receiving information on only 60% of eligible patients. It appears that this was mainly due to doctors electing not to seek consent.[97]

Cancer registries in Germany

In the 1980s, obtaining informed consent was made a statutory requirement for inclusion of data in cancer registries in two German regions. In the years following, it was reported that cancer registries in these regions were unable to collect more than 70% of cancer cases. The Hamburg registry, which had collected cancer data for over 50 years, broke down and was no longer able to add its results to international cancer indexes. These disastrous results led to new guidance from the Federal Government in 1994, which relaxed this requirement in all regions.[98]

The practical process of obtaining informed consent from patients is very similar whether consent is being sought for participation in a study or for the use of patient records for research not involving active participation. Non-response rates of 30% are frequent. Higher non-response rates are common in 'hard to reach' populations, such as ethnic groups and areas of social disadvantage which, as a result, are poorly represented in the research literature.

95 Barker D (2003) *The midwife, the coincidence and the hypothesis*. British Medical Journal **327**, 1428–9.
96 Ingelfinger J R & Drazen J M (2004) *Registry research and medical privacy*. New England Journal of Medicine **350**, 1542–3.
97 Tranberg H & Rashbass J (2004) *Medical records: use and abuse*. Radcliffe Medical Press, Oxford.
98 Ingelfinger J R & Drazen J M (2004) *Registry research and medical privacy*. New England Journal of Medicine **350**, 1542–3.

Researchers point out that non-response rates of 30% do not indicate that a third of patients do not wish to participate. Other explanations include the practical problems of contacting all patients, as well as the resources available to contact and remind those groups who are hard to reach.[99] In some clinical studies, there may also be questions around the feasibility or appropriateness of contacting patients whose medical condition may have progressed to a more serious stage.

3. Seeking consent may cause distress or harm

In some cases patients may be inconvenienced or upset at being contacted for their consent to use their data for a research project, even if they do not subsequently object to the research going ahead. The likelihood of this is increased if the research in question relates to a particularly distressing condition or incident.

Occasionally patients may not wish to dwell upon a disease diagnosis, or may even deny it. These are understandable ways of coping with bad news such as a diagnosis of cancer. To help the patient cope, doctors may not wish to re-activate discussion. Gavin *et al*. describe a retrospective review of 2222 case notes, which showed that non-discussion of diagnosis was an active part of patient management in 14% of lung cancer patients, 9% of colorectal, 4% breast cancer and 7% of women with ovarian cancer.[100] These patients tended to be older than average. In one study, 4% of patients still denied they had cancer 6 months after diagnosis. They estimated that requiring informed consent for cancer registration would cause a loss of at least 4–14% of data.

To avoid causing distress, an important part of the preparation of many studies involving patient participation, is to ask the patient's personal doctor if there is any reason, based on their up to date knowledge of the circumstances of the patient, why an approach may be considered inappropriate.

Attempted suicide using analgesics

In 1998 legislation limiting the size of over-the-counter packs of painkillers (analgesics) was introduced in an attempt reduce the mortality and morbidity associated with deliberate overdose, particularly from paracetamol. Several years later, a large evaluation was carried out to determine the effect of this legislation.[101] Data on analgesic-related deaths and non-fatal overdoses were collected from a number of sources, including the Office for National Statistics, hospital liver units and general hospitals. Data on the sales of analgesics were also included in the analysis. No patients were contacted during the study and data were accessed without consent.

The study found a 22% decrease in the number of suicidal deaths from paracetamol and aspirin in the year after legislation. This reduction was maintained for the following two years. Related admissions to liver units and liver transplants were also reduced by 30%. Large overdoses were reduced by 20% for paracetamol and 40% for aspirin in the second and third years after legislation. The study concluded that the legislation had been effective and suggested that a further reduction in the size of analgesic packs would prevent more deaths.

4. Seeking consent may lead to bias

Seeking consent to use data for research raises the issue of self-selection bias amongst data subjects. There is good evidence of differences between individuals who consent to participate in observational records-based research and those who do not.[102] One survey suggests that

99 Iversen A, Liddell K, Fear N, Hotopf M & Wessely S (In Press) C*onsent, Confidentiality and the Data Protection Act: Epidemiological Research and Hard-to-Engage Cohorts.* British Medical Journal.

100 Gavin A T, Fitzpatrick D, Middleton R J & Coleman M P (2002) *Patients' denial of disease may pose difficulty for achieving informed consent.* British Medical Journal **324**, 974.

101 Hawton K, Simkin S, Deeks J, Cooper J, Johnston A, Waters K, Arundel M, Bernal W, Gunson B, Hudson M, Suri D & Simpson K (2004) *UK legislation on analgesic packs: before and after study of long term effect on poisonings.* British Medical Journal **329,** 1076–9.

102 Al-Shahi R, Vousden C & Warlow C (2005) *Bias from requiring explicit consent from all participants in observational research: prospective population based study*. British Medical Journal **331**, 942.

people from higher social groups, older adults and men tend to be more willing than other groups to give consent for researchers to access their medical records.[103]

The impact of consent bias on the outcome of research is well documented but understated in policy discussions about the use of personal data in medical research. Evidence from ***Royal College of Physicians of Edinburgh***

5. Seeking consent may prevent appropriately large studies

In addition to the problem of systematic errors (through the introduction of bias in the study sample) described above, an insistence on seeking consent may reduce the study size that can be carried out within available resources, leading to larger random errors and consequently a reduction in the reliability and generalisability of the research findings.

Abortion and breast cancer

Until to 2001, there was a great deal of controversy about a potential link between termination of pregnancy and an increased risk of breast cancer. Several studies gave conflicting results. Most studies until this point involved interviews with patients. A much discussed issue at the time was whether such studies were subject to reporting bias, i.e. that women with breast cancer might be more likely than control women (with no history of breast cancer) to tell the interviewer if they had had a termination. Such bias would greatly reduce the accuracy and validity of the results.

To circumvent potential reporting bias, researchers conducted a study based on linkage of independent records.[104] Data were analysed from NHS hospital admissions and death certificates without consent. The analysis showed no increase in breast cancer risk after termination of pregnancy. This conclusive result ended the previous speculation and provided more accurate information for patients.

Primary care of schizophrenic patients

Schizophrenic patients increasingly rely on primary care for their physical health care. These patients suffer from increased physical illness and excess mortality and it is essential to ascertain that the care they receive is comparable to that of patients who are not mentally ill. One study addressed this through a case-control retrospective review of primary care records.[105] A requirement for consent would have introduced unacceptable bias into the study findings for two reasons:

1. Those patients who felt more strongly about the standard of care received (whether good or bad) might be more likely to consent.
2. Schizophrenia is often associated with paranoid thoughts, which may affect the likelihood of consent, especially in patients with more severe symptoms. This might result in consenting cases being unrepresentative and potentially invalidating the research results.

The investigators persuaded the relevant RECs to waive the consent requirement. They emphasised that adopting a consent requirement would further stigmatise an already stigmatised group by the production of low quality research.

103 Lawlor D A & Stone T (2001) *Public health and data protection: an inevitable collision or a meeting of minds?* International Journal of Epidemiology **30**, 1221–5.

104 Goldacre M J, Kurina L M, Seagroatt V & Yeates D (2001) *Abortion and breast cancer: a case–control records linkage study*. Journal of Epidemiology and Community Health **55**, 336–7.

105 Roberts L & Wilson S (2001) *Argument for consent may invalidate research and stigmatise some patients*. British Medical Journal **322**, 858.

4.4 Re-use of data for a new research purpose

Although it is not the primary focus of this report, a sustained area of confusion and difficulty concerns the procedures that researchers are sometimes obliged to follow when seeking to address a question not foreseen in the original research proposal and therefore not explicitly included in the original consent. Such procedures can often be extremely resource-intensive, and have been known to discourage important investigations.

Demands for researchers to go back to research participants for renewed consent raise issues not just about practicability, but also about the appropriateness of going back many years later to participants or their families.

> *The valuable uses of routine patient data cannot always be anticipated in advance and hence cannot always be included in the research protocol, ethics submission or original consent, unless that is so blanket in its nature as to raise other problems. Evidence from* ***Professor Mark Haggard FMedSci***

> *You can't really know where your results are going to take you, so good quality follow-up studies are becoming more and more difficult since technically you would have to predict exactly when and how and what data you will want for say a 5 then 10 year follow at the very outset. Evidence from* ***Royal College of General Practitioners Research Group***

In such circumstances, some (but not all) RECs have given approval for the further use of research records, when seeking an extension of the original consent was considered impractical or likely to cause distress to the study participants or their relatives. When the information obtained is not going to be fed back to the participants, and would have no practical bearing on the management of patients, the risk of harm to individuals is extremely low.

To avoid undue obstacles and allow existing research data to be used to address rapidly and efficiently those important hypotheses that arise after the original consent has been obtained, the Working Group considers that it should be accepted as usually unnecessary to obtain retrospective consent in such circumstances.

4.5 Using patient records to identify potential study participants

Most of this report describes instances in which there is no need to contact the data subject. However, patient records may also be used to identify potential study participants, who are then contacted directly and invited to participate in a research study. This raises several issues including:

- access to the data in order to identify potential study participants
- who should contact these individuals and
- how their consent should be registered.

4.5.1 Identifying and contacting potential study participants

Records from hospital discharges, PCTs or GP practices are a useful source for identifying potential study participants. However, unresolved questions remain about whether, how, and by whom identifiable patient records may be accessed as part of preparations for a research study. Evidence submitted to the Working Group indicates that this confusion has created real and significant difficulties for researchers.

Primary care is the main point of entry to the NHS and is often the doorway for patients taking part in research. The GP list system, by including virtually 100% of the general population, provides a potential sampling frame for studies requiring representative samples of the general population, as well as studies of patients with a particular disease or condition.

Some attractions of this sampling frame are listed below:

- It can provide more representative study populations, and therefore more generalisable results, particularly where the balance of care for a condition is based mainly in primary care.
- Practice lists are generally up to date and allow identification of people in particular age groups (unlike the electoral roll).
- Practice populations generally live in a local area, served by the practice, and can be accessed with relative ease.
- Practice populations are currently still relatively stable, particularly in older age groups, allowing continuity of contact and re-contact over time.

The law has sometimes been interpreted as indicating that an individual's medical records should only be accessed by the doctor responsible for that individual's medical care. In that case, only medical practitioners (most often GPs or hospital consultants) would be able to identify participants from medical records. Even when funding is available to reimburse GP practices or hospital consultants, they may not wish to participate because of competing pressures on their time. This can lead to the exclusion of a large research population and, as described above, cause significant bias in the research results.

In some instances, researchers have resolved this problem through setting up a data controller/data processor arrangement. Such an approach, whereby the researcher acts as a data processor under an arrangement with the data controller (e.g. GP, PCT or hospital Trust), is a practicable way to invite participation in large-scale epidemiological surveys and RCTs. For example, the Heart Protection Study (see section 1.3) used a central coordinating centre to act as the data processor of hospital discharge records, identifying and inviting 130,000 potentially eligible participants on behalf of local investigators (acting for the data controllers in accordance with the DPA). Potential recruits were comfortable with this strategy, with fewer than 50 of the 130,000 invitees raising concerns about being approached. In addition, concerns were generally resolved by explaining the controls on data processing and the purpose of the research, with most individuals going on to participate in the study.[106]

4.5.2 Consent for consent

At the beginning of a study, prior checking of records is often needed to confirm that patients meet the clinical entry criteria (e.g. a particular age range and/or the presence/absence of a particular clinical condition). A health professional with up to date knowledge of the patients' circumstances can also check records in order to identify individuals whom it might be inappropriate to contact (e.g. recently deceased, major illness, mental or physical incapacity).

If the first approach to the patient is made after prior checking, it will be clear that the person making the approach must have had access to a patient's personal medical information. Some patients may be offended, not by the invitation to take part in research, but by the knowledge that someone has had access to their personal medical records.

These concerns have contributed to the promotion of a policy termed 'consent for consent', in which an initial approach is made by someone known to patients (such as a GP) to ask for their consent to be contacted by a researcher. 'Consent for consent' has several significant implications for research, including the time and resource costs involved in carrying out an extra recruitment stage and the capacity and willingness of the health provider to do extra work on behalf of the researcher. Moreover, there is some evidence indicating that patients feel more obliged to consent to participation if the approach comes from their own practitioner.[107] GPs must therefore

106 Submission from Professor Rory Collins FMedSci.

107 Sugarman J, Regan K, Parker B, Bluman LG & Schildkraut J (1999) *Ethical ramifications of alternative means of recruiting research participants from cancer registries.* Cancer **86**, 2707.

Recruiting patients for an ulcerative colitis study

Ulcerative colitis (UC) is a long-term inflammatory disorder that causes ulceration of the bowel leading to diarrhoea, tiredness, abdominal pain and poor appetite. A research team undertaking a trial of self-management intervention in UC patients needed to identify and recruit suitable participants[108]. To do this the researchers needed to identify patients with a confirmed diagnosis (e.g. from a hospital database) and contact them with a 'screening' questionnaire regarding the impact of the condition on their daily life, in order to identify the most eligible patients for the trial.

Because it had been suggested that recruitment through direct contact from a researcher could distress patients, who might express concern about how the researcher has identified them, the researchers compared two methods of recruiting patients:

A: They obtained a list of names from the responsible clinician and, with the clinician's permission, wrote to those patients directly explaining the study and enclosing the questionnaire. Reminder letters and duplicate questionnaires were also sent. Patients could opt out of the study either by explicitly refusing to take part or implicitly through not returning the questionnaires.

B: After compiling a list of names, the clinician contacted the patients, sending brief details of the study and asking those who are interested to 'opt in' by returning a reply slip to the research team. Only then was the questionnaire sent. No reminders were sent after the initial request to opt in, although reminders were sent in respect of the subsequent questionnaire survey.

Under protocol A, 63% of patients returned the questionnaire. Under B, only 26% of patients returned the questionnaire (~67% of the 38% who agreed to be approached by the researchers). Under scheme B, roughly two and half times as many potential participants must be identified and contacted in order to recruit sufficient numbers for a viable study. When questioned further, participants in protocol B appeared to 'fear the unknown', perceiving the subsequent questionnaire to be more demanding than was the case and refusing participation.

Recruiting participants for angina management study

All potential recruits for a general practice based study into the management of angina received information about the study. However, half of the patients were randomised to an 'opt in' method of recruitment, whereby they only took part if they got in touch with the researchers expressing a positive interest. The other half were randomised to an 'opt out' approach and were contacted directly by a researcher to discuss recruitment, unless they explicitly asked to be excluded following the initial invitation.

The 'opt in' route resulted in a lower participation rate (38% versus 49% of the 'opt out' patients) and a biased sample at lower cardiac risk with less functional impairment. The study authors concluded that that giving patients the opportunity to 'opt out' is a better method of recruiting a representative sample of patients.[109]

108 Submission to call for evidence from Elaine McColl, Director of the Newcastle Trials Unit, University of Newcastle upon Tyne.

109 Junghans C, Feder G S, Hemingway H, Timmis A D & Jones M M (2005) *Recruiting patients to medical research: A double-blind randomised trial of 'Opt In' versus 'Opt Out' strategies*. British Medical Journal **331**, 940.

be careful in their involvement as study recruiters, striking the right balance between enabling patients to make their own decisions concerning participation in research and intervening to avoid approaches that might be inappropriate.

4.5.3 Opting in versus opting out

Participation rates are strongly influenced by the way in which patients are invited. The two examples opposite compare 'opt out' versus 'opt in' approaches to recruitment.

These examples illustrate various kinds of selection, recruitment and participation bias. These are not a new feature of research, but are important potential causes of misleading findings. Participatory research is confined to about 70% of the population, with deprived areas and populations being notably under-represented in the research literature. Insisting on an opt-in mechanism of participation will clearly increase such biases. In addition to population biases within research programmes, an insistence on 'opting in' mechanisms of consent presents great difficulty for research into conditions that affect some of the most vulnerable members of society. Research in mental health, sexually transmitted diseases and neurodegenerative conditions are especially likely to be affected by low recruitment rates.

4.6 Public expectations and engagement

The ethical basis for accessing and using patient records for a research study, with and without their subsequent participation, depends greatly upon patient expectations about how their routine health record is used. PIAG makes this point in its submission to the call for evidence: '*When people go to their doctor or health professional, they are seeking treatment because they are unwell, and not to become the subjects of research*.'[110]

Although this statement is undoubtedly true it should not be used to imply that patients regard research using their medical records as unacceptable. The fact that patients' motivation for accessing health services is to seek treatment does not exclude the possibility that, in the course of such treatment, they might be eager to help others through direct or indirect participation in research. In discussion with the Working Group, patient representatives were only too aware that evidence-based treatments rely on data from previous research participants and stressed the importance they place on research for the public good.

However, available evidence suggests that the current level of public awareness in relation to the use of medical records in research is low (section 5). This is especially true for research in general practice, which may reflect the low volume of research in primary care and the relative infrequency with which research features as an issue within routine consultations.

> *Patients may have good reason to have, or come to have, an expectation that their records will be used for low-risk research without their consent under certain conditions. Where this is the case such expectations provide reasonable grounds for considering such research to be ethical. Evidence from* ***Professor Michael Parker***

A sustained effort is needed to increase public engagement concerning the value of research using health care records and the arrangements under which records are held, as well as the circumstances and procedures by which their records may be accessed for this purpose (see section 5). It is particularly important to clarify the role of the GP, primarily in protecting patient's confidentiality and privacy, but also in assisting the research process. Improved public engagement will increase the likelihood that patients have a reasonable expectation that their record may be accessed for certain research purposes without their consent, within strict safeguards.

110 Submission to call for evidence.

4.7 Discussion and conclusions

Consent and the use of identifiable data for research

The examples given in this section illustrate that an insistence on explicit consent for the secondary use of identifiable data can reduce the size of studies that can be undertaken (or may even prevent them entirely) and may sometimes introduce biases that make the results difficult to interpret or unreliable. Such biases can lead to the exclusion of some social groups from the research, thereby rendering them less likely to be considered appropriately in subsequent health policy recommendations.

The ethical arguments surrounding the importance of consent in protecting the right to privacy must always be balanced against these risks. Although each case must be judged on its merits, evidence presented to us clearly shows that undue emphasis on the need to obtain explicit consent has impeded research using personal data.

The impracticability of seeking consent has long been recognised within the context of routine health care, where implied consent is the basis for most sharing of patient information. This system works because patients give explicit consent at key points, such as when agreeing to attend hospital for cardiac surgery, and in so doing implicitly consent to the use of their data (by technicians nurses, doctors) to facilitate the provision of that surgery.

In their evidence to us, researchers have argued that, in the absence of further legal clarification, research using personal data should be considered an integral part of health care and, as such, consent for data to be used in research is implicit by those using the health service. Researchers argue that, from their experience, this would be acceptable to most members of the public, but it is not demonstrably clear that this is so. Although the Working Group therefore supports the spirit of this argument, we consider that extending the limits of implied consent to include research using personal data is not generally compatible with the DPA.

Many respondents to the call for evidence called for the development of a system of NHS-wide general consent for low-risk research using personal data, which would include an opt-out mechanism. In this way, patients could opt out generally from all research (including being approached to participate in research), and have their records flagged to indicate this. Some respondents raised the possibility that patients may wish to opt out of some studies, but not others. However, we support the view that it would be impracticable to operate a system of variable opt-outs.

> *We believe that the Department of Health, NHS and broader research community should work together to raise public awareness of the need for, and value of, epidemiological/public health research, as a way of reducing the likelihood of individuals opting out. Evidence from* ***Health Protection Agency***

Connecting for Health may provide an opportunity for the development of such a mechanism. However, several respondents to the call for evidence have emphasised the potential detriment to the Connecting for Health data set that would be caused by large numbers of participants (who may be disproportionate in health status) opting out. It is essential that such a mechanism is accompanied by a positive public engagement campaign from the Government and NHS illustrating the benefits of research and the safeguards in place, as well as encouraging participants not to opt-out.

The Academy considers that Good Practice Guidance should be developed concerning consent requirements for research using personal data with reference to the following criteria:

- **the risk of introducing bias that will endanger the validity of the results**
- **the overall financial and time burdens imposed**
- **the size of the study population and the proportion likely to be untraceable.**
- **the risk of inflicting harm or distress by contacting people.**

Using patient records to identify potential study participants

Patient concerns about confidentiality are likely to be highest when research concerns sensitive conditions or where the records contain sensitive information. In these instances, the case for patients first being approached by their own health provider is stronger. However, as a general rule, the Academy considers that it should not be necessary to seek prior consent for studies requiring a general sample of people (e.g. population surveys) or involving patients with non-sensitive conditions (e.g. diabetes or coronary heart disease). In these circumstances, it would be appropriate for any invitation to explain how the researcher obtained the person's contact details and the limits of the researcher's knowledge about their health details. Much will depend on patients' general understanding of how their records are accessed and used for research (see section 5).

The Academy considers that Good Practice Guidance should include:

- **the conditions and procedures by which health records may be accessed at the start of the research process**
- **the mechanism for contacting potential study recruits**
- **the mechanism for registering agreement or refusal to participate.**

5 Engaging the public

Summary

- *Research using personal data provides the evidence by which public health can be improved. Public involvement in this process is essential for success.*
- *There has been little research that gives reliable information on public awareness of, or attitudes to, research of this type. Research has largely been concerned with general attitudes to the confidentiality of health data in the context of care and treatment. Further research is needed around specific issues related to research using personal data.*
- *The use of data for research should be seen as a major benefit for individuals within society. Public confidence and enthusiasm can only be gained by better understanding of the value of such research and the demonstration of high standards of research practice.*

5.1 Introduction

As described in section 2, the laws and regulations concerning research using personal data must strike a balance between individual rights and the public benefit derived from research findings. The regulatory emphasis placed on privacy and consent stems in part from an interpretation of public opinion concerning data confidentiality in general (see section 1.5). However, evidence of public attitudes and opinions around the specific issue of the use of personal data for medical research is largely absent. The lack of dialogue between researchers, regulators and the public has left a void that has been filled by defensive and restrictive interpretations of the law, which may not represent the wishes of an informed and engaged public.

It was a priority for the Working Group to consult with a wide range of patient representatives about their attitudes and concerns relating to research using personal data.[111] Through these consultations we received a great deal of helpful opinion and advice that has made us question assumptions that the public places personal privacy above the societal benefits from research using personal data. The opinions we heard were limited in number but were well informed and based on a lengthy discussion of the nature of this type of research. The strong support from patient groups for this form of research has strengthened our view that engagement with the public is one of the most important, and urgent, tasks in developing future arrangements and appropriate governance for the use of health information in medical research.

5.2 Research into public attitudes towards the use of personal data

Several commentators refer to a deep public unease over the extent to which public and private bodies have access to personal information.[112] While this general assertion may be accurate, it is unclear how it translates to the specific use of health data for medical research. Two large projects in central Hampshire and south Staffordshire concluded that the public are relatively uninterested in this issue.[113,114] The Academy's consultation with the GMC and ICO revealed that they very rarely, if ever, receive calls from the public about the use of data for medical research.

There is only a small body of research on public and patient awareness of, and attitudes to, the use of health information for medical purposes. A literature review carried out by Shickle in 2002[115] identified only 44 relevant UK papers. Most of these involved the use of information in the clinical care context (e.g. attitudes to the number of personnel who see medical notes during a

111 See www.acmedsci.ac.uk.

112 Cayton H & Denegri S (2003) *Is what's mine my own?* Journal of Health Services Research Policy 8 (supplementary).

113 Adam T, Budden M, Hoare C, Sanderson H (2004) *Lessons from the central Hampshire electronic health record pilot project: issues of data protection and consent*. British Medical Journal **328**:871-4

114 *The south Staffordshire patient consent project: findings from the HER awareness campaign*. June 2002.

115 School of Health and Related Research (2002) Patient Electronic Record: Information and Consent (PERIC) *Public attitudes to protection and use of personal health information*. School of Health and Related Research (ScHARR), Sheffield, University of Sheffield.

hospital episode). Since 2002 there has been more research, including 2 studies undertaken by the NHS,[116,117] but the research base is still small.

In addition to the relative paucity of research in this area, the research base suffers from a lack of comparability and specificity.

Comparability

Research methodologies can differ markedly between studies, including: omnibus survey interviews; face-to-face interviews; postal surveys; and focus groups. Similarly, the composition of participants can range from members of the public, randomly selected patients, or patients with a specific condition (e.g. previous human immunodeficiency virus (HIV) or cancer diagnosis or termination of pregnancy). This lack of comparability makes it difficult to build up a reliable picture of research findings.

Specificity

Most studies in this area have seldom asked specific questions about the use of patient identifiable data in medical research. Quantitative research asking undifferentiated questions is not adequate for assessment of attitudes towards different types of research conducted by different groups for different purposes.

With the limitations of the research base noted, some tentative conclusions can be drawn from the literature. Firstly, it appears that patients are generally more concerned about *who* has access to their medical information, rather than the *purposes* for which it is used. As a general rule, patients grow increasingly concerned about access to their data as control moves away from their own GP.

Secondly, a chief concern appears to surround access to patient information by parties

Two large studies are especially noteworthy because of the rigorousness of the methodology and the focus of the questions:

Shickle *et al*.[118] conducted a study of public opinions of the use of electronic records in health care. The findings showed that there were social variations in willingness to share records for health care (men, older people and higher social groups being more willing), that anonymised data were preferred where possible and that the uses to which the data were put was not a strong determining factor in whether participants were happy with data sharing. Participants were more accepting of the need for doctors to see their records than receptionists and social workers. For research there was some definition of research purpose but the enquiry was not explicit with respect to methods of ensuring confidentiality or research regulation so the underlying knowledge of the participants in answering the questions cannot be assessed.

A large and well-designed recent study by Barrett *et al*. concentrated on the use of medical records and registration for cancer research[119]. Interviews were carried out by the Office of National Statistics Omnibus Survey in a large random sample of UK homes. Participants were given a full explanation of the purpose of the research before being asked their opinion. The great majority of participants supported the use of their personal data for cancer research and registration, provided confidentiality and security were assured. Predictably perhaps, the investigators found that only a small proportion of the public knew of the existence of cancer registries. However, when asked, the great majority supported a law to make cancer registration statutory, as it is in many other developed countries.

116 NHS Information Authority in conjunction with The Consumers' Association and Health Which? (2002) *Share with Care! People's views on consent and confidentiality of patient information*.

117 NHS National Programme for Information Technology in conjunction with Health Which? (2003) *The public view on electronic health records*. Available online at http://crawl04.archive.org/ukgov/20040216051237/http://www.doh.gov.uk/ipu/programme/research.pdf.

118 School of Health and Related Research (2002) Patient Electronic Record: Information and Consent (PERIC) *Public attitudes to protection and use of personal health information.* School of Health and Related Research (ScHARR), Sheffield, University of Sheffield.

119 Barrett G, Cassell JA, Peacock JL, Coleman MP (In press) *In the public interest, or an invasion of privacy? A national survey of public attitudes towards the use of identifiable medical data by the National Cancer Registry.*

outside the health service. In a Consumers' Association study, these parties were identified as insurance companies, employers and schools; researchers were not mentioned. The question of whether researchers are perceived to be within the health service was not raised.

Thirdly, there appears to be stratification according to gender, race, socio-economic and age lines regarding willingness for data to be used in research.[120] This has important implications for the potential bias introduced by non-responders and non-consenters in research studies.

There are considerable gaps in the evidence and issues related to research have almost always been considered as part of a wider study of patient attitudes. There has also been little exploration of patients' perceptions of the effect of research on health care and the processes governing research using personal data. The Working Group particularly welcomes recent initiatives by the Wellcome Trust and Cancer Reseach UK to undertake work in this area.

The public may also have different perceptions of NHS, university and commercial researchers. There are indications that distinctions are made between these groups regarding access to data, but the reasons underlying these are unclear.

5.3 Discussion, conclusions and recommendations

Standards for researchers

It has been stressed to us that the research community will obtain, and deserve, the support it needs only if it can demonstrate high standards of research practice involving personal data. We concur with this view. Demonstrable standards of data handling are also of concern to ethics committees, PIAG, hospital and Primary Care Trusts, Universities and research funders.

The Research Councils have a central role in setting standards of best practice in research training and have issued a joint statement regarding the skills that doctoral students would be expected to develop during their studentship.[121] The Council's statement on training provides a common overview of the skills and experience of a typical research student or post-doctoral fellow. It is expected that each Council will have additional requirements specific to their field of interest. The section on Research Environment mentions ethical principles, confidentiality and data protection and specifies that students should be able to: '*demonstrate awareness of issues relating to the rights of other researchers, of research subjects, and of others who may be affected by the research e.g. confidentiality, ethical issues, attribution, copyright, malpractice, ownership of data and the requirements of the Data Protection Act.'*

It is essential that researchers conducting research involving health information are fully aware of the relevant legislation and underlying ethical principles, including the NHS Duties of Care and Confidentiality. Researchers should also possess a thorough understanding of research governance policy and processes, (including those of ICO, RECs, and PIAG), not least to ensure the most efficient use of research time and resources.

Although the responsibility for training and keeping up to date with best practice rests with principal investigators, the Working Group considers university departments and NHS Trusts to have a central role in ensuring that such training is available and publicised. The Working Group believes that research funders also have a responsibility in ensuring that policies on minimum standards and best practice relating to the use of personal data in research are known by researchers, regulators and the public. We recommend that funders publish such policies on their websites.

120 School of Health and Related Research (2002) Patient Electronic Record: Information and Consent (PERIC) *Public attitudes to protection and use of personal health information*. School of Health and Related Research (ScHARR), Sheffield, University of Sheffield.

121 Available online at http://www.bbsrc.ac.uk/funding/training/skill_train_req.pdf.

Improving public engagement and confidence

The Working Group's consultation with patients and patient representatives revealed over-whelming support for research using personal data and confidence in the integrity of research practices. However, evidence of public attitudes and opinions on the specific issue of research using personal data is largely lacking. The absence of such knowledge and the lack of public debate forces regulatory and advisory bodies to make defensive assumptions about what the public might find acceptable. Development of good practice should be informed, as far as possible, by empirical evidence on public and patients' awareness and attitudes.

The ethical basis for accessing and using patient records for a research study, with or without consent, depends greatly upon public expectations about how routine health records are used. Urgent work is needed to increase public engagement around the value of research using health care records and the arrangements under which records are held, as well as the circumstances and procedures by which their records may be accessed for research purposes.

Research funders should encourage and fund research into public awareness and attitudes towards medical research using personal data

Researchers know that medical research in general, and research using personal data in particular, is highly regulated. The public, however, is largely unaware of these controls and the way in which the standards of research are maintained. This is a matter of concern for both the public and researchers alike.

The advice we received indicates that the wider research community must itself engage with the public to raise awareness of the benefits of the research involving personal data and to demonstrate that high standards are consistently applied. Research funders, regulatory bodies and universities could do much in this area, and we encourage collaborative activity through the auspices of the UK Clinical Research Collaboration (UKCRC). Charities with strong patient/user input could also play a particularly important role in more actively advocating the value of research using personal data. Ultimately, there is a need for the UK Departments of Health to undertake a programme of public engagement around these issues.

Steps that could be taken quickly include:

- improving publicity in primary care and hospital settings of the value of research using personal data and the health benefits that have been derived
- taking advantage of forthcoming opportunities to engage the public through Connecting for Health and the introduction of the National Care Record
- ensuring the presence of public representatives on regulatory bodies for research using personal data.

The UK Departments of Health, working with the UK Clinical Research Collaboration, should develop public engagement programmes around the purpose and value of using personal data in medical research.

Appendix I Report Preparation

Working Group

This report was prepared by an Academy of Medical Sciences Working Group. Members participated in a personal capacity, rather than as a representative of their organisations.

Chair

Professor Robert Souhami CBE FMedSci
Emeritus Professor of Medicine, University College London

Members

Dr Sandy Chalmers
Director, Data Privacy Policy, GlaxoSmithKline

Professor Rory Collins FMedSci
Professor of Medicine and Epidemiology, University of Oxford

Professor Karen Luker FMedSci
Professor of Community Nursing, University of Manchester

Professor John Newton
Professor of Epidemiology and Public Health, University of Manchester

Professor Alan Silman FMedSci
Professor of Rheumatic Disease Epidemiology, University of Manchester

Professor Graham Watt FMedSci
Professor of General Practice and Primary Care, University of Glasgow

Professor Simon Wessely FMedSci
Professor of Epidemiological and Liaison Psychiatry, King's College London

Dr Ron Zimmern
Director, Public Health Genetics Unit, University of Cambridge

Legal Adviser to the Working Group

Dr Kathy Liddell
Lecturer in Law, University of Cambridge

Secretariat

Dr Helen Munn
Senior Policy Officer, Academy of Medical Sciences

Review Group

This report was reviewed by a group appointed by the Academy Council:

Professor Nick Wald FRS FMedSci (Chair)
Professor of Environmental & Preventative Medicine, St Bartholomew's & Royal London School of Medicine and Dentistry

Ms Mary Baker MBE
President, European Parkinson's Disease Society

Professor Roger Jones FMedSci
Wolfson Professor of General Practice, King's College London

Baroness Onora O'Neill PBA FMedSci
President, The British Academy

The Academy is also grateful to the following individuals for providing comments on the draft report:

Dr Richard Ashcroft
Head of Medical Ethics Unit, Imperial College, London

Sir Andrew Haines FMedSci
Dean, London School of Hygiene and Tropical Medicine

Professor Vivienne Harpwood
Professor of Medical Law, University of Cardiff

William Lowrance PhD
Consultant in Health Policy & Ethics, Geneva

Mrs Shirley Nurock
Alzheimer's Society

Dr Jem Rashbass
Clinical and Biomedical Computing Unit, University of Cambridge

Professor Genevra Richardson
Professor of Public Law, Queen Mary, University of London

Lord Turnberg FMedSci
Scientific Adviser, Association of Medical Research Charities

Professor Patrick Vallance FMedSci
Head, Division of Medicine, University College London

Appendix II List of consultees and respondents to the call for evidence

Responses to the call for evidence

Organisations

Cancer Research UK
Centre for Disease Surveillance and Control, Health Protection Agency
Genetic Interest Group
GlaxoSmithKline
Medical Research Council
Patient Information Advisory Group
Royal College of General Practitioners Research Group
Royal College of Pathologists
Royal College of Physicians
Faculty of Pharmaceutical Medicine, Royal College of Physicians
Royal College of Physicians of Edinburgh
Royal College of Obstetricians and Gynaecologists Academic Committee
Royal College of Surgeons
Society for Academic Primary Care
UK Faculty of Public Health Research Committee
The Wellcome Trust

Individuals

Mr Paul Affleck, Leeds Cancer Research UK Clinical Centre
Dr Rustam Al-Shahi, University of Edinburgh
Professor Jeanne Bell FRSE FMedSci, University of Edinburgh
Sir Walter Bodmer FRS FMedSci, University of Oxford
Dr Iain Buchan, University of Manchester
Professor Bruce Campbell, Royal Devon & Exeter Healthcare NHS Trust
Professor Michel Coleman, London School of Hygiene and Tropical Medicine
Sir Richard Doll OBE CH FRS FMedSci, University of Oxford
Professor Tom Fahey, University of Dundee
Mr Lester Firkins, Human BSE Foundation
Dr Anna Gavin, Northern Ireland Cancer Registry
Dr Steve George, University of Southampton
Sir John Grimley Evans FMedSci, University of Oxford
Professor Mark Haggard FMedSci, Addenbrooke's Hospital, Cambridge
Professor Pali Hungin, NHS R&D Forum
Professor Roger Jones FMedSci, King's College London
Professor Rick Kaplan, National Cancer Research Network
Dr Moira Malfroy, National Blood Service
Dr Elaine McColl, University of Newcastle upon Tyne
Professor Tom Meade FRS FMedSci, London School of Hygiene & Tropical Medicine
Professor David Menon FMedSci, University of Cambridge
Professor Mike Parker, The Ethox Centre, University of Oxford
Professor John Parkinson, University of Dundee
Sir John Pattison FMedSci
Sir Denis Pereira Gray OBE FMedSci, Chairman, The Nuffield Trust

Professor Julian Peto FMedSci, London School of Hygiene & Tropical Medicine
Professor Peter Pharoah, University of Liverpool
Professor Mike Pringle CBE FMedSci, University of Nottingham
Dr Cathy Ratcliffe, National Translational Cancer Research Network
Professor Stephen Sacks, Chair, King's College London Research Ethics Committee
Dr Peter Singleton, Cambridge Health Informatics
Professor W Cairns Smith, University of Aberdeen
Mr Peter Stephens, IMS Health
Dr Allan Sudlow, MRC
Professor Frank Sullivan, University of Dundee
Mr C Marc Taylor, Department of Health
Professor Stephen Tomlinson FMedSci, Provost, Wales College of Medicine
Professor Douglas Turnbull FMedSci, University of Newcastle upon Tyne
Dame Margaret Turner-Warwick DBE FMedSci
Professor Nicholas Wald FRS FMedSci, St Barts & Royal London School of Medicine & Dentistry
Professor Charles Warlow FMedSci, University of Edinburgh
Professor John Warner FMedSci, University of Southampton

Meetings with members of the Working Group

Mr Harry Cayton, National Director for Patients and the Public, Department of Health
Sir Graeme Catto FRSE FMedSci, Chairman, General Medical Council
Professor Sally Davies FMedSci, Director of Research & Development, Department of Health
Ms Jane Durkin, Assistant Commissioner Public Sector Compliance, Information Commissioner's Office
Sir Andrew Haines FMedSci, Dean, London School of Hygiene & Tropical Medicine
Professor Joan Higgins, Chair, Patient Information Advisory Group
Dr William Lowrance, Consultant in Health Policy & Ethics, Geneva
Ms Rebecca Mussell, Senior Ethics Adviser, British Medical Association
Mrs Shirley Nurock, Alzheimer's Society
Dr Jane O'Brien, General Medical Council
Dr Liam O'Toole, Chief Executive, UK Clinical Research Collaboration
Dr Jem Rashbass, Clinical and Biomedical Computing Unit, University of Cambridge
Dr Phil Walker, NHS National Programme for IT

Delegates attending the patient consultation meeting*

Ms Yaa Adjei, Islington PCT Patient and Public Involvement Forum
Mr Arthur Brill, Royal Free Patient and Public Involvement Forum
Mr Paul Bull, Multiple Sclerosis Society
Mrs Elaine Davies, National Kidney Research Fund
Ms Marian Dmochowska, Cystic Fibrosis Trust
Dr Lee Dunster, Multiple Sclerosis Society
Ms Wendy Fisher, South East London Strategic Health Authority
Ms Alison Forbes, University College London Hospitals NHS Foundation Trust
Mr Neil Formstone, Royal College of Pathologists - Lay advisory Committee
Ms Meg Gilpin, Alzheimer's Society
Mrs Christobel Hargraves, National Confidential Enquiry into patient outcome and death
Mr Nick Jolliffe, The Stroke Association
Ms Annabel Kanabus, Avert
Dr Joanne Knight, The Stroke Association

* For full details see http://www.acmedsci.ac.uk.

Mrs Christine Lavery, Society for Mucopolysaccharide Diseases
Mr John Marriott, RCP Patient and Carer Network
Mr Tom Mcloughlin, Cystic Fibrosis Trust
Mrs Shirley Nurock, Alzheimer's Society
Mr Simon O'Corra, Diverse Identities
Ms Linda Partridge, WellChild
Dr Sophie Petit-Zeman, Association of Medical Research Charities
Professor Naomi Pfeffer, Consumers for Ethics in Research
Mrs Margaret Ponder, Genetic Interest Group
Mrs Barbara Woodward-Carlton, Alzheimer's Society
Ms Denise Vaughan, Meningitis Research Foundation
Mr Jas Weir, Epsom and St Hellier Patient and Public Involvement Forum
Mr Mike Williams, The Kings Fund
Ms Jessie Winyard, West Hertfordshire Hospitals NHS Trust

The **Academy** *of*
Medical Sciences

Academy of Medical Sciences
10 Carlton House Terrace
London, SW1Y 5AH

Tel: +44(0)20 7969 5288
Fax: +44(0)20 7969 5298

Email: apollo@acmedsci.ac.uk
Web: www.acmedsci.ac.uk