NUFFIELD COUNCIL ON BIOETHICS

BACKGROUND PAPER

Genomics, health records, database linkage and privacy

Ruud H.J. ter Meulen, Ainsley J. Newson,
Mari-Rose Kennedy, Behnaz Schofield
Centre for Ethics in Medicine, School of Social and Community Medicine, University of Bristol

Forward Look
19-20 May 2011

Note
The authors were commissioned by the Nuffield Council on Bioethics to write this paper in order to inform the Council's discussions about possible future work on this topic. The paper is intended to provide an overview of key clinical, ethical, social, legal and policy issues, but is not intended to offer any conclusions or recommendations regarding future policy and practice. Any views expressed in the paper are the views of the authors and not those of the Nuffield Council on Bioethics.

NUFFIELD COUNCIL ON BIOETHICS

Genomics, health records, database linkage and privacy

Background Paper

Ruud H.J. ter Meulen, Ainsley J. Newson, Mari-Rose Kennedy, Behnaz Schofield
Centre for Ethics in Medicine, School of Social and Community Medicine
University of Bristol

Summary

1 This paper outlines the current state of the art, future developments and main ethical issues that arise in the context of genomics, health records, database linkage and privacy. Current oversight and potential questions that might need to be addressed are highlighted.

Introduction

2 Medical research is guided by a range of ethical and legal norms which in general aim to protect the rights and well-being of participants. Important safeguards include the principle of informed consent, the protection of confidentiality and privacy, protection from harm and the fair distribution of benefits and burdens.

3 Although the principles of consent, confidentiality and privacy are widely recognised and supported, debate continues as to whether these principles should be applied equally to all kinds of medical research. Epidemiological researchers are increasingly frustrated by what they see as the hampering of research by strictly applied ethical norms and legal procedures.[1,2] Some have argued that obtaining the relevant permissions for research that waives consent are arduous and a poor use of public funds,[3] while others have called on research governance bodies to better recognise that it is lawful to use identifiable data for research without consent provided that the use is necessary and proportionate.[4] However, these perspectives are also subject to criticism by other stakeholders.

---

1 Iversen A, Lidell K, Fear N, Hotopf M, Wessely S. (2006) Consent, confidentiality and the Data Protection Act. *BMJ* 332:165-169.

2 Vandenbroucke JP. (1998) Maintaining privacy and the health of the public. Should not be seen as in opposition. *BMJ* 316:1331-2.

3 Metcalfe C, Martin RM, Noble S, Lane JA, Hamdy FC, Neal DE, Donovan J. (2008) Low risk research using routinely collected identifiable health information without informed consent: encounters with the Patient Information Advisory Group. *J Med Ethics* 34:37-40.

4 Academy of Medical Sciences (2006) *Personal data for public good: using health information in medical research*. London: Academy of Medical Sciences.

4 This paper will present the main ethical arguments arising in the debate over whether the ethical and legal standards of medical research should be applied in the same degree to all kinds of medical research or whether there should be exemptions of such regulations for some kinds of medical research like epidemiology and other research on public health. The ethical issues in three areas of epidemiological research will be considered:

- Genome Wide Association Studies (GWAS)
- Health Records Research
- Data Linkage Research

5 The paper will provide an overview of the state of the art in each of these three areas and will then synthesise the ethical and legal issues arising from the tension between the public good and the protection of individual rights in each research area. These ethical issues are:

- The role of informed consent (including broad versus specific consent)
- Confidentiality and privacy
- Public benefits and the public good as opposed to individual rights
- Justice and solidarity
- The role of public trust
- Ownership of data and the right [not] to know

Scientific State of the Art

Genome-Wide Association Studies (GWAS)

6 In GWAS, researchers systematically search the entire human genome to find associations between variations in multiple short DNA sequences (termed single nucleotide polymorphisms or SNPs) and the prevalence of an observable disease or trait (such as blood pressure or weight).[5]

7 Some features of GWAS are that:[6,7]

- This kind of research tends to identify several polymorphisms that are of significance, rather than a single gene;
- The exact contribution (penetrance) of any one polymorphism identified to a particular condition, disease or trait is likely to be uncertain;
- Taken collectively, all the polymorphisms identified for a trait will not necessarily explain all of the heritability (the proportion of population variation that can be explained due to genetic effects) of a condition, disease or trait;
- This research requires collecting large numbers of biological samples and creating large data sets for statistical (bioinformatic) analysis; and

---

http://www.acmedsci.ac.uk/download.php?file=/images/publication/Personal.pdf (accessed 21/04/2011).

5 Donnelly P. (2008) Progress and challenges in genome-wide association studies in humans. *Nature*. 456(7223):728-31.

6 McCarthy MI, Abecasis GR, Cardon LR, et al. (2008) Genome-wide association studies for complex traits: consensus, uncertainties, and challenge. *Nature Reviews Genetics* 9:356-369.

7 Pearson, TA; Manolio, TA. (2008) How to Interpret a Genome-wide Association Study, *JAMA* 299(11):1335-1344.

- Studies that do find associations between polymorphisms and a condition or trait will require replication.

8 GWAS has led to some high-profile publications, such as those from the Wellcome Trust's case control consortium which has, among others, identified genes associated with obesity.[8]

9 Some have criticised the GWAS approach as being incapable of ever identifying all of the 'missing heritability' that influences human health and behaviour. One response to this type of criticism is that there is not yet any better method for identifying genes influencing complex human disease and traits and that GWAS continues to generate a large number of high-impact publications.

10 As methods of genetic sequencing continue to improve, future research may involve sequencing participants' entire genomes rather than concentrating on known SNP sites.

Data linkage

11 Advances in information technology now enable epidemiological research that is based on linkage of individuals' administrative health and social data (for example, educational and criminal records) as well as the enhancement and validation of existing research data.[9,10] Linkage to administrative databases is particularly useful for social epidemiological research exploring issues about the social distribution and social determinants of health. To be linkable, data *must be identifiable*; at least at the point of linkage.

12 Data linkage can help in follow-up studies of cohorts or other groups to determine factors such as residential status or health outcomes. It can be used for follow-up of various sorts of cohorts, clinical trials, and longitudinal surveys to identify the causes of diseases such as cancer.

Health Records Research

13 Epidemiological surveys significantly rely on data collected from the patient's medical records or from the patient through a questionnaire (either self completed or nurse/doctor led). These are negligible risk 'fact finding' studies; usually accompanied by taking a series of anthropometric measurements. Often a blood sample is required to ensure the measurement under study is examined in a systematic, standardised way.

Key ethical issues

---

[8] For a list of all published genome-wide association studies, see: http://www.genome.gov/26525384 (accessed 19/04/2011).

[9] Tate R, Calderwood L, et al. (2006) Mother's consent to linkage of survey data with her child's birth records in a multi-ethnic national cohort study. *International Journal of Epidemiology* 35:294-298.

[10] Huang N, Shih S-F, et al. (2007) Record linkage research and informed consent: who consents? *BMC Health Services Research* 7(18).

Genome-wide association studies

14 Kaye *et al* have pointed out some of the key aspects of GWAS that give rise to new ethical challenges. These include that GWAS research and data:

- Provides very detailed genetic information about individuals;
- Enables information about a variety of diseases or traits to be obtained from one biological sample; and
- Gives rise to a huge amount of data that can be shared relatively easily.[11]

Consent

15 The principle that research participants should have the right to make an informed decision about their participation in research or clinical treatment is entrenched in ethical principles, international instruments and some laws. Providing fully informed consent protects participants as it allows them to weigh the possible benefits and harms of an intervention and to exercise some control over their data, thus promoting their individual autonomy. With appropriately informed consent, those administering an intervention are also protected from legal sanctions, such as an action for negligence or assault. There are, however, limitations of informed consent common to all research involving human participants, including those under discussion here. This can make it difficult to define when consent is truly informed. Barriers to informing individuals include difficulties with language, communication and understanding as well as differing requirements for information that individuals may require before consenting.[12]

16 The challenges to consent posed by GWAS primarily involve determining a model of consent that appropriately balances the interests of participants and researchers and can account for the specific aspects of this research. Full informed consent is challenged by five factors: (1) Uncertainty over what the research may find or what harms it may give rise to; (2) The sheer volume of research data that is generated, including unexpected findings and the prospect of whole genome sequencing; (3) The benefit to be gained from sharing research data among different research groups, including beyond jurisdictional boundaries; (4) Future advances having a retrospective effect on existing data; and (5) Determining what level of detail to report back to participants (discussed further below). Given this, the traditional notion of 'fully informed consent' has been suggested to be impossible to obtain for GWAS.[13]

---

[11] Kaye J, Boddington P, de Vries J, Gowans H, Hawkins N, Heeney C, Melham K. (2009) *Ethical, Legal and Social Issues Arising from the Use of GWAS in Medical Research*. Report for the Wellcome Trust, p3.

[12] Manson, N. C. and O. O'Neill (2007). *Rethinking Informed Consent in Bioethics*. Cambridge, Cambridge University Press

[13] Caulfield, T., McGuire, A. L. *et al*. Research ethics recommendations for whole-genome research: Consensus statement. *PLoS Biol* 2008;6:e73, cited by Kaye J, Boddington P, de Vries J, Gowans H, Hawkins N, Heeney C, Melham K. (2009) *Ethical, Legal and Social Issues Arising from the Use of GWAS in Medical Research*. Report for the Wellcome Trust, at p8.

17 Kaye *et al* also observe that consent to GWAS studies also does not easily fit the traditional 'individual' model of informed consent, given that samples from populations play such an important role in this research and that results from GWAS can have implications for whole communities or populations.[14] There is also scant evidence examining how participants understand how their data is used in this research.

18 Some of the parameters of the debate over consent to GWAS include: whether to obtain consent to a broad or narrow range of uses of the sample donated, whether to see consent as a 'one-off' or ongoing process (to account for subsequent uses, or new developments), how to manage data sharing (also discussed below), how to explain complex genetic concepts to participants, how to handle commercial interests and the implications of research participation for a person's family, community or wider population (including family-centred or community consent). As yet there is no consensus in the literature as to which model of consent is best for these kinds of studies.

Confidentiality and Privacy

19 Pooled sample sets used in GWAS are anonymised, but the sharing of these (which is common and indeed important for progress and publication in research) may still give rise to implications for confidentiality and privacy. Modern bioinformatics techniques can allow researchers to detect a single person's profile from pooled data representing over 1,000 DNA samples.[15,16] This does require dense SNP data from the individual from another source. However with this information it is then possible, for example, to determine whether 'Person X' provided a sample for 'Study Y'. A concern from this is that GWAS datasets could be compromised, in that 'Person X' could be associated with the characteristics of a particular dataset; which may be alcoholism or mental disorder.

20 When this finding came to light, organisations such as the US National Institutes of Health and the Wellcome Consortium removed aggregate data from the public domain.[17] Tools have also been developed to help researchers balance the need for valuable data with participant privacy.[18]

---

[14] Kaye J, Boddington P, de Vries J, Gowans H, Hawkins N, Heeney C, Melham K. (2009) *Ethical, Legal and Social Issues Arising from the Use of GWAS in Medical Research*. Report for the Wellcome Trust.

[15] Homer N, Szelinger S, Redman M, Duggan D, Tembe W, Muehling J, Pearson JV, Stephan DA, Nelson SF, Craig DW. Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays. *PLoS Genet*. 2008; 4(8):e1000167.

[16] Lumley T, Rice K. (2010) Potential for revealing individual-level information in genome-wide association studies. *JAMA* 303:659-60.

[17] NIH Background Fact Sheet on GWAS Policy Update. August 28, 2008. http://grants.nih.gov/grants/gwas/background_fact_sheet_20080828.pdf (accessed 8/04/2011)

[18] SecureGenome software: http://securegenome.icsi.berkeley.edu/securegenome/ (accessed 8/04/2011).

21 Developments such as this also highlight how future discoveries can affect existing datasets, suggesting that controls on data sharing [19] should also account for exemplary record-keeping. The increasing complexity of data and various sources from which data can be obtained have been suggested as requiring privacy safeguards that take account of the 'whole data environment' rather than a discrete data set.[20]

22 Another challenge to individual confidentiality and privacy is that it can sometimes be difficult or resource-intensive to extract individual data from a GWA study once data is pooled. This is because while it may be possible to identify an individual from pooled data (that is, confirming that Person X is present in pooled dataset Y), this does not mean that this individual's data can be easily removed from this pool. This can give rise to challenges for research participants who wish to withdraw from a project.

Justice

23 GWAS research is undertaken in a global context, as variations between populations can be very informative to gene discovery. Justice issues can arise regarding consent; for example applying a model of consent from one culture that is inappropriate in another; or determining what may constitute an undue inducement to participation in a particular population (such as the tension over providing health care in return for research participation). Justice considerations are also relevant to the outcomes of the research; that is ensuring that benefits of research are shared in a fair way in a global context.

Trust

24 Promoting trust is an obvious goal of medical research. Trust implies mutual respect between researchers and the wider community. A lack of trust on the part of potential participants may impact recruitment rates and could lead to negative publicity. A particular challenge to trust in the domains under discussion in this paper is that while trust might traditionally be considered to involve mutual openness and transparency, the practicalities of full disclosure (given the inherent uncertainties and volume of data) may impact this. The challenge, therefore, is to find a mechanism to promote trust without this becoming an unreasonable burden for the research community.

Ownership of data and the right to know (or not to know)

25 Research data obtained during a study may (with appropriate approvals in place) be shared with other researchers. The main ethical tensions in data sharing are whether consent procedures can be tailored to account for data

[19] See, for example, the in-progress *Oxford Statement on Data Sharing*: http://helexoxford.com/content/oxford-statement-data-sharing (accessed 8/04/2011)

[20] Heeney C, Hawkins N, de Vries J, Boddington P, Kaye J. (2011) Assessing the Privacy Risks of Data Sharing in Genomics. *Public Health Genomics* 14(1):17-25.

sharing[21] and whether anonymisation is an appropriate safeguard of privacy given the inherently identifiable nature of GWAS data.

26 It is common practice in large genomic studies to not return research results to participants. This is because of the logistical processes this would require (and the resultant increase in research costs), the possibility for errors given that this information was obtained under a research protocol rather than in a clinical setting, the often uncertain significance of the research results and the possibility that their interpretation will change with time (and any 'duty of re-contact' this may give rise to), the observed poor recall of patients of the results of genetic testing and the possible need for genetic counselling to mitigate any psycho-social concerns.

27 Despite this, empirical research with the lay public has suggested that research participants strongly desire to obtain individual information arising from their research participation and that such an information exchange should be ongoing.[22] Debate on this issue focuses on whether research participants are 'partners' or 'donors,' whether not returning results is paternalistic or acknowledging the altruistic act of donation, and the obligations of researchers either way. The challenge is to either prove that such return of results is inappropriate (while not eroding trust) or to derive a protocol to enable return of results within the resource and scientific limits of a study. Ravitsky and Wilfond have suggested a 'results based approach' to determining whether to disclose, which involves evaluating the nature of the data and its possible impact.[23] Yet it is uncertain how this could apply to large-scale GWAS research. Other factors such as the mental capacity of research participants are also relevant.

28 Another aspect of returning results is 'incidental findings.' Given its large scale and open-ended nature, it is almost certain that GWAS studies will identify genetic associations that were not previously anticipated. A study may also disclose information that is tangential to health, such as misattributed biological relationships (for example, paternity). The significance of these results is often uncertain at the time they are identified and further studies are often needed to replicate results in different populations. If returning results to participants does become more common than it is now, researchers may need to consider incidental findings when developing protocols for returning results, including any duty of care that may be owed. This will also have an impact on consent processes.

29 A 'right not to know' one's genetic information, while contested in the literature, does tend to be accepted in practice in both research and clinical contexts. That is, if a participant does not wish to know his or her genetic status following research participation, that wish will be respected. How this right could or should be construed in GWAS remains uncertain. For example, a research

---

21 Lunshof JE, Chadwick R, Vorhaus DB, Church GM. (2008) From genetic privacy to open consent. *Nat Rev Genet* 9(5):406-11.

22 See, for example: Murphy J, Scott J, Kaufman D, Geller G, LeRoy L, Hudson K. (2008) Public Expectations for Return of Results from Large-cohort Genetic Research. *Am J Bioeth* 8(11):36–43.

23 Ravitsky V, Wilfond B. (2006) Disclosing Individual Genetic Results to Research Participants. *Am J Bioeth* 6(6):8-17.

participant may express a desire not to know her research results but a finding may come to light (whether anticipated or incidental) that has serious health implications and for which there is a proven intervention. In this kind of scenario researchers may have to determine whether disclosure against a participant's previously expressed with can be justified.

Data linkage

Consent

30 In data linkage research, data needs to be identifiable to enable linkage. Studies with the lay public have indicated that individuals are less happy for identifiable data to be shared without their consent compared with sharing anonymous data.[24,25] Buckley *et al* also found that younger people were less willing to engage with their research and the ones who did were more wary of their medical data being used.

31 Researchers conducting large-scale epidemiological studies involving data linkage where data is identifiable are often required to obtain individual informed consent from research participants.[26] The need for this level of consent has been argued to be difficult and expensive, absorbing study resources.[27] There is also an ethical tension between the obligations of researchers to obtain informed consent from research participants and their duty to conduct cost-effective, high quality research.

32 Data linkage research is often described as "low risk",[28] particularly in comparison to interventional research.[29] This description raises the issue as to whether research being "low risk" means that it is more ethically justifiable to use forms of consent other than opt-in informed consent; or even whether it is acceptable to waive consent altogether.

33 The requirement for study-specific consent, stemming from a "narrow" conception of autonomy, has been argued as hindering epidemiological research and a broader conception, based at the level of the institution, has been proposed as an alternative.[30] This would involve balancing the risks of participation with the potential benefits of the research, and could also take account of other factors such as research participants' attitudes towards and awareness of sharing of data and administrative records between studies.

---

[24] Buckley B, Murphy AW, et al. (2011). Public attitudes to the use in research or personal health information from general parcticioners' records: a survey of the Irish general public. *Journal of Medical Ethics* 37:50-55.

[25] Robling, M. R., K. Hood, et al. (2004). Public attitudes towards the use of primary care patient record data in medical research without consent:a qualitative study. *Journal of Medical Ethics* 30:104-109.

[26] Walley, T. (2006). Using personal health information in medical research. *BMJ* 332:130-131.

[27] Walley, T. (2006). Using personal health information in medical research. *BMJ* 332:130-131.

[28] Parker, M. (2005). When is research on patient records without consent ethical?*Journal of Health Services Research & Policy* 10(3):183-186.

[29] Walley, T. (2006). Using personal health information in medical research. *BMJ* 332:130-131.

[30] Hansson M. (2010). Do we need a wider view of autonomy in epidemiological research? *BMJ* 330: 1172-1174.

Other issues arising are the extent to which research consent should be specific and whether individuals should be able to use the consent process to control the use of different aspects of their administrative data.

34 The notion of opt-out consent is also relevant here; that is, whether anyone should be able to opt out the use or linkage of all or some of their administrative information in research. Considerations to this issue include the possibility of the data pool being skewed (for example by large opt-out rates from certain cultural or ethnic groups), research into sensitive topics being hampered and the challenges to public education about the right to opt out.

Confidentiality

35 A significant issue in considerations of confidentiality in data linkage is whether data collected in a confidence should ever be allowed to be shared for research purposes. Those arguing against records being used for research base their reasoning on the link between confidentiality and respect for patient autonomy. [31] Arguably, patient expectations about confidentiality have considerable moral weight and need to be recognised in addition to consequentialist arguments emphasising potential benefits of research.[32]

36 Some studies have shown that people have concerns about "sensitive" information in their medical records being shared for research purposes, such as mental health.[33] Willison *et al* also found that their participants were more restrictive about research that linked their health information to data about occupation, income or education than they were about biological samples.[34]

37 Most research about linkage to administrative data has been with regard to medical records. Little is known about public opinion regarding expectations of confidentiality with respect to research usage of other types of administrative data.

Privacy

38 Data linkage requires personal information to identify and link an individual's records successfully. The question arising here is how this process should be done, and by whom. Nissenbaum argues that aggregation of multiple sources of information can enable a picture of an individual to be constructed. She also points out that individuals do not just have privacy concerns about personal or sensitive information, but also the "contextual integrity" of their information

---

31 Parker M. (2005) When is research on patient records without consent ethical?*Journal of Health Services Research & Policy* 10(3):183-186.

32 Parker M. (2005) When is research on patient records without consent ethical?*Journal of Health Services Research & Policy* 10(3):183-186.

33 Robling, M. R., K. Hood, et al. (2004). Public attitudes towards the use of primary care patient record data in medical research without consent:a qualitative study. *Journal of Medical Ethics* 30:104-109.

34 Willison, D. J., V. Steeves, et al. (2009). Consent for use of personal information for health research: Do people with potentially stigmatizing health conditions and the general public differ in their opinions? *BMC Medical Ethics* 10:10.

(which relates to the appropriateness of information to a specific context and an individual's control over information sharing).[35]

39 Research focussing on specific populations, social groups or communities rather than individual people raises questions of the privacy of these social groups or communities;[36] more specifically questioning how much of the information used in research is already generally available to researchers and the public.

Public interest/public good

40 Proponents of data linkage argue that the methods are cost-effective, facilitating external validity and offering potential benefits to public health at a low risk to research participants.[37, 38] Data linkage methods could enable research budgets to go further because the utilisation of pre-existing data is cost-effective, arguably less burdensome for participants and allowing greater numbers of participants to be incorporated into studies.

41 An ethical tension exists between the public health emphasis of data linkage research (which strives for knowledge and benefits at population level) and the fact that using data from individual research participants has participation risks and burdens. For research participants, individual risks and benefits exist alongside potential population gains and these benefits and burdens are incommensurable in the ways different populations and individuals are affected.

Justice

42 Research indicates that individuals from certain social groups are potentially more likely to refuse to consent to data linkage research than others. This could lead to problems with consent bias and validity of research.[39] This raises an ethical dilemma between the just distribution of risks and benefits in public health and the requirement for informed consent to protect the individual autonomy of research participants.[40]

43 Relevant to this dilemma are potential variations in the volume of data between individuals, as those individuals who make greater use of statutory services will have more information held about them. Conversely, there will also be individuals who are not integrated into these types of records at all, raising the

---

35 Nissenbaum, H. (1998). Protecting Privacy in an Information Age: The Problem of Privacy in Public. *Law and Philosophy* 17(5):559-596.

36 Glanz, K., M. C. Keigler, et al. (2009). "Ethical Issues in the Design and Conduct of Community-Based Intervention Studies." In: Coughlin, S. S., T. L Beauchamp & D. L. Weed. *Ethics and Epidemiology*, 2nd Ed. New York: Oxford University Press, pp.112

37 Walley T. (2006).Using personal health information in medical research. *BMJ* 332:130-131.

38 Buckley, B., A. W. Murphy, et al. (2007). Selection bias resulting from the requirement for prior consent in observational research: a community cohort of people with ischemic heart disease. *Heart* 93: 1116-1120.

39 Tate R, Calderwood L, et al. (2006). Mother's consent to linkage of survey data with her child's birth records in a multi-ethnic national cohort study. *International Journal of Epidemiology* 35: 294-298.

40 Huang, N., S.-F. Shih, et al. (2007). Record linkage research and informed consent: who consents? *BMC Health Services Research* 7(18).

prospect that they could fail to benefit from the research. Considerations of existing health inequalities will also influence the distribution of the benefits of epidemiological research.

Trust

44 There seems to be less trust in confidentiality being maintained in research involving the use of medical records where there was a lack of understanding of research processes and fear of personal sensitive information being obtained by insurance companies and pharmaceutical companies.[41] A recent Irish study with the lay public found that 90% trusted their GP's to keep their health records confidentially and securely, although 67.5% were unwilling to allow their GP's to decide when researchers could access identifiable information about them. [42] Another earlier study has found that people had greater expectations of confidentiality and felt more in control within primary care than they were in other domains; however individuals were again concerned about the unauthorised access of information (especially sensitive information) by insurance and pharmaceutical companies.[43] This study also observed a lack of understanding of research processes and the authors suggest that raising public awareness of research could reduce anxieties.[44] Both studies found that anonymous data sharing was far more acceptable to individuals than identifiable data. Similar trust issues arise in the context of health records research and are discussed briefly below.

Ownership of information/right to know

45 The main questions regarding ownership are who administrative data belongs to and who should decide whether or not to share information. A more fundamental question is the extent to which individuals should be able to be involved in decision-making about data sharing. A further issue is whether individuals should be informed when administrative or study data about individuals is shared without their consent.

Health Records Research

Consent

[41] Robling, M. R., K. Hood, et al. (2004). Public attitudes towards the use of primary care patient record data in medical research without consent:a qualitative study. *Journal of Medical Ethics* 30: 104-109.

[42] Buckley, B., A. W. Murphy, et al. (2011). Public attitudes to the use in research or personal health information from general parcticioners' records: a survey of the Irish general public. *Journal of Medical Ethics* 37: 50-55.

[43] Robling, M. R., K. Hood, et al. (2004). Public attitudes towards the use of primary care patient record data in medical research without consent:a qualitative study. *Journal of Medical Ethics* 30: 104-109.

[44] Robling, M. R., K. Hood, et al. (2004). Public attitudes towards the use of primary care patient record data in medical research without consent:a qualitative study. *Journal of Medical Ethics* 30: 104-109.

46 Obtaining explicit informed consent to access medical records from all potential participants in large-scale epidemiological studies is considered by researchers to be beyond the scope of the funding of most studies, particularly where retrospective cases are being reviewed. The perspective of researchers is also that the richness of the data obtained from a set of life-long medical health records can be scientifically very relevant in answering important health related questions.

47 There is some debate as to whether participation in health records research should be based on 'opt-out' system. This would mean that instead of asking potential research participants whether they would like to give explicit and informed consent for the use of their data (an 'opt-in' approach), the use of data is allowed unless the participants have made an explicit statement that they do not give consent to such use. The ethical tension here is between the potential of coercion in the opt-out approach and the risk of obtaining only skewed data due to sample bias via an opt-in approach.

Confidentiality

48 Data sharing and linkage raise many concerns in relation to confidentiality. This is because even with explicit informed consent in place, it may not be possible to inform potential participants of how their data will be handled and for what outcomes, once they are linked and shared.

49 Currently, examining medical records without consent is perceived as an unacceptable breach of confidentiality. For example, a study of public attitudes towards the use of primary care records for medical research without consent showed that participants felt more anxious about sensitive data if research was conducted by third parties.[45] Another study found that research participants prefer to be informed about data collection, both because this can be seen as a courtesy to them but also to enable them to opt out. Failing to inform them was perceived as harmful as it removed their personal choice and demonstrated a lack of respect for their human rights.[46]

50 The concept of 'minimal risk' is relevant when considering the need for confidentiality. A 2006 study found that 72% of the British public did not consider the confidential use of personal, identifiable information by the national cancer registry for the purposes of research to be an invasion of privacy and 81% of the public would support statutory cancer registration.[47] Another study

[45] Robling, M. R., K. Hood, et al. (2004). Public attitudes towards the use of primary care patient record data in medical research without consent:a qualitative study. *Journal of Medical Ethics* 30: 104-109, cited by Noble 2009

[46] Noble S, Donovan J, Turner E,Metcalfe C, Lane A, Rowlands M-A, Neal D, Hamdy F, Ben-Shlomo Y,Martin R (2009) Feasibility and cost of obtaining informed consent for essential review of medical records in large-scale health services research. *J Health Serv Res Policy* 14:77-81.

[47] Barrett G, Cassell JA, Peacock JL, *et al*. (2006) National survey of British public's views on use of identifiable medical data by the National Cancer Registry. *BMJ*;**332**:1068–72, cited by Noble S, Donovan J, Turner E, Metcalfe C, Lane A, Rowlands MA, Neal D, Hamdy F, Ben-Shlomo Y, Martin R. (2009) Feasibility and cost of obtaining informed consent for essential review of medical records in large-scale health services research. *J Health Serv Res Policy*. Apr;14(2):77-81.

found evidence to support the claim that consent should not be required for accessing medical records for non-commercial research with no effect on the group whose records were being accessed so long as the research has been approved by a Research Ethics Committee.[48]

The public interest and the public good

51 As with GWAS and data linkage research, the debate in health records research is about how to balance the tensions between protection of the individual's rights and the public good. This debate is about the moral importance of a collective approach to health research and the relevance of the public perception of the benefits of participation in this research; both to them individually and collectively as a population.

52 Current UK requirements to obtain informed consent when accessing medical records for research purposes have been reviewed in relation to selection bias in observational studies. Kho et al undertook a systematic review of observational studies.[49] They concluded that significant differences existed between participants and non-participants, which could threaten the validity of results from these types of studies when consent is required.

Justice and solidarity

53 One can consider the nature of the health care system in the UK from the perspective of social values like solidarity. The ethical question is whether this means that, if we expect benefit through participating in a system like the NHS, there might be legitimate expectations or even obligations to participate in various ways in that system.

Trust

54 Patients' trust in the researchers was the most powerful determinant of the kind of control they want over their medical records in a study by Damschroder et al.[50] They highlighted that patients with the lowest trust in researchers were more likely to recommend a more stringent process for obtaining individual consent.

55 In June 2006 the Medical Research Council (MRC) commissioned Ipsos MORI to examine public attitudes to and awareness of the use of personal health

---

48 Noble S, Donovan J, Turner E, Metcalfe C, Lane A, Rowlands MA, Neal D, Hamdy F, Ben-Shlomo Y, Martin R. (2009) Feasibility and cost of obtaining informed consent for essential review of medical records in large-scale health services research. *J Health Serv Res Policy*. Apr;14(2):77-81.

49 Kho ME, Duffett M, Willison DJ, Cook DJ, Brouwers MC. (2009) Written informed consent and selection bias in observational studies using medical records: systematic review. *BMJ*. 12;338:b866.

50 Damschroder LJ, Pritts JL, Neblo MA, Kalarickal RJ, Creswell JW, Hayward RA. (2007) Patients, privacy and trust: patients' willingness to allow researchers to access their medical records. *Soc Sci Med* 64(1):223-35.

information in research.[51] The report stated that the vast majority (87%) trust GPs to have access to their personal health information and over half trust other health professionals – such as consultants or hospital doctors (59%). Medical researchers working in the public sector i.e. for Government and universities (both trusted by 11%) were stated as being more trusted in than their counterparts working for private companies (4%)

Current oversight and regulation

56 The legal framework controlling access to patient data in England and Wales is based on statutory legislation, common law decisions (although there are no known UK decisions that deal with anonymity and identifiablity in research)[52] and various EU Directives. Despite multiple sources of guidance, not one body is responsible for overseeing decisions relating to the use or management of patient data for research use or for managing data arising from GWAS.

57 The Data Protection Act applies to all personal data and, relevant to this discussion paper, sensitive personal data (which attracts an additional layer of protection). Administrative data that is anonymous or not personal/sensitive (for example by being pooled or aggregated) does not fall under the ambit of the Data Protection Act; however data linkage requires use of identifiers to "link" the data, rendering it identifiable. Section 33 of the Data Protection Act 1998 is an exemption allowing researchers to utilise personal data for research without consent of the individual, subject to the following conditions, as set out below.[53]

58 The Administrative Data Liaison Service (ADLS) is funded by the Economic and Social Research Council (ESRC) to support research using administrative data in the UK and is managed by St Andrews University and the Universities of Oxford and Manchester. The ADLS outline that under Section 33 of the Data Protection Act, personal data can be utilised for research purposes without the consent of the individual concerned if: 1) the data is not processed in a way that would "support measures or decisions affecting particular individuals" and 2) the processing does not cause or is unlikely to cause any data subject "substantial damage or distress. The data can then be further processed and used for purposes other than originally intended and collected for, kept indefinitely and data is exempt from rights of access by the "data subject" if it does not identify or cause harm to them. However, researchers still need to be complicit with remaining data protection principles.

59 The National Information Governance Board for health and social care (NIGB) is an "independent statutory body established to promote, improve and monitor information governance in health and adult social care".[54] The NIGB "protects

---

51 MRC/IPSOS Mori (2007) Keeping it Confidential. The Use of Personal Health Information in Medical Research. General Public Consultation.

52 Curren L, Boddington P, Gowans H, Hawkins N, Kanellopoulou N, Kaye J, and Melham K (2010) Identifiability, genomics and U.K. data protection law. Eur J Health Law, 17(4):329-44

53 Administrative Data Liaison Service (ADLS) http://www.adls.ac.uk/important-guidance/ (accessed 10/03/2011)

54 National Information Governance Board for Health and Social Care (NIGB) http://www.nigb.nhs.uk/ (accessed 07/04/2011).

the interests of individuals data which is stored, shared and used in the NHS and adult social care" and also "advises on the use of powers under Section 251 of the NHS Act 2006 to permit the duty of confidentiality to be set aside, where other legal routes are not available" such as researchers seeking to utilise patient information (ibid). The Ethics and Confidentiality Committee (ECC) at the NIGB who deal with applications for support under section 251 only have jurisdiction in England and Wales.

60 Section 251 of the NHS Act 2006 "allows the common law duty of confidentiality to be set aside in specific circumstances where anonymised information is not sufficient and where patient consent is not practicable. For example a research study may require access to patient identifiable data to allow linkages between different datasets where the cohort is too large for consent. This would require time limited access to identifiable information where gaining consent from a large retrospective cohort would not be feasible and would require more identifiable data than would be necessary for linkage purposes".[55] Applications for approval to use section 251 are dealt with by the ECC (Ethics and Confidentiality Committee) and are made by assessing whether the public benefits are significant enough to dismiss confidentiality.

61 The current regulatory and governance pathway requires application to the Integrated Research Application System (a single system for applying for the permissions and approvals for health and social care / community care research in the UK). Then a series of assessments are undertaken both nationally and locally. Following the UK-wide single ethics opinion of the National Research Ethics Service, access to patient data is permitted by either the Caldicott guardian (at the level of a health organisation such as a NHS Trust) or the Ethics and Confidentiality Committee (or both). Finally NHS R&D permissions are sought from each NHS Trust where the research takes place.

62 Article 8 of the The Human Rights Act 1998 deals with the right for respect to private and family life, home and correspondence could be relevant in relation to data linkage research. To meet human rights requirements, conformity with the Data Protection Act 1998 and the common law of confidentiality is required but not always sufficient by itself.[56]

63 Common law relevant to data linkage research is the common-law duty of confidentiality whereby a breach would constitute unauthorised sharing of information that was thought to be given in confidence. Exceptions to the duty of confidentiality, whereby confidential information can be disclosed are for public interest (where the public interest of disclosure must outweigh the public interest of maintaining confidentiality, Information Commissioner's Office ICO) or where allowed or required by statute or court order (NHS Wales).[57]

---

55 National Information Governance Board for Health and Social Care (NIGB) http://www.nigb.nhs.uk/ (accessed 07/04/2011).

56 Thomas, R. Walport, M. 2008. *Data Sharing Review Report*. London: Ministry of Justice. http://www.justice.gov.uk/docs/data-sharing-review-report.pdf (accessed 21/04/2011).

57 Thomas, R. Walport, M. 2008. *Data Sharing Review Report*. London: Ministry of Justice. http://www.justice.gov.uk/docs/data-sharing-review-report.pdf (accessed 21/04/2011).

64 In the context of GWAS, Curren et al have recently suggested that the UK Data Protection Act may offer greater coverage of research data than previously thought.[58]

65 The Medical Research Council (MRC) and Wellcome Trust both state a commitment to high quality ethical research, encouraging national and international data-sharing between research studies through data preservation and sharing strategies and exploitation developments in information and computer technologies.[59] The Joint Data Standards Study, outlines that original consent for data use and confidentiality by research participants must be respected when seeking to re-use data.[60]

66 The World Medical Association Declaration of Helsinki states that the interests of research participants should take precedence over all other interests.[61] Researchers should do their utmost to protect the privacy of participants and informed consent should be gained. However, section 25 outlines that research using identifiable data where obtaining consent would be "impossible or impractical" or would threaten the validity of the research, can be conducted only after consideration an approval by a research ethics committee.

Questions of ethics and policy that might be addressed

- What model(s) of consent will best uphold the various interests arising in these kinds of research? Relevant factors might include the parameters of: research complexity and scale, cultural context, future uses and privacy?

- Should participants in large genomic studies be allowed to have access to individual data found during the course of the research?

- Should cultural values and differences be considered when applying models of consent and confidentiality in genomic and epidemiological studies?

- How can the privacy of specific populations in data linkage research be protected?

- Can individuals make a claim to ownership of their individual samples and medical records? If so, should they have ultimate say on what can and cannot be accessed?

- Should participation in health records research be based on an opt-out or an opt-in system?

---

58 Curren L, Boddington P, Gowans H, Hawkins N, Kanellopoulou N, Kaye J, and Melham K (2010) Identifiability, genomics and U.K. data protection law. Eur J Health Law, 17(4):329-44

59 MRC, Medical Research Council.

60 UK Research Councils (2005). *Joint Data Standards Study*. Available at: http://www.mrc.ac.uk/Utilities/Documentrecord/index.htm?d=MRC002552 (accessed 21/04/2011).

61 World Medical Association. *Declaration of Helsinki* (2008) http://www.wma.net/en/30publications/10policies/b3/17c.pdf (accessed 07/04/2011)

- How should we define ‘minimal risk’ in the context of epidemiological research? If breaching of confidentiality of individual health data is harmful to the individual, does that mean that such a breach should never be allowed? Or should we allow the use of some data without consent if there is only minimal harm?

- To what extent should we allow public values about consent and confidentiality of individual data in medical research to determine policy-making on the accessibility of these data?

- As there are inconsistencies between the various legal frameworks and professional regulations, both nationally and internationally, how can we reach a more consistent framework that allows both further progress of medical research and an acceptable protection of the rights of individuals?

- How should all of these considerations be handled in the context of global health, specifically health inequalities?

References

Academy of Medical Sciences (2006) *Personal data for public good: using health information in medical research*. London: Academy of Medical Sciences. http://www.acmedsci.ac.uk/download.php?file=/images/publication/Personal.pdf (accessed 21/04/2011).

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

Buckley, B., A. W. Murphy, et al. (2007). Selection bias resulting from the requirement for prior consent in observational research: a community cohort of people with ischemic heart disease. *Heart* 93: 1116-1120.

Buckley B, Murphy AW, et al. (2011). Public attitudes to the use in research or personal health information from general parcticioners’ records: a survey of the Irish general public. *Journal of Medical Ethics* 37:50-55.

Caulfield, T., McGuire, A. L. *et al*. Research ethics recommendations for whole-genome research: Consensus statement. *PLoS Biol* 2008;6:e73.

Curren L, Boddington P, Gowans H, Hawkins N, Kanellopoulou N, Kaye J, and Melham K (2010) Identifiability, genomics and U.K. data protection law. Eur J Health Law, 17(4):329-44.

Damschroder LJ, Pritts JL, Neblo MA, Kalarickal RJ, Creswell JW, Hayward RA. (2007) Patients, privacy and trust: patients' willingness to allow researchers to access their medical records. *Soc Sci Med* 64(1):223-35.

Donnelly P. (2008) Progress and challenges in genome-wide association studies in humans. *Nature*. 456(7223):728-31.

Glanz, K., M. C. Keigler, et al. (2009). “Ethical Issues in the Design and Conduct of Community-Based Intervention Studies.” In: Coughlin, S. S., T. L Beauchamp & D.

L. Weed. *Ethics and Epidemiology*, 2nd Ed. New York: Oxford University Press, pp.112

Hansson M. (2010). Do we need a wider view of autonomy in epidemiological research? *BMJ* 330: 1172-1174.

Heeney C, Hawkins N, de Vries J, Boddington P, Kaye J. (2011) Assessing the Privacy Risks of Data Sharing in Genomics. *Public Health Genomics* 14(1):17-25.

Homer N, Szelinger S, Redman M, Duggan D, Tembe W, Muehling J, Pearson JV, Stephan DA, Nelson SF, Craig DW. Resolving individuals contributing trace amounts of DNA to highly complex mixtures using high-density SNP genotyping microarrays. *PLoS Genet*. 2008; 4(8):e1000167.

Huang N, Shih S-F, et al. (2007) Record linkage research and informed consent: who consents? *BMC Health Services Research* 7(18).

Iversen A, Lidell K, Fear N, Hotopf M, Wessely S. (2006) Consent, confidentiality and the Data Protection Act. *BMJ* 332:165-169.

Kaye J, Boddington P, de Vries J, Gowans H, Hawkins N, Heeney C, Melham K. (2009) *Ethical, Legal and Social Issues Arising from the Use of GWAS in Medical Research*. Report for the Wellcome Trust.

Kho ME, Duffett M, Willison DJ, Cook DJ, Brouwers MC. (2009) Written informed consent and selection bias in observational studies using medical records: systematic review. *BMJ*. 12;338:b866.

Lumley T, Rice K. (2010) Potential for revealing individual-level information in genome-wide association studies. *JAMA* 303:659-60.

Lunshof JE, Chadwick R, Vorhaus DB, Church GM. (2008) From genetic privacy to open consent. *Nat Rev Genet* 9(5):406-11.

Manson, N. C. and O. O'Neill (2007). *Rethinking Informed Consent in Bioethics*. Cambridge, Cambridge University Press.

McCarthy MI, Abecasis GR, Cardon LR, et al. (2008) Genome-wide association studies for complex traits: consensus, uncertainties, and challenge. *Nature Reviews Genetics* 9:356-369.

Metcalfe C, Martin RM, Noble S, Lane JA, Hamdy FC, Neal DE, Donovan J. (2008) Low risk research using routinely collected identifiable health information without informed consent: encounters with the Patient Information Advisory Group. *J Med Ethics* 34:37-40.

MRC/IPSOS Mori (2007) Keeping it Confidential. The Use of Personal Health Information in Medical Research. General Public Consultation.

Murphy J, Scott J, Kaufman D, Geller G, LeRoy L, Hudson K. (2008) Public Expectations for Return of Results from Large-cohort Genetic Research. *Am J Bioeth* 8(11):36–43.

Nissenbaum, H. (1998). Protecting Privacy in an Information Age: The Problem of Privacy in Public. *Law and Philosophy* 17(5):559-596.

Noble S, Donovan J, Turner E,Metcalfe C, Lane A, Rowlands M-A, Neal D, Hamdy F, Ben-Shlomo Y,Martin R (2009) Feasability and cost of obtaining informed consent

for essential review of medical records in large-scale health services research. *J Health Serv Res Policy* 14,77-81.

Parker M. (2005). When is research on patient records without consent ethical?*Journal of Health Services Research & Policy* 10(3):183-186.

Pearson TA; Manolio, TA. (2008) How to Interpret a Genome-wide Association Study, *JAMA* 299(11):1335-1344.

Ravitsky V, Wilfond B. (2006) Disclosing Individual Genetic Results to Research Participants. *Am J Bioeth* 6(6):8-17.

Robling MR, Hood K, et al. (2004). Public attitudes towards the use of primary care patient record data in medical research without consent:a qualitative study. *Journal of Medical Ethics* 30:104-109.

Tate R, Calderwood L, et al. (2006) Mother's consent to linkage of survey data with her child's birth records in a multi-ethnic national cohort study. *International Journal of Epidemiology* 35:294-298.

Thomas, R. Walport, M. 2008. *Data Sharing Review Report*. London: Ministry of Justice. http://www.justice.gov.uk/docs/data-sharing-review-report.pdf (accessed 21/04/2011).

UK Research Councils (2005). *Joint Data Standards Study*. Available at: http://www.mrc.ac.uk/Utilities/Documentrecord/index.htm?d=MRC002552 (accessed 21/04/2011).

Vandenbroucke JP. (1998) Maintaining privacy and the health of the public. Should not be seen as in opposition. *BMJ* 316:1331-2.

Walley, T. (2006). Using personal health information in medical research. *BMJ* 332:130-131.

Willison, D. J., V. Steeves, et al. (2009). Consent for use of personal information for health research: Do people with potentially stigmatizing health conditions and the general public differ in their opinions? *BMC Medical Ethics* 10:10.

World Medical Association. *Declaration of Helsinki* (2008) http://www.wma.net/en/30publications/10policies/b3/17c.pdf (accessed 07/04/2011).