Clause 67, Medical Research and Privacy: the Options for the NHS

Ross Anderson[1], Rudolf Hanka[2], and Alan Hassey[3]

[1] University of Cambridge Computer Laboratory,
New Museums Site, Pembroke Street, Cambridge CB2 3QG, UK
Ross.Anderson@cl.cam.ac.uk
[2] University of Cambridge Medical Informatics Unit
Institute of Public Health, University Forvie Site
Robinson Way, Cambridge CB2 2SR
hanka@medschl.cam.ac.uk
[3] Fisher Medical Centre
Millfields, Skipton, N Yorks BD23 1LP
alan.hassey@btinternet.com

1 Executive Summary

Over the least few years there has emerged a consensus on the legitimate research uses of medical records that balances patient privacy, professional autonomy, public health effectiveness, and the needs of science. This consensus is becoming widely accepted, with small local variations, in most developed countries. Now the Department of Health has introduced legislation that will disturb this balance and give patients in Britain very much less protection than elsewhere. This may be thought helpful in making medical research more competitive, but is almost certainly misguided. By missing the opportunity to modernise UK research databases, undermining international collaboration, condemning British researchers to work with low-quality data, and risking adverse public reaction that will further undermine the standing of research, the proposed clause 67 is not merely a setback for patient rights; it could be extremely damaging to the research community.

2 Introduction

By the early 1990s, medics and healthcare IT people, in Britain and elsewhere, had become aware of a growing gap between what the public get told about records-based research and what actually happens. The public by and large believed that their medical records were only shared by the GPs, hospitals and other care providers directly concerned with treating them, while in reality more and more information was being harvested for a growing range of research databases.

By late 1996, it was realised that these data collection exercised not only tested the limits of medical ethics, but in some cases might contravene the criminal law. The obvious example is the collection of data on HIV and AIDS: in

Britain, the Venereal Diseases Act of 1917 prohibits the sharing of information about sexually transmitted diseases without the patient's consent, yet this was ignored by the Public Health Laboratory Service and others in the rush to track the progress of the epidemic. Similar issues arise with human fertilisation and mental health.

Exactly the same problems arose in other developed countries. In 1996, the BMA sponsored a conference on personal medical information in Cambridge at which medical researchers and informatics people shared their experiences. The picture that emerged was that, under pressure from public opinion and data protection law, medical research databases had started either to use data from volunteers or to de-identify the data collected from clinical management systems.

- In Germany, following reunification, it was found that the former East Germany had an extensive cancer registry that was of considerable research potential but which had almost no privacy protection. The database was given a temporary exemption from data protection law while the necessary de-identification mechanisms and access controls were fitted [5].
- In New Zealand, there is a central research database with copies of all the nation's medical records, but with personal identifiers such as names and addresses removed. This can be accessed by approved researchers, but through software that only answers a query if the answer is drawn from the data in six or more peoples' records [11]. Denmark has something similar.
- The US Healthcare Finance Administration, which administers Medicare, has 'beneficiary-encrypted' databases of records whose personal identifiers have been encrypted, and which are available to approved researchers; there are public-access databases that have been further scrubbed to remove even circumstantial data that might identify a patient. There is extensive public debate about privacy, and the protection mechanisms are subject to public audit [8].
- There are also transnational disease registers. An example is Diabcare, used to monitor the quality of care for diabetics in Europe. The data collected by this system is de-identified in such as way that hospitals can see their relative performance compared to national averages for various outcome measures, but can see no data about other hospitals' patients (or even the performance of other identifiable hospitals) [7].
- Some transnational research programmes involve getting the consent of patients and their families to the collection of very detailed data. An example is given by amyotrophic lateral sclerosis (ALS), for which a number of countyries share research data. Even here, the identities of patients are not visible outside their country of residence. The effect of this combination of patient consent and privacy protection was a research network, whose members included doctors from both Serbia and the USA, continued to function even while the USAF was bombing Belgrade.

The point is that medical researchers, privacy advocates and governments worldwide have, over the last 5–10 years evolved a working compromise on the

use of personal health information in medical research, and this has taken much the same shape in most countries.

1. There are some databases compiled with the active consent and assistance of the patients; as the data can be comprehensive and of high quality, these are generally the most valuable.
2. Then there are also some databases containing data that have been largely or wholly anonymised; being abstracted automatically from operational data, their quality can be highly variable, but they can be useful in some public health applications.

The protection technologies involved are relatively mature, having evolved to meet the requirements of national census bureaux in the 1970s, and are well described in textbooks on security engineering (e.g., [4]). This compromise is also being implemented in the UK private sector. In 1996, the BMA had reached agreement with a number of health data companies to process data only in approved and ethical ways [2]. Thus, for example, when a company called Source Informatics (now IMS) wanted to build a system to collect prescription data from pharmacies for commercial resale, they had the system vetted by a BMA expert; significant changes were made to their original design to ensure that neither doctors nor patients could be identified against their will in the collected data [10].

3 Options for the NHS

The obvious way forward for the NHS would be to fall in line with the rest of the world: to clean up the existing medical research databases by either moving them to a basis of informed consent, or by removing personal identifiers and adding such further access controls as are required to prevent patients being identified.

Instead, the Department of Health proposes, in Clause 67 of the Health and Social Care Bill currently before parliament, to grant the Secretary of State very sweeping powers to legitimise and entrench the existing systems.

This is unlikely to gain much support from practising doctors, since law and ethics are largely separate systems; legalising an unethical practice does not make it ethical. It will merely put doctors (and in particular GPs) in a vice between patient expectations of privacy on the one hand, and civil service demands for data on the other. The most likely reaction of GPs could be to switch off their connections to the NHS network, in order to prevent their data flows being "collected by the secretary of state". This could undo much of the effort invested in building patient and professional confidence in the application of e-commerce techniques to healthcare. Overall, Clause 67 threatens to undermine much of the "New NHS"/"NHS Plan" agenda by discouraging GPs from collecting routine health data for research, audit, health needs assessment and clinical governance activities.

From the point of view of the research community, there are some specific risks and costs of the proposed legislation.

3.1 Opportunity costs

One direct cost will be the loss of an opportunity to modernise Britain's record-based research. Existing disease register systems are very diverse, and although some are well-designed and competently operated, many are ad-hoc affairs run on a shoestring. The government's belated recognition of the potential value of these systems is most welcome; but the Secretary of State should make the funds available to modernise the systems rather than simply legitimising the current ramshackle arrangements by ukase and passing on to the next firefighting task.

The action of the Medical Research Council in issuing guidelines on privacy and consent, as well as that of the General Medical Council and the Information Commissioner in giving researchers a year's grace before tightening up ethical enforcement, were a welcome push in the right direction. It is regrettable that the Department saw fit to undermine them by introducing Clause 67.

While it may be argued that some disease registers will take more than a year to reimplement, we feel it important that any derogation from data protection law and ethical medical practice should be strictly time-limited. This is what happened in Germany and Switzerland [12], and is now happening in the USA. A three year grace period should be ample; more than that, and the necessary work will simply be postponed.

3.2 International issues

More and more research is international; the examples mentioned above, from Diabcare to the ALS register, are quickly becoming the norm rather than the exception. If Britain becomes the odd man out on medical data protection – in effect, a 'data haven' where third world practice is tolerated – then British researchers risk being frozen out of these networks.

Ostracism has already affected researchers in Iceland, whose government went ahead with a national genetic database over the protests of local medics and researchers [3]. The power that Clause 67 grants the Secretary of State to seize any medical data within the jurisdiction will have a similar effect here. For example, participants in the ALS network may be unwilling to share data with researchers in Britain if it may be impounded at any time. European information and privacy commissioners roundly denounced the Iceland database [6]; abolutely the last thing UK research needs is a similar contretemps. (In fact, data protection laws in countries such as Germany will probably make it illegal for doctors there to share data with British colleagues.)

There will be other direct effects. For example, international medical journals may be expected to ask detailed ethical questions before agreeing to publish work from the UK. There will also be less tangible but no less damaging indirect effects from loss of reputation and standing.

3.3 Data quality issues

Data collected from operational medical record systems tend to be fragmentary, conflicting and of highly variable quality. It is common to find the same heart

patient described by three different physicians as 'This 53-year old gentleman who claims to be almost teetotal but suffers high levels of work-related stress', 'This ex-smoker who drinks 40 units of alcohol a week' and 'this obese sedentary worker'.

For most research purposes, the goal is to have a sufficient quantity of high-quality information rather than a mass of low-quality, noisy data that was collected for a different purpose. Quality is more important than quantity, and this means acquiring not just the trust but also the cooperation of patients.

3.4 Threats to existing research

The Primary Care Information Services (PRIMIS) project was designed to help primary care organisations improve patient care through the effective use of their clinical computer systems. PRIMIS uses a technology that allows GPs to interrogate their practice clinical databases. These tools help assure the reliability of health data, and also facilitate the analysis of aggregated anonymised datasets. They are built into the current generation of GP clinical systems and provides a powerful tool to support research.

Much effort and expense has been directed at gaining patient and professional confidence in the ability to use these tools without jeopardising patient trust and confidentiality. Clause 67 threatens this activity and GPs may decline to share these datasets in future. Although the Secretary of State will have powers to compel, the use of these powers would have extremely grave effects on medical morale and public trust. The casualties would not just be research, but audit, clinical governance, performance monitoring and health needs assessment.

3.5 Public reaction issues

The recent Alder Hey scandal shows how public opinion has changed in the last decade, and the Secretary of State is to be commended for declaring that the days of the old, paternalistic NHS are over. But patient consent extends to medical data as much as to medical samples, and many public opinion surveys show that patients' views on privacy are even more robust than doctors' views. Patients do not want their personal health information shared, and the further the sharing is away from the clinicians who are treating them, the less they like it [9].

The risks of disregarding public opinion become clear when one considers the controversy over experimental animals. In the 1960s and 1970s, laboratory animals were often mistreated to an extent that would be illegal today, and people who complained were derided as cranks. The protest movement grew and grew, until nowadays it poses a direct and present threat to the personal safety of researchers and the ability of universities to raise charitable funding. This mistake simply must not be repeated.

In the shorter term, there is the direct risk that Clause 67 will make it more difficult to recruit volunteers for various kinds of research projects. This could significantly push up the cost of doing certain types of research.

3.6 Political issues

Although technical mechanisms such as de-identified research databases may deal with 80% of the problem, the residual 20% is going to depend on political compromise. The risk here is that both confidentiality risks, and research benefits, are often presented out of context. There is no effective public consultation about the tradeoffs, and side issues – such as easier personal access to personal records – tend to be ignored.

When presenting the benefit to society of electronic medical records and of professional access to epidemiological data, those who are, or should be, consulting with the public would do well to present a more comprehensive and balanced picture of the key risks and benefits. The Government seems to underestimate the importance of gaining public ownership of the optimal middle ground. This will leave the debate open to capture by extremists from both sides.

Acknowledgement: We are grateful for constructive input from Dr Iain Buchan.

References

1. *'Personal Medical Information – Security, Engineering and Ethics'*, RJ Anderson (editor), Springer-Verlag (1997) ISBN 3-540-63244-1
2. "An Update on the BMA Security Policy", RJ Anderson, in [1] pp 233–250; `http://www.cl.cam.ac.uk/users/rja14/bmaupdate/bmaupdate.html`
3. "The DeCODE Proposal for an Icelandic Health Database", RJ Anderson, in The Icelandic Medical Journal v 84 no 11 (Nov 98) pp 874–5; `http://www.cl.cam.ac.uk/users/rja14/#Med`
4. *'Security Engineering – a Guide to Building Dependable Distributed Systems'*, RJ Anderson, Wiley (2001) ISBN 0-471-38922-6
5. "Clinical record Systems in Oncology. Experiences and Developments on Cancer Registers in Eastern Germany", B Blobel, in [1] pp 39–56; also at `http://www.cl.cam.ac.uk/ftp/users/rja14/blobel.pdf`
6. Two statements, made by the Data Protection Commissioners of EU and EES countries and Switzerland, 20th International Conference on Data Protection, Santiago de Compostela, 16-18 September 1998; available at `http://www.dataprotection.gov.uk/20dpcom.html`
7. `http://www.diabcare.de`
8. *'Medicare – Improvements Needed to Enhance Protection of Confidential Health Information'*, General Accounting Office, USA, GAO/HEHS-99-140; `http://www.gao.gov/AIndexFY99/abstracts/he99140.htm`
9. "Clinical Systems Security – Implementing the BMA Policy and Guidelines", A Hassey, M Wells, in [1] pp 79–94
10. "Protecting the identity of doctors in drug prescription analysis", V Matyáš, in *Health Informatics Journal* v 4 nos 3–4 (Dec 1998) pp 205–209
11. "Managing Health Data Privacy and Security", R Neame, in [1] pp 225–232
12. "Datenfluesse im Gesundheitswesen", M Schnyder, in *in Symposium für Datenschutz und Informationssicherheit*, Zürich, Oct 98