Consultation on the linking and use of biological and health data

Submission by Patrick Finlay, pfinlay@medimaton.com

This note addresses question 2 of the Nuffield consultation document, namely "What are the new privacy issues". The opinions expressed here are in a personal capacity and do not necessarily reflect the views of any organisation with which I am connected.

Summary:
There is a conflict between the public benefits of sharing biological and health data and an individual's right to privacy. Analysis of the relative merits indicates that patients should have the right to data confidentiality. Three types of data use are recognised:

- Personal Care: it is reasonable to assume that a patient implicitly consents to the disclosure of data needed for their own treatment, on a need-to-know basis
- Medical Research: It is reasonable that biological and health data can only be used in medical research if (a) the patient has given formal written consent and (b) the researcher has ethical committee agreement for use of the specific personal information involved. It is not practical to restrict such data once published, nor can anonymity be guaranteed. Therefore it is preferable that a patient gives generic consent for their data to be used for medical research, similar to the organ donor system. This consent can then be shown on their records.
- Commercial: Commercial companies can have legitimate reasons for needing biological and health data. This can be made available to them on agreed commercial terms by willing patients, who would sell their details to an approved database company.

Discussion and Rationale

1. The privacy of all types of personal data is a sensitive issue, as illustrated by recent disclosures following the actions of Edward Snowden. These revealed widespread harvesting of private communications and personal data by state authorities in several countries. There is a public perception that ordinary people are routinely being spied on by government and commercial organisations, and there is a general sense of insecurity that this personal data may somehow be used to the detriment of the individual, either deliberately or through unintended consequence.

2. Even if the secret harvesting of someone's personal data has no negative consequences, many people still resent the idea that they are being covertly monitored and analysed. There is an argument that "only the guilty have reason to hide", but this itself is offensive to many – it implies that any desire for privacy must be the result of some shameful secret. The perceived wisdom used to be that personal privacy was a value for older people, whilst the younger social media generation were happy to share their personal details widely. But current impressions are that the social media-savvy are actually quite careful with what they choose to reveal on Facebook or Twitter.

Many choose not to register or log in to social media sites that enable their identity to be tracked and cross-linked

3. Conversely, many people welcome the benefits of the "personal touch" and targeted messages that result from sharing personal information on the internet. Whether the question is supermarket purchases, career choices or finding a partner, people appreciate a filter that presents only compatible candidates.

4. For personal biological and health data, there is a wide range of additional concerns. There are some issues that arise from sharing personal data and others that arise from not sharing them. I have listed some of these in the Appendix under various headings.

5. In general, there is a conflict between the individual's right to privacy of biological data and society's need to know in order to provide public benefits. Comparing the relative benefits and costs shown in the Appendix indicates that the balance is heavily weighted towards the individual's rights to privacy. In other words, the potential benefit to society of disclosing someone's biological data is outweighed by the potential disbenefits to the individual. Therefore as a general principle **personal biological data should be considered private by default**, and should not be disclosed unless the individual has given consent.

6. It is normal in clinical trials and similar studies for data to be anonymised. However, it has repeatedly been shown that anonymisation is not reliable. Personal identity can be inadvertently revealed for example when two data sets are merged. And for a determined hacker it is not usually difficult to trace an identity. Whereas it is clearly good practice to anonymise data to the most rigorous level, it is unrealistic to expect this will be watertight. **If someone is asked to take part in a trial in which their data will be anonymised, it should be explicitly made clear that this process cannot be guaranteed.**

7. It is possible to distinguish **three types of disclosure** of data that an individual might choose to make:

    a. <u>Personal Care</u> personal health and biology data required for clinical diagnosis and treatment of my own condition, made available to those who need to know.

    b. <u>Medical Research</u> Personal health and biology data that is useful for medical research, whether for a specific project or more generally

    c. <u>Commercial</u> Personal health and biology data made available for commercial purposes on commercial terms.

8. In <u>Personal Care</u> It is reasonable to assume **implicit consent by a patient for health professionals to access the individual health and biology data they need for providing that individual's care**. This is generally the current

position in the NHS. There are loopholes however. Clinical papers routinely show images of specific patients and/or detailed medical and social personal data. This is contrary to the principle of "need to know". There are obvious scenarios where a reader of a journal may recognise an "anonymised" patient, leading to the patient's social or economic harm. Publication should therefore be treated as Medical Research, considered below.

9. In medical research it is impossible to guarantee that data released for one purpose is not used in another. For example meta-analyses of multiple studies are commonplace, and by definition this re-uses original data. The difficulties of guaranteeing anonymity have also been noted. It is therefore disingenuous to invite a patient to submit data that will only be used anonymously for one study. If personal data are released for medical research, the patient should be advised that their data may be used and published widely and that there is a risk that their anonymity cannot be guaranteed. It is therefore appropriate that **patients give formal written consent before their data are used in medical research**. To avoid even the impression of evil, this consent should not be sought until after any associated clinical procedure is complete, so that there is no impression that treatment is somehow linked to consent.

10. In practice as patient records are made more widely accessible, it should be possible for a patient to give generic consent for their data to be used in medical research, similar to the organ donor system. The patient's records can then be marked accordingly, and made available to authorised researchers.

11. People undertaking medical research should be required to obtain specific ethics committee consent to access particular patient records for a given piece of research. If a patient's personal data are published or used (even though anonymised), this should be considered a prima facie breach of the rules unless the publisher can show (a) that the patient has consented to their data being used and (b) that specific ethics committee consent for the publication has been granted. The question of secondary publications (e.g. meta-analyses) requires further consideration.

12. Finally, Commercial companies may be interested in acquiring personal health and biological data. At one end of the ethical scale, a healthcare company may need epidemiological data to improve the development of a new treatment. At the other end, an insurance company may wish to filter out high-risk customers. In either case, it is preferable that this data is made available by consenting patients in the form of a database that is anonymised as far as practical. Unlike medical research, there is no moral obligation on the individual to disclose this information – it is a simple commercial transaction. In practice, a system could be devised whereby a willing individual sells their health and biology data to a database company for an agreed price. There may be agreed restrictions on how the data will be used, with different price bands for different applications.

APPENDIX
Issues with disclosing personal biomedical data

Stigmatism	social media defamation reputational damage with family, friends, community, workplace
Discrimination:	insurance and pension Job and career Hobby, activity, holiday Elected or civic office Voluntary post or office Medical Treatment Adoption and fostering
Coercion	Threat of exclusion from benefits if data withheld Threat of sensitive data exposure if co-operation withheld
Exploitation	Personal data sold commercially Data harvested for targeted commercial advertising Individual or collated data sets used for a controversial cause
Misrepresentation	Misinterpreted data resulting in personal disadvantage Wrongful assumption of guilt-by-association Use of selective data that misrepresents the true position
Abuse	Targeted data eavesdropping by government agencies Illegal harvesting of personal data for scurrilous publication Random or disaffected hacking and publication of data Altering of data
Proportionality	Required to disclose excessive data in order to access a benefit Required to give unreasonable access rights to access a benefit

Issues with not disclosing personal biomedical data

Exclusion	Missed chance to screen, diagnose or treat Missed prophylaxis Missed access to comparative study results
Avoidable Risk	exposure to undesirable side effects of treatment Ignorance of the implications of lifestyle choices Misdiagnosis
Discrimination	Refusal to provide goods or services if data withheld Additional “risk premiums” charged if data withheld
Public interest	Other individuals are disadvantaged when data is withheld Society is disadvantaged when data is withheld
Freeriding	Access to benefits gained from data shared by others