

This response was submitted to the consultation held by the Nuffield Council on Bioethics on *The linking and use of biological and health data* between 17 October 2013 and 10 January 2014. The views expressed are solely those of the respondent(s) and not those of the Council.

British Medical Association

bma.org.uk
BMA House, Tavistock Square, London, WC1H 9JP
T 020 7383 6286
E ethics@bma.org.uk



Mr Tom Finnegan
Nuffield Council on Bioethics
28 Bedford Square
London WC1B 3JS

Ethics

9 January 2014

Dear Mr Finnegan

The linking and use of biological and health data

Thank you for seeking the BMA's views on this important but complex and far-reaching consultation document. We are not able to respond in detail to all of the questions raised but have set out the guiding principles against which the BMA believes any new developments in technology or research methodology should be judged. We have also included some specific comments on data linkage and on the scope of medical confidentiality in relation to shared genetic data. The BMA would be very happy to look again at this issue once your work has progressed to the stage of developing clear proposals and/or recommendations.

Guiding principles

The BMA supports the revised Caldicott principles for data usage set out in Dame Fiona Caldicott's report, *To Share Or Not To Share? The Information Governance Review*¹:

1. Justify the purpose(s)

Every proposed use or transfer of personal confidential data within or from an organisation should be clearly defined, scrutinised and documented, with continuing uses regularly reviewed, by an appropriate guardian.

2. Don't use personal confidential data unless it is absolutely necessary

Personal confidential data items should not be included unless it is essential for the specified purpose(s) of that flow. The need for patients to be identified should be considered at each stage of satisfying the purpose(s).

3. Use the minimum necessary personal confidential data

Where use of personal confidential data is considered to be essential, the inclusion of each individual item of data should be considered and justified so that the minimum amount of personal confidential data is transferred or accessible as is necessary for a given function to be carried out.

4. Access to personal confidential data should be on a strict need-to-know basis

Only those individuals who need access to personal confidential data should have access to it, and they should only have access to the data items that they need to see. This may mean introducing access controls or splitting data flows where one data flow is used for several purposes.

¹ Department of Health (2013) *To Share Or Not To Share? The Information Governance Review*. DH, London, pp.20-21

5. Everyone with access to personal confidential data should be aware of their responsibilities

Action should be taken to ensure that those handling personal confidential data — both clinical and non-clinical staff — are made fully aware of their responsibilities and obligations to respect patient confidentiality.

6. Comply with the law

Every use of personal confidential data must be lawful. Someone in each organisation handling personal confidential data should be responsible for ensuring that the organisation complies with legal requirements.

7. The duty to share information can be as important as the duty to protect patient confidentiality.

Health and social care professionals should have the confidence to share information in the best interests of their patients within the framework set out by these principles. They should be supported by the policies of their employers, regulators and professional bodies.

The BMA believes that developments in technology, new research methodologies and the linkage of data, both within the health sector and between health and other sectors, should conform to these guiding principles. The principles should not be amended or changed to accommodate new developments. Rather, being clear that these principles must be at the heart of all activities allows developers to find ways to facilitate research in an ethical way.

As noted in the Caldicott review, researchers have devised robust solutions to aspects of information governance so that they can access the data they need, while ensuring confidentiality is protected.² These provide mechanisms which permit the appropriate use of data for research without a need for a change in the law or weakening of well established and understood principles, for example the Confidentiality Advisory Group (CAG) of the Health Research Authority (HRA) has introduced a fast track approval process for certain s251 applications. In Scotland, the Privacy Advisory Committee has adopted a mechanism, including a fast-track process, for scrutinising applications to link data without explicit consent. A further example is the use of privacy enhancing technologies whereby computer software can select only those patients who are eligible for a specific study and only reveal the identity of a potential participant to someone who has a direct clinical relationship with the patient. Once selected, patients can be contacted by a member of their care team inviting them to take part in the study or ask for their consent to be contacted by a researcher. The Clinical Practice Research Datalink Service (CPRD) is an example of this approach highlighted in the Caldicott review. The review is clear that such privacy enhancing technologies should be used to minimise the need for access to identifiable information.³

The BMA believes that much research can be undertaken using anonymised information or secure pseudonymisation and that this should always be the default position. This is in line with rule 3 of the Health & Social Care Information Centre's recently published guide to confidentiality.⁴ Where access to identifiable data is crucial to the work, patient consent should be sought or some other legal avenue explored (eg section 251 approval). The BMA believes that patients should control what happens to their data but also believes that more work needs to be done to inform the public about the potential benefits that can arise from the use of data for research. Steps to encourage people to consent to the use of their data are entirely appropriate but the BMA would not go so far as to say that there is a moral obligation on those using public health services to allow the use of their identifiable data for research.

² Department of Health (2013) *To Share Or Not To Share? The Information Governance Review*. DH, London, pp.62,73

³ Department of Health (2013) *To Share Or Not To Share? The Information Governance Review*. DH, London, p.71.

⁴ Health & Social Care Information Centre (2013) *A guide to confidentiality in health and social care. Treating confidential information with respect*. HSCIC, Leeds, pp. 17-23.

Genetic data

The BMA is aware that tensions can arise when data are not unique to one individual or when information about one individual has implications for other people. The most obvious example is genetic data (available either from clinical testing or from research), which has personal relevance for more than one individual. It has been suggested, for example, that genetic test results should not be considered as personal information – and therefore should not be subject to the usual rules of confidentiality. Rather, the results should be viewed as familial information, such that information may be shared with other family members unless there is a good reason why it should not be divulged.⁵ The BMA does not accept this view and believes that, as with other areas of healthcare, information about or provided by one patient should not be shared with others unless consent has been obtained or there is a legal requirement or an overriding public interest to justify disclosure. Individuals should, however, be informed of the relevance and importance of the information for other family members and should be encouraged to consent to the sharing and use of that information for their benefit. Seeking consent to the use of data for the benefit of others who are, or may be, directly affected should be a routine part of the consent process for genetic testing.

Where patients consent to the sharing of information with family members, the issue is still not straight forward. Many individuals will be aware of a problematic family history but there are likely to be some cases where family members are unaware of their at-risk status. Giving them this information allows them to make their own decisions about whether to seek testing for themselves but also denies them a so-called 'right not to know'. This could lead to considerable anxiety but may also have significant practical implications, for example, in relation to insurance; although genetic test results need not be provided to insurance companies, those applying for cover are required to declare information about family history. This issue, about the right to know and the right not to know, is one that that requires careful consideration.

Currently genetic information can give a predictive analysis about some patients and their close relations in relation to certain medical conditions. In the future it is highly likely that the predictive tools will become more comprehensive and sophisticated. This will have a significant effect on how medicine is practised and how the insurance and pharmaceutical industries, in particular, function. This will raise the value of genetic data considerably and therefore strict adherence to the current law must be followed and it may be that additional protection might be required in the future. A regular review of the issue will be important.

BMA views on data linkage

The BMA has considered various aspects of data linkage. The points below summarise the BMA's views.

- The BMA recognises the value in linking data from medical records from different healthcare settings to present a joined up view of care or to compare care pathways. These de-identified linked⁶ data can be extremely useful to researchers, commissioners and others.
- Whilst the BMA is not in a position to advise on the technicalities in relation to linkage, the starting point must be to consider privacy enhancing technology solutions prior to assumptions that identifiers are required for linkage, for example pseudonymisation at source or use of closed system technology.⁷
- The Health and Social Care Act 2012 (HSCA) empowers the Health and Social Care Information Centre (HSCIC) to require identifiable data from healthcare providers in certain circumstances so that it can be linked to data from other sources and made available in anonymised or pseudonymised form. As a secure and controlled environment (or 'safe haven') with statutory

⁵ Parker M, Lucassen A. (2004) Genetic information: a joint account? *BMJ* 329, 165–7.

⁶ As set out in the Caldicott Review, there are two categories of de-identified data: de-identified data for limited access ie data which are deemed to have a high risk of re-identification if published, but low risk if held in an accredited safe haven and subject to contracts to prevent re-identification; and anonymised data for publication.

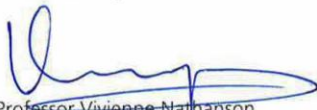
⁷ A closed system processes data automatically so that no human sees identifiable data.

powers, it is appropriate for the HSCIC to carry out linkage as detailed discussions have taken place to ensure that appropriate controls and safeguards are applied to protect patient confidentiality.

- Other than the powers available to the HSCIC in the HSCA there is currently no legal basis in England and Wales for other bodies to extract or hold multiple identifiers to carry out data linkage.
- The recent Caldicott review states that linking data which contain a single identifier⁸ must only take place within the HSCIC or an accredited safe haven (ASH) and where there is a legal contract in place. Whilst the BMA is committed to ensuring that researchers have access to high quality de-identified linked data, we would be very concerned about the creation of multiple ASHs with powers to link data. The number of organisations with powers to handle data which contain identifiers must be limited and subject to robust accreditation and audit processes. We are very concerned that it is as yet unclear how ASHs will operate.
- Whilst it may not be possible to identify individuals from a single database, it may be possible to identify individuals by linking multiple datasets. There must be clear and robust terms in all access agreements prohibiting attempts to deduce identities. Consideration should not only be given to published datasets but the possibility of linkage between a published dataset and an unpublished dataset held by an individual organisation. If the government publishes information, for example, it should not be possible for an organisation to infer further information about individuals by linking the identifiable and anonymised published datasets. This is a particular risk with rare conditions, for example when publishing prescribing data at an organisational level. There must be agreed information governance processes to ensure that data are appropriately anonymised or pseudonymised prior to release and measures put in place to ensure patients cannot be inadvertently identified by opening up access to information.
- The HSCIC have published de-identification standards which helps those releasing data assess the risks involved.⁹ The Information Commissioner's Office has also published an anonymisation code of practice which provides practical advice about how to reduce the risk of inappropriate sharing and the consequent sanctions by the ICO.¹⁰
- The Clinical Practice Research Datalink¹¹ (CPRD) service is an example of a system which links data from GP practices with data from other NHS settings. This is done by the use of a dual pseudonymisation process and involves sending certain identifiers to the HSCIC. This release of identifiers from GP practices is covered by approval under s251 of the NHS Act 2006.

We look forward to seeing the outcome of your discussions in due course.

Yours sincerely



Professor Vivienne Nathanson
Director of Professional Activities

⁸ These type of data are called 'de-identified data for limited access'. As referenced above, these data do not identify individuals on their own but there is a high risk that re-identification could occur outside of the secure and controlled safe haven environment.

⁹ Health and Social Care Information Centre (2013) *Anonymisation Standard for Publishing Health and Social Care Data Specification*

¹⁰ Information Commissioner's Office (2012) *Anonymisation: managing data protection risk code of practice*

¹¹ www.cprd.com