Patient Information Governance and the Caldicott 2 Review

Introduction
Patient information governance is an increasingly important topic because of the urgent need to share and integrate patient data to improve care, power research, empower individual patients, and in England to commission care. Dame Fiona Caldicott led an independent panel of experts which reviewed patient information governance practice in the English NHS. The objective of the final report “Information: To share or not to share”, known colloquially as Caldicott2 and published at the end of April 2013, was to “to ensure that there is an appropriate balance between the protection of the patient or user’s information, and the use and sharing of such information to improve care.” This paper makes extensive use of ‘Fair Shares for All’, a review of patient information governance prepared by the BCS Primary Healthcare Group for BCS Health in 2011-12.

Whilst the BCS, Chartered Institute for IT1 welcomes the report’s recommendations2, there are issues which we believe need to be addressed.

BCS calls for
➢ a greater role for patients in information governance
➢ increased use of privacy enhancing technologies to avoid the need to disclose identifiable data for secondary purposes outside its origins without consent, unless it is essential for the intended purpose
➢ tighter and more transparent governance for safe havens that store identifiable individual patient data
➢ ultimately, more granular dissent options for patients to minimise the impact of patient dissent on data availability
➢ a single body involving all stakeholders to generate policy and guidance on, and to monitor the practice of, NHS patient information governance.

While these recommendations are generally applicable, we are particularly concerned here with their relevance to England where the Health and Social Care Act 2012 has significantly reduced the legal protection afforded to patient data. Before looking at our concerns in more detail, we define a few important terms and summarise the evolution of patient information governance in the NHS.

Definitions
Patient Information Governance – the process of ensuring that patient information is used in a way that complies with patient expectations, the relevant legislation, the common law duty of confidentiality that care professionals owe to their patients, and any relevant codes of practice.

Identifiable data - data about dead or living people or groups of people, such as families, that either explicitly identifies them or offers a significant risk of doing so in conjunction with other information that the person holding the data knows or has access to3. The opposite of identifiable data is non-identifiable data4.

---

1 Hereinafter referred to as ‘we’ and ‘our’. ‘Patient’ should be taken to mean either an NHS patient or a Social Services client: this document covers both, as Caldicott2 did.
2 We identify the recommendations by R and the number given to them in Caldicott2 section 14.4, e.g. “R7”.
3 Free text items in patient records may well contain information that would significantly help to identify someone, and non-identifiable data containing free text should be treated as identifiable data.
4 Almost all data about an individual person offers a risk of identification even where it doesn’t include any explicit identifiers, such as name, address, dob or NHS number. So can aggregate data. The risk ranges from minute with sparse data about a small random sample from a large population to high where the data is comprehensive, describes when and where specific events happened and is from a known small population, e.g. a GP practice.
**Personal data** - defined by the Data Protection Act 1998 (DPA) as identifiable data that applies to living people. The DPA categorises personal data about a person’s health and healthcare as **sensitive personal data**, a kind of personal data which the DPA protects more tightly.

**Secondary purpose or use** - the use of patient data for purposes other than personal care, e.g. de-identification, linking, health service management, care commissioning, education and research.

### Background to NHS patient information governance

Patients provide their personal data to clinicians for their own care with implied rather than explicit consent, and patients trust their clinicians to share relevant parts of it with others caring for them as necessary. Most patients support the use of their non-identifiable data for health-related research and other worthy secondary uses, but wish to be asked for consent before their personal data is used\(^5\). A person’s health, social condition and care data is highly confidential, but its full power is only realised when it is shared as needed for personal care, health service management and research. This introduces an inevitable tension between privacy and data sharing, particularly for secondary purposes. Unconstrained and non-consensual data sharing risk damaging the trust at the heart of the doctor/patient relationship.

Whilst the DPA defines how personal data may be legally processed, it does not override patient privacy, as embodied in the common law duty of confidentiality.

Since 2000 there have been several attempts to facilitate the sharing of patient data:

- **NHS Health Act 2002.** This legalised the operations of the Cancer Registries and some kinds of data collections for public health purposes, and created the Patient Information Advisory Group\(^6\) to advise the Secretary of State on applications to use identifiable data for secondary purposes.
- **The 2002 legislation continues in force as section 251 of the NHS Act 2006.** S251 may require or permit the use of identifiable data, but the Confidentiality Advisory Group\(^2\) intends to operate in permissive mode as far as possible, as did its predecessor.
- **The research lobby attempted to reduce patient data protection further using legislation within the Coroners and Justice Bill 2009 based on the proposals in the Data Sharing Review\(^7\), but this never reached the statute book\(^8\).**
- **The Health and Social Care Act 2012 (HaSC Act 2012) applies only to England.** It abolished the National Information Governance Board for Health and Social Care (NIGB), the body responsible for the development and oversight of NHS information governance. It enables the Secretary of State, NHS England, Monitor, Care Quality Commission, NICE and other bodies specified in regulations to direct the new Health and Social Care Information Centre (HSCIC) to collect and process identifiable patient data\(^9\). Patient data controllers can refuse to comply with such a direction, which would probably lead to legal action to enforce it. It is not clear that such legal action would succeed in the face of the Human Rights Act 1998.
- **The Secretary of State stated at the launch of Caldicott2 in April 2013 that patients have the right to stop their data leaving their GP practice, and/or being disclosed by the HSCIC.** As we understand it this is different to the DPA section 10 patient right to object to its processing, and is to be guaranteed by an NHS Constitution pledge rather than the law. However the exact meaning is still not clear – for example:
  - does the dissent apply to data held by anyone providing care to the patient, or just that held by GPs?
  - does it apply to de-identified data as well as identifiable data?

\(^5\) The DPA Schedules 2 and 3 list other grounds in on which sensitive personal data can be processed without patient consent, but for brevity’s sake we do not mention them in this document.

\(^6\) The Patient Information Advisory Group became the Ethics & Confidentiality Committee (ECC) and then the Confidentiality Advisory Group (CAG) in the HaSC Act 2012.

\(^7\) The review, published in July 2008, was led by the UK Information Commissioner, Richard Thomas and Mark Walport, Director of the Wellcome Trust, now Sir Mark Walport and UK Government Chief Scientific Advisor.

\(^8\) The relevant clauses were removed from the Bill after opposition by the BMA, Privacy International and the BCS, Chartered Institute for IT.

\(^9\) It does this by creating new statutory gateways that set aside the duty of confidentiality clinicians owe to their patients and most of the protections and patient rights that the DPA gives to sensitive personal data. The HSCIC may not publish identifiable patient information, and only disclose it to specific persons or organisations under certain circumstances or with the patient’s consent (see HaSC Act sections 260-262).
The Secretary of State published his response to Caldicott2 - “Information: to share or not to share” - in September 2013. It is very supportive, saying that “while it will not always be possible to implement the recommendations to the letter, the aim is to do so within the spirit intended by the review’s findings” (14.1) and sees a key role for the Department of Health in delivering its recommendations (1.14). It emphasises a patient’s right to object to the sharing of data collected for their personal care with others (7.11, 7.16). 7.11 suggests that this applies to all care providers and not just GPs – “the Department has made it clear in relation to the provision in the Health and Social Care Act 2012 for the HSCIC to obtain confidential information that patients and service users are able to block their own information from being disclosed” and emphasises that “professionals’ duty of confidentiality is not undermined by any of the actions in this response. It still remains the case that, for direct care, where it is in the best interest of the individual for information not to be shared, it won’t be shared” (3.10).

On the other hand it does not give data subjects a role in developing and monitoring information governance policy and practice equivalent to the one they had in NIGB. It notes that “until a revised NHS Constitution and associated documents are published, the rights and duties will be gathered together in the Confidentiality Code of Practice” applicable to health and social care (5.16) and acknowledges that “some of the actions outlined in the response require significant further work” (14.9).

On the same day the HSCIC published “A Guide to confidentiality in health & social care”. This is not the Code of Practice that the HSCIC is charged with delivering. The most significant material is in the references document, q.v.

Key Issues with Caldicott2 recommendations, and our position

1. The patient’s role

Caldicott2 rightly advocates greater patient awareness of their data, and how it may be and has been used (R1,7,9,19) but does not recommend greater patient involvement in collective information governance tasks.

We believe that accountability to the patient in patient data management is crucial practically and ethically. This requires involvement at the collective and individual level. Patients should have non-token involvement in mechanisms that:

- provide checks and balances to the use of the statutory gateways allowing the sharing of identifiable patient data in the Health and Social Care Act 2012,
- develop patient IG policy and practice, and that monitor and regulate it.

2. Patient consent & dissent to sharing

Caldicott2 recognises that patients should be able to consent/dissent to the sharing of all or some of their personal data (R8) for secondary use, and to object to its use (R11). However further clarification is necessary.

The BCS position is that:

- very few secondary uses need identifiable data, and that patients should be able to dissent to any combination of the following:
  - sharing their identifiable data without explicit consent for secondary uses (including collection by safe havens such as the HSCIC and via other gateways in the HaSC Act

10 Caldicott also used the term ‘indirect care for some secondary purposes’. While some data uses, such as risk prediction & stratification, serve both direct care and secondary purposes, we feel that the use of “indirect care” to describe “health services management, preventative medicine, and medical research …. service evaluation, needs assessment, financial audit” (Glossary, page 129) is likely to confuse rather than clarify, particularly when used in discussions about patient information governance. We have avoided using the term.

11 At the launch of the Caldicott2, the SoS stated that “GPs will not share information with the HSCIC if people object”… “people will have a veto on that information being shared in the wider system”
2012), unless approved by the CAG or a similar mechanism. To minimise the degree of dissent more granular dissents should ideally also be available to patients, for example for any combination of:

- risk stratification and other forms of healthcare targeting
- administration and management (including commissioning)
- service audit, other than by their care provider(s)
- research (including public health functions not permitted by law)
- transfer of data to any other data controller, e.g. the HSCIC.

O the sharing of specific items in their care record(s) for any purpose without their explicit consent, as implied by Caldicott in section 3.2-3. As the Care Record Guarantee says, dissent to sharing for direct care should be discussed beforehand by the patient and the care professional holding the record involved.

- Implicit consent with an explicit dissent option is only credible where patients are aware of the latter and the possible consequences of using/not using it before any sharing starts, and where it is easy to register dissent. NHS England’s proposal to publicise its care.data collection of identifiable patient data solely by providing posters and leaflets to GP practices is insufficient to achieve this.

3. Why not use privacy enhancing technology (PET)?

Caldicott considers (6.4) that "enterprise wide" privacy enhancing technology (PET) to de-identify (i.e. pseudonymise or anonymise) data at source requires further evaluation (6.4) before widespread use, but does not recommend this.

The BCS position is that:

- proven techniques exist to de-identify data at source, to link pseudonymised data from multiple sources and to re-identify it at source if necessary. Linking is not significantly less successful than with identifiable data. Those providing pseudonymised data to third parties (such as the HSCIC) must take any additional steps needed to make the risk of re-identification insignificant. THIN and Q Research, both respected safe havens of data for research, collect pseudonymised patient data from GPs and then link it with data from other sources.

- proven open source and commercial software is available to do this.

- for most secondary purposes (including care commissioning), PET removes the requirement to use identifiable data. PET is especially relevant to safe havens (e.g. the HSCIC) and others directed / wishing to collect and link data, particularly where the data may subsequently be used by a variety of bodies and for purposes that are not known at the outset. The culture should encourage the use of de-identified data whenever it satisfies the intended purpose.

- PET has reached a point where any planned processing of identifiable data for secondary purposes without patient consent should have to demonstrate why it cannot use de-identified data, as is the case with applications to the Confidentiality Advisory Group. This should include intended disclosure via the new statutory gateways created by the HaSC Act 2012.

- full exploitation of de-identification techniques will

---

12 Although not a legal requirement, a research-specific opt-out to the use of de-identified data appears to have been offered in David Cameron’s speech on life sciences research of 5.12.2011.
13 This facility is available in most GP systems, but is not common in others.
14 Patients need to be made aware that personal data they conceal may be inferable from other unconcealed data recorded about them, e.g. attendance at an AIDS clinic would suggest a diagnosis of AIDS.
15 We understand that the Information Commissioner has requested that the start of the care.data collection is postponed until he is satisfied that patients have been properly informed about it, and that NHS England has just agreed to send a leaflet explaining care.data to every household in England in January 2014.
16 What Caldicott2 means by "enterprise wide" is not clear: much pseudonymisation is not required NHS-wide. It is relevant to note that more elaborate versions of the same techniques can safely deal with data pseudonymised at source for data that is collected incrementally over time about the same population.
17 Such as data blurring (e.g. holding an age band rather than date of birth, or only the first part of a postcode), random sampling of the study cohort and ensuring that it is only used for specific purposes in a closed environment by those owing the patient a duty of confidentiality equal to that of a clinician.
18 These techniques are described in more detail in Annexes H & I of ‘Fair Shares for All’
- reduce the need to seek patient consent / make s251 applications to use identifiable patient data without patient consent
- greatly reduce or obviate the use of the new statutory gateways in the HaSC Act 2012 for sharing identifiable data
- avoid the administrative burden generated by holding personal data.
- encourage patients to share their data, including sensitive content.

- although we recognise that current central data collection tools require upgrading to cater for pseudonymisation at source and that careful key management is required for incremental collections, PET is ready for deployment in high profile applications, such as NHS England’s care.data collection, and the HSCIC should take up the challenge.

- de-identification and re-identification at source should be as automated as possible across NHS care providers. Requirements for it should be embedded in system procurement agreements, such as GP Systems of Choice.

5. Safe havens
Caldicott2 recommends (R10) that only specialist, well governed independently scrutinised and accredited Safe Havens should link identifiable data.

The BCS position is that:
- We agree, but the need to link identifiable data is rare (see 3). While there are other sound justifications for safe havens19, they raise the risk of confidentiality breaches when they hold copies of identifiable data held elsewhere, create enriched datasets of linked identifiable data and disclose identifiable data to third parties.
- To merit the description ‘safe haven’, we believe that an organisation should:
  - include substantial representation of data source controllers and their data subjects in their own governance arrangements.
  - make the maximum possible use of de-identification at source (qv 3 above) for data collection when they collect data.
  - only collect and disclose20 identifiable data for secondary purposes without consent after prior approval by the CAG or an equivalent body. Rich de-identified individual data that poses a significant risk of re-identification should be treated in the same way, i.e. as though it were identifiable.
  - investigate with their data sources whether it would be feasible and practical to act as a data processor for them rather than as the data controller of the identifiable data that it collects from them. While this may prove more acceptable to data sources, it would require the local controllers, e.g. GPs, to engage in further data governance and would impact their workload. Clear national guidance would be needed to support it.
  - minimise replication of patient data by encouraging end users to process it at the safe haven, rather than taking a copy for their own use.
  - maintain an easily accessible public register of the data it collects holds and discloses, showing at least the details of the data involved, from/to whom it is collected/disclosed, how often this occurs or occurred and when it started, the purpose(s) of the collection/disclosure, and the legal basis for doing either.
  - hold encrypted data behind firewalls to prevent unauthorised access, and ensure that its governance policies and practices are regularly and independently audited.

- As one of the largest would-be safe havens, it is appropriate that HSCIC should contribute to the development of accreditation criteria for safe havens but not that it should be responsible for their development. The CAG or an equivalent body (see 5) would be a more suitable

---

19 The commonest one is to enable data which has already been collected, cleansed and/or linked to be used for purposes (including further linking) and by organisations not envisaged (and so notified) to patients at the time of collection. For identifiable data this may require seeking additional patient consent or CAG approval.

20 Many, probably most, patients would be concerned to learn that the HaSC IC sells identifiable patient data, vide its recent approval of BUPA as a suitable customer for it, and that there is an expectation that it will attempt to cover its costs by selling data that it holds.
candidate for this role. HSCIC’s duty to deliver advice and a code of practice for health and social care information governance is similarly questionable.

6. The need for a national information governance body
Cadicott2 recommends allocating responsibility for major interrelated IG topics in England (e.g. R3, 8, 14, 18, 20, 22-24, 26) to several organisations, many of which have conflicting interests and little or no patient or care provider involvement.

We strongly believe that this will hinder the evolution of a coherent English information governance policy that recognises the expectations of patients and their clinicians, and the Government’s stated intention to give patients a greater role in the governance of their data21.

BCS calls for a single independent organisation involving all patient information governance stakeholders to handle these tasks.

Published 21 October 2013

---