Effective Sharing of Health Records, Maintaining Privacy: A Practical Schema

Roderick Neame[1]

1 University of Queensland St Lucia Campus, Brisbane QLD Australia

Abstract

A principal goal of computerisation of medical records is to join up care services for patients, so that their records can follow them wherever they go and thereby reduce delays, duplications, risks and errors, and costs. Healthcare records are increasingly being stored electronically, which has created the necessary conditions for them to be readily sharable. However simply driving the implementation of electronic medical records is not sufficient, as recent developments have demonstrated (1): there remain significant obstacles.

The three main obstacles relate to (a) record accessibility (knowing where event records are and being able to access them), (b) maintaining privacy (ensuring that only those authorised by the patient can access and extract meaning from the records) and (c) assuring the functionality of the shared information (ensuring that the records can be shared non-proprietorially across platforms without loss of meaning, and that their authenticity and trustworthiness are demonstrable). These constitute a set of issues that need new thinking, since existing systems are struggling to deliver them.

The solution to this puzzle lies in three main parts. Clearly there is only one environment suited to such widespread sharing, which is the World Wide Web, so this is the communications basis. Part one requires that a sharable synoptic record is created for each care event and stored in standard web-format and in readily accessible locations, on 'the web' or in 'the cloud'. To maintain privacy these publicly-accessible records must be suitably protected either stripped of identifiers (names, addresses, dates, places etc.) and/or encrypted: either way the record must be tagged with a tag that means nothing to anyone, but serves to identify and authenticate a specific record when retrieved. For ease of retrieval patients must hold an index of care events, records and web locations (plus any associated information for each such as encryption keys, context etc.). For added security, as well as for trustworthiness, a method of verifying authenticity, integrity and authorship is required, which can be provided using a public key infrastructure (PKI) for cryptography (2). The second part of the solution is to give control over record access and sharing to the patient (or their identified representative), enabling them to authorise access by providing the index and access keys to their records. This can be done using a token (fe.g. smart card) or a secure online index which holds these details: this serves to relieve the formal record keeper of responsibility for external access control and privacy (internal access control and privacy can remain an institutional responsibility). The third part of the solution is to process the content of the stored records such that there is a 'plain English' copy, as well as an electronic copy which is coded and marked up using XML tags for each data element to signify 'type' (e.g. administrative, financial, operational, clinical etc.) and sub-types (e.g. diagnosis, medication, procedure,

investigation result etc.). This ensures that the recipient can always read the data using a basic browser, but can readily manipulate and re-arrange the data for display and storage if they have a more sophisticated installation.

Correspondence: roddyneame@taskcare.com

Background

Record keepers are faced with increasing demands for widespread access to stored information, often involving providers who are unknown to them and not part of their system or network. Whilst endeavouring to meet these demands, they are obligated to ensure that ethical and legal privacy requirements are met. In addition they must make provision for patients to exert control over their own information and to determine who may see what of their records. Finally they must endeavour to ensure that data shared with third parties can readily be understood and displayed meaningfully on their desktop – whatever the nature of their technology.

A decade ago Mandl et al. (3) published an article with a similar title to this, in which they recognised many of the same issues as outlined below, and proposed some general approaches to their solution. Mandl's work was speculative and raised numerous issues, but did not propose any formed solution. More recently Steinbrook (4) in 2008 indicated that personally controlled online health data might be coming, based on centralised commercial web services (e.g. Microsoft HealthVault, Dossia and Google Health) providing the data protection and storage: he raised issues as to whether the public would have confidence that their confidential and personal records were safe, secure and private in the hands of such third party commercial service providers, and whether these providers would be bound by HIPAA, but again suggested no solutions to the issues. Steinbrook's paper has little practical detail as to how sharing with assured privacy might work.

Another paper seemingly addressing a similar topic, by Liu et al. (5) proposes an approach to sharing patient data between hospitals using XML tags associated with each significant element: the tags would allow different systems to identify the data elements sent from remote systems and store them in the equivalent places and formats that they use for data generated locally. This environment in Taiwan has been developing slowly, and now has a basic smart health token acting as an index and pointer: however it does not address the need for robust security, nor for record readability across an entire range of possible user workstations. Taiwan's development has been made possible by the fact that there are not significant numbers of stakeholders with investments and positions to protect, and that the government has exerted considerable pressure, political and legal, to force the solution through: observers do not believe that a similar approach would be successful elsewhere.

This current paper proposes a practical and workable solution, based on tried and tested technology all of which works in similar environments. It builds upon some of the ideas developed in the above publications. It does not threaten any existing investment in the sector, and is entirely platform independent. Whilst it could operate using commercial storage and data protection services (e.g. HealthVault), the uncertain level of privacy and security they offer suggests that a decentralised but encrypted system would be wiser and more secure, with the patients holding the index and keys to their own data. The main focus of the paper is to offer a practical solution to the sharing of medical data where privacy and security are robust and where the records can be trusted as being unaltered and unchanged as they pass between providers.

Outline Plan

The principal issues are outlined above. In this section we address the proposed solution in operational terms.

Privacy. One of the more important issues for acceptability (and legality) is that the issue of maintaining confidentiality is fully addressed. Maintaining the privacy of patient information is becoming increasingly onerous in the face of changing legal requirements and public expectations, increasing penalties for errors and failures, and changes in information technology, networking and data storage. All data which is associated with a readily recognisable identifier (e.g. name, address and date of birth) is confidential and complex rules have to be observed for keeping it generally secret whilst at the same time making it readily accessible to those who are authorised. The simple step of disassociating the human-meaningful identifiers (names, clinics, addresses, dates etc. – "the context") from the remainder of the event data ("the content") can solve much of the privacy issue, since now the stored content reveals nothing on its own unless associated with the context. However that separation may not always be possible, and it may be preferable in some cases to leave the full record marked up with HTML, and to encrypt it before web storage. The privacy challenge becomes how to manage the association of context with the stored content as and when required for legitimate purposes, and the management of encryption keys. Whilst the stored data may be identifier and context free, it must still have some sort of embedded key or tag that serves to identify that block of data itself and which can be used to confirm the data block identity for linking and retrieval. However such tags will have no other meaning.

The literature shows that privacy concerns are not a trivial issue, as sometimes represented: there is a significant section of the community who have concerns about the confidentiality of their records (6) and who will act accordingly – either not seeking care or withholding important personal/health details from their consultations. Public concern about privacy in the context of computerised information is high, especially in the context of the almost daily losses of personal information from both private and government records, the statistics of which are quite alarming (7): disclosed reports detail over 120 million records that were affected in 2011, with almost one third of the incidents occurring in the health sector; undisclosed losses are in all probability considerably greater.

Trustworthiness: The user of the data, typically a clinician, needs to know that they have the correct block of data (from the embedded key), but also that it has not been changed from when it was written and can be trusted – regardless of how that data makes its way to their desktop. Without that there will not be sufficient confidence in the data for it to be useful. Therefore the block of data will contain a 'checksum' (mathematical computation based on the entire contents of the block), and that checksum will be encrypted by a key: decryption using the relevant key and subsequent comparison with the material will permit an automatic check that the data has not been altered since being written and checksummed. In many environments a public key infrastructure is being implemented (2): this offers an additional verification capability. Where the author encrypts the checksum with their private key, decryption with the author's (web-listed) public key not only confirms that the data is unchanged, but also verifies the identity of the purported author. This confirmation is of medico- legal significance, as well as reducing the risk of misadventure.

Access Control: The issue of who should control patient records has at times been contentious. A paternalistic attitude has prevailed widely amongst health professionals that they knew best who should be privy to the records, but this has been supplanted in recent times with the increasing weight given to ethical issues and privacy legislation. Privacy as a concept aligns with the ethical principle of 'autonomy'(8): this principle holds that an individual has a right to self-determination and to control and make their own decisions about personal matters, including who may know what about them. There are numerous situations in which an individual discloses their secrets to a professional (e.g. doctor, priest, accountant, lawyer). The professional receiving the secrets has a clear ethical responsibility to keep that information private, and in addition has a fiduciary (based on trust, confidence and good faith) duty in common law. In terms of related statutory legislation, there are also in most jurisdictions specific bodies of law requiring personal information to be kept private (e.g. Data Protection, Privacy and HIPAA legislation etc.) as well as provisions under Human Rights legislation for ensuring that no organs of the state interfere with the fundamental right to privacy and family life of the individual. All these ethical and legal constraints make clear that as far as possible the person who should be in control is the data subject (in this instance the patient), or their authorised care providers and/or representative(s).

This raises the issue of how to give data sharing/disclosure control to the patient such that they may exercise this right from anywhere at any time leaving a trail that is auditable. Again there are alternatives. One solution would involve the use of a unique token (e.g. smart card) and/or biological identifier (e.g. scan of fingerprint, iris, face), and a previous publication by this author (9) detailed such a schema. Harrison & Booth (10) presented views that are entirely consistent with this schema. Both of these publications address the need to manage multiple aliases and identifiers that individuals may be assigned or choose to use in healthcare settings as well as in other contexts. An alternative solution is that individuals could choose to store their records in their own personal vault (or on provided by a commercial service, if they are comfortable with the security it offers). A third solution is for the patients to store just the index to their personal records, leaving them where they are on the web. All of these can comfortably co-exist.

The key to management of the patients personal records is an index of care encounters, which comprises a series of dates and event/encounter descriptors: these can simply be text files which are dated and assigned an appropriate descriptor, and can therefore be ordered using any folder/file viewing application. That index can be stored on a smart card, or any other similar secure memory token, or it can be kept in a personal online storage location. Whichever technology is used, the index requires several components in addition to the list of care encounters/events and their chronology: it must include the location (URL) of the stored sharable record and its embedded tag, any associated data such as encryption methods/keys and checksum verification information, and any stripped context (e.g. people and identifiers).

Assuring Meaningfulness: The remaining obstacle to sharing relates to the usability of the data passed to the recipient, who must be able to access and use the data. The reader will typically be a browser-based, although some institutions may use sophisticated point-of-care systems that manipulate and store the data in different ways. The lowest common denominator is a browser which predicates that the data must be stored as 'raw' text in HTML, which can be read (after decryption) by any generic browser: this follows the scheme recently adopted by the HL7 organisation (11) in their clinical document architecture (HL7v3). However to better support those with more sophisticated systems, a further structure based on XML (eXtended Markup Language) can be incorporated, along the lines of the approach adopted by Liu et al. (5): this would have data elements tagged according to their nature (e.g. financial, administrative, clinical; diagnostic, medication, investigation etc.). Based on the tags, the reading system can process, store and display the data intelligently: however for this to function optimally there would need to be agreements in place about data classification and coding.

Procedure: Operationally the record-related steps involved in a care event/encounter would be:

1. Care encounter/event commences, automatically assigned an ID; patient activates their smart card token, or opens their online deposit, and authorises clinician access to their records
2. Clinician retrieves relevant records from local and remote/web storage using the index to guide access to locations and provide required context, keys etc.
3. Notes/records for the present encounter are written and stored locally as normal
4. At the conclusion of the encounter, the desktop software prepares an identifier-free encounter/event abstract and exhibits this summary to the clinician for correction and addition of a descriptive title
5. Approved summary is processed into HTML, data are automatically type-tagged with XML, an encrypted checksum added, and the document is assigned a linktag and uploaded to a selected URL: the whole document may be encrypted (see discussion)
6. A text file is generated, named according to the event descriptor, and dated: this file contains URL, linktag, encryption keys, context etc. The file can be copied to the patient index, whether on a smart card or on the web, or could be placed into secure storage that can be accessed by the patient using a password (provided): where the patient is not present concurrently (e.g. blood analysis report, image report) this index

file can be forwarded to the requesting clinician to pass to the patient at their next encounter.

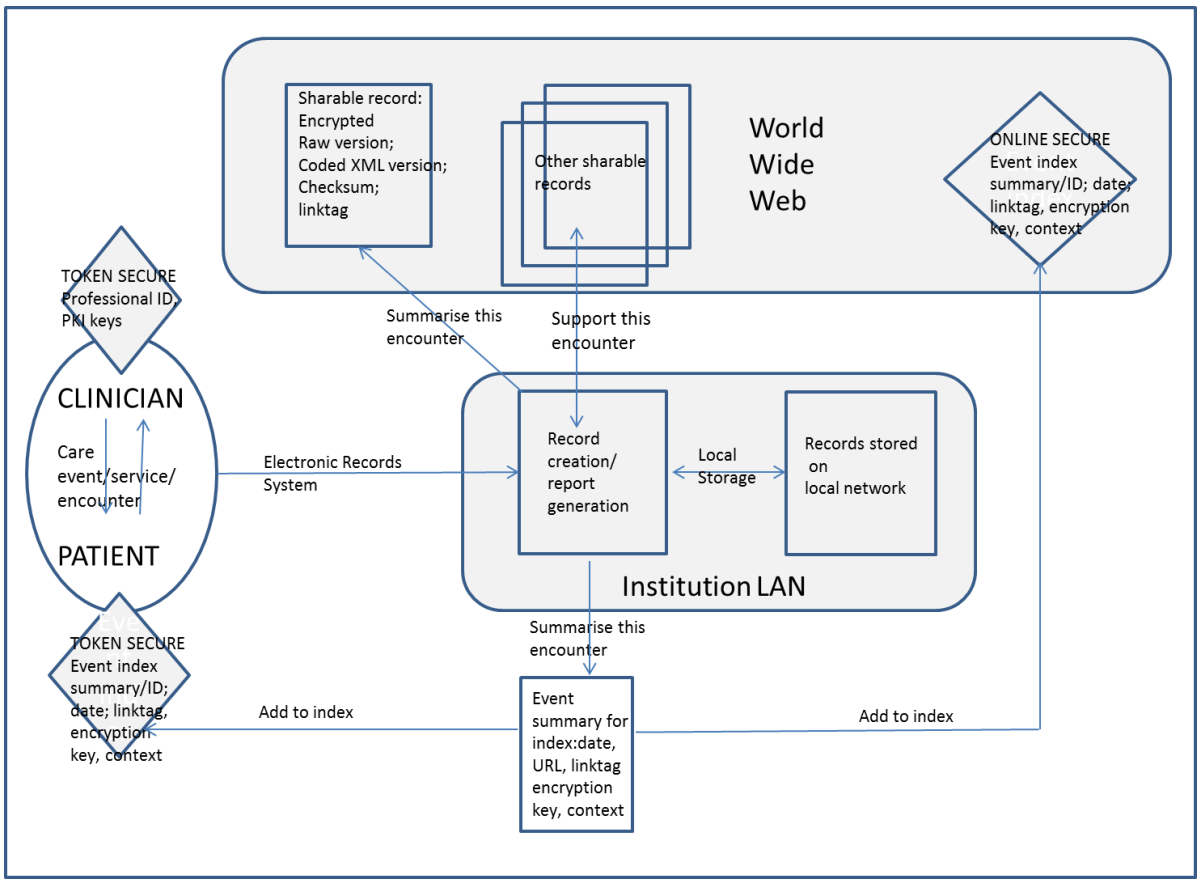

Figure 1: Steps involved in care/ event encounter

Discussion

The aim of this work is to outline a method for joining up care services for the patient such that their care records are fully portable and can follow them wherever they go, so reducing delays, duplications, errors and risks of care, and cost. Access to care records is currently often difficult or impossible since where records contain personal identifiers they have to be secured behind services that greatly impede ready access except to those who are local (i.e. registered users of the host system). This schema frees them of this impediment and allows them to be stored in readily accessible locations on the web. The user then has to be guided to the URL of the relevant record, enabled and authorised to access it, and provided with the context and keys to be able to use it: these are all functions that can be managed by the patient using a token which stores URLs, tags, keys and any required context of their records, or a suitable secure web service.

The role of the record keeper is to make information available as and when required by a duly authorised person in order to support shared care, to ensure that records can follow patients on their journeys through the healthcare system, but at the same time to keep those records secure from unauthorised users. Accessing patient records electronically on demand from a

workstation that is not part of the data keepers network is at present generally impossible, and shared care is thus effectively limited to within a single institution, excepting where external parties are included by prior agreement and records are forwarded ('pushed') to them. However the real need is for any professional authorised by the patient to be able to 'pull' records from storage to their desk top – and this is exactly how the world wide web works, but brings with it issues of access and privacy.

The above scheme is based on technology that is tried, tested and for the most part universally available: being web-based it is built on applications such as generic browsers, and is independent of user computing platforms or applications. Infrastructure components such as PKI, smart health cards and XML tags are already in widespread use, and progressively being implemented more widely: but the system proposed can work whatever the level of sophistication of the local users. Most existing records systems are proprietary, storing data within their own structures and formats: the use of the simple transform of creating HTML outputs from these renders them universally readable and readily 'pulled' from an online repository. The use of encryption ensures they cannot be read other than by those with the keys; and the use of checksums enables the records to be validated and their authorship verified. Enabling the patient to control access to their own data makes it possible to greatly simplify existing access control infrastructures and transfers the burden of privacy management from record keepers to patients.

The scheme also manages multiple identities. Harrison & Booth (10) noted that individuals do not have just a single 'identity' but many – e.g. for healthcare, tax, finance, driving, insurance and many more – and that there is no requirement that all these 'identities' are the same although for certain purposes there may be a requirement to provide evidence to prove an identity (e.g. bank accounts, national insurance etc.). An individual may even choose to have multiple identities or personae for different healthcare encounters, despite the risk that using aliases may incur risks due to fragmentation of their records. For example the patient may choose to keep some records (e.g. psychiatric, sexuality or obstetric care events) separated from their general medical records, just as they may choose to keep their court records under a separate alias from their employment or financial records. Controlling the linkage between personal information kept in different folders is a right. However individuals will normally have only one 'official' identity at any one time, for 'public' purposes (e.g. passport, welfare, tax etc.). This identity provides the authority for claims by care providers for care services from public insurers: another insurance identity may authorise private insurance claims. Once these 'top-level' identities are established, a patient can create or manage sub-identities as suits their requirements, and distribute their care (and other) records between them as they choose.

There is an issue of what happens when the patient loses their card/token or access to their web repository. In the absence of the patient index, providers will have to rely as they do currently on their previous records for that patient, which of course they hold on their own system in addition to having placed a sharing copy on the web, and use their clinical acumen in eliciting a relevant history: the patient is no worse off than at present, except that in the absence of their identifiers for claims purposes they may be required to pay for their care

rather than automatically having it billed to an insurer. This will act as an incentive to take care of their card, as it has real monetary and convenience value. Facilities to back-up the index, whether on a card or in a deposit box, can be widely provided using public (e.g. using bank ATMs) or private services so that data should never be lost. And if all else fails, the patient should know where their recent records were created and be able to re-populate the data onto a new token, albeit at some personal inconvenience. So there really is no downside.

An obstacle often raised is that of the costs for implementing and maintaining such an arrangement. Having the right information in the right place at the right time has the potential to avoid a substantial proportion of the adverse care events that consume so much of the health budget (12), so providing resources to fund required infrastructure. Issuing card tokens is cheap: the task could even be outsourced to an organisation (e.g. bank) that does this all the time. The record identifier stripping, checksum insertion and data type tagging can readily be undertaken automatically and simply by a routine, which then writes the event summary, tag details and URL to the patient card. Storing data on the web is cheap: such organisations as YouTube and Facebook store vast quantities of web-based data and have easy upload arrangements in place. Without undertaking a detailed financial analysis, the system should quickly pay for itself.

Conclusion

There are issues that make the sharing of relevant information between those caring for the same patient difficult. One issue is that of ensuring records are readily accessible whilst at the same time ensuring personal privacy; a second relates to passing control over access and sharing to the patient; and a third concerns storing these records in a form that can readily be imported (‘pulled’) and displayed flexibly by the reader irrespective of whether they have a simple web browser or a more sophisticated electronic records system.

This paper outlines a simple schema whereby healthcare data can be shared flexibly. The data for sharing is encrypted, tagged, and made verifiable and secure against alteration before storage on the web, so creating trustworthiness and assuring privacy. The clinician requiring the data can ‘pull’ it to their workstation, using the linking data provided by the patient, who thus controls access to their personal information. The key to the schema is the use of an index held either on a token (e.g. smart card) or in a secure web location, and controlled by the patient. All the required technology already exists and is in routine and widespread use: it is envisaged that such a system could more than pay for itself through savings arising out of reduced delays and duplications, and avoided adverse events. The system can be implemented without the need for standardisation across electronic records systems, although such agreements would be useful.

Future Work

Patient identification tokens are in use in some locations and in the process of being distributed in many others: these would need to have an application added (if not already present) to hold an index of care events and any associated data (e.g. keys, locations etc.).

There is no special need for the token to be new: any token with sufficient memory (e.g. a bank card) could be registered and used. Card readers are already widely distributed, and where individuals do not have their own at home, they could make use of public terminals or even of the banking system. The application to process the records for posting to the Internet is of no great complexity as an add-on to a clinic system. Developing and agreeing the record data types and sub-types, and their associated XML tags will be necessary, as well as developing a browser add-on that can import and display flexibly the XML marked-up records.

Corresponding Author

Roderick Neame
Adjunct Professor Health Informatics,
University of Queensland St Lucia Campus, Brisbane QLD Australia
Email: roddyneame@taskcare.com
URL: www.health-informatics.co

References

[1] Kellermann AL, Jones SS. What It Will Take To Achieve The As-Yet-Unfulfilled Promises Of Health Information Technology**.** Health Affairs, 2013; 32 (1): 63-68 DOI: 10.1377/hlthaff.2012.0693. Available from: http://networkingdefinition.bringthegame.info/153/more-changes-in-health-care-needed-to-fulfill-promise-of-health-information-technology/

[2] Wikipedia. Public Key Cryptography. [Internet] (cited 17 January 2013) http://en.wikipedia.org/wiki/Public-key_cryptography

[3] Mandl KD, Szolovits P, Kohane IS. Public standards and patients' control: how to keep electronic medical records accessible but private (2001). BMJ 2001;322:283. Available from: http://www.bmj.com/content/322/7281/283

[4] Steinbrook R. Personally controlled online health data – the next big thing in medical care? NEJM 2008 358:16. Available from: http://www.ipacohio.org/Websites/ipac/Images/LinksResources/NEJMelectronicrecords.pdf

[5] Liu CT, Long AG, Li YC, Tsai KC, Kuo HS. Sharing patient care records over the world wide web. Int J Med Inform 2001 May: 61 (2-3): 189-205. Available from: http://libir.tmu.edu.tw/bitstream/987654321/30231/1/2001-Sharing+Patient+Care+Records+Over+The+World+Wide+Web.pdf

[6] Sankar P, Mora S, Merz J, Jones N. Patient Perspectives of Medical Confidentiality. J Gen Intern Med. 2003 August; 18(8): 659–669. Available from: http://www.ncbi.nlm.nih.gov/pmc/articles/PMC1494903/

[7] Roiter N. Latest wave of healthcare data breaches symptomatic of sloppy security practices. Security Bistro, April 28, 2012 . [Internet] (cited 26 July 2012) Available from: http://www.securitybistro.com/blog/?p=1450

[8] Medindia. An introduction to biomedical ethics. [Internet] (cited 17 January 2013) http://www.medindia.net/education/familymedicine/biomedical-ethics-autonomy.htm

[9] Neame R. Privacy and Health information: health cards offer a workable solution. Inform Prim Care. 2008;16(4):263-70. Available from: http://www.ingentaconnect.com/content/rmp/ipc/2008/00000016/00000004/art00003

[10] Harrison, J and Booth N. Applying new thinking from the linked and emerging fields of digital identity and privacy to information governance in health informatics. Inform. Prim. Care 2003; 11: 223-8 Available from: http://www.ingentaconnect.com/content/rmp/ipc/2003/00000011/00000004/art00007

[11] HL7 Clinical Document Architecture. [internet] (cited 28 December 2012) http://www.hl7.org.uk/version3group/cda.asp

[12] Vincent C. Adverse events in British hospitals: preliminary retrospective record review. BMJ 2001; 322:517. Available from: http://www.bmj.com/content/322/7285/517