Between Scylla and Charybdis: Charting the Wicked Problem of Reusing Health Data for Clinical Research Informatics

Nathan C. Lea[1], Jacqueline Nicholls[2], Natalie K. Fitzpatrick[1]
[1] Institute of Health Informatics, University College London, England
[2] Institute for Women's Health, University College London, England

Summary

Objectives: Recognising dilemmas posed by the sharing and reuse of health data as a classic wicked problem and uncover some current key challenges to clinical research informatics.

Methods: A modified thematic review process including identification of agreed critical research questions, appropriate query terms and search strategy, identification of relevant papers in accordance with inclusion criteria, and authors' co-review of full text papers.

Results: Queries returned 4,779 papers published between January 2014 and November 2017. A shortlist of 197 abstracts was analysed and 18 papers were finally selected for review. Thematic assessment of findings revealed four key challenges: (1) uncertain reliability of consent as a cornerstone of trust due to the limits to understanding and awareness of data sharing; (2) ethical challenges around equity and autonomy; (3) ambitious overly theoretical governance frameworks lacking practical validity; and (4) a clear desire for further public and individual engagement to achieve clearer and more nuanced knowledge dissemination around data sharing practice and governance frameworks.

Conclusions: Understanding the wicked problem of reusing clinically acquired health data for research purposes is essential if clinical research is to benefit from informatics advances. A lack of understanding around the context of data acquisition and sharing undermines the foundations of patient-professional trust. Efforts to protect privacy, where tailoring to specific contexts is a key driver, should support the development of solutions which more adequately honour privacy needs, justify access, and protect equity and autonomy.

Keywords

Clinical research; informatics; privacy; ethics; access and data sharing; trust; consent; public engagement and involvement.

Yearb Med Inform 2018:170-6
http://dx.doi.org/10.1055/s-0038-1641219

Introduction

Advances in research capability heralded by Electronic Healthcare Records (EHRs) and more recently genomics, the Big Data paradigm, and richer data gathering through personal mobile devices bring to the fore new challenges for health data sharing that lie between privacy protection and access to those data. Appropriate balancing of personal privacy with the potential benefits of health data reuse to improve care represents a wicked dilemma.

In early 2016, an editorial in the British Journal of General Practice [1] explored the importance of considering the confidential therapeutic relationship as the context of data acquisition to guide how to govern the uses of data for purposes beyond that context of confidence and trust. The primacy of the therapeutic relationship helps us to appreciate the essence of the wicked problem: it is one of an individual desire for privacy which can be set aside when the individual chooses to do so – which one controls as part of the trusting relationship with his/her healthcare provider. But data reuse beyond that relationship cedes control to unknown others with whom the individual has no relationship at the time the individual originally shared the data. In approaching this survey, we have decided to apply this lens to develop an understanding of the current and evolving challenges.

Nearly two years on, the literature re-emphasises how data is becoming richer, its capture more personal and more revealing. It has also revealed the extent to which mechanisms by which the clinical research informatics community can assure privacy and justify access are not as robust as expected or assumed. Whilst public awareness is increasing and transparency in research and protection are becoming more paramount and mandated under an international law with the forthcoming enforcement of the EU General Data Protection Regulation (GDPR) [2], the objective of this review is to uncover current key challenges to clinical research informatics and emphasise how understanding of these issues for both data subjects and data sharing beneficiaries is a particular challenge so that the clinical research informatics community can start to develop ways of thinking meaningfully about the requirements of solutions.

Methods

As the aim of this review was to survey current approaches to the wicked problem of appropriate reuse of clinical information, the authors adopted a modified thematic review approach focused on understanding the dilemma space in a pragmatic way in order to provide a useful platform to support reasoned debate about the development of solutions. Neither meta-analysis nor extensive critique of the shortlisted literature formed part of the methodology. Rather, the research team focused on the identification of a tractable research question within the 'wicked' dilemma space, the development of a robust search strategy, data gathering, and the synthesis of key findings.

Research Question and Scoping Questions

The authors agreed the core research question would focus on the dilemmas in balancing privacy and access needs when re-using health data for clinical research. Inherent in this question is the recognition that individuals have many changing roles and identities, and that the dilemma goes way beyond a binary view of private individual vs. public citizen. We agreed to consider scoping questions that would help us to narrow search terms, query articulation, and points of relevance as we retrieved and reviewed the literature.

These scoping questions assessed how key concepts were understood in the literature, how well privacy risks and access needs were balanced, and the extent to which consented or unconsented clinical research informatics endeavours had any impact. We started by asking "what is understood by privacy in the legal, ethical, and societal contexts?" and "what is understood by access to data shared to support health research?". Questions further included "how are privacy needs met and do they honour expectations when health data is accessed?", "how well are privacy and access needs balanced in health–related research and is there a trade-off?", and "what bearing does consented and non-consented research have?" to help guide our review.

Search Terms

In considering our research questions and focus, we listed a series of query terms related to the scoping questions and manually cross referenced them with MeSH terms. We first developed a scoping query that would narrow our search for literature on MEDLINE to sharing and reuse of data to support clinical research informatics that included genomics, mobile health, Big Data, clinical data warehouses, and repositories in scope. We focused on clinical, epidemiological, bioinformatics, omics, cohort, and evaluation studies.

In combination with the scoping query, we developed queries that were related to privacy, consent, data protection, information security, trust and trusting relationships, and ethics. We also agreed that public engagement and involvement were an important area to consider in our scoping questions given their significance to transparency, the understanding of public and individual expectations, and trust. These concepts were further expanded using indexed and MeSH terms, and the final set of terms was agreed by each author. These queries are available on request.

Paper Selection, Data Extraction, and Analysis

The queries were run on MEDLINE on 20th November 2017 for manuscripts published between 1st January 2014 and 20th November 2017 in line with the 2018 IMIA Yearbook scope for review. In total, 4,779 publications were returned. Titles were reviewed in the first instance to shortlist according to their relevance to the scoping and research questions, and an agreement was reached on which to discard. No language restrictions were placed on publications as the authors were prepared to seek translators to help translate the papers or a translation may have been otherwise available online, though no publication fell into this category.

Of the 197 manuscripts shortlisted for review on the abstract, it was agreed that the full text article of 18 of the manuscripts should be reviewed. For the first stage of manuscript selection, the lead reviewer (NL) screened articles based on titles and abstracts and only eliminated articles clearly not relevant to the research question. The authors also agreed that the cyber / information security manuscripts retrieved in particular appeared predominantly to be descriptive of technical solutions tested in vitro and provided little insight for the research questions.

After the identification and removal of duplicates, all reviewers independently screened titles and abstracts against the inclusion criteria to identify potentially relevant papers. Articles identified for potential inclusion by the three reviewers were obtained and a final assessment was made by two authors to confirm their eligibility. Consensus between the reviewers was reached by discussion and a final list of articles to be read in detail was made. Where there was disagreement between authors, the authors agreed they would go with the majority view, though no disagreement arose.

The manuscripts selected for potential inclusion were divided and assigned to each author. At least two authors reviewed each of the manuscripts independently to extract data and confirm relevance to the research questions. Then, authors reviewed the data and agreed on findings.

To minimise the risk of bias, data were extracted using a structured approach that included the following information: methods, study design, duration, and setting; aims and inclusion criteria; number of participants eligible, included, and evaluated; participant characteristics (age, gender, career stage); outcomes on which authors collected measurements, including qualitative evidence analysis of participant views where we will report results relevant to this review; and notes on any important limitations of the study.

In synthesising results from the data, we drew broadly on methods for thematic research described by Thomas and Harden [3]. Qualitative data was extracted and the main findings grouped together in order to identify themes which reflect similarities and differences in content to identify the key dilemmas for health data reuse in clinical research informatics.

Results

We describe the 18 papers that were reviewed in detail. The findings from the data analysis have been grouped into the following five themes to identify the key dilemmas for sharing health data to support clinical research. These include Ethical Concerns, Consent and Control, "Hamstrung" Trust: Low Awareness vs. Appetite to Learn and Engage, Legal and Practical Concerns, and Societal Values.

Ethical Concerns

De Lusignan et al. use their literature review and consensus building approach to develop a framework to honour privacy and ethical

concerns around data access [4] where they identify the challenges in bringing together enough experts with the capabilities to honour the requirements of the ethical framework. From the ethical perspective, they promote a series of questions derived from Willison et al.[5]. These questions focus on identifying potential burdens and harms and who must bear them and whether they are justified, whether participant selection is fair and appropriate, whether informed consent is warranted and feasible, what level of engagement is appropriate and what the social implications are of the data access, and what the potential longer-term consequences are.

Research facilitated by mobile apps raise the profile of ethical concerns around data sharing. Moore et al. [6] identify concerns about whether fundamental ethical principles as espoused in the Nuremberg Code are being followed across the board. App-mediated research poses a unique risk to privacy which participants may not be aware of (e.g. around GPS data use). This is a point that was also picked up by Nebeker et al. who reported that participants were concerned about wearing devices that could track their locations along with concerns about impacts on lifestyle privacy and security [7].

Arora et al. further illustrate the ethical concerns with the issues of limited understanding of technology used for mHealth research. They highlight that if people don't understand the technology they won't be able to meaningfully grasp the risks versus benefits. Since it takes time to help educate and develop awareness, mobile health researchers should focus on developing systems that enhance participant privacy [8].

Consent and Control

Nebeker et al. emphasise that "creating a meaningful informed consent process is critical and will likely require involving participants as partners who are willing to review and modify consent language and processes to increase access and understanding" [7].

The qualitative study performed by Spencer et al. [9] illustrates the nuances of consent where participants did not talk about the time varying nature of consent, but rather spoke about the value of using the dynamic consent system to make a one-off decision if they wanted to opt-out.

Dynamic consent also features in the discussion piece by Williams et al. [10], which recognises that dynamic consent implies an e-Infrastructure and whilst it may provide a portal for research participants to consent or revoke consent, it does not provide an ethical framework to account for variances in rights, needs, and understanding of participants.

This raises a key issue around consent related to the points raised by Arora et al. [8] which is also illustrated in the qualitative study performed by Audrey et al. [11] that identifies uncertainties around effectiveness of anonymisation. This causes further doubt around the validity of informed consent as a cornerstone of good governance and the extent to which research participants understand different types of consent and what they are/are not consenting to.

Big Data supported health research casts further scrutiny on consent: Rothstein recognises consent as being something that shows respect for autonomy and dignity of the individual, despite citing other work that consent for Big Data is too expensive and that consent is fetishised. He argues that a robust opt out is ethically equivalent to opt in and that ticking an opt out box is no different than ticking an opt in box in terms of effort [12].

Balas et al. in their literature review of Big Data clinical research identify public concerns around the reuse of data in terms of security and a preference for research on aggregate data only. They find that reuse by clinical professionals and others for the benefit of others (i.e. community health) directly impacts autonomy whether participants would have consented or not. They emphasise that confidentiality is a mandate to protect anything shared in a trusting relationship, but that this can be overridden if sharing the data mean benefits for society that outweigh the individual's interests in keeping it private [13].

In the systematic review by Aitken et al. [14] that focused on public responses to health data sharing and linkage, a large number of studies were identified reporting that assurances of individuals' confidentiality were crucial for public supporting the reuse of health data for research. This was largely associated with anonymisation of data, although the review notes that anonymisation was not an absolute guarantee of confidentiality. The review also reported a recognition that the anonymisation process is imperfect and does not adequately protect privacy.

Autonomy and control were identified in the review as key factors, but there was no clear consensus found on what the control looked like. Respondents also wanted to strike a balance between control and efficiency of research, where consent was seen as a key requirement for research acceptability, but there was some acceptability for practicality and not to hinder research.

Grande et al. [15] highlight the importance of purpose and make the point that whilst consent is emphasised in research, it is important to consider purpose as well because that has a great deal of impact on people's opinions.

"Hamstrung" Trust: Low Awareness vs. Appetite to Learn and Engage

Discussions around consent and participation must be considered while taking into account the identified low awareness of data sharing and use. Spencer et al. highlight that participants reported low levels of awareness of how their personal information is stored and shared for research where there was some fear of the unknowns, and whilst patients were supportive of sharing their anonymised EHR for research, they noted a lack of transparency and awareness around the use of the data making it difficult to secure public trust [9].

The systematic review from Aitken et al. also point to a discernible appetite for more information around research and its governance. Assurances of safeguards were identified as essential for support but there was low awareness of the research practice and ethical processes in place. Knowledge about what was involved with research helps to increase acceptance [14].

The findings by Nebeker et al. in the context of mobile apps add further weight to the need to properly engage with participants and for efforts to educate individuals

so that they can develop an improved ability to make informed decisions around studies that involve mobile sensing technology. Such education about technologies used in research "can reduce barriers associated with a lack of familiarity and, subsequently, increase trustworthiness of the research enterprise" [7].

Legal and Practical Concerns

There are numerous examples of guidelines and codes of practice around legal and practical good practice, perhaps reflecting both the inherent wickedness of the space as well as deficiencies both in the current articulation and use of law and of its role in addressing a wicked issue. De Lusignan et al. provide a set of practical questions from their consensus work for data protection, focusing on identifying responsibility for the accountability of the data and where it will be stored, who will have access to it, whether there is an audit trail to indicate the data were legally obtained, whether it is sufficiently anonymised, restrictions for secondary processing of the data, whether their accuracy be verified and any processing is documented, and whether individuals can opt out of its processing [4].

Eagleson et al. define an 11-point set of guidelines derived from literature pertaining to information security assurance mechanisms, including authentication, audits, access controls based on roles and patient preferences by way of consent, and a patient-controlled privacy policy [16]. They also highlight how transparency can keep researchers accountable and that sharing security dilemmas and solutions can help to keep projects from having to implement the same security mechanisms from scratch.

A qualitative study by Shabani et al. provides insight into concerns around the practical and legal considerations in the US. The study conducted interviews and presents responses from data access regulators (Data Access Controllers or DAC). The study raises key points about practical challenges for managing access to data [17], in particular, the study questions the adequacy of access reviews which look at whether access should be granted for processing.

Qualifications of applicants to process data are assessed but the checks are inconsistent and have a "low bar" where the qualification criteria are fragmented or poorly defined. This involves "googling" applicants and requesting numbers of publications, but this is only to check their affiliation with an organisation. The study raises the point of a bona fide researcher but reports that what this means is not clear. Nebeker et al. also point out that "those who donate their data for research probably expect that the research will be carried out by those with demonstrated competence in the field" [7].

DACs discussed the monitoring of the uses of data once accessed. However, this is limited to checking on publications and intellectual property registrations, while they can request reports from data accessors, but this is hard to regulate since there are many international collaborators. The overriding point was the study "...showed the DAC members and experts are ambivalent about the effectiveness and consistency of the current access review procedures, and oversight process...".

Sanctions were identified as important but there is no professional code that governs researchers. The DACs favour Codes of Practice where sanctions could be specified and they feel that holding the employer responsible for the activities of his/her staff would provide an appropriate basis to apply sanctions.

Aitken et al. identified in their systematic review that security was an important factor, where fallibility of IT systems and human errors are understood to be key risks. There appears to be a tolerance for security risks that participants identified as always being possible because they also valued the benefits of research [14].

Genomics research also raises particular concerns around security. Wang et al. in the US illustrate that genomic data for which identifiers have been removed is unprotected, yet re-identification risks are significant. De-identified data are often obtained using cursory broad consent but these risks can clearly increase as more 'side' data become accessible. Risks are therefore not communicated and are not static so re-identification risk increases with both technological advances in linkage plus increased knowledge about a specific individual offering a more complete picture. This is illustrated by the work of Gymrek and colleagues who managed to identify several male participants of the 1000 Genomes Project using their Y-chromosome markers, their family structure, and online genetic genealogy databases [18].

Societal Values

Liyanage et al. [19] identify an increased risk to privacy inherent with the increased use of data from multiple sources and they illustrate the risks to privacy that need to be balanced with the benefits of data sharing. They work towards the development of a privacy and ethics framework to help strike a balance between those risks and the right to a higher attainable level of health. They recognise the legitimate concerns of citizens to protect their ethics and privacy where their health data is used to improve health sector performance.

The systematic review by Aitken et al. [14] identified concerns of misuses and proliferation of data to third parties and for surveillance, with lower concern over political use. There were concerns over stigma and discriminatory treatment and insurance premium issues for research participants, but they found that private company uses and profitability were acceptable if it was for the public good. Many studies reported that concerns to personal privacy were balanced with the recognition of the importance of societal benefits anticipated. Two studies reported that some participants prioritised societal benefits over personal privacy.

Spencer et al. found that 98% considered the altruistic benefits of sharing data outweighed the risks of data falling in wrong hands [9]. Additionally, they emphasise the importance of trust, finding that most participants felt confident NHS Trusts manage records securely and anonymity was preserved when data was used for research. A majority expressed satisfaction towards governance arrangements in the UK NHS but acknowledged no system can be completely secure. A small minority described concerns about risks to their privacy speculating that those with more sensitive health conditions may be more guarded with what happens to their health information.

Discussion

Table 1 summarises the challenges the review has uncovered related to the themes that the review identified and the solutions we propose for how to tackle them based on the review and developments, particularly around preparations for the new General Data Protection Regulation in the European Union (GDPR).

The review has highlighted some of the challenges for sharing and reusing health data across the evolving landscape of clinical informatics research which can be distilled into four key challenges: (1) uncertain reliability of consent as a cornerstone of trust due to limits to understanding and awareness of data sharing; (2) ethical challenges around equity and autonomy; (3) ambitious overly theoretical governance frameworks lacking practical validity; and (4) a clear desire for further public and individual engagement to achieve clearer and more nuanced knowledge dissemination around data sharing practice and governance frameworks.

The challenges we have identified are not new concepts but the landscape has increased in complexity and uncertainty with the advances in omics and mobile health research, and even key messages with regards

Table 1 Summary of challenges and proposed solutions

Item no.	Challenge identified in literature	Theme	Proposed solutions
1.	Are fundamental rights protected, particularly if participants are unaware of the risks involved where they don't understand the underlying technology (e.g. location data and privacy)?	Ethical Concerns	Engagement with participants and ensuring they understand the nature of the research and new technologies being studied.
2.	Are there enough experts to oversee ethical concerns and dilemmas, particularly as data reuse and management increase and evolve?	Ethical Concerns	Share ethical oversight with greater reliance on informed and engaged professionals and lay representations (see item 6 below).
3.	Meaningfulness of consent if participants do not understand what they are consenting to.	Consent and Control / "Hamstrung" Trust	Closely follow the developments for GDPR preparation across EU, particularly with regards appropriateness of consent as a legal basis for processing data where other bases may be more appropriate; where consent is sought to participate in research data reuse must be understood, particularly with regards transparency.
4.	Purpose of data reuse is not always clear and awareness of reuse is limited	Consent and Control / "Hamstrung" Trust	Ensure that engagement with participants and public clearly articulates purposes and benefits in a way that can be understood and challenged, as required by GDPR for transparency; be clear on the rights of participants, particularly with the new GDPR requirements.
5.	Demonstrating accountability and transparency around data reuse in a provable and clear manner	Legal and Practical	GDPR now requires proof of compliance and accountability as well as transparency with regards data use so preparations to support this are key to handling these challenges within and outside the EU. Approval mechanisms must be rigorous and must be independently verifiable, and should always be transparent.
6.	Competence of data users as a means to demonstrate trustworthiness and assure compliance	Legal and Practical	Training and education around data protection must focus more on contextual understanding and less on "tick-box" learning, where understanding must be demonstrated in line with GDPR and other developments; assessment of competence must also have formal, independently verifiable processes that are clearly defined and transparent.
7.	Reliability of anonymity is more limited than often expected, particularly given it has been proven false	Consent and Control / Legal and Practical / "Hamstrung" Trust	Claims around anonymity of datasets must be verified more clearly and should take into account the context of data processing and the time-limited nature of anonymity.
8.	Risks to privacy are often more severe than expected, particularly with regards the fallibility of IT systems and where more data sets are processed for reuse purposes	Societal Values	Clearly articulate the risks involved so individuals can make an informed decision about balancing their contribution to societal benefits and risks against their own privacy and wellbeing. Ensure preparations for GDPR compliance are in line with its requirements around transparency and accountability for handling data, and its emphasis on the need for security.
9.	Some contexts, particularly where sensitive health issues are concerned, carry greater anxieties around privacy and more reluctance to participate.	Societal Values	Understand the context of data acquisition and reuse, and ensure the concerns and anxieties of the studied population are taken into account.

to using basic EHR data for research are not getting across. A recurrent theme was that our ability to share data has outstripped our ability to clearly and appropriately articulate ways of managing it. An analysis by van Staa et al. highlights the importance of transparency, public involvement, evidence of credible science and clear benefit, and public confidence that data is held securely and appropriately anonymised are critical to the success of Big Data research [20].

These points are consistent with our findings across the spectrum of clinical research informatics, but we propose that the lack of understanding around health data sharing, lack of clarity on what people are consenting to and the resulting problems in earning individual trust lie at the heart of the wicked problem. For this reason, we elected to review the literature with a high-level lens, focused as it is on the relationship of trust between the private individual and healthcare providers when the individual chooses to relinquish some of that privacy.

The review returned several publications that claim to handle technical security issues that threaten privacy, propose methods to limit risks to privacy and provide a basis for offering other technical or procedural solutions, but do not take into account the problems of consent and understanding risks in context, particularly with respect to genomics. This represents a significant ethical challenge in terms of there being a very strong case for participants to be informed. The review recognises that the threats to data security vary according to context and data type, and highlights the special issues raised by the integration of genomic data.

Rather than focussing on which consent mechanisms are most favoured by the public, it may be more valuable to focus on how relationships are built up (and conversely eroded) and how trust can be facilitated within research and data reuse through data linkage processes including through public engagement and involvement. Studies overwhelmingly suggest an appetite for more information about current research practices and uses of data. The public should not be conceived only as subjects of information provision but there remains a public interest and enthusiasm for more meaningful forms of public engagement or involvement.

With the new GDPR coming into force across the European Union and garnering wider international interest, the clinical research informatics community now has a mandatory goal but also an opportunity to tackle some of the wickedness. GDPR enhances the need for transparency and meaningful dialogue around data processing with data subjects, a better understanding of the role of consent in research, and offers the basis for a more meaningful discussion with participants based on new and clearer rights. Those who rely on data sharing should embrace these as opportunities and prepare to be more accountable for their practices as guided by the regulation and engage more meaningfully and more publicly. This will go a long way in navigating the problem between sharing data and advancing care through clinical research informatics.

We suggest that our findings are the basis for a much-needed full systematic review aimed at gaining further traction on developing a far more nuanced understanding of the wicked problem(s) we are grappling with than has hitherto been the case. It seems reasonable that this be performed after GDPR has been enforced, as this is likely to have an impact on our understanding of the challenges and how they can be handled. It further seems pertinent to incorporate changes in protocols, attitudes, concerns, and security risks as genomics research continues to thrive.

Conclusions

The review has spotlighted some key dilemmas that need to be addressed by the clinical and health informatics community. Whilst elaborate security and ambitious ethical frameworks exist, there remains a lack of understanding and knowledge for patients which undermines the ethical basis under which such frameworks operate.

Without real transparency or understanding, trust is harder to earn, and despite a recognition of the importance of consent and confidentiality, consent cannot be the cornerstone of trust if people do not understand what they are consenting to and cannot meaningfully assert control over what they choose to share when they subsequently and inadvertently relinquish it later on without their knowledge.

No one perspective – be it technical, legal, societal, or ethical– holds the answer to this wicked problem. It is a combination of perspectives that we have focused on so that we can start to understand how to make the data sharing problem less wicked.

Clinical research informatics advances rely on tackling the wicked problems of data reuse. A lack of understanding around the context of data acquisition and sharing undermines the foundations of trust and attempts to protect privacy, but there is however a clear public enthusiasm to learn more about the research technologies and governance, and to engage with the endeavours. This, with the additional legal requirements of GDPR and the challenges set by the new advances in genomics research, provides an opportunity for enhanced collaborations between agencies and an engagement tailored according to specific contexts and individual needs so that the research community will learn to honour personal privacy more meaningfully and effectively, justify access with more transparency and detail, and better protect equity and autonomy.

References

1. Lea NC, Nicholls J. Are patient relationships the driver for information governance? Br J Gen Pract 2016;66(648):342-3.
2. European Regulation (EU) 2016/679 of 27 April 2016 on the protection of natural persons with regard to the processing of personal data and on the free movement of such data, and repealing Directive 95/46/EC (General Data Protection Regulation), art. 12
3. Thomas J, Harden A. Methods for the thematic synthesis of qualitative research in systematic reviews. BMC Med Res Methodol 2008;8:45.
4. De Lusignan S, Liyanage H, Di Iorio CT, Chan T, Liaw ST. Using routinely collected health data for surveillance, quality improvement and research: Framework and key questions to assess ethics, privacy and data access. J Innov Health Inform 2016;22(4):426-32.
5. Willison DJ, Ondrusek N, Dawson A, Emerson C, Ferris LE, Saginur R, et al. What makes public health studies ethical? Dissolving the boundary between research and practice. BMC Med Ethics 2014;15:61.
6. Moore S, Tasse AM, Thorogood A, Winship I, Zawati M, Doerr M. Consent Processes for Mobile App Mediated Research: Systematic Review. JMIR

MHealth UHealth 2017;5(8):e126.
7. Nebeker C, Murray K, Holub C, Haughton J, Arredondo EM. Acceptance of Mobile Health in Communities Underrepresented in Biomedical Research: Barriers and Ethical Considerations for Scientists. JMIR MHealth UHealth 2017;5(6):e87.
8. Arora S, Yttri J, Nilse W. Privacy and Security in Mobile Health (mHealth) Research. Alcohol Res 2014;36(1):143-51.
9. Spencer K, Sanders C, Whitley EA, Lund D, Kaye J, Dixon WG. Patient Perspectives on Sharing Anonymized Personal Health Data Using a Digital System for Dynamic Consent and Research Feedback: A Qualitative Study. J Med Internet Res 2016;18(4):e66.
10. Williams H, Spencer K, Sanders C, Lund D, Whitley EA, Kaye J, et al. Dynamic consent: a possible solution to improve patient confidence and trust in how electronic patient records are used in medical research. JMIR Med Inform 2015;3(1):e3.
11. Audrey S, Brown L, Campbell R, Boyd A, Macleod J. Young people's views about consenting to data linkage: findings from the PEARL qualitative study. BMC Med Res Methodol 2016;16:34.
12. Rothstein MA. Ethical Issues in Big Data Health Research: Currents in Contemporary Bioethics. J Law Med Ethics 2015;43(2):425-9.
13. Balas EA, Vernon M, Magrabi F, Gordon LT, Sexton J. Big Data Clinical Research: Validity, Ethics, and Regulation. Stud Health Technol Inform 2015;216:448-52.
14. Aitken M, de St Jorre J, Pagliari C, Jepson R, Cunningham-Burley S. Public responses to the sharing and linkage of health data for research purposes: a systematic review and thematic synthesis of qualitative studies. BMC Med Ethics 2016;17(1):73.
15. Grande D, Mitra N, Shah A, Wan F, Asch DA. The importance of purpose: moving beyond consent in the societal use of personal health information. Ann Intern Med 2014;161(12):855-62.
16. Eagleson R, Altamirano-Diaz L, McInnis A, Welisch E, De Jesus S, Prapavessis H, et al. Implementation of clinical research trials using web-based and mobile devices: challenges and solutions. BMC Med Res Methodol 2017;17(1):43.
17. Shabani M, Thorogood A, Borry P. Who should have access to genomic data and how should they be held accountable? Perspectives of Data Access Committee members and experts. Eur J Hum Genet 2016;24(12):1671-5.
18. Wang S, Jiang X, Singh S, Marmor R, Bonomi L, Fox D, et al. Genome privacy: challenges, technical approaches to mitigate risk, and ethical considerations in the United States. Ann N Y Acad Sci 2017;1387(1):73-83.
19. Liyanage H, Liaw ST, Di Iorio CT, Kuziemsky C, Schreiber R, Terry AL, et al. Building a Privacy, Ethics, and Data Access Framework for Real World Computerised Medical Record System Data: A Delphi Study. Contribution of the Primary Health Care Informatics Working Group. Yearb Med Inform 2016(1):138-45.
20. van Staa T-P, Goldacre B, Buchan I, Smeeth L. Big health data: the need to earn public trust. BMJ 2016;354.

Correspondence to:
Nathan C. Lea
UCL Institute of Health Informatics
222 Euston Road
London NW1 2DA
Tel: +4477 33 117 359
Fax: +44 20 7679 8002
E-mail: n.lea@ucl.ac.uk