International Journal of Medical Informatics 97 (2017) 43–51

Contents lists available at ScienceDirect

International Journal of Medical Informatics

journal homepage: www.ijmijournal.com

The other side of the coin: Harm due to the non-use of health-related data

Kerina H. Jones[a,*], Graeme Laurie[b], Leslie Stevens[b], Christine Dobbs[a], David V. Ford[a], Nathan Lea[c]

[a] Data Science, Swansea School of Medicine, Swansea University, Swansea SA2 8PP, UK
[b] School of Law, University of Edinburgh, Old College, South Bridge, Edinburgh EH8 9YL, UK
[c] Centre for Health Informatics, University College London, Gower Street, London WC1E 6BT, UK

ARTICLE INFO

Article history:
Received 31 March 2016
Received in revised form
15 September 2016
Accepted 22 September 2016

Keywords:
Health data
Non-use
Omission
Harm

ABSTRACT

Introduction: It is widely acknowledged that breaches and misuses of health-related data can have serious implications and consequently they often carry penalties. However, harm due to the omission of health data usage, or data non-use, is a subject that lacks attention. A better understanding of this 'other side of the coin' is required before it can be addressed effectively.
Approach: This article uses an international case study approach to explore why data non-use is difficult to ascertain, the sources and types of health-related data non-use, its implications for citizens and society and some of the reasons it occurs. It does this by focussing on issues with clinical care records, research data and governance frameworks and associated examples of non-use.
Results and discussion: The non-use of health-related data is a complex issue with multiple explanations. Individual instances of data non-use can be associated with harm, but taken together, they can describe a trail of data non-use that may complicate and compound its impacts. There is ample indirect evidence that health data non-use is implicated in the deaths of many thousands of people and potentially £billions in financial burdens to societies.
Conclusions: Harm due to the non-use of health data is difficult to attribute unequivocally and actual proven evidence is sparse. Although it can be elusive, it is nevertheless a real problem with widespread and serious, if largely unquantifiable, consequences. The most effective initiatives to address specific contexts of data non-use will be those that: firstly, understand the pertinent sources, types and reasons for data non-use in a given domain in order to meet the challenges and create appropriate incentives and repercussions; and secondly, are cognisant of the multiple aspects to this complex issue in other domains to keep benefits and limitations in perspective, to move steadily towards socially responsible reuse of data becoming the norm to save lives and resources.

1. Introduction

In 2014, the (UK) Nuffield Council on Bioethics Working Party on Biological and Health Data and the Wellcome Trust's Expert Advisory Group on Data Access commissioned a review of evidence relating to harm resulting from uses of health and biomedical data [1]. The main focus was on harms resulting from data uses, but harms due to omission of data usage (or non-use of data) were also considered. The review took a multi-faceted approach, using hard and soft evidence strands from legal sources, the grey literature and social media, and merging the evidence for analysis [1]. However, little/no actual evidence of harm due to the non-use of data was revealed by the searches used in the review, except that two cases concerning interference in human rights were uncovered in the legal strand. One focussed on a request for access to children's health data and the other on the non-disclosure of data on exposure to radiation on Christmas Island in 1958 [1]. This lack of evidence was no real surprise because the searches were focussed on finding proven cases of harm, and it would be challenging at best to determine with a high degree of confidence whether an instance of harm was truly due to the non-use of data, or if its causes were otherwise.

It is commonly acknowledged that health-related data routinely collected as part of everyday practice, or generated as part of a

* Corresponding author.
E-mail address: k.h.jones@swansea.ac.uk (K.H. Jones).

http://dx.doi.org/10.1016/j.ijmedinf.2016.09.010

research study, have great potential to improve patient care, citizens' lives and professional services. However, this potential is often thwarted by data non-use. There is much publicity about data losses and misuses of data, but what about the other side of the coin – harms due to the non-use of health-related data? There is a lack of evidence about the reasons these data are not being used, an associated lack of appreciation of the harms arising from their non-use, and little indication of the scale of the problem. Therefore, this paper focusses on the non-use of health-related data. It sets out to identify and discuss some important types of data non-use to help clarify why harms due to this issue are not more evident. By using international case study examples it will illustrate the types of problem that can arise through the non-use of health-related data and the implications of this for citizens and society. It will touch upon some of the current initiatives to address data non-use, and by exploring the apparent reasons for non-use, will highlight the remaining challenges for effective data use.

2. Approach

A crucial part of the approach used in this article is to focus on harm due to the non-use of health-related data, as distinct from the benefits due to proper data usage. This is because it would be inaccurate to invert the latter and effectively equate the two. That is, it could be dangerously misleading to postulate that benefits resulting from the use of data would not have been realised, and that the opposite outcomes would have occurred, if those data had not been used. Whilst the complexities and constraints of the legislative and regulatory governance frameworks are undoubtedly relevant here, some other, rather fundamental, factors are at play in the non-use of data. Using a case study approach, issues with clinical care records and research data will be considered before moving on to governance frameworks. This will illustrate sources and types of non-use, their implications, possible reasons they occur and challenges to be addressed. It is appreciated that clinicians, researchers and governance professionals will have an awareness of issues, particularly within their own domains, but one of the novel aspects of this study is to consider a more holistic view of data non-use and its impacts. This is the first known study to address the topic of health-related data non-use in this way. The article is written from a UK standpoint, but this is a global issue, and therefore includes international case studies.

3. Results and discussion

3.1. Clinical records

In 1995, the UK National Audit Office (NAO) published a report entitled '*Setting the Records Straight*'[2] which noted numerous problems with the keeping of paper case-notes. Among the hospitals studied, 12 of the 16 kept multiple sets of case-notes for some patients, which could lead to confusion in administering care. Among 121 clinics, only two-thirds of case-notes were at hand for immediate use, and although most were located in adequate time, on some occasions (up to 3%) the search was fruitless and the patient was unable to receive their consultation. This has serious implications for the continuity of patient care, and may force a delay in surgical procedures or other interventions because patient history cannot be verified. It carries professional risks to the duty of care of the clinical team in not being able to make informed decisions. It also imposes an unnecessary financial burden due to wasted time for staff and patients. Missing case-notes can bias clinical audit, thus skewing the information used to monitor and advance clinical practice. For example, an audit of antenatal risk factors found that 6.4% of the case-notes were missing. Although this sounds like a relatively small proportion, the suggestion was that this was non-random due to clinicians holding onto interesting cases for research or further discussion [3]. As well as entire case-notes going missing, individual test and procedure results, and sometimes episodes of care can be missing from the file. This again may delay timely care leading to poorer outcomes and subject the patient to duplicate risky invasive processes. It also wastes public money and staff time [2].

Since the publications of this NAO report two decades ago, there have been considerable advances in the use of electronic clinical systems in healthcare. Nevertheless, the UK is still a long way from having a comprehensive electronic patient record, let alone being able to use and share it effectively. Swansea University hosts the UK Multiple Sclerosis (MS) Register [4]. When the Register was being established in 2009, a survey of clinical recordkeeping methods in NHS Neurology clinics across the UK was carried out. Of the 47 respondent clinics (N = 83), 5 still used paper records only, 8 stated that they used a word processor package, and only 10 reported using an MS-specific clinical IT system [5]. Purposive action was needed in order to facilitate data collection for the MS Register, and an open-source clinical system was adapted and made available to participating sites. There does not seem to be any reason to assume that Neurology is vastly different to most other disciplines, and so the significance of this is that the pace of change towards electronic systems is slow, and that without the use such systems, the effective use of data is hampered. Although there are moves towards increasing the use of clinical systems, data sharing and research opportunities in the UK NHS [6,7], these remain largely aspirational for many care practitioners. The main focus for health practitioners is on delivering optimal patient care, and with the current high demands, staff are inevitably limited in the effort they can dedicate to other pursuits without strategic-level decisions to provide sufficient funding, training and time [8,9]. Until these issues can be addressed effectively, problems due to the non-use of case-note data will continue.

However, even when a clinical IT system is in place, there are issues that impact on data availability for use. The traditional position is of data in silos, such that primary care data are generally available only within the practice and not systematically shared with hospitals. Indeed, even within a given hospital, data are often still held on administrative or departmental systems that may or may not communicate with each other. A team of healthcare professionals may be involved in an episode of care, some of whom may enter data into the system, and some of whom may record data on paper to be transcribed later by an administrator. Although it may never be intended that every piece of information should reside in the electronic system, this does introduce the possibilities of error and non-entry of important data. It is also the case that when an electronic system is implemented, a judgement call has to be made on the bases of relevance and resources as to how much back data are entered into the system. Furthermore clinical systems may have usability deficiencies and data entry constraints that prove difficult to work with resulting in major problems such as were seen with the EPIC system at Cambridge University Hospitals NHS Foundation Trust (England) [10,11]. Thus there are issues of data quality and completeness within individual systems to contend with before we consider interoperability problems which can limit combining information from different systems. Without this, the data are still in silos, albeit now electronic ones. A case in point is that of a vulnerable little boy who died in 2011 following systematic abuse. There were multiple visits and reports to the GP practice, health visitor, community paediatrics and emergency departments. Each instance was isolated as the data on presentations elsewhere were not available. Because of this, instances were not seen in context and the problems were not identified in time to save the life of the child [12].

Among the classic difficulties in data compatibility are differing formats and data structures which may inhibit data integration, and different coding systems which limit semantic interoperability. The power of data is only realisable if they can be made accessible across clinical and scientific communities and national borders. Challenges such as these are already the focus of multi-disciplinary, multi-national initiatives [13]. The problem is widespread and long-standing as often data systems are not created with sharing in mind at the outset. For example, in the UK primary care services use Read codes to record diagnoses [14], whereas hospital settings often use the International Classification of Disease (ICD) and OPCS nomenclatures [15,16]. These differences require a means of translation in order to interpret information from one system to another when seeking to provide the best patient care. There is a move to standardise coding systems to promote interoperability by means of the Systematized Nomenclature of Medicine Clinical Terms (SNOMED CT). This system provides a comprehensive clinical terminology and it has been adopted in over 50 countries [17]. It takes considerable financial resource, time and effort to implement a standard coding system, but it is essential that information is consistent and transferable if data are to be used optimally for individual patient care and, beyond that, to maximise their usefulness in studies for wider benefits.

A major problem that ensues from the non-use of clinical data in all its guises is misdiagnosis. The incidence of delayed, missed and incorrect diagnoses has been estimated at between 10 and 20% [18] Although the majority of these errors are of little consequence or instances are rectified in good time, it has been estimated from autopsies in the US that 40,000 to 80,000 deaths per year are accounted for by misdiagnosis, with many more suffering injury [19]. Many reasons may contribute to this problem, but data non-use is likely to be causal in at least some of these cases. As well as reasons already described, lack of proper interpretation of available data presents another form of non-use. It has been noted that most medical training programs have no dedicated curriculum on diagnosis, assuming instead that the skill will be passed from senior to junior staff [20]. With the increasing volume and complexity of data in medical records and charts, it is an even more pressing issue that optimal use of available data is enabled by ensuring that clinicians are trained in the necessary interpretive skills. This is critical for patient care and to rein in the rising costs of healthcare [20]. Somewhat ironically, insufficient use is being made of available clinical data at the bedside, at least partly due to lack of analysis skills, and in the wider research arena, where the skills might be more readily available, effective use of data is hampered by lack of availability.

3.2. Research data

Though the impacts of research are not immediate, in that it takes time for findings to be translated into policy and practice, the non-use of data in research can have far reaching effects on patient care, the healthcare profession and the economics of health services. There are a variety of reasons why this occurs: some related to the conduct of research and some to the failure to make the findings of research available.

3.2.1. Conduct of research

The issue of data absence within clinical IT systems is an obvious cause of data non-use through non-availability for research. A study of 10,000 electronic health records in New York found that the selection process whereby researchers naturally aim for data completeness can result in systematic bias in analysis. This is because sicker patients tend to have a higher degree of data sufficiency within their records [21]. This is an enduring challenge for studies relying on the reuse use of data, since data items resulting from tests and procedures are simply not present for healthy individuals. Although selection bias is a well-known problem in research, this particular case was chosen because it was shown to have occurred in the clinical records, that is, before the data would reach researchers who would therefore not know the extent of bias before they begin. But unless it is taken into account, the findings may not represent the population from which the sample was drawn, since they will over-estimate the problem and limit external validity [21]. This is the converse of the problem identified in the antenatal clinical audit earlier (Section 3.1) [3], where some of the problematic cases were excluded, thus underestimating the extent of issues to be addressed. Either way, the findings could be misleading and lead to suboptimal recommendations. However, unless individual cases could be obtained, examined and followed up, it would be difficult to assess the extent to which harm due to the non-use of data occurs. Even then, it would be challenging to show cause and effect in this type of non-use.

Within the non-commercial sector, the majority of substantial health and social care research takes place in academia. Researchers often invest considerable time, energy and intellectual effort into gathering, collating and analysing datasets, and there are still few incentives for the onward sharing of data [22]. Researchers may also be under intense pressure to produce high impact outputs for the UK's Research Excellence Framework (REF) and similar initiatives in other countries [23]. Some have argued that this is an unethical practice, since it influences the research that is carried out and what is ultimately published [24].

3.2.2. Availability of research findings

Sometimes although data are used, the research findings are not made available or they are disseminated selectively. As a result, the benefits of the data are not realised and harm may occur as if the data had not been used. Publication bias is a well-known phenomenon, with weaker or negative findings less likely to make it in to the journals. Biomed Central has established the *Journal of Negative Results in Biomedicine* to publish '*unexpected, controversial, provocative and/or negative results in the context of current tenets*' [25]. But an impact factor of 1.47 offers little kudos, and as long as current expectations persist, it is likely that many researchers will have to concentrate their efforts on hitting their targets. As a result, publication bias will remain a source of data non-use, biasing the direction of further work and allowing studies with undesirable findings to be repeated unnecessarily. The resulting waste of time and effort represents an opportunity cost as public money could be better utilised and needless duplicative intrusion into patients' lives be avoided.

The commercial sector invests billions in drug development and clinical trials. Taking into account the high failure rate in drug creation, it is estimated that it costs approximately $5 billion to bring a new drug to market [26]. As well as seeking the best treatments for patients, pharmaceutical companies are, of course, concerned with generating income and protecting their intellectual property. As a result, the information they release about drugs may be biased by being selected to maintain and extend their market share. In his book '*Bad Pharma: how medicine is broken and how we can fix it*', Ben Goldacre states that it is beyond doubt that '*industry-funded trials are more likely to produce positive, flattering results than independently-funded trials*' [27]. For example, a 2010 review of trials suggested that 85% of industry-funded findings were positive, compared to only 50% of government-funded trials. A variety of reasons are proposed for this higher rate of apparent success including: not publishing unflattering results; comparing a new drug against a placebo, or against an inadequate drug at too low a dose; selecting patients without proper randomisation; or using small, specific sample groups [27]. Though these practices are well-known, they represent a form of data non-use with serious consequences as we illustrate with examples here.

There have been some devastating examples where non-use of data due to non-publication of research findings has been linked to harm to individuals. A particularly high-profile example occurred in the UK in 2006, in a first-in-man commercial trial of an immune-modulatory drug referred to as TGN1412 [28]. Six healthy volunteers were administered with the drug and within an hour they began suffering serious side-effects. The Department of Health convened an Expert Advisory Group to investigate the situation and develop recommendations to try and prevent similar occurrences. In response to the question of whether the situation could have been avoided, it transpired that there had been some experience with a similar intervention ten years previously. A researcher presented the inquiry with unpublished data relating to the use of an antibody molecule with parallel effects in a single human subject. No one could have foreseen the significance of this unpublished piece of information, but the final report recommended that the results of every first-in-man trial should be made available to avoid a repeat of the ordeal to which the six volunteers were subjected. However, a review conducted in 2009 showed that the majority of these Phase 1 trials were still not being published, allowing this form of data non-use to continue [29]. At the time of writing, an investigation is underway into a phase 1 commercial trial conducted in France in 2015 which proved fatal for at least one of six participants, with the others hospitalised. After the 2006 UK trial, recommendations were that participants in phase 1 trials should receive the test drug sequentially not simultaneously. Investigations continue, with more information needed, but this recommendation does not appear to have been followed properly [30].

Harm to patients through non-publication of data is exemplified in the use of an antiarrhythmic drug administered in the 1980s to patients who had suffered a heart attack. In this case it is estimated that over 100,000 patients died of a heart attack after taking the drug before it was realised that it was not appropriate for people who did not have arrhythmia. As to whether this disastrous situation was avoidable, it transpired that a small study had been carried out in 1980 in which 9 of 48 men who took the antiarrhythmic drug (Lorcainide) died, compared to 1 of 47 taking the placebo. The drug was dropped for commercial reasons and the findings were not published. Over a decade later the researchers did publish and stated that their results might have provided an early warning [31].

Sometimes there are direct accusations that data have been withheld to the detriment of patients. Dabigatran is a drug developed by Boehringer Ingelheim as one of the new generation oral anti-coagulants for the prevention of stroke. It was approved and adopted in the UK, other parts of Europe, Japan, Canada and the US (2008–2011) at least partly because it was marketed as a single dosage drug. That is, it did not require dose level adjustments based on plasma levels or anticoagulant activity, thus making it simpler to administer and avoiding the monitoring needed in older drugs like warfarin. By 2012 it had achieved blockbuster status (global annual turnover >$1 billion). However, the number of fatal bleeds accumulated, with the US Food & Drug Administration (FDA) receiving reports of 542 deaths and 2367 reports of haemorrhage among patients on Dabigatran; at the time, there was no antidote. Bleeds are a known risk factor with anti-coagulants, but by comparison, warfarin accounted for 72 deaths in the same period. It transpired that the company knew there were serious risks with single dose level administration, but that it withheld this information from regulators for market advantage [32]. Boehringer Ingelheim has since received approval to market a Dabigatran antidote and the drug continues to be used [33].

Lack of transparency and failure to use clinical trial data fully continue to be major concerns [34]. February 2016 saw the call for an investigation into the use of another new oral anticoagulant, Rivoraxaban, developed by Bayer. It was evaluated against warfarin in a clinical trial and it has become evident that data about the device used to monitor the warfarin arm of the trial are unreliable. In a response to requests to share data for independent analysis, Bayer stated that this is not within the current scope of their clinical trial data sharing [35]. Once a drug is on the market, regulators such as the FDA and the European Medicines Agency (EMA) do not have a mandate to take action unless there are safety concerns. Again the risk is serious bleeds if the dose is too high or stroke if the dose is too low. If independent reanalysis goes ahead and findings point to changes in the safety profile of Rivoraxaban, regulators should be able to act and endorse tailored dosing [35]. These examples demonstrate the harms associated with data non-use in commercial trials and the challenges of bringing these to light so action can be taken. Extended delays and failures to disseminate trial results effectively are not limited to the commercial sector, but also occur in academia to the detriment of potential recipients [36–38]. The All Trials campaign, set up in 2013, is a widely-supported international initiative calling for all clinical trials to be registered and their results reported so that standards can be raised, greater benefits achieved and harms avoided in future [39]. But questions remain about adequate repercussions for wilful non-use of trial data as even large fines may not approach the annual revenue that companies gain from drug sales [32], and in the meantime the problem continues and patients are harmed.

3.3. Governance frameworks

This brings us to consider the impact of governance frameworks on the non-use of data. 'Governance frameworks' in this context refers to the body of legislation and regulations relevant to the sharing and use of data within a jurisdiction. They are in place to ensure responsible and appropriate data use, to protect individual privacy and safeguard professionals. However, some of these, such as the Research Governance Frameworks in the UK have long been blamed for hampering the use of data and hindering research [40,41]. Consequently much work is underway to improve the situation through increased data sharing, open data initiatives and streamlining approval processes [7,42–46]. Criticisms have also been levelled at subjective interpretations resulting in over-cautious implementation and unnecessary bureaucracy. These may include lengthy forms and approval processes, complicated steps and numerous parties involved in approval procedures, over-stringent rules on data access, and the lack of clear responsibilities delaying permissions [40,41]. Despite on-going initiatives [7,42–46], there remain huge challenges to be addressed in the use of data, not only for research, but also for service and care planning. It is a commonly-held belief among the public that healthcare and government administrative data are already linked and shared across services [47]. However, this is of course not the case at least in the UK, since data are not even routinely shared and linked across different sectors of the health service for clinical and management purposes. While it is important to protect individual privacy and rights in data access [48], it does mean there is great potential for harm through lack of joined-up information. For example, observationally, UK hospital patients are asked if they are taking any medication before they are treated in hospital. A similar process occurs with other practitioners such as community pharmacists and dentists before they prescribe. A more reliable approach would be to review their primary care record if it were accessible. Medication errors are the single most common preventable cause of adverse events in medical practice [49], and it is easy to argue that more joined-up information could circumvent at least some of these occurrences.

However, both individuals and their practitioners may not find combining and sharing identifiable data with other practitioners acceptable without their respective agreements [50]. The

governance frameworks that exist rightly serve to safeguard individual privacy, human rights, privacy and the duties of care and confidence, but even so, legal authority is not the same as social licence [51]. Our understanding of public expectations of therapeutic relationships remains limited, and this is another factor that may exacerbate the data non-use conundrum [52]. But what is known is that most people are willing for their records to be used for medical research providing that data are anonymised [53]. While this is a valuable insight, more work needs to be done to increase trustworthy data sharing.

In order to make use of personal data for research, it is routinely necessary to obtain regulatory approvals, often including participants' informed consent; this is an established and proper part of research ethics and governance frameworks to safeguard individual human rights [54]. However, it can be argued that, in some cases the pursuit of informed consent can disadvantage certain groups, particularly those who are hard to reach or on the edges of society [55]. This is almost certainly the case when addressing '*wicked*' problems associated with abuses in childhood, which may only become apparent in young adults. These include psychosocial issues, school failures and drop-outs, risk-taking behaviours, substance misuse and juvenile crime. A strong argument can be made that, as such problems require the best data, insisting on consent for observational studies using existing data is a failure of duty [56]. A similar argument in relation to bias due to consent was evidenced by comparing baseline and follow-up data from GP and hospital records on patients who did, with those who did not, consent to an intracranial malformation study. The results showed that consenters were different in ways that could not have been estimated in advance. The authors concluded that those who oversee medical research are harming public health by imposing greater constraints on patient data than those required by the law [57]. However, in considering the requirement for consent, it is important to distinguish between non-interventional studies that rely on observational data only, compared to invasive studies such as phase I trials where new medications or procedures are being tested. But even in apparently low-risk studies there are risks to be mitigated, such as re-identification of individuals with the possible repercussions that may bring. Achieving the optimum balance where rights and benefits are in accord is challenging and unlikely to be constant over time, with dependencies on factors such as the state of the art in research, infrastructure and societal perceptions.

Within the UK the specific regulatory mechanism that permits the use of identifiable data without consent, by over-riding the common law duty of confidentiality for important medical and research purposes, comes under section 251 of the NHS Act 2006 [58]. Applications are administered by the Confidentiality Advisory Group of the Health Research Authority [59], but anecdotally, the success rate is low and applicants are strongly encouraged to pursue the consent route or to use anonymous data where at all possible. In some cases, this does not compromise the purpose, but in others it does. Nevertheless there are many success stories; for example, the lead author is engaged in a study of vulnerable young mothers and their children that has successfully obtained s251 support. Without this, the study would have been biased because the participants could not be followed-up reliably. Even so, the waiver was granted for matching purposes only: that is, so that the study data could be linked to hospital and education data, and it was a condition of the approval that the resulting de-identified data would be accessed via a Safe Haven [60,61].

The UK Information Governance Review published in 2013, commonly referred to as Caldicott 2, proposes that the duty to share information can be as important as the duty to protect patient confidentiality [6]. This principle is embodied in the UK Care Act of 2014 with a new legal requirement for research regulators to work together both to protect research participants and promote research [62]. Caldicott 2 also includes considerable discussion on Safe Havens as '*specialist, well governed, independently scrutinised accredited environments*' as the sole location where the linkage of personal confidential data from more than one organisation for any purpose other than direct care should take place. This helps to highlight the innovative work that has been underway for some years on the development of Safe Havens for access to de-identified linked data for research. The Secure Anonymised Information Linkage) SAIL and (Scottish Health Informatics Programme) SHIP systems are examples where approved researchers can access data for research within a secure environment [63–65]. However, although Safe Havens hold great promise for using a wealth of valuable, extensive health-related datasets, they are still subject to limitations [66]. Not least among these is the constraint of using anonymised data only. Many types of study do not require identifiable data to produce benefits, as proven by the rich array of important research outputs produced via anonymous data linkage research; a good example being work conducted via the long-established Western Australia data linkage unit [67]. However, sometimes anonymised data are simply not sufficient, particularly in studies aiming to benefit individuals directly and to be used to support clinical decisions.

Moreover, anonymised data still require control measures because of re-identification risks: it has been established that individuals can sometimes be re-identified from data purported to have been anonymised [68]. Because of this, it is good governance practice that such data are curtailed before being made available for research. This may take the form of aggregation or suppression of records, or in some cases perturbative methods may be employed [69]. But this can create another form of data non-use and produce bias in research findings, because individual records or items within those records, where they occur in unique or low-copy numbers, may be amended or omitted to mitigate perceived risks of re-identification. Often, the more unusual records and extreme data items are the most interesting for research, since they may underlie pressing health problems. So this well-intentioned practice can limit external validity as the application of results will gravitate to treating the mean characteristics and phenotypes in the population. Furthermore, conditions for the use of anonymised data often include no reversal of the process to lead back to individuals. This, therefore, precludes the opportunity to highlight a worrying indicator in their data, since to do this would require permission to hold identifiable information, to the frustration of medical researchers and clinicians and the detriment of patients. So although the use of anonymised data is proving to be invaluable in data linkage research, they are still subject to forms of data non-use.

As with the vulnerable little boy who died following abuse (Section 3.1), it can be argued that too much emphasis is placed on the risks of implementing data sharing initiatives, rather than on the potentially enormous risks of not making data available [12]. A powerful example of this is that in 2013 in the US, the Centres for Medicare and Medicaid Services began withholding claims with a substance-use disorder from research datasets because of privacy concerns. It is estimated that this disorder is associated with over 60,000 deaths in the US each year, that it costs society hundreds of billions of dollars, as well as untold emotional damage to families. Yet, harm due to breaches of these data is unknown and there is no evidence that the data cannot be properly secured for use in research. Without these data it is felt that research is *'flying blind'* [70]. The perspective that regulatory regimes can result in greater harm than the risks they seek to address is widely known, and this led Australian researchers to coin the term *'privacy protectionism'* [71]. Furthermore, a recent report from the Canadian Council of the Academies states that *'the risk of potential harm resulting from access to data is tangible but low. The level of risk can be further lowered through effective governance mechanisms.'* And *'timely access to data is hindered by variable legal structures and differing*

interpretations of the terms identifiable and de-identified across jurisdictions' [72].

Non-use of health-related data in relation to governance frameworks takes a variety of forms, and can result in wasted time and lost opportunities. The case studies suggest misunderstandings among those administering frameworks may be at play, and that there may be issues inherent within some frameworks. The extent of public support for health-related data usage is an important factor, but one about which knowledge is incomplete and fragmented. As a result, greater constraints may be imposed out of a concern for privacy without necessarily providing better safeguards. Sometimes approval processes can be laborious and complex possibly due to over-caution and a fear of making mistakes in the use of personal data. Even when using anonymised data, concerns about re-identification may result in over-stringent controls being imposed, such that data utility is compromised without necessarily demonstrating improved data security and privacy protection. These issues raise difficult questions about the balance between ensuring privacy and seeking better patient care.

Furthermore, governance frameworks are not static, and new regulations and their interpretations can have positive and negative consequences for data use. A review of consent for data use has been published (UK) which proposes a new model for the use of personal confidential data beyond direct care [73]. (This is part of a wider report referred to as Caldicott 3.) Such data are seen as essential for the running of the UK's health and social care system, to support the provision of high quality care, for improving the safety of care and for protecting public health through research and innovation. The report recommends that (alongside the existing models for consent to research studies) individuals should be given the opportunity to opt out of their identifiable data being used for these purposes, and that this preference would be respected across health and social care settings. Individuals would also have the opportunity to change their mind. This mechanism aims to enable greater data sharing and onward data use, but there are already questions around the practicalities of recording and enacting preferences and the unknown impact of bias due to incomplete datasets. Also, as proposed, it would still not allow individuals to choose which aspects of their data they are willing to share and which not. This may result in individuals who wish to keep only some details of their records private opting out of data sharing completely. It is early days and the proposal is to be put to consultation. Obviously, robust, proportionate governance is imperative to balance data protection and utility, but when it comes to solutions there are no magic bullets [74].

At the same time as great efforts are being made to streamline and simplify governance procedures and to encourage greater data accessibility, other moves may pose a serious threat to current research practice. During its development, concerns were widely expressed that the European Commission's new General Data Protection Regulation (GDPR) would prohibit much medical, epidemiological and social science research due to amendments proposed by the European Parliament's Committee on Civil Liberties, Justice and Home Affairs (LIBE committee) [75,76]. The GDPR was the subject of high profile campaigns and sterling work coordinated by the European Data in Health Alliance [77]. The agreed text was passed at the end of 2015, and it was welcomed because a reasonable compromise has been reached on many of the more worrying issues for research [78]. However some concerns remain, for example, Recital 23 suggests that pseudonymised data should be considered personal data, with pseudonymisation defined in the text (A.4(3b)) as '*the processing of personal data in such a way that the data can no longer be attributed to a specific data subject without the use of additional information, as long as such additional information is kept separately and subject to technical and organisational measures to ensure non-attribution to an identified or identifiable person*' [79]. Depending upon how this is interpreted, it could have practical implications in relation to participant consent for certain types of studies and for data usages that are currently exempt. The GDPR will bring increased financial and other penalties for data breaches (A79(3) [79], and while this move is welcomed from a safety perspective, it places greater pressure on organisations to avoid the penalties. This could feasibly result in an unwanted by-product in the form of greater privacy-protectionism, which could increase data non-use. The GDPR comes into force in 2018 and the work is not over yet[1]: intelligent interpretation and implementation of the Regulation will be essential to ensure that it does not introduce new regulatory burdens for research and undermine sophisticated data linkage and sharing infrastructures, such as Safe Havens. Otherwise its administration could herald a new era of data non-use to the detriment of health and wellbeing.

In combination, the non-use of health-related data has far-reaching but often unquantifiable consequences. It is a deadly serious issue sometimes eclipsed by outrage over privacy violations and security breaches, which may only be due to a vociferous and possibly badly-informed minority. A strong statement on this line comes from a US report which states that the non-use of patient clinical data is a greater risk than misuse, such that: *'[T]he greatest threat, the biggest risk to people with diabetes, or heart disease, or cancer, or HIV/AIDs or any other chronic disease or disability seems not to be from un-authorized sharing or use of their personal health information, rather it is from the failure to share or the inadequate use of that information, and sometimes even valuing protecting privacy over protecting an individual's life, their health, and the health of their families, friends and neighbors*'[80].

4. Recommendations and future research

4.1. Recommendations

A number of recommendations arise from the findings of this study.

- Plans providing additional funding to implement clinical IT systems in healthcare must also include a commitment to support culture changes in clinician practice and recognise the scope for human error. This includes the need for improved training opportunities in data interpretation and analysis for clinicians to enable more effective use of the data already available.
- Clinical IT systems must be fit for setting but also be interoperable to enable communication and data sharing, and this requires time and effort as well as money.
- Initiatives promoting the onward sharing of clinical and research data need to provide incentives for reuse, taking into account the personal and organisational investments in creating datasets and the intellectual property they may represent.
- Within the research domain, there are penalties for data misuse and abuse, but there appear to be few repercussions for wilful or negligent data non-use. This raises questions about appropriate measures, as in the case of commercial companies annual revenue from drug sales can far outweigh current fines imposed [32].
- Governance frameworks and their administration need to be proportionate, so that they protect but do not impose excess burdens in the name of caution. This includes appropriate interpretation and implementation requiring staff training and the expertise of responsible officers.

[1] Questions remain regarding the adoption of GDPR in the UK following the Brexit vote in the EU referendum (23/6/16). However, it is unlikely that there would be significant divergence from the standards of data protection in Europe, and highly likely that there would be equivalence in legislation.

- Governance regimes need to be tailored and proportionate so that they are robust and at the same time do not limit opportunities and data utility for research and patient care. For example, emerging forms of data for reuse such as free-text, image and omic data may present different privacy challenges, and be subject to different public perceptions, than structured micro-data.
- The tension between data privacy and better patient care needs to be rebalanced through greater support for trustworthy initiatives for data reuse. Since it is known that anonymised data reuse has majority support, further work with the public should focus on how data should be shared, with whom and for which purposes with mechanisms to allow individual choice, not simply one-size fits all.
- Although large-scale anonymised data are proving invaluable in research, more should be done towards facilitating the use of identifiable data where it would add particular value.
- More awareness of the many factors that contribute to the trail of health data non-use would be highly beneficial for all concerned with data use, so that healthcare practitioners, clinical coders, clinical system developers, multi-disciplinary researchers, governance professionals and the public understand more about the potential impacts of data non-use *en route* to decision-making for individual care and wider society.

It is imperative that the many factors that contribute to the non-use of health-related data are addressed to move towards socially responsible reuse of data becoming the path of least resistance to save lives and resources.

4.2. Future research

It has not been possible to cover all aspect of health data non-use in this article, but the way has been opened for further in-depth studies, such as:

- The impacts of data non-use on: the individual patient (such as diagnosis, treatment, prescribed medications and health maintenance); the clinical organisation (including hospital performance in various key areas such as infectious disease control, surgical error or readmission rates); a disease-specific or general population across a large region or nation (including the impact on clinical research by regulatory frameworks and their administration imposing inappropriate restrictions on data sharing).
- A comparison of governance frameworks in the UK *versus* a selection of other countries, identifying the most and least productive aspects of such frameworks with specific recommendations for improving existing frameworks or constructing new ones.
- An exploration of what the goal should be in reducing data non-use regardless of its present political or technical feasibility, and to determine if zero non-use is always the right goal or if there is some level or type of non-use that is actually beneficial.
- Working with the public to raise awareness of the various types, causes and impacts of data non-use to gain their perspectives on how to move forward in a socially acceptable way.
- Estimating the contribution of health informatics and information technology to reducing health data non-use. This could include systems that have been developed to address this problem at the patient, organisation, and governmental levels, including the trade-off between costs and benefits, strengths and weaknesses, applications and unintended consequences.

5. Conclusions

This article has used a case study approach to focus on the non-use of health-related data. As such, it has not attempted an exhaustive review nor does it purport to account for all types of health data non-use. By highlighting issues with clinical care records, research data and governance frameworks it has shown why instances of health-related data non-use are not more evident, and by using international case studies it has illustrated the types of data non-use that occur, with their consequences for citizens and society. It is the first known study to have addressed health data non-use in a wider perspective.

Instances of health-related data non-use can be associated with harm, but when taken together the effects may be magnified via a trail of data non-use barriers and bottlenecks, thus posing greater threats by impacting multiple settings. An example of this would be clinical care records to be used for research being subject to factors such as missing elements, transcription errors and clinical system constraints, followed by selection bias, failures in onward sharing, lengthy governance processes with their perceived barriers, lack of data transparency and publication bias before the findings are taken forward to inform policy and patient care.

The most effective initiatives to address specific contexts of data non-use will be those that: firstly, understand the pertinent sources, types and reasons for data non-use in a given domain, in order to meet the challenges and create appropriate incentives and repercussions; and secondly, are cognisant of the multiple aspects to this complex issue in other domains to keep benefits and limitations in perspective to move steadily towards socially responsible reuse of data becoming the norm.

Looking for verified evidence of harm due to the non-use of health-related data is like the blind man in a dark room looking for a large, agile, polymorphic, lethal, black cat that most certainly is there. A better understanding of its nature is required before it can be captured and successfully tamed.

Summary points

What was already known

- Breaches and misuses of health-related data can have serious implications and consequently they can be high-profile
- Harm due to the omission of health-related data usage, or data non-use, has lacked attention
- Previous work has dealt with some instances of data non-use in specific settings or sectors
- This is the first known study to address the wider perspective of health-related data non-use

What this study adds

- Harm due to health-related data non-use can be difficult to ascertain and attribute unequivocally
- There is ample indirect evidence that it is implicated in the deaths of many thousands of people and potentially £billions in financial burdens to societies
- Health-related data non-use can occur as a trail across settings and sectors, compounding its effects
- Better understanding of the types of, and reasons for, health-related data non-use is needed to inform initiatives to address this problem

Author contributions

The study was led by KHJ who also drafted the manuscript. All the authors were involved in (1) the conception and design of the study, literature review, information gathering and interpretation of the findings; (2) revising the article critically for important intellectual content; (3) final approval of the version to be submitted.

References

[1] G. Laurie, K.H. Jones, L. Stevens, C. Dobbs, A Review of Evidence Relating to Harm Resulting from Uses of Health and Biomedical Data, Nuffield Council on Bioethics, 2015 http://nuffieldbioethics.org/wp-content/uploads/FINAL-Report-on-Harms-Arising-from-Use-of-Health-and-Biomedical-Data-30-JUNE-2014.pdf.

[2] National Audit Office, 1995. http://webarchive.nationalarchives.gov.uk/20111203034031/http://www.audit-commission.gov.uk/SiteCollectionDocuments/AuditCommissionReports/NationalStudies/SettingtheRecordsStraight.pdf.

[3] A. Yoong, C. Hudson, T. Chard, Medical audit – the problem of missing notes, Health Trends 25 (13) (1993).

[4] D.V. Ford, K.H. Jones, R.M. Middleton, H. Lockhart-Jones, I.D.C. Maramba, J.G. Noble, L.A. Osborne, R.A. Lyons, The feasibility of collecting information from people with Multiple Sclerosis for the UK MS Register via a web portal: characterising a cohort of people with MS, BMC Med. Inform. Decis. Mak. 12 (1) (2012) 73, http://dx.doi.org/10.1186/1472-6947-12-73.

[5] R.M. Middleton, K. Tuite-Dalton, D.V. Ford, J.G. Noble, L.A. Osborne, C. Maramba, H. Lockhart-Jones, K.H. Jones, Clinical System Usage in NHS Specialist Neurology Sites (poster Presentation) RIMS Conference, Brighton, June, 2014, p. 2014.

[6] The Information Governance Review, 2013. https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/192572/2900774_InfoGovernance_accv2.pdf.

[7] Academy of Medical Sciences, 2012. https://www.acmedsci.ac.uk/viewFile/51dd839626e27.pdf.

[8] K. Cresswell, A. Sheikh, Organizational issues in the implementation and adoption of health information technology innovations: an interpretative review, IJMI 82 (5) (2013) 73–86.

[9] M. Lluch, Healthcare professionals' organisational barriers to health information technologies-a literature review, IJMI 80 (12) (2011) 849–862.

[10] J. Viitanin, H. Hyppönen, T. Lääveri, J. Vänskä, J. Reponen, I. Winblad, National questionnaire study on clinical ICT systems proofs: physicians suffer from poor usability, IJMI 80 (10) (2011) 708–725.

[11] Digital Health Intelligencei, EPR implementation led to 'catastrophic loss of confidence http://www.digitalhealth.net/clinical_software/47217/epr-implementation-led-to-'catastrophic-loss-of-confidence.

[12] G. Iacobucci, Sharing care data could save lives of vulnerable children, hospital leader says, BMJ 348 (2014) g1953.

[13] L. Field, S. Suhr, J. Ison, W. Los, P. Wittenberg, D. Broeder, A. Hardisty, S. Repo, A. Jenkinson, Realising the full potential of research data: common challenges in data management, sharing and integration across scientific disciplines, 2013 https://zenodo.org/record/7636/files/Common_Challenges-v3.pdf.

[14] Health and Social Care Information Centre. http://systems.hscic.gov.uk/data/uktc/readcodes.

[15] World Health Organisation. http://www.who.int/classifications/icd/en/.

[16] Health and Social Care Information Centre. http://systems.hscic.gov.uk/data/clinicalcoding/codingstandards/opcs4.

[17] Health and Social Care Information Centre. http://systems.hscic.gov.uk/data/uktc/snomed.

[18] E.S. Berner, M.L. Graber, Overconfidence as a cause of diagnostic error in medicine, Am. J. Med. 121 (suppl. 5) (2008) S2–S23.

[19] L. Leape, D. Berwick, D. Bates, Counting deaths from medical errors [letter reply], JAMA 288 (19) (2002) 2405.

[20] M.L. Graber, R.M. Wachter, C.K. Cassel, Bringing diagnosis into the quality and safety equations, JAMA 308 (12) (2012) 1211–1212.

[21] A. Rusanov, N.G. Weiskopf, S. Wang, C. Weng, Hidden in plain sight: bias towards sick patients when sampling patients with sufficient electronic health record data for research, BMC Med. Inform. Decis. Mak. 14 (2014) 51, http://dx.doi.org/10.1186/1472-6947-14-51.

[22] Expert Advisory Group on Data Access, Establishing Incentives and Changing Cultures to Support Data Access, 2014 http://www.wellcome.ac.uk/stellent/groups/corporatesite/@msh_peda/documents/web_document/wtp056495.pdf.

[23] P.-B. Brutscher, S. Wooding, J. Grant, Health Research Evaluation Frameworks: An International Comparison, RAND, Europe, 2008 http://www.rand.org/content/dam/rand/pubs/technical_reports/2008/RAND_TR629.pdf.

[24] D. Shaw, Unethical framework. Times Higher Education, 2012 https://www.timeshighereducation.com/news/unethical-framework/418830.article.

[25] Journal of Negative Results in Biomedicine. http://www.jnrbm.com/about.

[26] Forbes, The Cost Of Creating A New Drug Now $5 Billion, Pushing Big Pharma To Change, 2013 http://www.forbes.com/sites/matthewherper/2013/08/11/how-the-staggering-cost-of-inventing-new-drugs-is-shaping-the-future-of-medicine/.

[27] B. Goldacre, Bad Pharma: How Medicine Is Broken and How We can Fix It, Fourth Estate publishers, 2013, ISBN-13: 978-0007498086.

[28] H. Attarwala, TGN:1412 From discovery to disaster, J. Young Pharm. 2 (3) (2010) 332–336 http://www.ncbi.nlm.nih.gov/pmc/articles/PMC2964774/.

[29] E. Decuilier, A.-W. Chan, F. Chapuis, Inadequate dissemination of phase 1 trials: a retrospective cohort study, PLoS Med. 6 (February (2)) (2009) e10000034, 17.

[30] D. Butler, E. Callaway, Scientists in the dark after French clinical trial proves fatal, Nature 529 (2016) 263–264, http://dx.doi.org/10.1038/nature.2016.19189.

[31] A.J. Cowley, A. Skene, K. Stainer, J.R. Hampton, The effect of lorcainide on arrhythmias and survival in patients with acute myocardial infarction: an example of publication bias, Int. J. Cardiol. 40 (2) (1993) 161–166.

[32] D. Cohen, Dabigatran: how the drug company withheld important analyses, BMJ 349 (2014) g4670, http://dx.doi.org/10.1136/bmj.g4670.

[33] BoehringerIngleheim, Investigational Antidote for Pradaxa (Dabigatran Etexilate Mesylate) Receives FDA Breakthrough Therapy Designation. Press release, 2014 http://us.boehringer-ingelheim.com/news_events/press_releases/press_release_archive/2014/06-26-14-boehringer-ingelheim-investigational-antidote-pradaxa-dabigatran-etexilate-mesylate-fda-breakthrough-therapy-designation.html.

[34] B. Goldacre, Are clinical trial data shared sufficiently today? BMJ 347 (2013) f1880, http://dx.doi.org/10.1136/bmj.f1880.

[35] D. Cohen, Rivoraxaban: can we trust the evidence? BMJ 352 (2016) i575, http://dx.doi.org/10.1136/bmj.i575.

[36] F. Godlee, Data transparency is the only way, BMJ 352 (2016) i1261, http://dx.doi.org/10.1136/bmj.i1261.

[37] M.L. Anderson, K. Chiswell, E.D. Peterson, A. Tasneem, J. Topping, R.M. Califf, Compliance with results reporting at ClinicalTrials.gov, NEJM 372 (March) (2015) 1031–1039, http://dx.doi.org/10.1056/NEJMsa1409364, 12.

[38] STAT. https://www.statnews.com/2015/12/13/clinical-trials-investigation/.

[39] All Trials. http://www.alltrials.net/.

[40] G. Elwyn, A. Seagrove, K. Thorne, W.Y. Chung, Ethics and research governance in a multicentre study: add 150 days to your study protocol, BMJ 330 (2005) 847, http://dx.doi.org/10.1136/bmj.330.7495.847.

[41] J.C. Dumville, J. Watson, P. Raynor, D.J. Torgeson, Research governance: a barrier to ethical research? QJM (2004) 113–114, http://dx.doi.org/10.1093/qjmed/hch026.

[42] The Farr Institute of Health Informatics Research. http://www.farrinstitute.org/.

[43] The Global Open Data Initiative. http://globalopendatainitiative.org/.

[44] Open Data Roadmap for the UK – 2015. http://theodi.org/roadmap-uk-2015.

[45] MRC. Data Sharing Policy. http://www.mrc.ac.uk/research/research-policy-ethics/data-sharing/policy/.

[46] Global Alliance for Genomics and Health. https://genomicsandhealth.org/.

[47] Ipsos MORI Social Research Institute, 2014. Dialogue on data. https://www.ipsos-mori.com/DownloadPublication/1652_sri-dialogue-on-data-2014.pdf.

[48] HM Government, Data Protection Act 1998. http://www.legislation.gov.uk/ukpga/1998/29/contents.

[49] European Medicines Agency. http://www.ema.europa.eu/ema/index.jsp?curl=pages/special_topics/general/general_content_000570.jsp.

[50] G. Perera, A. Holbrook, L. Thabane, G. Foster, D.J. Willison, Views on health information sharing and privacy from primary care practices using electronic medical records, IJMI 80 (2) (2011) 94–101.

[51] P. Carter, G.T. Laurie, M. Dixon-Woods, The social licence for research: why *care.data* ran into trouble, J. Med. Ethics (2014), http://dx.doi.org/10.1136/medethics-2014-102374.

[52] N. Lea, J. Nicolls, Are patient relationships the driver for information governance? BJGP (2016), http://dx.doi.org/10.3399/bjgp16X685753, July 2016.

[53] Wellcome Trust, Wellcome Trust Monitor Summary Report Wave 3: Tracking Public Views on Science and Biomedical Research, 2016 https://wellcome.ac.uk/sites/default/files/monitor-wave3-summary-wellcome-apr16.pdf.

[54] Health Research Authority. http://www.hra.nhs.uk/.

[55] G.T. Laurie, E. Postan, Rhetoric or reality: what is the legal status of the consent form in health-related research? Med. Law Rev. (2012), http://dx.doi.org/10.1093/medlaw/fws031.

[56] Stanley, F., 2010. Privacy or public good? Why not obtaining consent may be best practice. Volume 7. Issue 2. June 2010. DOI: 10.1111/j.1740-9713.2010.00423.x.

[57] C. Dyer, Stringent constraints on use of patients' data are harming research, BMJ 2007 (2007) 335, 1114.3.

[58] NHS Act, 2006. http://www.legislation.gov.uk/ukpga/2006/41.

[59] Health Research Authority, Confidentiality Advisory Group http://www.hra.nhs.uk/about-the-hra/our-committees/section-251/what-is-section-251/.

[60] Building Blocks 2-6: Evaluating the long-term effectiveness, and the cost and consequences of the Family Nurse Partnership parenting support programme in reducing maltreatment in young children http://medicine.cardiff.ac.uk/clinical-study/building-blocks-2/.

[61] K.H. Jones, D.V. Ford, S. Thompson, R.A. Lyons, The UK Secure eResearch Platform for public health research: a case study, Lancet (2016), accepted for publication 26th August 2016.

[62] H.M. Government, 2014. Care Act http://www.legislation.gov.uk/ukpga/2014/23/contents/enacted.

[63] Secure Anonymised Information Linkage databank. http://www.saildatabank.com/.

[64] K.H. Jones, D.V. Ford, C. Jones, R. D'silva, S. Thompson, C.J. Brooks, M.L. Heaven, D.S. Thayer, C.L. McNerney, R.A. Lyons, A case study of the Secure Anonymous Information Linkage (SAIL) Gateway: a privacy-protecting remote access system for health-related research and evaluation, J. Biomed. Inform. (2014), http://dx.doi.org/10.1016/j.jbi.2014.01.003.

[65] N. Sethi, G.T. Laurie, Delivering proportionate governance in the era of eHealth: making linkage and privacy work together, Med. Law Int. 13 (2–3) (2013) 168–204 http://www.ncbi.nlm.nih.gov/pmc/articles/PMC3952593/.

[66] N.C. Lea, J. Nicholls, C. Dobbs, N. Sethi, J. Cunningham, J. Ainsworth, M. Heaven, T. Peacock, A. Peacock, K.H. Jones, G. Laurie, D. Kalra, Data safe havens

and trust: towards a common understanding of trusted research platforms for governing secure and ethical health research, JMIR Med. Inform. 4 (2) (2016) e22.
[67] University of Western Australia, staff profile of CDJ Holman: https://www.socrates.uwa.edu.au/Staff/StaffProfile.aspx?Person=D%27ArcyHolman&tab=publications.
[68] P. Ohm, Broken Promises of Privacy: Responding to the Surprising Failure of Anonymization, 2009 http://epic.org/privacy/reidentification/ohm_article.pdf.
[69] Information Commissioner's Office, Anonymisation: Managing Data Protection Risk Code of Practice, 2012 http://ico.org.uk/~/media/documents/library/Data_Protection/Practical_application/anonymisation-codev2.pdf.
[70] A.B. Frakt, J.D. Bagley, Protection or harm? Suppressing substance-use data, NEJM 372 (2015) 1879–1881, http://dx.doi.org/10.1056/NEJMp1501362.
[71] J. Allen, C.D.J. Holman, E.M. Meslin, F. Stanley, Privacy protectionism and health information: any redress for harms to health? J. Law Med. 21 (2) (2013) 473–485.
[72] Council of Canadian Academies, Accessing Health and Health-Related Data in Canada, 2015 http://www.scienceadvice.ca/uploads/eng/assessments%20and%20publications%20and%20news%20releases/health-data/healthdatafullreporten.pdf.
[73] National Data Guardian, Review of Data Security, Consent and Opt-Outs, 2016 https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/535024/data-security-review.PDF.
[74] G. Laurie, J. Ainsworth, J. Cunningham, C. Dobbs, K.H. Jones, D. Kalra, N. Lea, N. Sethi, On moving targets and magic bullets: can the UK lead the way with responsible data linkage for health research? Int. J. Med. Inf. 84 (11) (2015), 08/2015.
[75] M.C. Ploem, Proposed EU data protection regulation is a threat to medical research, BMJ 346 (2013) f3534.
[76] L. Stevens, The Proposed Data Protection Regulation and Its Potential Impact on Social Sciences Research in the UK. European Data Protection Law Review, EPDL 2/15, 2015 http://www.lexxion.de/images/pdf/edpl_2015_02-005.pdf.
[77] European Data in Health Alliance. http://www.datasaveslives.eu/.
[78] Wellcome Trust, Vote by European Parliament on Data Protection Welcomed by Research Community, 2015 http://www.wellcome.ac.uk/News/Media-office/Press-releases/2015/WTP060073.htm.
[79] European Commission, General Data Protection Regulation, 2015 http://www.statewatch.org/news/2015/dec/eu-council-dp-reg-draft-final-compromise-15039-15.pdf.
[80] D. St. Clair, Non-Use of Patient Clinical Data a Greater Risk than Misuse, Managed Healthcare Executive, 2008 http://managedhealthcareexecutive.modernmedicine.com/managed-healthcare-executive/news/non-use-patient-clinical-data-greater-risk-misuse?page=full.