

DEBATE

Open Access

# Data Access Committees

Phaik Yeong Cheah<sup>1,2,3\*</sup>  and Jan Piasecki<sup>4</sup>



## Abstract

**Background:** Sharing de-identified individual-level health research data is widely promoted and has many potential benefits. However there are also some potential harms, such as misuse of data and breach of participant confidentiality. One way to promote the benefits of sharing while ameliorating its potential harms is through the adoption of a managed access approach where data requests are channeled through a Data Access Committee (DAC), rather than making data openly available without restrictions. A DAC, whether a formal or informal group of individuals, has the responsibility of reviewing and assessing data access requests. Many individual groups, consortiums, institutional and independent DACs have been established but there is currently no widely accepted framework for their organization and function.

**Main text:** We propose that DACs, should have the role of both promotion of data sharing and protection of data subjects, their communities, data producers, their institutions and the scientific enterprise. We suggest that data access should be granted by DACs as long as the data reuse has potential social value and provided there is low risk of foreseeable harms. To promote data sharing and to motivate data producers, DACs should encourage secondary uses that are consistent with the interests of data producers and their own institutions. Given the suggested roles of DACs, there should be transparent, simple and clear application procedures for data access. The approach to review of applications should be proportionate to the potential risks involved. DACs should be established within institutional and legal frameworks with clear lines of accountability, terms of reference and membership. We suggest that DACs should not be modelled after research ethics committees (RECs) because their functions and goals of review are different from those of RECs. DAC reviews should be guided by the principles of public health ethics instead of research ethics.

**Conclusions:** In this paper we have suggested a framework under which DACs should operate, how they should be organised, and how to constitute them.

**Keywords:** Data Access Committee, Data sharing, Ethics, Research ethics, Research ethics committees

## Background

Expectation from health research funders, regulatory agencies, and journals for greater sharing of de-identified individual-level health research data is now increasing [1–5], but the volume of data shared remains low [6]. Arguments in favour of data sharing include maximising the utility of the data, improving research transparency and allowing confirmation of the interpretation of results, with the overall goal of improving science and health [1, 2, 7, 8]. However, many have cautioned that there are potential harms with

regards to data sharing, such as misuse of data, breaching participant confidentiality, group harms of discrimination and stigmatization, as well as exacerbating existing inequalities between researchers in low- and high-income settings [2, 9, 10].

There are ongoing discussions on how best to approach data sharing. The landscape of research ethics is evolving and after the eras of researcher and regulatory paternalism [11], a new paradigm shift is being discussed: an ethical framework for learning healthcare systems [12]. This paradigm brings a few new ideas. It puts into question the difference between research and routine clinical practice: new information technologies alter the nature of medical practice, which has recently become a learning activity where its goals are not limited to benefiting an individual patient, but also embraces the generation of knowledge. The main aim

\* Correspondence: [phaikyeong@tropmedres.ac](mailto:phaikyeong@tropmedres.ac)

<sup>1</sup>Mahidol Oxford Tropical Medicine Research Unit (MORU), Faculty of Tropical Medicine, Mahidol University, 420/6 Rajvithi Road, Bangkok, Thailand

<sup>2</sup>Centre for Tropical Medicine and Global Health, Nuffield Department of Clinical Medicine, University of Oxford, Old Road Campus, Roosevelt Drive, Oxford, UK

Full list of author information is available at the end of the article



of an ethical framework for learning is not to protect individuals, but to promote the common good of an efficient and safe healthcare system. Therefore it has been argued that an ethical framework for learning healthcare systems is public health ethics [13]. By the same token, it has been suggested that data sharing should be governed by the principles of public health ethics, rather than research ethics [14]. This is because public health ethics focuses on public benefit, proportionality, accountability, equity and trust while research ethics tends to focus on consent and individual interests [14]. There are significant similarities between public health activities, epidemiological research and data sharing in terms of goals, benefits and risks associated with these activities. Therefore in this paper we defend a position that data sharing should be guided by the principles of public health ethics than by those of research ethics. We provide reasons for this position in the later part of the paper.

Health researchers and ethicists have proposed that one way to promote potential benefits of data sharing and ameliorate its potential harms would be through the adoption of a managed access approach where requests are channeled through a Data Access Committee (DAC), rather than making data openly available without restrictions [15]. DACs, a formal or informal group of individuals who have the responsibility of reviewing and assessing data access requests [15], have only been developed relatively recently. Many group, consortia, institutional and independent DACs have been set up but there is currently no widely accepted framework under which DACs operate.

This paper aims to suggest a framework for DACs and to advance discussions on what the functions of these DAC should be, how they should be organised, and how to constitute them. While many previous discussions have centred around genomic data, our paper discusses DACs which operate as custodians of all types of health data generated from public funded health research. The sharing of data collected in a clinical context or health administrative data, insurance data is beyond the scope of this paper.

We propose that DACs should have the role of both promotion of data sharing and protection of data subjects, their communities, data producers and their institutions. With these roles in mind, we then discuss the organisation and composition of a DAC. We suggest that DACs should not be modelled after research ethics committees (RECs) and give reasons why. In addition, we suggest that DAC reviews should be guided by the principles of public health ethics instead of research ethics.

## **Main text**

### **Functions of Data Access Committees**

#### ***Promotion of data sharing in the interest of science, data producers, data subjects and their communities***

We think that given DACs central role in data sharing, they have an important role in promoting data sharing.

The primary question for a DAC should be whether the nature, degree, and likelihood of possible public benefits from the data reuse outweigh the nature, degree and likelihood of possible harms to the data subjects, relevant communities, or other stakeholders e.g. primary researchers, their institutions, or countries.

The features of ethical primary research studies are well established. Human subject research guidelines identify that the first question for any research should be whether the proposal is likely to generate scientifically sound, socially valuable knowledge [7, 16]. However, unlike primary research, data access should be granted as long as the data reuse fulfils the criterion of having even a minimal social value [7], and minimal risk to data subjects and their communities (we discuss risks to primary researchers and their institutions later in the paper). A critical and skeptical approach to existing knowledge is an important element of the scientific enterprise. Therefore, reuse of data does not always have to generate new knowledge; it is of significant social value when it verifies results of previous research, for example. Moreover, the analysis of already existing data could spur new scientific hypothesis and guide new research projects [8].

Although data research carries risks, it is not equivalent to enrolling subjects in new clinical or observational research studies. The nature and magnitude of risks arising from secondary data use is different from the nature and magnitude of risks of original studies. This fact is reflected in many existing international and national regulations. For instance, in the European Union General Data Protection Regulation (GDPR) (Art 5(1)(b), Art 89), data-driven health research is given a special derogation and, if researchers demonstrate that a study is “in the public interest” a REC may grant the study a waiver of informed consent [17, 18]. Similar standards are present in the United States Common Rule (45 CFR §46.116, 45 CFR §164.512) and in the Health Insurance Portability and Accountability Act (HIPAA) [19], both regulations grant several derogation for data-driven research involving public benefit, posing minimal risk and allow access to identifiable health information [20, 21].

For data sharing to be successful, there must be a win-win situation for both data producers and secondary data users, their wider teams and institutions. DACs should encourage secondary uses that promote the interests of data producers, such as research that contributes to the goals of their institution [22]. For example, the DAC for an institution with the goal to improve the treatment of malaria should encourage data reuse that ultimately contributes to malaria treatment improvement. However, it does not mean that data should not be shared if the objective of the secondary use conflicts with institutional priorities.

In addition, researchers could reap greater benefits from their data when others are involved in the secondary data analysis (e.g. mathematical modellers). Data sharing can increase scientific productivity. In most cases, it is better to collaborate with others than doing it alone, or planning to conduct the secondary analyses but never getting to it. These extra outputs that arise as a direct result of data sharing will help maximise the utility and cost-effectiveness of studies and in turn increase the overall output and visibility of the institution and its members. This serves as a powerful internal incentive for data producers. If they benefit from data sharing, then data sharing would not be perceived by researchers as another box to tick or an obligation imposed by funders and journals.

Many secondary analyses involve research in the same disease or topic, which could directly benefit the data subjects and their communities, but there are also many secondary data uses that will not have these direct benefits. Some data research will confirm existing results, and others will advance the knowledge of the disease or topic thereby potentially benefiting future patients in downstream research. Other data uses could be for teaching purposes or for shaping a new project. Irrespective of how the data are used, provided that there are some benefits to society, we are respecting the altruism and commitment of data subjects by reusing their data.

#### ***Protecting data subjects, their communities, data producers, their institutions and the scientific enterprise***

Data sharing has triggered concerns for the privacy of the data subjects. Data scientists have proved on multiple occasions that datasets that were thought to be anonymized i.e. “personally identifiable information (PII) is irreversibly altered in such a way that a PII principal can no longer be identified directly or indirectly” [23], could be linked with other public health data to identify the specific data subject [24, 25]. However, there is scarce evidence that these individual data breaches have resulted in individual harms [26]. In fact, many of the potential harms apply to the data community, rather than individuals. These potential harms apply even when data are anonymized, because of potential group harms be it by geography, disease or ethnicity. Some have voiced concerns over potential harmful uses that could result from stigma and discriminatory uses by employers or insurance companies [14, 27]. Hence, although the primary role is to promote data sharing, DACs should also be aware of potential group harms and when such risks are more than minimal, data reuse should not be allowed. These potential group harms sometimes depend on the economic and cultural context of data subjects. They are more worrying in places that have a history of

targeting and discriminating against minorities, countries that do not provide their citizens with universal access to healthcare and where access to healthcare depends on private insurance or ability to pay. We recognize that that DACs do not necessarily know what these group harms will be. There will be risks of underestimating or misidentifying potential harms to data subjects and their communities. These risks can be minimized by careful engagement with research communities during primary research [28, 29], e.g. consultation with community leaders and community advisory boards [30–32].

Protection of data subjects also entails protecting their rights. DACs should make sure that the shared data do not contain any personally identifiable information, and that data will be used within the scope of broad consent provided by subjects. In the case of old datasets where broad consent for sharing had not been obtained, DACs should adhere to the criteria set out by CIOMS 2016: the secondary use offer important otherwise unobtainable information, has social value, and poses minimal risks to the subjects, and that it would be impractical or prohibitively expensive to contact subjects for their consent for secondary use [7].

There have also been worries that data sharing can disadvantage data producers and potentially disincentivise primary research [10]. This would be detrimental to the research enterprise and scientific progress. In order to prevent this, DACs should provide guidelines, within constraints of funder and regulatory requirements, on when specific conditions of access should be put in place. These could include recognition requirements such as authorships, acknowledgements or standard citations. In some cases, collaborations may be necessary, especially where interpretation of the data requires the experience of the primary researchers and an in-depth understanding of the context. In addition, an institution may have exclusive access periods, requirements for benefit sharing, preferential access provisions (e.g. to collaborators) and embargo periods. In addition, DACs should mandate when formal data access agreements to specify terms of access should be signed and if a cost recovery or cost sharing mechanisms should be put in place.

#### **Establishment, composition and procedures of a Data Access Committee**

##### ***Establishment***

The influential Council for International Organization of Medical Sciences (CIOMS) 2016 guidelines recommend that “when data are stored, institutions must have a governance system to obtain authorisation for these data in research” [7]. In addition, the guidelines state that “the ethical acceptability of broad informed consent relies on proper governance”. Governance of data, which includes

data access mechanisms is ideally outlined within the institutional, group or departmental data sharing policy. We suggest that DACs should be established within institutional and legal frameworks with clear lines of accountability, terms of reference and membership.

Some suggest that DACs should be independent of the institution to avoid any conflicts of interest. Indeed many independent DACs exist for this reason such as the MalariaGEN and Managing Ethico-social Technical and Administrative DACs [33, 34]. To motivate data sharing, we must recognise that sharing data might reveal sensitive information not only about data subjects but also about researchers, healthcare providers and/or their institutions which might cause harm or embarrassment [35]. However, the argument for DACs to be institutional, instead of completely independent is ultimately a practical one. If institutions reserve the final authority regarding data sharing decisions, then they will be more willing to share their data. Institutions are the custodians of data and they should act on behalf of research participants, who have consented to broad research reuse of their data. Institutional DACs are then accountable to both their home institution and their research participants. It is unclear who independent DACs are accountable to.

### **Composition**

In order to fulfill the functions of a DAC as described in the previous section, a DAC should consist of a reasonable number of members, each covering multiple relevant areas of expertise. For DACs of large research groups, departments or institutions, ideally there should be members representing senior management, data management, ethics, relevant research areas and potentially a data sharing advocate. It is also desirable to have independent members to address the issue of conflicts of interest and to prevent “data hoarding” by researchers internal to a study.

The CIOMS guidelines state that DAC should have “representation from the original setting” [7]. In the context of large clinical studies or institutional DACs that review multiple studies, it is not feasible to have “representatives of the original setting” serve as members on the DAC. However, having members who are familiar with the context or contexts where the research is conducted is necessary. Some data reuse may require consultation with study investigators, country or community representatives. DACs can also consult on an ad-hoc basis with people familiar with the community or data subjects where necessary.

### **Application procedures**

In order to promote data sharing, there should be transparent, consistent, simple and clear procedures for data

request and data access. The approach to the review of applications should be proportionate to the potential risks involved and streamlined because DACs approve or disapprove data already collected, rather than new research studies. Reviews should be guided by the data sharing policies of institutions or pre-agreed terms in the case of independent DACs; and DAC reviews, as we argue in the next section, should not be guided by review criteria adopted by RECs. Elements of a DAC review should include among other things who is applying, what are the objectives of data reuse, exactly what data are requested, anticipated benefits and potential risks.

### **Why Data Access Committees should not be modelled after research ethics committees?**

One can ask a question, should DACs be modelled after the RECs (in the US context, known as Institutional Review Boards) system that reviews new human subject research? There are a few organizational and ethical reasons for devising a different *modus operandi* for DACs and these reasons pertain to: organizational culture; goals of review; ethical framework of review; accountability to the host institution.

#### **Organizational culture**

Conceptual analyses and empirical studies reveal that there is an inbuilt adversarial relationship within the system of a research ethics review [36–38]. If the main function of RECs is to protect research subjects, then there is an implicit assumption that research poses risks, and puts the burden of proof that this is not the case on researchers. As such, researchers do not perceive RECs members as ethics advisors, but as judges and punishers: researchers have to prove they have good intentions [37].

Regulations give RECs a function of protecting research participants by ensuring that they are provided with proper information and that research projects have a favourable risk-benefit ratio. However, there is no a universal standard of how exactly the RECs should function. Some suggest that RECs should check, if research protocols are consistent with certain ethical or legal codes rather than with abstract ethical principles [39]; while others argue that RECs should “perform ethically informed code consistency review” [40]. As result of these disagreements, over time the function of protection has evolved and RECs have acquired new functions, some of them intended, some were a consequence of institutional logic and gradual legalisation of ethics reviews [36]. Hence RECs guard research integrity and quality by filtering “bad science” projects, and protect community by screening wasteful and dangerous studies [37]. In this sense, members of RECs often perceive themselves as acting on behalf of communities. Moreover, when one

looks at a REC from a wider perspective and takes into consideration the fact that their function is grounded in democratically enacted law, these institutions could also have a function of providing political and ethical legitimacy either to a single study or to biomedical research in general [41]. Proliferation of functions coupled with inconsistencies among RECs in multicenter studies have resulted in growing bureaucratic and financial burden on researchers. This has led to a recognition that an ethics review might be in certain instances an overprotection of research subjects that leads to underprotection [12, 42]. A REC delaying a low-risk study, may at the same time expose patients to substantial risks, for instance in a situation when a post-marketing study would provide evidence for serious adverse effects of already approved drugs [43]. This recognition has been reflected in the regulations that do not require a full-ethics review of low-risk studies and the introduction of central reviews for multicenter studies.

Although an adversarial relationship between RECs and researchers may be legitimate in riskier clinical trials, an adversarial relationship between DACs and primary researchers or secondary users is inconsistent with the desired goals and functions of DACs. As we have argued, DACs should promote reuse of research data. DACs should be part of new research culture that helps in promoting scientific progress. Therefore, the main function of DACs is not defined in adversarial terms of “protection”, and DACs should be conceived as an institution’s tool for realising its goals. Those who apply for access to data should not be perceived as a potential danger, but as potential collaborators.

#### **Goals of review**

A REC comprising of researchers, lawyers, ethicists, nurses, patient representatives and community representatives is intended to bring a diverse perspective on scientific and ethical aspects of a study involving human beings. The goals of an ethics review are to discern those aspects of a study that were overlooked or misjudged by a researcher whose perspective might be skewed by conflicting interests, and to ensure that the research complies with specific laws and research ethics guidelines.

A REC may work in accordance with either a panel of professionals or a jury board [44]. In the case of a professional panel, a REC adjudicates from the position of a professional and objective judge; in the case of a jury board, a REC makes its decision from the point of view of a reasonable layperson. The main goal of ethics review is to protect research participants, and although it is the principal investigator and sponsor who bear the ultimate responsibility for the well-being of research subjects, RECs are at least morally responsible too.

RECs review new research studies and any secondary data research that requires ethical approval which can differ from jurisdiction to jurisdiction. DACs review data access requests for secondary uses. These uses may be for secondary data research but could also be for teaching purposes, to confirm the findings of the original analyses or other purposes. DACs roles should not include full review of the secondary research such as methodology of the secondary research and the statistical approaches. That is the job of RECs.

The secondary use of data already collected differs significantly from conducting clinical research. A secondary data user does not interact with research subjects; data research does not require additional diagnostic tests or examinations and the possible risks to an individual are often limited to privacy breaches and group harms. DACs have different objectives from RECs. DACs are custodians of research data, but this function cannot be understood as a protection from intruders, who may want to have a peek into their treasure trove, but promoters of the beneficial use of data. Institutional DACs should also review the consistency of data use with institutional data sharing policies. The goal of this review should include maximising data research utility either by confirming previously tested results or generating new data, and assessing if there are any potential harms to all relevant stakeholders.

#### **Ethical framework for review**

In its origins, research ethics has been informed by the scandals and tragedies of research, such as atrocities of Nazi doctors, the Willowbrook School Experiment, the Jewish Chronic Disease Hospital and the Tuskegee syphilis studies [11]. Due to these historical events, the main goal of research ethics framework and resulting guidance documents was protection of individual research participants. Moreover, those who were conducting medical research were physicians themselves, whose professional identity was intrinsically connected with an obligation to protect and promote individual patient interests. This is why many research ethics guidelines, such as the Oviedo Convention [45] and the Declaration of Helsinki [16], contain some version of the principle of precedence of individual interests, for example, “the interests of an individual should prevail over the sole interest of society or science” [46]. However, it has become clear recently, that the principles of research ethics cannot be universalised, and not all kinds of research can be held to the same ethical and procedural standards. For instance, an ethical review and an informed consent procedure may seriously impede multicentre epidemiological studies [47]. The goals of epidemiology and public health are different than those of clinical research. The main focus of epidemiology and public health is not an individual

patient, but promotion of population health [48]. Moreover, in public health research it is much more difficult to distinguish research from routine clinical practice [13, 49, 50].

Similar problems of inadequacy of the research ethics approach to multicenter data-driven research has been recently discussed in regards to learning healthcare systems (LHS), in which conducting research is embedded into healthcare practice [12, 42]. The learning process is driven by data that are produced in healthcare practice and then collected and analysed in a search for generalisable knowledge. Efficiency of LHS requires a different ethical approach. Here again protection of an individual is not a priority, since an individual patient is not exposed to risks other than those inherently associated with healthcare practice. The conceptual and ethical framework for LHS also applies to principles of public health ethics rather than research ethics - to weigh public benefits against possible infringements of individual rights [12, 13, 48].

Epidemiology, public health research, and LHSs have a few common characteristics: the benefits and risks pertain to groups rather than to individuals; in many cases the research activity consists of collecting and analyzing vast volumes of data; research ethics standards (e.g. full-informed consent, full-ethics review) are not feasible and can either hamper research or put extra bureaucratic burden on researchers. In all three cases it is clear that the public health ethics approach is applied. Data sharing, at least in two important respects, has similar characteristics to epidemiology, LHS and public health research: the benefits and risks of data sharing pertain to groups rather than to individuals, and data sharing is about accessing and processing vast volumes of data and there are minimal additional risks to a data subject. Taken together, the more suitable ethical approach to data sharing should be that of public health ethics instead of research ethics [14].

#### **Accountability to host institution**

There are at least two models of a REC: independent bodies established by private or public actors, or RECs that are established by research institutions such as universities. Both models provide independent review of research studies. Independence does not mean that there is no institutional links between a REC and a research institution, but it means that the research institution does not influence the workings of its REC which should be independent in its judgments of the ethical standards of studies. Importantly, RECs do not implement any institutional research policies nor are they tasked to promote research.

We think that DACs should play a central role in implementing institutional policies on data sharing. This

is yet another reason for DACs to be institutional rather than independent. The task of a DAC is to balance the goals and policies of its institutions, the goals and interests of those who apply for data and the public good. RECs protect research subjects by applying ethical principles and rules of law; DACs should promote data sharing while mitigating any potential risks, and should be a mechanism to implement institutional data sharing policies. While it is challenging to evaluate the efficacy and efficiency of RECs, it is possible to evaluate those of a DAC, that is by measuring the realisation of data sharing goals and policies.

#### **Strengths and limitations**

Our normative proposal is supported by experience of establishing and coordinating the Mahidol Oxford Tropical Medicine Research Unit (MORU) DAC, which has reviewed over 40 applications since its establishment in January 2016 [51–53]. The MORU DAC has reviewed many types of data requests including data in real time from an ongoing clinical study, from historical trials done without participant consent for data sharing, and from pharmaceutical companies for data from trials conducted in low-resource settings for registering products in developed countries [52, 54].

Our suggestions are primarily focused on DACs of publicly funded large research groups, departments or institutions conducting clinical research. We acknowledge that some research groups may be too small and may not have the resources or skills to establish and run their own DACs. Efforts are underway to provide support for research groups in low-resource settings to establish their own data sharing policies and DACs. We think that future empirical research is needed to verify the feasibility and efficacy our suggestions, and compare them with other models of DACs such as DACs established to review requests for health system data.

#### **Conclusions**

In this paper we have suggested a framework for the functions and establishment of DACs and demonstrated that sharing de-identified health data should be governed by a different conceptual and ethical framework than clinical research involving human beings. DACs should promote the beneficial use of data, while mitigating any potential harms in line with the ethical framework for public health instead of research ethics. We have also argued that the system of ethics review as it is operated by RECs is not suitable for the realizing the ideals of data sharing and therefore should not be a model for DACs.

## Abbreviations

DAC: Data Access Committee; LHS: Learning healthcare system; MORU: Mahidol Oxford Tropical Medicine Research Unit; REC: Research ethics committee

## Acknowledgements

We would like to thank Gaye Proctor for help with editing. PYC gratefully acknowledges the support of the Rockefeller Foundation which provided an ideal setting for her to write this paper (March to April 2019).

## Authors' contributions

PYC conceived the idea of the paper. Both authors contributed equally to the intellectual content of the paper and approved the final version of the paper.

## Authors' information

PYC is a bioethicist, Associate Professor at University of Oxford and the coordinator of the Mahidol Oxford Tropical Medicine Research Unit (MORU) Data Access Committee. She led the development of the MORU data sharing policy and the establishment of the MORU Data Management department. This paper was written during PYC's month-long academic residency at the Rockefeller Foundation Bellagio Centre in Italy.

JP is a philosopher and bioethicist, assistant professor at the Department of Philosophy and Bioethics, Faculty of Health Sciences, Jagiellonian University Medical College, Krakow. His research interests include research ethics and public health ethics.

## Funding

This work is funded in part by a "Wellcome Trust Strategic Award (096527)" and a "Wellcome Trust Research Enrichment, Open Research (106698/Z/14/J)" grant. The Mahidol Oxford Tropical Medicine Research Unit is funded by the Wellcome Trust (106698/Z/14/Z). The funder had no role in the conception, design, analyses and writing of the manuscript.

## Availability of data and materials

Not applicable.

## Ethics approval and consent to participate

Not applicable.

## Consent for publication

Not applicable.

## Competing interests

PYC is an Associate Editor of BMC Medical Ethics. JP declares no competing interests.

## Author details

<sup>1</sup>Mahidol Oxford Tropical Medicine Research Unit (MORU), Faculty of Tropical Medicine, Mahidol University, 420/6 Rajvithi Road, Bangkok, Thailand. <sup>2</sup>Centre for Tropical Medicine and Global Health, Nuffield Department of Clinical Medicine, University of Oxford, Old Road Campus, Roosevelt Drive, Oxford, UK. <sup>3</sup>The Ethox Centre, Nuffield Department of Population Health, University of Oxford, Oxford, UK. <sup>4</sup>Department of Philosophy and Bioethics, Faculty of Health Sciences, Jagiellonian University Medical College, ul. Michalowskiego 12, Krakow, Poland.

Received: 16 April 2019 Accepted: 28 January 2020

Published online: 03 February 2020

## References

- Taichman DB, Sahni P, Pinborg A, Peiperl L, Laine C, James A, et al. Data sharing statements for clinical trials: a requirement of the International Committee of Medical Journal Editors. *Lancet*. 2017;389(10086):e12–e4.
- Bull S, Cheah PY, Denny S, Jao I, Marsh V, Merson L, et al. Best practices for ethical sharing of individual-level health research data from low- and middle-income settings. *J Empir Res Hum Res Ethics*. 2015;10(3):302–13.
- Wellcome Trust. Policy on data, software and materials management and sharing. 2017. <https://wellcome.ac.uk/funding/managing-grant/policy-data-software-materials-management-and-sharing>. Accessed 14 Apr 2019.
- European Medicines Agency. Policy on publication of clinical data for medicinal products for human use (EMA/240810/2013). 2014. [http://www.ema.europa.eu/ema/index.jsp?curl=pages/special\\_topics/general/general\\_content\\_000556.jsp](http://www.ema.europa.eu/ema/index.jsp?curl=pages/special_topics/general/general_content_000556.jsp). Accessed 14 Apr 2019.
- Bill & Melinda Gates Foundation. Open access policy. 2015. <https://www.gatesfoundation.org/How-We-Work/General-Information/Open-Access-Policy>. Accessed 14 Apr 2019.
- Terry RFLK, Oliaro PL. Sharing health research data – the role of funders in improving the impact [version 1; referees: 3 approved with reservations]. *F1000Research*. 2018;7:1641 <https://doi.org/10.12688/f1000research.16523.1>.
- Council for International Organizations of Medical Sciences. International ethical guidelines for health related research involving humans. 2016. <https://cioms.ch/wp-content/uploads/2017/01/WEB-CIOMS-EthicalGuidelines.pdf>. Accessed 14 Apr 2019.
- Bauchner H, Golub RM, Fontanarosa PB. Data sharing: an ethical and scientific imperative. *JAMA*. 2016;315(12):1237–9.
- Cheah PY, Tangseefa D, Somsaman A, Chunsuttiwat T, Nosten F, Day NP, et al. Perceived benefits, harms, and views about how to share data responsibly: a qualitative study of experiences with and attitudes toward data sharing among research staff and community representatives in Thailand. *J Empir Res Hum Res Ethics*. 2015;10(3):278–89.
- Serwadda D, Ndebele P, Grabowski MK, Bajunirwe F, Wanyenze RK. Open data sharing and the global south—who benefits? *Science*. 2018;359(6376):642–3.
- Emanuel E, Grady C. Four paradigms of clinical research and research oversight, the Oxford textbook for clinical research; 2008. p. 222–30.
- Kass NE, Faden RR, Goodman SN, Pronovost P, Tunis S, Beauchamp TL. The research-treatment distinction: a problematic approach for determining which activities should have ethical oversight. *Hast Cent Rep*. 2013;43(S1):S4–S15.
- Piasecki J, Dranseika V. Research versus practice: the dilemmas of research ethics in the era of learning health-care systems. *Bioethics*. 2019;33(5):617–24.
- Ballantyne A. Adjusting the focus: a public health ethics approach to data research. *Bioethics*. 2019;33(3):357–66.
- Kaye J, Hawkins N. Data sharing policy design for consortia: challenges for sustainability. *Genome Med*. 2014;6(1):4.
- World Medical Association. Declaration of Helsinki. 2013. <https://www.wma.net/policies-post/wma-declaration-of-helsinki-ethical-principles-for-medical-research-involving-human-subjects/>. Accessed 14 Apr 2019.
- Regulation (EU) 2016/679 of the European Parliament and of the council 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), Mai 2018. <http://eur-lex.europa.eu/legal-content/EN/TXT/PDF/?uri=CELEX:32016R0679&from=EN>. Accessed 31 Jan 2020.
- Rumbold JM, Pierscionek B. The effect of the general data protection regulation on medical research. *J Med Internet Res*. 2017;19(2):e47.
- USA Health Insurance Portability and Accountability Act of 1996. <https://www.govinfo.gov/content/pkg/PLAW-104publ191/pdf/PLAW-104publ191.pdf>. Accessed 31 Jan 2020
- Gostin LO, Halabi SF, Wilson K. Health data and privacy in the digital era. *JAMA*. 2018;320(3):233–4.
- Cohen IG, Mello MM. HIPAA and protecting health information in the 21st century. *JAMA*. 2018;320(3):231–2.
- Cheah PY. Institutions must state policy on data sharing. *Nature*. 2019; 565(7739):294.
- Joint Technical Committee ISO/IEC JTC 1, Information technology, Subcommittee SC 27, IT Security techniques. ISO (the International Organization for Standardization) and IEC (the International Electrotechnical Commission). 2011. <https://www.iso.org/obp/ui-iso:std:iso-iec:29100:ed-1:v1:en>. Accessed 17 Nov 2019.
- Sweeney L. Simple demographics often identify people uniquely. 2000. (Data privacy working paper 3). Available at <https://dataprivacylab.org/projects/identifiability/paper1.pdf>. Accessed 14 Apr 2019.
- Simon GE, Shortreed SM, Coley RY, Penfold RB, Rossom RC, Waitzfelder BE, et al. Assessing and minimizing re-identification risk in research data derived from health care records. *EGEMS (Wash DC)*. 2019;7(1):6.
- El Emam K, Jonker E, Arbuckle L, Malin B. A systematic review of re-identification attacks on health data. *PLoS One*. 2011;6(12):e28071.
- Rothstein MA. Is deidentification sufficient to protect health privacy in research? *Am J Bioeth*. 2010;10(9):3–11.

28. Barnes KI, Canario JA, Vernekar SS, et al. Equitable data sharing: challenges and suggestions for ways forward [version 1; peer review: awaiting peer review]. *Wellcome Open Res.* 2019;4:172 <https://doi.org/10.12688/wellcomeopenres.15425.1>.
29. Participants in the Community E, Consent Workshop KKM. Consent and community engagement in diverse research contexts. *J Empir Res Hum Res Ethics.* 2013;8(4):1–18.
30. Cheah PY, Lwin KM, Phaiphun L, Maelankiri L, Parker M, Day NP, et al. Community engagement on the Thai-Burmese border: rationale, experience and lessons learnt. *Int Health.* 2010;2(2):123–9.
31. Maung Lwin K, Cheah PY, Cheah PK, White NJ, Day NP, Nosten F, et al. Motivations and perceptions of community advisory boards in the ethics of medical research: the case of the Thai-Myanmar border. *BMC Med Ethics.* 2014;15:12.
32. Pratt B, Lwin KM, Zion D, Nosten F, Loff B, Cheah PY. Exploitation and community engagement: can community advisory boards successfully assume a role minimising exploitation in international research? *Dev World Bioeth.* 2015;15(1):18–26.
33. Murtagh MJ, Blell MT, Butters OW, Cowley L, Dove ES, Goodman A, et al. Better governance, better access: practising responsible data sharing in the METADAC governance infrastructure. *Hum Genomics.* 2018;12(1):24.
34. Parker M, Bull SJ, de Vries J, Agbenyega T, Doumbo OK, Kwiatkowski DP. Ethical data release in genome-wide association studies in developing countries. *PLoS Med.* 2009;6(11):e1000143.
35. Simon GE, Coronado G, DeBar LL, Dember LM, Green BB, Huang SS, et al. Data sharing and embedded research. *Ann Intern Med.* 2017;167(9):668–70.
36. Pettit P. Instituting a research ethic: chilling and cautionary tales. *Bioethics.* 1992;6(2):90–112.
37. Guillemin M, Gillam L, Rosenthal D, Bolitho A. Human research ethics committees: examining their roles and practices. *J Empir Res Hum Res Ethics.* 2012;7(3):38–49.
38. Klitzman R. The ethics police?: IRBs' views concerning their power. *PLoS One.* 2011;6(12):e28773.
39. Moore A, Donnelly A. The job of 'ethics committees'. *J Med Ethics.* 2018; 44(7):481–7.
40. Holm S. The job of 'ethics committees' should be ethically informed code consistency review. *J Med Ethics.* 2018;44(7):488.
41. Piasecki J, Dranseika V, Waligora M. Should epidemiological studies be subject to ethics review? *Public Health Ethics.* 2018;11:213–20.
42. Faden RR, Kass NE, Goodman SN, Pronovost P, Tunis S, Beauchamp TL. An ethics framework for a learning health care system: a departure from traditional research ethics and clinical ethics. *Hast Cent Rep.* 2013;43(S1): S16–27.
43. Gulmez SE, Lignot-Maleyran S, de Vries CS, Sturkenboom M, Micon S, Hamoud F, et al. Administrative complexities for a European observational study despite directives harmonising requirements. *Pharmacoepidemiol Drug Saf.* 2012;21(8):851–6.
44. Levine RJ The institutional review board, in: Steven Scott Coughlin, Tom L. Beauchamp, and Douglas L. Weed, eds. *Ethics and epidemiology.* 2nd. Oxford: Oxford University Press, 2009.
45. Council of Europe. 1997a. Convention for the protection of human rights and dignity of the human being with regard to the application of biology and medicine: Convention on human rights and biomedicine. <http://conventions.coe.int/Treaty/en/Treaties/Html/164.htm>. Accessed 14 Apr 2019.
46. Piasecki J, Waligora M, Dranseika V. Non-beneficial pediatric research: individual and social interests. *Med Health Care Philos.* 2015;18(1):103–12.
47. Rothman KJ. The rise and fall of epidemiology, 1950–2000 A.D. *N Engl J Med.* 1981;304(10):600–2.
48. Holland S. *Public health ethics.* 2nd ed. Cambridge: Polity Press; 2015.
49. Fairchild AL. Dealing with humpty dumpty: research, practice, and the ethics of public health surveillance. *J Law Med Ethics.* 2003;31(4):615–23.
50. Fairchild AL, Johns DM. Beyond bioethics: reckoning with the public health paradigm. *Am J Public Health.* 2012;102(8):1447–50.
51. Mahidol Oxford Tropical Medicine Research Unit. MORU tropical network data sharing policy 2016. [www.tropmedres.ac/data-sharing](http://www.tropmedres.ac/data-sharing). Accessed 14 Apr 2019.
52. Cheah PY, Day NPJ, Parker M, Bull S. Sharing individual-level health research data: experiences, challenges and a research agenda. *Asian Bioeth Rev.* 2017;9(4):393–400.
53. Cheah PY, Jatupornpimol N, Hanboonkunupakarn B, Khirikoekong N, Jittamala P, Pukrittayakamee S, et al. Challenges arising when seeking broad consent for health research data sharing: a qualitative study of perspectives in Thailand. *BMC Med Ethics.* 2018;19(1):86.
54. Cheah PY, Day NPJ. Data sharing: experience from a tropical medicine research unit. *Lancet.* 2017;390(10103):1642.

## Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

**Ready to submit your research? Choose BMC and benefit from:**

- fast, convenient online submission
- thorough peer review by experienced researchers in your field
- rapid publication on acceptance
- support for research data, including large and complex data types
- gold Open Access which fosters wider collaboration and increased citations
- maximum visibility for your research: over 100M website views per year

**At BMC, research is always in progress.**

Learn more [biomedcentral.com/submissions](https://biomedcentral.com/submissions)

