



# Views of Ethical Best Practices in Sharing Individual-Level Data From Medical and Public Health Research: A Systematic Scoping Review

Journal of Empirical Research on  
Human Research Ethics  
2015, Vol. 10(3) 225–238  
© The Author(s) 2015   
Reprints and permissions:  
sagepub.com/journalsPermissions.nav  
DOI: 10.1177/1556264615594767  
jre.sagepub.com  


Susan Bull<sup>1</sup>, Nia Roberts<sup>1</sup>, and Michael Parker<sup>1</sup>

## Abstract

There is increasing support for sharing individual-level data generated by medical and public health research. This scoping review of empirical research and conceptual literature examined stakeholders' perspectives of ethical best practices in data sharing, particularly in low- and middle-income settings. Sixty-nine empirical and conceptual articles were reviewed, of which, only five were empirical studies and eight were conceptual articles focusing on low- and middle-income settings. We conclude that support for sharing individual-level data is contingent on the development and implementation of international and local policies and processes to support ethical best practices. Further conceptual and empirical research is needed to ensure data sharing policies and processes in low- and middle-income settings are appropriately informed by stakeholders' perspectives.

## Keywords

biomedical research ethics, data sharing, data release, data access, research data, research governance, low-income countries, middle-income countries, clinical research, health policy, privacy, systematic review

Policies mandating the sharing of individual-level data from biomedical and public health research are becoming widespread and commanding increasing support from large funding bodies, regulatory agencies, and the pharmaceutical industry (Medical Research Council, 2011; National Institutes of Health, 2003; Nisen & Rockhold, 2013; Research Information Network, 2008; Toronto International Data Release Workshop Authors 2009; UK Data Archive, 2011; Walport & Brest, 2011; Wellcome Trust, 2009). Discussions of data release in the literature highlight the importance of taking seriously both ethical arguments for sharing individual-level data from health research and the need to develop appropriate governance and protections (Antman, 2014; Eichler, Petavy, Pignatti, & Rasi., 2013; Greenhalgh, 2009; White, 2013; Zarin, 2013).

The increasing amount of clinical and public health research being conducted in low- and middle-income settings has the potential to generate datasets of significant value to researchers seeking to address disease burdens in such settings (Manju & Buckley, 2012). Consequently, there is a pressing need to determine how best to develop effective, ethical, and sustainable approaches to data sharing in such contexts. Experiences of data release for genomic research suggest that challenges raised by individual-level data sharing in low- and middle-income settings

will be different in important and morally significant ways from those arising in high-income settings (Parker et al., 2009). In particular, although timely data sharing may be particularly important in low- and middle-income settings to inform effective and urgently needed public health interventions, it is important that data sharing is conducted in a way that does not disadvantage or harm researchers, research institutions, communities, and participants in such settings. Potential benefits and harms of data sharing are discussed in more detail below and summarized in Table 1.

## Potential Advantages of Data Sharing

Sharing individual-level data from clinical and public health research can be valuable in multiple ways. Sharing data allows for independent scrutiny of research results to ensure they are reliable and reproducible, and increases the accountability of researchers (Estabrooks & Romyn, 1995; Godlee & Groves, 2012; Kuntz, 2013; Manju & Buckley, 2012;

<sup>1</sup>University of Oxford, UK

### Corresponding Author:

Susan Bull, Ethox Centre, Nuffield Department of Population Health, University of Oxford, Old Road Campus, Oxford OX3 7LF, UK.  
Email: susan.bull@ethox.ox.ac.uk

**Table I.** Summary of Potential Benefits of and Concerns About Data Sharing.

Reasons to share individual-level data	Concerns about sharing individual-level data
To improve science	May hamper science
<ul style="list-style-type: none"> <li>• Enable verification, replication, and expansion of research results</li> <li>• Address biases, deficiencies, and dishonesty in research</li> <li>• Enable novel analyses and increase study power</li> <li>• Improve meta-analyses</li> <li>• Maximize data use, particularly for datasets that cannot be replicated</li> <li>• Inform research design and research funding</li> <li>• Improve teaching resources</li> <li>• Increase primary data producers' academic profiles and collaboration opportunities</li> </ul>	<ul style="list-style-type: none"> <li>• Reputational harms of critical secondary analyses</li> <li>• Consequences of flawed/poor quality secondary analyses</li> <li>• Reduction of incentives for primary research</li> <li>• Increased incentives to conduct short-term research rather than long-term research</li> <li>• Opportunity costs of curating and sharing data</li> </ul>
To improve health	May hamper health
<ul style="list-style-type: none"> <li>• Inform health care planning and allocation</li> <li>• Inform regulatory review</li> <li>• Improve evidence base for clinical decision making</li> <li>• Improve use of health care resources</li> <li>• Improve patient care</li> </ul>	<ul style="list-style-type: none"> <li>• Effects of flawed secondary analyses on scientific evidence base</li> <li>• Burden of evaluating validity of secondary analyses</li> <li>• Effects of second-guessing regulatory procedures, policies, and processes</li> </ul>
Explicit moral claims	Explicit ethical issues
<ul style="list-style-type: none"> <li>• Importance of maximizing the value and utility of data</li> <li>• Promotion of scientific values</li> <li>• Promotion of best practices in research conduct, analysis, and reporting</li> <li>• Demonstration of respect for research participants</li> <li>• Promotion of the public good</li> </ul>	<ul style="list-style-type: none"> <li>• Protection of participants' privacy and confidentiality</li> <li>• Validity of consent, including broad consent</li> <li>• Potential harms of secondary research for research participants including discrimination and stigma</li> <li>• Researchers' ability to fulfill commitments made to research participants during data collection</li> <li>• Effects of moral distance and limited awareness of the context in which data were collected</li> <li>• Potential impacts on public trust and confidence of conflicting analyses</li> <li>• Balancing the interests of differing stakeholders in data sharing</li> <li>• Making best use of limited research resources</li> </ul>
	Barriers to sharing
	<ul style="list-style-type: none"> <li>• Costs of developing and maintaining appropriate expertise and infrastructure</li> <li>• Curation costs</li> <li>• Ownership, intellectual property rights, and commercial confidentiality</li> <li>• Lack of policies and processes</li> </ul>

Mello et al., 2013; Sieber, 2006). This may be particularly important where there are differing approaches to analyses (Smith, 1994) or where there are concerns that reports of research have been selective, biased, or dishonest (Doshi, Goodman, & Ioannidis, 2013; Gotzsche, 2011b; Rathi et al., 2012; Rodwin & Abramson, 2012; Ross, Gross, & Krumholz, 2012). Sharing data also enables identification of gaps in research and can inform both future research priorities and

research design (Eichler et al., 2013; Gotzsche, 2011a; Sandercock, Niewada, Czlonkowska, & International Stroke Trial Collaborative Group, 2011; Strech & Littmann, 2012). Some datasets of particular value may not be able to be re-collected due to changes in available treatment and disease incidence, and may be useful as reference datasets, particularly in different contexts, such as low- and middle-income settings (Eichler et al., 2013; Sandercock et al., 2011).

Many commentators have discussed the value of conducting novel analyses with shared datasets, including testing innovative statistical methods and alternative analytical approaches (Coady & Wagner, 2013; de Wolf, Sieber, Steel, & Zarate, 2005; Hrynaszkiewicz & Altman, 2009; Pisani & AbouZahr, 2010; Toronto International Data Release Workshop Authors, 2009; Vickers, 2006; Whitworth, 2010). Meta-analyses combining individual-level datasets may provide more reliable results than those based on summary data (Chan et al., 2014; Pisani & AbouZahr, 2010). Meta-analyses may also provide different results from the primary studies and permit examination of topics such as the heterogeneity of treatment effects, subgroup effects, temporal and geographical effects, and identification of rare safety events (Anderson & Merry, 2009; Chan et al., 2014; Dawson & Verweij, 2011; Manju & Buckley, 2012; Mello et al., 2013).

Additional arguments in favor of sharing data are that it can be an efficient and cost-effective means of maximizing the utility of a dataset for research purposes and for teaching and methodology development (Gotzsche, 2011b; Manju & Buckley, 2012; Smith et al., 2014; Walport & Brest, 2011). Increasing use of collected data can also reduce unnecessary duplication of research, which in turn limits potential harms to and burdens on research participants (Eichler et al., 2013; Rani, Bekedam, & Buckley, 2011; Strech & Littmann, 2012).

These claims suggest that data sharing can make a very important contribution to public health, by improving the evidence base used to make regulatory, funding, and clinical decisions, and to make the best use of available resources (Hrynaszkiewicz & Altman, 2009; Hughes, Wells, McSorley, & Freeman, 2014; Rathi et al., 2012; Rodwin & Abramson, 2012; Ross, Lehman, & Gross, 2012). As a routine best practice in research, it may contribute to improving public faith in research and drug regulation, particularly by promoting accountability and transparency in processes where there are potential conflicts of interest (Haines & Gabor Miklos, 2011; Hampton, 2011; Rani et al., 2011; Zarin, 2013).

In addition to the potential of advancing scientific development and health, commentators have discussed ethical imperatives for promoting data sharing. Principles of fairness and reciprocity require data be shared to benefit communities that fund research indirectly and that provide the data on which research relies (Langat et al., 2011; Pisani & AbouZahr, 2010; Strech & Littmann, 2012; Tangcharoensathien, Boonperm, & Jongudomsuk, 2010; Walport & Brest, 2011). In addition, respect for research participants requires that their contributions to research be maximized by making the best use of their data. In particular, expectations that the results of research will be disseminated to advance science must be honored (Gotzsche, 2011b; Mello et al., 2013; Pisani & AbouZahr, 2010; Walport & Brest, 2011).

## Potential Disadvantages of Data Sharing

Numerous concerns and issues about sharing individual-level health research data have been discussed in the literature in addition to potential benefits. A core concern is to ensure that the privacy of participants is protected during secondary uses of data (Castellani, 2013; de Wolf, Sieber, Steel, & Zarate, 2006a; Eichler et al., 2013; Nisen & Rockhold, 2013; Savage & Vickers, 2009; Walport & Brest, 2011; Zarin, 2013). Processes for de-identifying data must be not only robust but also proportionate if the utility of the data is to be preserved (Antman, 2014; de Wolf, Sieber, Steel, & Zarate, 2006b; Eichler et al., 2013). Concerns have been raised about the ability of primary researchers to guarantee that re-identification will not take place (Mello et al., 2013), particularly when reverse engineering and/or the combination of datasets may increase chances of identifying specific participants (Estabrooks & Romyn, 1995; Geller, Sorlie, Coady, Fleg, & Friedman, 2004; Nisen & Rockhold, 2013; Rabesandratana, 2013; Wieseler, McGauran, Kerekes, & Kaiser, 2012).

Although curating and sharing data may make the most efficient and effective use of datasets, preparing data for research and implementing appropriate policies and processes require significant effort, expertise, and resources (Anderson & Merry, 2009; Mello et al., 2013; Rathi et al., 2012; C. T. Smith et al., 2014; Walport & Brest, 2011). Lack of resources needed to share data has been identified as an impediment to data release in empirical research in higher income settings (Mello et al., 2013; Rathi et al., 2012; Reidpath & Allotey, 2001; Savage & Vickers, 2009; C. T. Smith et al., 2014) and as a serious obstacle in low- and middle-income settings (Manju & Buckley, 2012; Pisani & AbouZahr, 2010; Rani et al., 2011; Whitworth, 2010).

Concerns have been raised that if sufficient safeguards are not in place, inappropriately prepared or shared data may hamper, rather than promote, public health (Nisen & Rockhold, 2013; Pearce & Smith, 2011; Piwowar, Becich, Bilofsky, Crowley, & on behalf of the caBIG Data Sharing and Intellectual Capital Workspace, 2008; Spertus, 2012). Data may be misinterpreted, or the subject of biased, inappropriate, or poorly designed studies (Greenhalgh, 2009; Kirwan, 1997; Pisani, Whitworth, Zaba, & Abou-Zahr, 2010a; Rathi et al., 2012; Spertus, 2012; Wieseler et al., 2012). The results of such studies may mislead health care providers and regulators, lead to false hopes or unfounded concerns about treatments, reduce public confidence in research, and result in litigation (Anderson & Merry, 2009; Castellani, 2013; Kuntz, 2013; Mello et al., 2013; Nisen & Rockhold, 2013; Ross & Krumholz, 2013). In addition, incentives for novel biomedical research may be reduced, if secondary data users can “free-ride” on the efforts of those collecting the data (Castellani, 2013; Langat et al., 2011;

Rabesandratana, 2013; Rathi et al., 2012; Ross & Krumholz, 2013; Zarin, 2013).

## **Stakeholders' Interests in Data Sharing**

Sharing individual-level research data will affect the interests of stakeholders in different ways. Primary researchers have interests in conducting initial analyses of data they have collected, and in receiving appropriate acknowledgment for dataset production (Castellani, 2013; Lopez, 2010; Pisani & AbouZahr, 2010; Tangcharoensathien et al., 2010; Whitworth, 2010). Research participants and the communities from which they are drawn have interests in understanding that data may be shared, the consequences of sharing, and ways in which potential harms of sharing can be minimized (Mello et al., 2013; Pearce & Smith, 2011; Piwowar et al., 2008). Research funders have interests in promoting the utility of datasets and may also have interests in commercial exploitation of research results (Anderson & Merry, 2009; Castellani, 2013; Eichler et al., 2013; Kmietowicz, 2013; Mello et al., 2013).

Data sharing policies and process must recognize and respond to the differing interests of stakeholders appropriately if they are to effectively promote the benefits of data sharing and minimize potential harms. Calls have been made for policies and processes for data sharing to be informed by, and developed in consultation with, relevant stakeholders (Manju & Buckley, 2012; Vallance & Chalmers, 2013; Whitworth, 2010). This scoping review sought to map evidence about stakeholders' experiences of data sharing and their perspectives of best practices, particularly in low- and middle-income settings, with the aim of informing future policy development and research agendas (Parker & Bull, 2015).

## **Method**

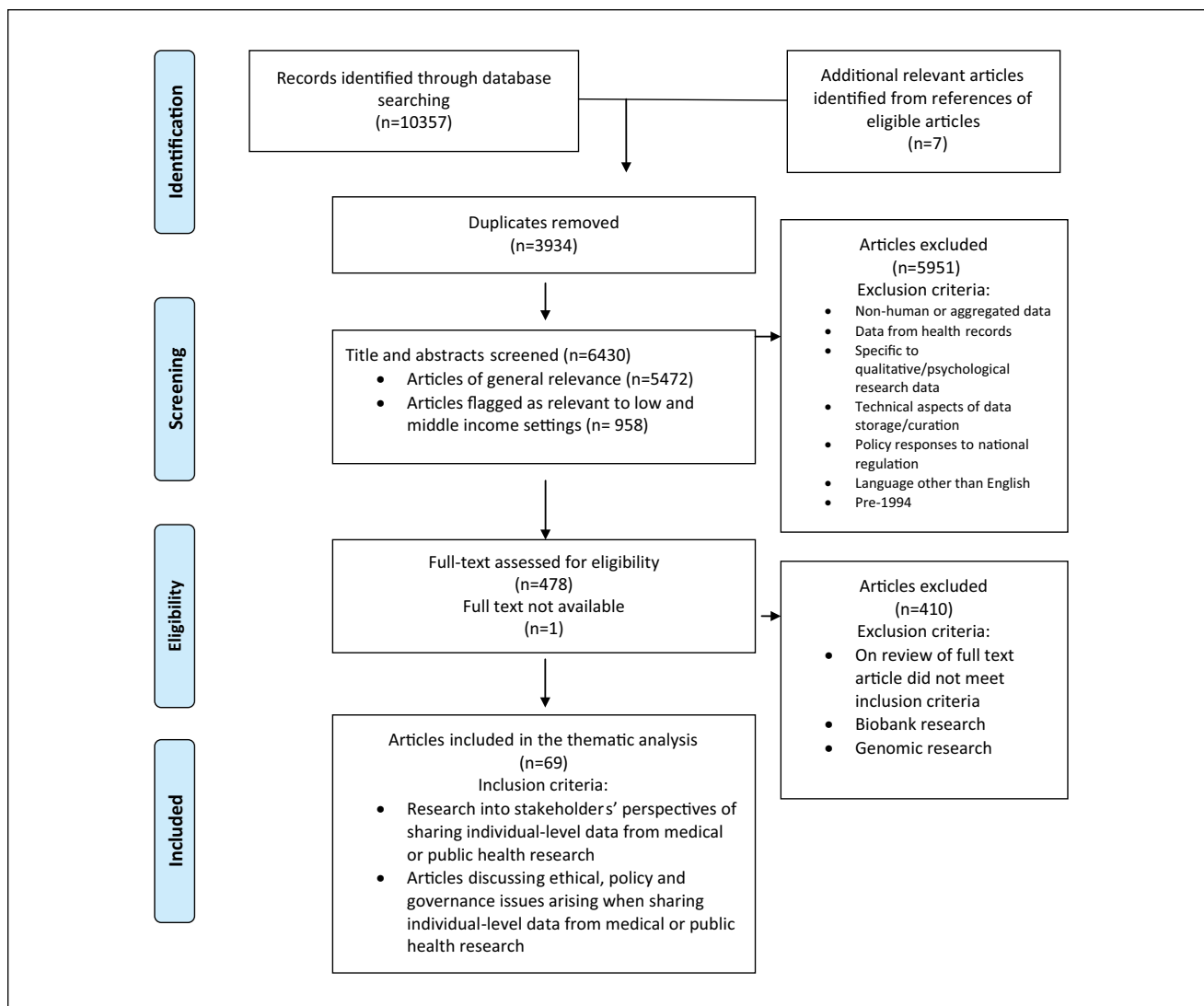
Scoping reviews seek to identify literature relevant to the research objective and may include a variety of research formats and conceptual literature (Arksey & O'Malley, 2005; Armstrong, Hall, Doyle, & Waters, 2011). This study sought to review published literature on stakeholders' experiences of sharing individual-level data from medical and public health research and views of ethical best practices reported in peer-reviewed journals. Inclusion criteria for the study encompassed a broad range of article types, including empirical studies, news articles, opinion pieces, features, editorials, reports of practice, and theoretical articles. The initial search strategies for capturing views in this range of article formats were developed through an iterative process and used a combination of text words and subject headings (see Online Supplementary Materials 1 at <http://jre.sagepub.com/supplemental>).

The following databases were searched for relevant studies: Embase (OvidSP)[1974-present], Global Health (OvidSP) [1973-present], Global Health Library–Regional Databases (Virtual Health Library) [<http://www.globalhealthlibrary.net>], MEDLINE(R) In-Process & Other Non-Indexed Citations and MEDLINE(R) (OvidSP) [1946-present], ABI Inform (Proquest) [1971-current], PAIS International (Proquest) [1977-current], Science Citation Index (Web of Science Core Collections, Thomson Reuters) [1945-present] and WHOLIS (Virtual Health Library) [<http://www.globalhealthlibrary.net>]. The original search was conducted on June 24, 2013, and searches were repeated on December 9, 2013, and June 27, 2014, to update findings. No language or publication date limits were applied. Research relevant to low- and middle-income countries was isolated and grouped using a geographic search filter; however, all references were screened. (The full search strategy for Medline is available in Online Supplementary Materials 2).

The total of 6,430 abstracts identified by the strategy were screened, 958 of which were flagged as being particularly relevant in low- and middle-income settings. A matrix of inclusion and exclusion criteria was developed to inform screening (see Figure 1).

All abstracts were reviewed by a single researcher, with sample of 20% of abstracts being co-reviewed by additional researchers using a trial outline of inclusion and exclusion criteria. After co-reviewing 20% of abstracts, the value of a dual review was assessed. Given the complexity of consistently determining from abstracts which articles contained discussions of relevant ethical, policy, and governance issues, and the large number of abstracts to be screened, multiple review of all the abstracts was considered inefficient. Instead, a single reviewer applied revised inclusion and exclusion criteria consistently, marking articles that were potentially relevant (228), and additional articles that required full text review to determine relevance (246). References from these two categories were imported into bibliographic software (Endnote X6), which was then used to track decisions during a detailed review (King, Hooper, & Wood, 2011).

In scoping reviews, to ensure appropriate identification of the literature, it may be important to adopt an iterative approach to study selection (Arksey & O'Malley, 2005; Armstrong et al., 2011). Following screening of full text articles, five empirical studies of stakeholders' perspectives of sharing individual-level data from clinical and public health research were identified, all of which reported views and practices of researchers and research institutions from high-income settings. During full-text screening, articles focusing on samples and individual-level data from biobanks and genomic research were not routinely excluded, particularly when they reported on perspectives from data subjects or from stakeholders in low- and middle-income settings. A subsequent review of ethical, policy, and governance issues



**Figure 1.** PRISMA 2009 flow diagram of the scoping review.  
Source: Moher, Liberati, Tetzlaff, Altman, and the PRISMA Group (2009).

raised in such papers demonstrated some important differences with issues raised by sharing data from clinical and public health research, and they were subsequently excluded from the review.

The full text of the final 69 shortlisted papers was imported into qualitative data analysis software (NVIVO 10; see Table 2). Descriptive codes were developed to chart perceived advantages of data sharing, barriers and concerns about data sharing, and recommendations for best practices in governing data sharing.

## Results

This section begins by reviewing empirical research into stakeholders' experiences of, and views about, best practices in sharing individual-level data from medical or public

health research. It then outlines the views expressed in articles focusing on ethical, policy, and governance issues arising when sharing such data. It concludes by focusing on issues identified as particularly relevant to best practices when sharing data from low- and middle-income settings.

### Empirical Research

There is very limited empirical research into stakeholders' experiences of sharing individual-level data from clinical or public health research, and their views about best practices when doing so. This review identified five empirical studies, all of which sampled researchers and reviewers from high-income settings. Details of the studies, including the primary findings reported in the original articles, are set out in Table 3.

**Table 2.** Articles Included in the Scoping Review.

Type of Article	Articles of general relevance	Articles of particular relevance in low-and middle-income settings
Empirical research articles	Kirwan, 1997; Rathi et al., 2012; Reidpath & Allotey, 2001; Savage & Vickers, 2009; C. T. Smith et al., 2014.	
Articles focusing on ethics, policy, and governance issues	Anderson & Merry, 2009; Antman, 2014; Brewer, Potterat, & Muth, 2010; Castellani, 2013; Chan et al., 2014; Coady & Wagner, 2013; Dawson & Verweij, 2011; de Wolf, Sieber, Steel, & Zarate, 2005, 2006a, 2006b; Doshi, 2013; Doshi, Goodman, & Ioannidis, 2013; Editorial, 2014; Eichler, Petavy, Pignatti, & Rasi, 2013; Estabrooks & Romyn, 1995; Geller, Sorlie, Coady, Fleg, & Friedman, 2004; Godlee & Groves, 2012; Goldacre, 2013; Gotzsche, 2011a, 2011b, 2012; Greenhalgh, 2009; Haines & Gabor Miklos, 2011; Hampton, 2011; Harris, 2011; Hawkes, 2012, 2013; Hede, 2013; Hrynaszkiewicz & Altman, 2009; Hughes, Wells, McSorley, & Freeman, 2014; Kmietowicz, 2013; Kuntz, 2013; Langat et al., 2011; Lopez, 2010; Mello et al., 2013; Nisen & Rockhold, 2013; Pearce & Smith, 2011; Piwowar, Becich, Bilofsky, Crowley, & on behalf of the caBIG Data Sharing and Intellectual Capital Workspace, 2008; Rabesandratana, 2013; Rodwin & Abramson, 2012; Ross, Gross, & Krumholz, 2012; Ross, Lehman, & Gross, 2012; Ross & Krumholz, 2013; Sandercock, Niewada, Czlonkowska, & International Stroke Trial Collaborative Group, 2011; Sieber, 2006; G. D. Smith, 1994; Sommer, 2010; Spertus, 2012; Strech & Littmann, 2012; Toronto International Data Release Workshop Authors, 2009; Vallance & Chalmers, 2013; Vickers, 2006; Walport & Brest, 2011; White, 2013; Wieseler, McGauran, Kerekes, & Kaiser, 2012; Zarin, 2013	Manju & Buckley, 2012; Pisani & AbouZahr, 2010; Pisani, Whitworth, Zaba, & Abou-Zahr, 2010a; Pisani, Whitworth, Zaba, & AbouZahr, 2010b; Rani, Bekedam, & Buckley, 2011; Sankoh & Ijsselmuiden, 2011; Tangcharoensathien, Boonperm, & Jongudomsuk, 2010; Whitworth, 2010

### *Best Practices in Data Sharing*

In the introduction to this article, stakeholders' views about potential benefits and harms of sharing individual-level data were outlined. When considering the implications of such potential benefits and harms for best practices in data sharing, the fundamental importance of protecting the privacy of research participants was universally acknowledged in the reviewed literature. Some authors went further and set out additional specific principles and considerations for best practices in ethical data sharing (see Table 4).

### *Governed Data Sharing*

To maximize the potential benefits of sharing de-identified data, some stakeholders recommended that de-identified datasets should typically be made available publicly, with minimal restrictions (Doshi et al., 2013; Eichler et al., 2013; Gotzsche, 2011a, 2011b; Haines & Gabor Miklos, 2011; Harris, 2011; Ross, Gross, & Krumholz, 2012; Strech & Littmann, 2012; Vallance & Chalmers, 2013). In contrast, in the majority of reviewed papers, a governed approach to data release was considered valuable to minimize potential harms and maximize potential benefits. Some authors discussed specific advantages of adopting a governed approach to data sharing, as outlined in Table 5.

To guide governed data sharing, stakeholders made a number of recommendations about appropriate policy

development. The current lack of policies or inconsistent policies in some settings was considered both frustrating and inefficient, as well as providing loopholes for researchers who did not want to share data (Manju & Buckley, 2012). A number of papers recommended that harmonized policies with broad applicability be developed, following consultation with a broad range of stakeholders, including policy makers, researchers, patients, patient advocates, privacy experts, funders, research institutions, journal editors, ethicists, NGOs, and governments (Estabrooks & Romyn, 1995; Hrynaszkiewicz & Altman, 2009; Manju & Buckley, 2012; Mello et al., 2013; Vallance & Chalmers, 2013; Whitworth, 2010). These could be complemented by institutional policies where appropriate (Manju & Buckley, 2012; Piwowar et al., 2008). Areas to be addressed in the policies are outlined in Table 6. Some commentators questioned the effectiveness of guidelines and policies encouraging data sharing to date (Gotsche, 2012; Mello et al., 2013; Savage & Vickers, 2009; Vickers, 2006) and suggested legal requirements for data sharing be implemented (Gotsche, 2012; Vickers, 2006).

### *Best Practices in Sharing Data From Low- and Middle-Income Settings*

In both the discussion of potential advantages and disadvantages of data sharing in the introductory section of this article,

**Table 3.** Empirical Research Into Stakeholders' Experiences and Perspectives.

Author, publication year	Study aim	Sample	Methods	Key findings
Kirwan, 1997	Determine opinions of making original data available for alternative analysis.	21 (84%) of 25 invitees from pharmaceutical companies with an interest in rheumatology. Locations of respondents not discussed.	Letter with brief rationale for sharing data and asking whether it would be useful for trial data related to a publication to be shared via a databank and if so, what issues might arise.	5 (24%) of respondents were in favor of sharing data via a databank and 16 (76%) were not. The four most common reasons for not sharing data were: it is inappropriate to conduct analyses not defined in the protocol, authors may be willing to share data for a specific purpose on request instead, other researchers may conduct inappropriate analyses due to a lack of awareness of aspects of the original study, and further analysis is <i>post hoc</i> and data dredging. 12 (57%) participants thought difficulties might arise with sharing trial data due to commercial sensitivity extending beyond the point of publication.
Reidpath and Allotey, 2001	Investigate the preparedness of researchers to share their data.	21 (72%) of 29 inquiries sent to corresponding authors of research articles in the <i>British Medical Journal</i> . Locations of respondents not discussed.	Emailed specific request to reanalyze the data used in a published study (15) or a general inquiry about willingness to share data from a published study (14).	9 (60%) of authors receiving specific requests and 12 (86%) of authors receiving general requests responded. Of the 21 responding authors, one shared the dataset and one was prepared to share the dataset without further conditions, 10 were prepared to release the data in principle subject to further discussions/conditions, three would not release data (suggesting they conduct new analyses themselves or that sufficient data were available in the articles), and six were ultimately non-committal. Authors wanted to know more about why the data were being requested and the proposed analyses. Some authors wanted more time to conduct their own analyses before sharing data, some required conditions to be met (such as payment or contracts to be completed) and others needed to consult with co-investigators prior to release.
Savage and Vickers, 2009	To test effects of journal policies requiring data to be shared.	10 requests for datasets from corresponding authors of research articles in PLoS Medicine or PLoS Clinical Trials. Locations of respondents not discussed.	An emailed request for a dataset to test a pre-specified hypothesis about prediction modeling.	Two authors had changed institution and could not be contacted. One author asked for further details and then shared the dataset. Four authors declined to share the data. When reminded of the journal policy, one said that a formal request to the research group was required, two said it was too much work, the fourth said that he was not permitted to share the data and wouldn't have published in the journal if he'd known it was a requirement. Three authors didn't respond to the first email, two of whom didn't respond to a second email and the third declined to share data because more analyses were proposed.

(continued)

**Table 3. (continued)**

Author, publication year	Study aim	Sample	Methods	Key findings
Rathi et al., 2012	To investigate clinical trialists' opinions and experiences of sharing data with non-collaborating investigators.	317 (46%) of 683 corresponding authors of clinical trials published in 2010 or 2011 in one of the six highest impact general medical journals. Respondents were from the United States or Canada—167 (53%) Western Europe—113 (36%) Other—37 (12%).	38 item adaptive-response online survey.	236 (74%) of respondents supported sharing de-identified data via repositories, and 229 (72%) thought investigators should be required to share data on request. 56 (18%) were required to deposit trial data in a repository by funders and 149 (47%) had received an individual request to share data. Concerns about sharing data through repositories included: potentially misleading secondary analyses, ensuring appropriate data use, ensuring clarity of data elements, indirect costs associated with sharing, colleagues' abilities to publish original research, ability to publish own research, scientific or academic recognition, direct costs associated with sharing, protecting commercially sensitive information, obtaining consent, and maintaining confidentiality. Reasons for denying individual requests for data related to: potentially misleading analyses, potential for misinterpretation of data, potential mistrust of requestor's intent, ability to publish own research, colleagues' ability to publish original research, indirect costs associated with sharing, potential prohibition by formal agreement, potential lack of recognition, protecting commercially sensitive information, direct costs associated with sharing, being unsure of employer or funder policy, protection of patient confidentiality, and potential lack of consent. Reasons for sharing data in response to individual requests included: promoting new research and open science, enhancing robustness of previous research, facilitating student or fellow opportunities, avoiding redundant data collection, increasing impact of own research, professional or personal relationship with the requestor, additional academic recognition, compliance with employer or funder policy, compliance with journal policy.
C. T. Smith et al., 2014	Evaluate support and identify major issues for establishing a central repository of individual participant data.	30 (42%) of 71 reviewers affiliated with the Cochrane Collaboration's individual participant data meta-analysis method group. Respondents were from the United Kingdom—22 (73%) Other European Countries—6 (20%) Australia—1 (3%) Canada—1 (3%).	Synopsis and link to 16 question online survey.	25 (83%) of respondents thought a central repository would be valuable, 25 (83%) would be willing to deposit data in such a repository provided conditions met. The five most commonly suggested conditions that needed to be met for deposition were: approval from primary data source, appropriate acknowledgment of primary data source, involvement of original investigators in the process, reassurance about who would access the data, and the presence of a scientific committee to review data access requests. The five most commonly expected governance arrangements were: restricted access requiring approval, data security, an oversight committee, appropriate recognition for data owners, and anonymised data.



**Table 4.** Principles and Considerations to Inform Best Practices in Ethical Data Sharing.

Principles and considerations	Reference
Ensure sufficiently broad access to realize the benefits to scientific innovation and public health, which are the main justification for sharing.	Mello et al., 2013
Ensure data are used responsibly so that poor quality analyses do not harm public health.	
Treatment of researchers qualified to access data must be evenhanded.	
Data sharing processes must be accountable and transparent.	Mello et al., 2013; Rabesandratana, 2013
Equitable: The needs of researchers, secondary users, communities, and funders should be recognized and balanced.	Walport & Brest, 2011
Ethical: The privacy of individuals and dignity of communities should be protected and public health promoted by productive data use.	
Efficient: Proportionate approaches should build on existing practice to improve the quality and value of research.	
Ensure fair trade and not free trade in data.	Pisani, Whitworth, Zaba, & Abou-Zahr, 2010a; Walport and Brest, 2011
Ensure the rights and responsibilities of researchers generating data and data accessors are balanced.	Sankoh and Ijsselmuiden, 2011
Ensure the benefits of data sharing outweigh the harms, and consider whether restricting the flow of information to avoid rare adverse events is appropriate.	Vickers, 2006
Clearly specify public interests in data sharing and clearly specify any legitimate reasons to restrict access to research data (following market approval of an intervention).	Strech and Littmann, 2012
Ensure that the analytic value of the data is preserved during the protection of privacy and confidentiality.	Sieber, 2006; Vallance and Chalmers, 2013
Ensure data sharing processes are responsive to the context within which datasets were collected.	Pearce and Smith, 2011
Honor the altruism of research participants.	Zarin, 2013

**Table 5.** Potential Benefits of a Governed Approach to Data Sharing.

Potential benefits of curation	Reference
Adequate safeguards can be established, bona fide access restrictions can be put in place.	Pisani, Whitworth, Zaba, and Abou-Zahr, 2010a; Walport and Brest, 2011
Patient privacy is increased.	Doshi, Goodman, and Ioannidis, 2013; Hawkes, 2012; Hughes, Wells, McSorley, and Freeman, 2014; Nisen and Rockhold, 2013
Poor quality research, which may lead to erroneous conclusions, can be prevented following review and requirements to adhere to a rigorous analytical plan.	Doshi et al., 2013; Eichler, Petavy, Pignatti, and Rasi, 2013; Hughes et al., 2014; Manju and Buckley, 2012; Mello et al., 2013; Nisen and Rockhold, 2013; Rathi et al., 2012
Permits compliance with legislation and or regulation.	Nisen and Rockhold, 2013
Promotes adherence to commitments made during the consent process.	Hughes et al., 2014
Enables researchers to fulfill responsibilities to ensure data are used ethically.	Pearce and Smith, 2011
Curation can be responsive to the types of data being shared. Differing approaches can be taken to aggregate and individual-level data, particularly valuable or sensitive datasets, and analyses that require detailed data that could potentially identify participants.	Geller, Sorlie, Coady, Fleg, and Friedman, 2004; Hrynaszkiewicz and Altman, 2009; Rabesandratana, 2013; Toronto International Data Release Workshop Authors, 2009; Vallance and Chalmers, 2013

and the discussion of perspectives about best practices above, the views of authors discussing data sharing in low- and middle-income settings were similar to those expressed in the more substantial body of literature from higher income settings. In contrast to lower and middle-income settings, articles from higher income settings had more discussion about

ways in which these issues had been addressed and data shared to date. When discussing how to provide resources for best practices in data sharing, and how to balance the interests of stakeholders in the data sharing process (particularly those generating datasets), views remained similar, but some different emphases also emerged.

**Table 6.** Priority Areas for Policy Development.

Areas for policy development	References
Appropriate analytic methods, data and meta-data standards, including means of preserving privacy	Kuntz, 2013; Mello et al., 2013; Pisani & AbouZahr, 2010; Rani, Bekedam, and Buckley, 2011; Vickers, 2006
Determining where, how, when, and which data are archived and made available	Manju and Buckley, 2012
Determining for which trials data will be shared, which data and supporting documents will be available, the process for data sharing, how transparent the process will be, who will get access, what types of analyses are permitted, who will decide, what criteria will be used, and what ongoing role the trial sponsor might have.	Zarin, 2013
Methods to permit evaluation of individual applications, including to ensure that the use does not harm participants and is in conformity with ethical approvals	Eichler, Petavy, Pignatti, and Rasi, 2013; Toronto International Data Release Workshop Authors, 2009
Transparent, explicit, and reasonable criteria for case by case decision making	Mello et al., 2013
Requirements and rewards for the collection and curation of datasets for sharing	Pisani, Whitworth, Zaba, and Abou-Zahr, 2010a; Walport and Brest, 2011

The importance of balancing the interests of primary researchers and secondary data users has received considerable attention in the reviewed literature. Stakeholders from higher and lower income settings commented on the importance of ensuring that researchers received appropriate recognition for producing datasets in the subsequent publications by secondary analysts, in professional assessments, and in funding applications (Kuntz, 2013; Manju & Buckley, 2012; Pisani & AbouZahr, 2010; Pisani et al., 2010a; Piwowar et al., 2008; Rani et al., 2011; Rathi et al., 2012; Ross & Krumholz, 2013; G. D. Smith, 1994; C. T. Smith et al., 2014; Walport & Brest, 2011; Whitworth, 2010). Perspectives on authorship differed. Some commentators suggested that co-authorship or at least the chance to publish an associated response or commentary should be offered to the researchers who produced the dataset (Pearce & Smith, 2011; Savage & Vickers, 2009; Vickers, 2006). Others noted that the contribution of data creators may not be sufficient to warrant co-authorship of the secondary analysis (Anderson & Merry, 2009; Gotzsche, 2011b).

Although some commentators considered the value of releasing data prior to publication (Toronto International Data Release Workshop Authors, 2009), others noted the value of exclusive fair use periods for researchers in higher and lower income settings (Geller et al., 2004; Gotzsche, 2011b; Manju & Buckley, 2012; Pearce & Smith, 2011; Pisani & AbouZahr, 2010; Pisani et al., 2010a; Rathi et al., 2012; Ross, Lehman, & Gross, 2012; Savage & Vickers, 2009; Tangcharoensathien et al., 2010; Vickers, 2006). Such periods ranged from 12 months from the end of data collection to unspecified lengths of time, which were, in some cases, linked to the publication of an article with primary findings.

Although limited resources may be a hindrance to data sharing in higher income settings, they were identified as a very significant barrier in lower income settings (Manju &

Buckley, 2012; Pisani & AbouZahr, 2010; Pisani et al., 2010a; Pisani, Whitworth, Zaba, & AbouZahr, 2010b; Rani et al., 2011; Sankoh & Ijsselmuiden, 2011; Tangcharoensathien et al., 2010; Walport & Brest, 2011; Whitworth, 2010). For high-quality individual-level data to be shared in databases with long-term sustainability, significant investment in human resources, technology, and infrastructure will be required. Training, mentoring, and career pathways need to be provided for a range of specialist support staff who will document and curate datasets and manage data release processes. Where data archives are hosted within low- and middle-income settings, expertise in managing biomedical information will be required in addition to the development of storage infrastructure.

Commentators have noted that it would be unfair to develop capacity to share data in low- and middle-income settings without also developing the capacity for data generators and secondary users from such settings to analyze that data (Pisani et al., 2010a; Sankoh & Ijsselmuiden, 2011; Walport & Brest, 2011; Whitworth, 2010). Collaboration between primary and secondary data users was discussed as a potential means of improving the quality of analyses in both higher and lower income settings (Geller et al., 2004; Kuntz, 2013; Pearce & Smith, 2011; Spertus, 2012). Stakeholders from lower income settings also focused on the value of such collaborations to build capacity among researchers generating datasets (Manju & Buckley, 2012; Pisani et al., 2010a; Tangcharoensathien et al., 2010; Whitworth, 2010).

## Discussion

The reviewed literature demonstrated considerable support for sharing individual-level data from clinical and public

health research. As discussed above, numerous recommendations have been made for best practices in governing such sharing, to ensure that potential benefits are promoted and potential harms are managed appropriately. Although significant consensus about some aspects of best practice is evident, such as the need to protect the privacy of research participants, there are differences of opinion about practical achievement of these, such as the measures needed to protect privacy and the extent to which privacy can be assured (Gotzsche, 2011b; Mello et al., 2013; Nisen & Rockhold, 2013). In other areas, there is less consensus about best practices. Opinions differ, for example, about the need for and length of protected time primary researchers should have with data before they are shared (Geller et al., 2004; Toronto International Data Release Workshop Authors., 2009), and the nature of consent required, if any, for sharing de-identified data (de Wolf et al., 2005; Pearce & Smith, 2011).

Commentators have suggested that gaps and inconsistencies in policies and practices for data sharing are frustrating and inefficient, and have recommended that consensus be sought on developing harmonized policies and processes for sharing individual-level data which are informed by stakeholders' views (Manju & Buckley, 2012; Whitworth, 2010).

This review identified just five examples of empirical literature into stakeholders' experiences of and views about sharing individual-level data, all of which focused on the views of data producers and reviewers, primarily from higher income settings (Kirwan, 1997; Rathi et al., 2012; Reidpath & Allotey, 2001; Savage & Vickers, 2009; C. T. Smith et al., 2014). Four of the five studies have sample sizes of 30 or less, and three are five or more years old. Although the findings from these articles provide interesting insights into researchers' opinions and practices of sharing data, some of the perspectives are dated, and differences in the research questions and approaches mean that views of best practices have not been systematically elicited.

This review was unable to identify any empirical research into research participants' perspectives about sharing individual-level data from clinical and public health research that does not involve genetic, genomic, or biobank research. In addition, no research into stakeholders' experiences and perspectives of best practices in sharing clinical data in low- and middle-income settings was found. To develop best practices in data sharing that are appropriate in low- and middle-income settings, empirical research into the perspectives of stakeholders from such settings is needed. We suggest that research into the perspectives of research participants, community representatives, researchers, research ethics committees, and data managers be made a priority to inform current policy development initiatives. The following five articles in this special issue begin to address this gap in the literature and report on the results of empirical studies of stakeholders'

perspectives in India, Kenya, Thailand, South Africa, and Vietnam (Cheah et al., 2015; Denny, Silaigwana, Wassenaar, Bull, & Parker, 2015; Hate et al., 2015; Jao et al., 2015; Merson et al., 2015).

Eight of the conceptual articles in this scoping review focused on the perspectives of stakeholders from low- and middle-income settings (Manju & Buckley, 2012; Pisani & AbouZahr, 2010; Pisani et al., 2010a, 2010b; Rani et al., 2011; Sankoh & Ijsselmuiden, 2011; Tangcharoensathien et al., 2010; Whitworth, 2010). These articles suggest that challenges raised by sharing individual-level data from low- and middle-income settings can differ in important and morally significant ways from those arising in high-income settings. An example is the critical importance of building capacity to generate, curate, share, and analyze high-quality datasets if data are to be shared effectively and fairly. Further theoretical analysis will be valuable to evaluate additional issues arising when sharing individual-level data in low- and middle-income settings, and to inform how best to address them (Bull, Cheah et al., 2015).

### *Limitations of the Review*

Although double screening of all materials is desirable in systematic reviews, it was not possible in this case due to the volume of potential references identified and the complexity of determining the relevance of papers from the supplied abstracts. To minimize error and bias, 20% of abstracts were co-reviewed, and the strategy for a structured approach to analysis was discussed by the co-authors with the collaborating partners in this study. A second limitation of this review is that it was confined to literature in peer-reviewed publications. A valuable addition to the findings of this review would be a review of policies and processes currently in place for curating and sharing individual-level data from clinical and public health research.

### **Acknowledgments**

The authors acknowledge and greatly appreciate the contributions of partners in the multi-site empirical study, particularly Spencer Denny, toward the initial co-review of abstracts.

### **Authors' Note**

The views expressed here are the opinions of the authors and not of the University of Oxford.

### **Declaration of conflict of interests**

The authors declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

### **Funding**

The authors disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: This

research was funded by the Wellcome Trust Strategic Award 096527, on behalf of the Public Health Research Data Forum. The Forum brings together more than 20 health research funders, who are committed to increasing the availability of health research data in ways that are equitable, ethical, and efficient, and will accelerate improvements in public health ([www.wellcome.ac.uk/publichealthdata](http://www.wellcome.ac.uk/publichealthdata)).

## References

- Anderson, B. J., & Merry, A. F. (2009). Data sharing for pharmacokinetic studies. *Pediatric Anesthesia, 19*, 1005-1010.
- Antman, E. (2014). Data sharing in research: Benefits and risks for clinicians. *British Medical Journal, 348*, g237.
- Arksey, H., & O'Malley, L. (2005). Scoping studies: Towards a methodological framework. *International Journal of Social Research Methodology, 8*, 19-32. doi:10.1080/1364557032000119616
- Armstrong, R., Hall, B. J., Doyle, J., & Waters, E. (2011). Cochrane update. "Scoping the scope" of a Cochrane review. *Journal of Public Health, 33*, 147-150.
- Brewer, D. D., Potterat, J. J., & Muth, S. Q. (2010). Withholding access to research data. *The Lancet, 375*, 1872; author reply 1873.
- Bull, S., Cheah, P. Y., Denny, S., Jao, I., Marsh, V., Merson, L., . . . Parker, M. (2015). Best practices for ethical sharing of individual-level health research data from low- and middle-income settings. *Journal of Empirical Research on Human Research Ethics, 10*(3), 302-313. doi:10.1177/1556264615594606
- Castellani, J. (2013). Are clinical trial data shared sufficiently today? Yes. *British Medical Journal, 347*(7916), f1881.
- Chan, A.-W., Song, F., Vickers, A., Jefferson, T., Dickersin, K., Gotsche, P. C., . . . van der Worp, H. B. (2014). Increasing value and reducing waste: Addressing inaccessible research. *The Lancet, 383*, 257-266.
- Cheah, P., Tangseefa, D., Somsaman, A., Chunsuttiwat, T., Nosten, F., Day, N., . . . Parker, M. (2015). Perceived benefits, harms and views about how to share data responsibly: A qualitative study of experiences with and attitudes towards data-sharing among research staff and community representatives in Thailand. *Journal of Empirical Research on Human Research Ethics, 10*(3), 278-289. doi: 10.1177/1556264615592388
- Coady, S. A., & Wagner, E. (2013). Sharing individual level data from observational studies and clinical trials: A perspective from NHLBI. *Trials, 14*, 201.
- Dawson, A., & Verweij, M. (2011). Could do better: Research data sharing and public health. *Public Health Ethics, 4*, 1-3.
- Denny, S. G., Silaigwana, B., Wassenaar, D., Bull, S., & Parker, M. (2015). Developing ethical practices for public health research data sharing in South Africa: The views and experiences from a diverse sample of research stakeholders. *Journal of Empirical Research on Human Research Ethics, 10*(3), 290-301. doi: 10.1177/1556264615592386
- de Wolf, V. A., Sieber, J. E., Steel, P. M., & Zarate, A. O. (2005). Part I: What is the requirement for data sharing? *IRB: Ethics & Human Research, 27*(6), 12-16.
- de Wolf, V. A., Sieber, J. E., Steel, P. M., & Zarate, A. O. (2006a). Part II: HIPAA and disclosure risk issues. *IRB: Ethics & Human Research, 28*(1), 6-11.
- de Wolf, V. A., Sieber, J. E., Steel, P. M., & Zarate, A. O. (2006b). Part III: Meeting the challenge when data sharing is required. *IRB: Ethics & Human Research, 28*(2), 10-15.
- Doshi, P. (2013). Transparency interrupted: The curtailment of the European Medicines Agency's policy on access to documents. *JAMA Internal Medicine, 173*, 2009-2011.
- Doshi, P., Goodman, S. N., & Ioannidis, J. P. A. (2013). Raw data from clinical trials: Within reach? *Trends in Pharmacological Sciences, 34*, 645-647.
- Editorial. (2014). Data sharing will pay dividends. *Nature, 505*(7482), 131.
- Eichler, H.-G., Petavy, F., Pignatti, F., & Rasi, G. (2013). Access to patient-level trial data—A boon to drug developers. *New England Journal of Medicine, 369*, 1577-1579.
- Estabrooks, C. A., & Romyn, D. M. (1995). Data sharing in nursing research: Advantages and challenges. *Canadian Journal of Nursing Research, 27*, 77-88.
- Geller, N. L., Sorlie, P., Coady, S., Fleg, J., & Friedman, L. (2004). Limited access data sets from studies funded by the National Heart, Lung, and Blood Institute. *Clinical Trials, 1*, 517-524.
- Godlee, F., & Groves, T. (2012). The new BMJ policy on sharing data from drug and device trials. *British Medical Journal, 345*, e7888.
- Goldacre, B. (2013). Are clinical trial data shared sufficiently today? No. *British Medical Journal, 347*, f1880. doi:10.1136/bmj.f1880
- Gotsche, P. C. (2011a). We need access to all data from all clinical trials. *Cochrane Database of Systematic Reviews, 12*, ED000035.
- Gotsche, P. C. (2011b). Why we need easy access to all data from all clinical trials and how to accomplish it. *Trials, 12*, 249.
- Gotsche, P. C. (2012). Strengthening and opening up health research by sharing our raw data. *Circulation: Cardiovascular Quality and Outcomes, 5*, 236-237.
- Greenhalgh, T. (2009). Sharing medical research data: Whose rights and who's right? *British Medical Journal, 338*(7700), 903.
- Haines, I. E., & Gabor Miklos, G. L. (2011). Time to mandate data release and independent audits for all clinical trials. *Medical Journal of Australia, 195*, 575-577.
- Hampton, T. (2011). European drug agency under fire: Critics charge that trial data are too inaccessible. *The Journal of the American Medical Association, 306*, 593-595.
- Harris, D. J. (2011). Sharing of research data. *The Lancet, 378*, 1994-1995.
- Hate, K., Meherally, S., Shah More, N., Jayaraman, A., Bull, S., Parker, M., & Osrin, D. (2015). Sweat, skepticism, and uncharted territory: A qualitative study of opinions on data sharing among public health researchers and research participants in Mumbai, India. *Journal of Empirical Research on Human Research Ethics, 10*(3), 239-250. doi: 10.1177/1556264615592383
- Hawkes, N. (2012). Full access to trial data holds many benefits and a few pitfalls, conference hears. *British Medical Journal, 344*, e3723.
- Hawkes, N. (2013). Drug companies should release data from all trials of licensed drugs, say MPs. *British Medical Journal, 346*, f321.
- Hede, K. (2013). Project data sphere to make cancer clinical trial data publicly available. *Journal of the National Cancer Institute, 105*, 1159-1160.
- Hrynaskiewicz, I., & Altman, D. G. (2009). Towards agreement on best practice for publishing raw clinical trial data. *Trials, 10*, 17.

- Hughes, S., Wells, K., McSorley, P., & Freeman, A. (2014). Preparing individual patient data from clinical trials for sharing: The GlaxoSmithKline approach. *Pharmaceutical Statistics, 13*, 179-183.
- Jao, I., Kombe, F., Mwalukore, S., Bull, S., Parker, M., Kamuya, D., & Marsh, V. (2015). Involving research stakeholders in developing policy on sharing public health research data in Kenya: Views on fair process for informed consent, access oversight and community engagement. *Journal of Empirical Research on Human Research Ethics, 10*(3), 264-277. doi: 10.1177/1556264615592385
- King, R., Hooper, B., & Wood, W. (2011). Using bibliographic software to appraise and code data in educational systematic review research. *Medical Teacher, 33*, 719-723.
- Kirwan, J. R. (1997). Making original data from clinical studies available for alternative analysis. *The Journal of Rheumatology, 24*, 822-825.
- Kmietowicz, Z. (2013). Drug firms take legal steps to prevent European regulator releasing data. *British Medical Journal, 346*, f1636.
- Kuntz, R. E. (2013). The changing structure of industry-sponsored clinical research: Pioneering data sharing and transparency. *Annals of Internal Medicine, 158*, 914-915.
- Langat, P., Pisartchik, D., Silva, D., Bernard, C., Olsen, K., Smith, M., & Upshur, R. (2011). Is there a duty to share? Ethics of sharing research data in the context of public health emergencies. *Public Health Ethics, 4*, 4-11.
- Lopez, A. D. (2010). Sharing data for public health: Where is the vision? *Bulletin of the World Health Organization, 88*, 467.
- Manju, R., & Buckley, B. S. (2012). Systematic archiving and access to health research data: Rationale, current status and way forward. *Bulletin of the World Health Organization, 90*, 932-939.
- Medical Research Council. (2011). *MRC policy on research data sharing*. London, England: Medical Research Council.
- Mello, M. M., Francer, J. K., Wilenzick, M., Teden, P., Bierer, B. E., & Barnes, M. (2013). Preparing for responsible sharing of clinical trial data. *New England Journal of Medicine, 369*, 1651-1658.
- Merson, L., Phong, T.V., Nhan, L.N.T., Dung, N.T., Ngan, T.T.D., Kinh, N.V., . . . Bull, S. (2015). Trust, respect and reciprocity: Informing culturally appropriate data sharing practice in Viet Nam. *Journal of Empirical Research on Human Research Ethics, 10*(3), 251-263. doi: 10.1177/1556264615592387
- Moher, D., Liberati, A., Tetzlaff, J., Altman, D. G., & The PRISMA Group. (2009). Preferred reporting items for systematic reviews and meta-analyses: The PRISMA statement. *PLoS Medicine, 6*(6), e1000097. doi:10.1371/journal.pmed1000097
- National Institutes of Health. (2003). *Final NIH statement on sharing research data*. Bethesda, MD: National Institutes of Health.
- Nisen, P., & Rockhold, F. (2013). Access to patient-level data from GlaxoSmithKline clinical trials. *New England Journal of Medicine, 369*, 475-478.
- Parker, M., & Bull, S. (2015). Sharing public health research data: Towards the development of ethical data-sharing practice in low- and middle-income settings. *Journal of Empirical Research on Human Research Ethics, 10*(3), 217-224. doi: 10.1177/1556264615593494
- Parker, M., Bull, S. J., de Vries, J., Agbenyega, T., Doumbo, O. K., & Kwiatkowski, D. P. (2009). Ethical data release in genome-wide association studies in developing countries. *PLoS Medicine, 6*(11), e1000143.
- Pearce, N., & Smith, A. H. (2011). Data sharing: Not as simple as it seems. *Environmental Health: A Global Access Science Source, 10*, 107.
- Pisani, E., & AbouZahr, C. (2010). Sharing health data: Good intentions are not enough. *Bulletin of the World Health Organization, 88*, 462-466.
- Pisani, E., Whitworth, J., Zaba, B., & Abou-Zahr, C. (2010a). Time for fair trade in research data. *The Lancet, 375*, 703-705.
- Pisani, E., Whitworth, J., Zaba, B., & AbouZahr, C. (2010b). Withholding access to research data—Authors' reply. *The Lancet, 375*, 1873.
- Piwowar, H. A., Becich, M. J., Bilofsky, H., & Crowley, R. S., & on behalf of the caBIG Data Sharing and Intellectual Capital Workspace. (2008). Towards a data sharing culture: Recommendations for leadership from academic health centers. *PLoS Medicine, 5*(9), e183.
- Rabesandratana, T. (2013). Europe. Drug watchdog ponders how to open clinical trial data vault. *Science, 339*, 1369-1370.
- Rani, M., Bekedam, H., & Buckley, B. S. (2011). Improving health research governance and management in the Western Pacific: A WHO expert consultation. *Journal of Evidence-Based Medicine, 4*, 204-213.
- Rathi, V., Dzara, K., Gross, C. P., Hrynaskiewicz, I., Joffe, S., Krumholz, H. M., & Ross, J. S. (2012). Sharing of clinical trial data among trialists: A cross sectional survey. *British Medical Journal, 345*, e7570.
- Reidpath, D. D., & Allotey, P. A. (2001). Data sharing in medical research: An empirical investigation. *Bioethics, 15*, 125-134.
- Research Information Network. (2008). *To share or not to share: Research data outputs*. London, England: Research Information Network.
- Rodwin, M. A., & Abramson, J. D. (2012). Clinical trial data as a public good. *The Journal of the American Medical Association, 308*, 871-872.
- Ross, J. S., Gross, C. P., & Krumholz, H. M. (2012). Promoting transparency in pharmaceutical industry-sponsored research. *American Journal of Public Health, 102*, 72-80.
- Ross, J. S., & Krumholz, H. M. (2013). Ushering in a new era of open science through data sharing: The wall must come down. *The Journal of the American Medical Association, 309*, 1355-1356.
- Ross, J. S., Lehman, R., & Gross, C. P. (2012). The importance of clinical trial data sharing: Toward more open science. *Circulation: Cardiovascular Quality and Outcomes, 5*, 238-240.
- Sandercock, P. A., Niewada, M., & Czlonkowska, A., & International Stroke Trial Collaborative Group. (2011). The International Stroke Trial database. *Trials, 12*, 101.
- Sankoh, O., & Ijsselmuiden, C. (2011). Sharing research data to improve public health: A perspective from the global south. *The Lancet, 378*, 401-402.
- Savage, C. J., & Vickers, A. J. (2009). Empirical study of data sharing by authors publishing in PLoS journals. *PLoS ONE, 4*(9), e7078.
- Sieber, J. E. (2006). Introduction: Data sharing and disclosure limitation techniques. *Journal of Empirical Research on Human Research Ethics, 1*(3), 47-50.

- Smith, C. T., Dwan, K., Altman, D. G., Clarke, M., Riley, R., & Williamson, P. R. (2014). Sharing individual participant data from clinical trials: An opinion survey regarding the establishment of a central repository. *PLoS ONE*, 9(5) e97886. doi:10.1371/journal.pone.0097886.
- Smith, G. D. (1994). Increasing the accessibility of data. *British Medical Journal*, 308(6943), 1519-1520.
- Sommer, J. (2010). The delay in sharing research data is costing lives. *Nature Medicine*, 16, 744.
- Spertus, J. A. (2012). The double-edged sword of open access to research data. *Circulation: Cardiovascular Quality and Outcomes*, 5, 143-144.
- Strech, D., & Littmann, J. (2012). Lack of proportionality. Seven specifications of public interest that override post-approval commercial interests on limited access to clinical data. *Trials*, 13.
- Tangcharoensathien, V., Boonperm, J., & Jongudomsuk, P. (2010). Sharing health data: Developing country perspectives. *Bulletin of the World Health Organization*, 88, 468-469.
- Toronto International Data Release Workshop Authors (2009). Prepublication data sharing. *Nature*, 461, 168-170.
- UK Data Archive. (2011). *Managing and sharing data: Best practice for researchers*. Essex, UK: UK Data Archive.
- Vallance, P., & Chalmers, I. (2013). Secure use of individual patient data from clinical trials. *The Lancet*, 382, 1073-1074.
- Vickers, A. J. (2006). Whose data set is it anyway? Sharing raw data from randomized trials. *Trials*, 7, 15.
- Walport, M., & Brest, P. (2011). Sharing research data to improve public health. *The Lancet*, 377, 537-539.
- Wellcome Trust. (2009). *Policy on data management and sharing*. London, England: Wellcome Trust.
- White, P. D. (2013). Sharing data from clinical trials: Is sharing data from clinical trials always a good idea? *British Medical Journal*, 346(7910), f3379.
- Whitworth, J. (2010). Data sharing: Reaching consensus. *Bulletin of the World Health Organization*, 88, 467-468.
- Wieseler, B., McGauran, N., Kerekes, M. F., & Kaiser, T. (2012). Access to regulatory data from the European Medicines Agency: The times they are a-changing. *Systems Review*, 1.
- Zarin, D. A. (2013). Participant-level data and the new frontier in trial transparency. *New England Journal of Medicine*, 369, 468-469.

### Author Biographies

**Susan Bull** is a senior researcher at the Ethox Center, University of Oxford. She led this multi-site study, which was conducted with collaborators from the KEMRI-Wellcome Trust Research Programme, Kilifi, Kenya; the School of Applied Human Sciences, University of Kwazulu Natal, South Africa; the Mahidol Oxford Tropical Medicine Research Unit, Mahidol University, Bangkok, Thailand; and the Society for Nutrition, Education, and Health Action (SNEHA), India. She contributed to the conceptual development of the study and the development of data collection methods, and led the analysis and writing of this article.

**Nia Roberts** is an information specialist and outreach librarian at the Bodleian Health Care Libraries, University Oxford, who regularly contributes to preparing systematic reviews. She contributed to the development of search strategies for this review, conducted the literature searches, and contributed to the writing of this article.

**Michael Parker** is a professor of bioethics and the director of the Ethox Center, University of Oxford. His main research interest is in the many practical ethical aspects of collaborative global health research, including those arising in the sharing of data and biological samples internationally. Together with partners in Kenya, Thailand, Malawi, South Africa, and Vietnam, he co-ordinates the Global Health Bioethics Network, which is a program to carry out ethics research and build ethics capacity. He contributed to the conceptual development of the study and the writing of the article.