

## RESEARCH

# Sharing of clinical trial data among trialists: a cross sectional survey



OPEN ACCESS

Vinay Rathi *medical student*<sup>1</sup>, Kristina Dzara *research associate*<sup>2</sup>, Cary P Gross *associate professor of medicine*<sup>2,3</sup>, Iain Hrynaszkiewicz *publisher*<sup>4</sup>, Steven Joffe *associate professor of pediatrics*<sup>5</sup>, Harlan M Krumholz *professor of medicine, professor of investigative medicine and public health*<sup>3,6,7</sup>, Kelly M Strait *statistician*<sup>6</sup>, Joseph S Ross *assistant professor of medicine*<sup>2,3,6</sup>

<sup>1</sup>Yale University School of Medicine, New Haven CT, USA; <sup>2</sup>Section of General Internal Medicine, Department of Medicine, Yale University School of Medicine; <sup>3</sup>Robert Wood Johnson Foundation Clinical Scholars Program, Department of Medicine, Yale University School of Medicine; <sup>4</sup>BioMed Central, London, UK; <sup>5</sup>Department of Pediatric Oncology, Dana-Farber Cancer Institute and Department of Medicine, Boston Children's Hospital, Boston MA, USA; <sup>6</sup>Center for Outcomes Research and Evaluation, Yale-New Haven Hospital, New Haven CT, USA; <sup>7</sup>Section of Cardiovascular Medicine, Department of Medicine, Yale University School of Medicine; Section of Health Policy and Administration, Yale University School of Epidemiology and Public Health

## Abstract

**Objective** To investigate clinical trialists' opinions and experiences of sharing of clinical trial data with investigators who are not directly collaborating with the research team.

**Design and setting** Cross sectional, web based survey.

**Participants** Clinical trialists who were corresponding authors of clinical trials published in 2010 or 2011 in one of six general medical journals with the highest impact factor in 2011.

**Main outcome measures** Support for and prevalence of data sharing through data repositories and in response to individual requests, concerns with data sharing through repositories, and reasons for granting or denying requests.

**Results** Of 683 potential respondents, 317 completed the survey (response rate 46%). In principle, 236 (74%) thought that sharing de-identified data through data repositories should be required, and 229 (72%) thought that investigators should be required to share de-identified data in response to individual requests. In practice, only 56 (18%) indicated that they were required by the trial funder to deposit the trial data in a repository; of these 32 (57%) had done so. In all, 149 respondents (47%) had received an individual request to share their clinical trial data; of these, 115 (77%) had granted and 56 (38%) had denied at least one request. Respondents' most common concerns about data sharing were related to appropriate data use, investigator or funder interests, and protection of research subjects.

**Conclusions** We found strong support for sharing clinical trial data among corresponding authors of recently published trials in high impact general medical journals who responded to our survey, including a willingness to share data, although several practical concerns were identified.

## Introduction

The conduct of a clinical trial requires substantial investment from research funders, demands considerable time and effort from investigators, and may expose human volunteers to health risks. For these reasons, many have advocated data sharing to enhance the scientific value of trial data. In addition, full access to trial data reduces the potential for incomplete reporting of study outcomes and improves the medical evidence base, which should ultimately improve clinical decision making.<sup>1</sup> In clinical research, data sharing involves a research team making trial data available to individuals with whom they are not collaborating. Data sharing generally takes place via either of two methods. First, investigators may share trial data by depositing it in a repository, an archive of data with terms of access defined by the organisation that maintains it. Second, investigators may share trial data on their own terms in response to individual requests.

In recent years, several major clinical research funders have adopted policies supporting or mandating data sharing. These include the US National Institutes of Health,<sup>2</sup> the UK Medical

Correspondence to: J S Ross, Section of General Internal Medicine, Yale University School of Medicine, PO Box 208093, New Haven CT 0520, USA joseph.ross@yale.edu

Supplementary material as supplied by author: Survey distributed to potential respondents in this study (see <http://www.bmj.com/content/345/bmj.e7570?tab=related#webextra>)

Research Council,<sup>3</sup> and the Bill and Melinda Gates Foundation.<sup>4</sup> Similarly, there are now major biomedical journals, such as *Annals of Internal Medicine*<sup>5</sup> and *BMJ*,<sup>6</sup> which require all authors of original research to state their willingness to share data in published articles, although the effectiveness and enforcement of these policies has been questioned.<sup>7</sup> Several other journals, such as *Nature*, have even made data sharing a condition of publication.<sup>8</sup> In the clinical trials community, the journal *Trials* encourages authors to publish the de-identified raw data supporting published trials as supplementary material,<sup>9</sup> although to date few authors have done so.<sup>10</sup> Secondary users of clinical trial data, such as the Cochrane Collaboration, advocate for stronger data sharing policies in the hopes of increasing access to clinical trial data to test the reliability of medical evidence and improve evidence based practice.<sup>11-17</sup>

Amid these calls for clinical trial data sharing, major regulators, most notably the European Medicines Agency,<sup>18</sup> are beginning to contemplate far reaching policies on open access to data, as are several companies in the pharmaceutical and medical device industries.<sup>19-21</sup> However, despite compelling arguments for mandating full access to trial data,<sup>13 22-24</sup> authors of clinical studies on human subjects are among the least likely to share their raw data.<sup>25</sup> Data suggest that these investigators, particularly clinical trialists, have historically claimed an exclusive right to collected clinical trial data and are largely opposed to data sharing policies due to concerns over research subject confidentiality,<sup>26</sup> resources required for additional data management,<sup>26 27</sup> inappropriate secondary data use,<sup>26 28</sup> and diminished rewards for conducting original research.<sup>26-28</sup> However, most of these studies are dated, and concerns were not systematically elicited.

The success of efforts to promote data sharing depends on cooperation from clinical trialists, who generate and maintain these data. Therefore, in an effort to inform future policies and initiatives and to better understand the trialist community's collective thoughts about and experiences of sharing data, we surveyed corresponding authors of clinical trials published in 2010 or 2011 in one of the six general medical journals with the highest impact factor in 2011. The perspectives of clinical trialists publishing studies in these high impact journals have not been studied previously, although their studies are likely to address important clinical questions that can potentially affect clinical decision making—exactly the type of data whose scientific value should be maximised for public benefit.

## Methods

### Study sample and design

We assembled a sample of clinical trialists through a review of the literature to identify individuals who had published clinical trial findings in 2010 or 2011 in the six highest impact general medical journals. Relevant trials were identified by searching Ovid Medline (1 January 2010 to 31 December 2011) using the terms “clinical trial as topic” and “clinical trial” as free text, limiting our search results to articles published by the six general medical journals with the highest impact factor in 2011 (Journal Citation Reports, Thomson Reuters; *New England Journal of Medicine*, *Lancet*, *JAMA*, *Annals of Internal Medicine*, *PLoS Medicine*, and *BMJ*) (n=903). All non-clinical publications (n=101) and retrospective studies (n=49) were excluded from the resulting list to limit our sample to prospective observational or interventional studies of human subjects (n=753). Primary and secondary analyses of clinical trials published as original articles or research letters were eligible. The first corresponding author named in each article was identified for participation,

except for those trials published in *Annals of Internal Medicine*, whereby the author named in the “reproducible research statement” was identified for inclusion. After accounting for studies with corresponding authors named in multiple publications (n=49) by selecting one publication randomly for inclusion, we identified 709 unique authors as potential survey respondents.

From the original article, we abstracted the following information in order to compare survey respondents and non-respondents: corresponding author location (US or Canada, Western Europe, other) and affiliation (medical school or hospital, private industry, government, other), trial funder (private industry, government, other, mixed), trial enrolment, and journal in which the article was published.

All potential survey respondents were sent an initial email describing the purpose of the study, requesting their participation, and providing a link to the survey in late July 2012, with three follow-up requests sent by email in early August, late August, and early September. Non-respondents were then contacted by telephone during the last two weeks of September 2012 to solicit their participation. Non-respondents were called up to three times and no more than once per day; those with whom we were unable to establish telephone contact were sent one last follow-up request by email in late September. The invitations to participate contained no information about specific hypotheses of the study. Participation was voluntary and included an opportunity to win one of five \$100 gift certificates for Amazon. All responses were rendered anonymous by the web based survey platform (Qualtrics Labs, Provo, UT, USA). Approval from the Yale University School of Medicine Human Research Protection Program was obtained before conducting the study, and consent was considered to be implied when participants completed the online survey.

### Survey instrument development

The design of our 38 item survey instrument was informed by previously published surveys of academic geneticists on data withholding,<sup>29-31</sup> a review of the literature on clinical trial data sharing,<sup>23 25-28 32 33</sup> and discussions with experts (including authors IH and SJ). The instrument was pretested by five clinical investigators unaffiliated with the research team and modified iteratively to improve clarity, face validity, and content validity. Adaptive questioning was used to decrease response burden. Items were presented in multiple response, Likert scale, and open ended formats. The complete instrument is provided as supplementary material on bmj.com.

### Survey domains

#### *Support for and prevalence of data sharing*

We used Likert-type questions to assess clinical trialists' support for data sharing in principle, through data repositories, and in response to personal requests. We used yes/no questions to ascertain whether respondents who were required by their research funder to deposit data from their published study in a repository had done so. We used multiple choice questions to ascertain the number of instances in which respondents who had received at least one individual data sharing request related to their published study had shared or withheld data.

#### *Concerns with and reasons for data sharing*

We used multiple response questions that allowed for open ended feedback to ask respondents about concerns with sharing data through repositories, about experiences with receiving and

making data sharing requests, about reasons for granting or declining individual requests, and about their beliefs on the right of first use of clinical trial data. For those respondents who had experience with depositing or sharing data, concerns and reasons were solicited by asking about actual experience; for those without data depositing or sharing experience, concerns and reasons were solicited by asking about hypothetical situations. For those questions related to concerns with and reasons for data sharing, respondents were first asked to select any or all overarching categories of concerns and reasons provided as multiple choice responses; when respondents selected a category, more detailed concerns and reasons were provided as non-exclusive multiple choice responses. From among the more detailed concerns listed, respondents were asked to indicate the magnitude of their concern (major, minor, or none). We did not differentiate between major or minor reasons for granting or denying individual requests to share data.

### Respondent characteristics

We also collected respondent sociodemographic characteristics, including age, sex, and primary employer, as well as professional characteristics, including academic rank, years since completion of highest degree, location of scientific training, academic productivity, and funding status.

### Statistical analysis

To compare characteristics of survey respondents and non-respondents, we used  $\chi^2$  tests for categorical variables (author employer and location, trial funder, and journal) and the Kruskal-Wallis test for continuous variables (trial enrolment), using two-sided tests with a type I error level of 0.05. Comparison data were analysed by KD using SPSS version 19.0 (IBM, Armonk, NY, USA).

Next, we conducted descriptive analyses, calculating the proportion of respondents who supported data sharing in principle, had engaged in data sharing, had identified major and minor concerns about data sharing, and had identified reasons for granting and denying data sharing requests. For Likert questions, we collapsed “strongly disagree” and “somewhat disagree” into one category (“disagree”) and collapsed “strongly agree” and “somewhat agree” into another (“agree”). No variables were missing >1% of responses. Survey data were analysed by KMS using SAS Version 9.3 (SAS Institute, Cary, NC, USA).

### Results

Survey participation requests were sent to 709 unique clinical trialists (fig 1); 26 were subsequently excluded from our study population because contact information was invalid (n=18), technical difficulties prevented the author from accessing the survey (n=4), or the author was a secondary user of clinical trial data generated by another research group (n=4). Of the remaining 683 trialists, 317 completed the survey either online (n=306) or by phone (n=11), yielding a response rate of 46%.

Survey respondents did not differ from non-respondents with respect to corresponding author location or affiliation, trial enrolment, or journal in which the article was published (table 1). However, trial funders differed among responders and non-responders ( $P=0.001$ ), as corresponding authors of trials funded solely by government sources responded more often than corresponding authors of trials funded solely by industry or by mixed funding sources.

More than 90% of respondents were corresponding authors of clinical trials (275 (87%) of primary reports and 20 (6%) of secondary reports of clinical trials), and 22 (7%) were corresponding authors of prospective observational studies. Most respondents were aged 50–64 years, male, received their scientific training in the US or Canada, and had completed their training 10–24 years ago (table 2). Most (83%) were employed by an academic institution, and two thirds of these had reached the rank of full professor. Respondents were professionally productive: in the past three years, 41% had published  $\geq 25$  articles, 46% had been awarded four or more grants, and 52% had received more than \$1million in direct research support.

### Support for and prevalence of data sharing

Overall, 278 (88%) of respondents supported data sharing. Specifically, 236 (74%) thought that, in principle, sharing de-identified data through a data repository should be mandatory. Furthermore, 229 (72%) thought that investigators should be required to share de-identified data upon individual request.

In practice, only 56 (18%) respondents were required by their research funder to deposit their trial data in a repository, and 32 of the 56 (57%) had thus far deposited the data. Similarly, 149 (47%) had received an individual request to share their clinical trial data, and 115 (77%) of these had granted at least one request and 56 (38%) had denied at least one request. The most common reasons for data requests were for systematic review or meta-analysis (n=85 (57%)), for subgroup analysis of the originally published study (n=58 (39%)), and to pursue novel research questions (n=47 (32%)). Among the respondents, 101 (32%) had made a request for clinical trial data of another investigator: 81 of 101 (80%) had at least one request granted, and 49 (49%) had at least one request declined.

Respondents varied in their beliefs about right of first use of clinical trial data. Five (2%) indicated that clinical trial data should be made available to investigators external to the study team immediately on trial completion, 109 (34%) stated that it should be within one to two years of trial completion, 97 (31%) stated that it should be within three to five or more years of trial completion, while 106 (33%) indicated that there should be no time limit and that the right of first use should extend until the main findings are accepted for publication.

### Concerns with data sharing through repositories

Respondents identified potential major and minor concerns with data sharing through repositories. The most common concerns related to appropriate data use (n=205 (65%)), but investigator and funder interests (n=129 (41%)) and protection of research subjects (n=91 (29%)) were also common concerns. Specific concerns, and whether they were indicated to be major or minor, are displayed in fig 2.

### Reasons for granting and denying individual data sharing requests

Respondents offered several reasons for sharing data in response to an individual request. The most common reasons cited were related to promoting open science (n=248 (78%)), although academic benefits and recognition (n=133 (42%)) were also often identified. More detailed reasons for granting data sharing requests are displayed in fig 3.

Similarly, respondents offered several reasons for denying individual requests to share data. The most common reasons cited were related to ensuring appropriate data use (n=233



(74%)). However, protecting investigator or funder interests (n=121 (38%)) and protecting research subjects (n=107 (34%)) were also often cited. More detailed reasons for denying data sharing requests are displayed in fig 4.

## Discussion

In this survey of clinical trialists who had recently published trials in high impact general medical journals, we found strong support among respondents, in principle, for sharing clinical trial data with investigators who are not directly collaborating with their research team. Nearly three quarters of respondents thought that sharing de-identified data through data repositories or in response to individual requests should be required. These findings contradict prior research which suggested that trialists are largely opposed to data sharing.<sup>26-28</sup> Our results may reflect changes in attitude among clinical researchers over time, since our survey was administered more than a decade after two of these earlier studies.<sup>26-28</sup> However, our survey also differed from the third, more recent, study by asking potential respondents to self report support for and data sharing behaviours, as opposed to testing actual willingness to share clinical trial data.<sup>27</sup> Nevertheless, a self reported willingness to engage in data sharing is a critical first step.

For the value of data sharing to be realised, support in principle needs to translate to actions, on the part of both potential data sharers and funders. Fewer than a fifth of trialists were required by funders to deposit their trial data in a repository. Additionally, fewer than half of the respondents had received an individual request for their data, and fewer than a third had requested data from others. Partially, this may be a consequence of an underdeveloped data sharing infrastructure, which requires robust repositories, data standards, copyright and licensing agreements, and patient consent.<sup>34</sup>

However, our findings suggest a genuine willingness among these respondents to engage in data sharing. Although less than half had received individual requests for data, more than three quarters of these had thus far shared data in response. Since the trials of surveyed investigators were all published in 2010 or 2011, willingness to share might increase as time passes. Moreover, even though only half of respondents had received a request for data, this suggests that there is a clear demand for shared data. Perhaps the low rate of data sharing requests was due to investigators external to the research team only recently becoming aware of the data. In addition, although slightly over half of respondents who were required by funders to deposit data had thus far done so, open ended comments suggested this was because the deadline date for deposit had not been reached. However, the fact that so few funders require deposition of de-identified data in a repository for use by other investigators suggests that a commitment among funders to require data sharing is a clear mechanism by which to promote and increase the practice.

## Generalisability of results

Despite general support for and willingness to engage in data sharing efforts, only 46% of potential participants completed our survey. That fewer than half of potential participants responded may suggest that our findings overestimate support for and willingness to engage in data sharing in the clinical trial community, limiting the external validity of our findings—as individuals who chose not to respond may have been less supportive of data sharing efforts than those who responded. Furthermore, even among survey respondents, our findings may have been biased by social desirability, as respondents may have

been less likely to self report beliefs and behaviours that may be negatively perceived by others and instead indicated stronger support for data sharing efforts.<sup>35-36</sup> However, there were few observed differences between survey respondents and non-respondents, with the exception that the response rate among corresponding authors of trials funded solely by government sources was higher than that among corresponding authors of trials funded solely by industry or by mixed funding sources. Because government funded trialists may be more oriented toward public health interests, such as data sharing initiatives, our findings may be biased toward that viewpoint. Nevertheless, a substantial number of trialists who had received industry funding participated in this survey.

In addition, we used several mechanisms to ensure prospectively the generalisability of our findings, prevent social desirability bias, and improve response rates. First, our solicitation letters consistently reminded potential participants that the purpose of the survey was to fairly represent their views in the growing international debate on whether to require investigators to release their clinical trial data to others. Second, we used a web based survey platform for ease of completion and limited the scope of the survey to reduce response burden. Finally, we offered financial incentives for participation and employed several reminder contacts, including three emails and at least one telephone contact. Although our response rate compares favourably with other surveys of physicians (a difficult group to engage in participation),<sup>37-39</sup> it was lower than that of other web based surveys of clinical trial investigators.<sup>40-42</sup>

## Implications of results

Respondents had several concerns that may need attention to ensure their engagement in data sharing efforts. Most concerns related to the integrity of the process and to the need for data sharing to lead to both public and private benefit. For instance, the most common concerns related to appropriate data use, such as preventing misinterpretation or misleading secondary analyses and ensuring clarity of data elements for other investigators. However, there are mechanisms to allay such concerns, including pre-registration and specification of secondary data analysis plans,<sup>43</sup> deposition of clear data dictionaries, and public posting of frequently asked questions about data sources to ensure that individuals learn from one another. Moreover, to ensure competency among both data sharers and requestors, training curriculums can be developed to teach best practices for preparing and using shared clinical trial data.<sup>23</sup>

Similarly, some trialists indicated that they were concerned about protecting either their own or their colleagues' interests, including the need to ensure that trialists receive sufficient academic or scientific recognition for sharing their clinical trial data, do not spend undue time or effort preparing data for sharing, and have sufficient opportunity to publish studies using the data they were responsible for collecting. Addressing these concerns will, in many respects, require a cultural shift in the clinical research community, including within academic medicine.<sup>44</sup> Academic institutions and promotions committees need to begin crediting investigators not just for publishing articles in high-impact journals, but also for creating data by designing and conducting clinical trials, sharing data with other investigators and enabling them to address important questions. Interestingly, academic recognition was not always a concern, as more than 40% of respondents reasoned that their decision to share trial data would increase their academic recognition by increasing the impact of their work and helping them develop professional relationships.

## Limitations of study

In addition to the question of generalisability raised above, our study was limited to corresponding authors of clinical trials published in the highest impact general medical journals. Our findings may not be applicable to the entire clinical trial research community, although these high impact studies are likely to address important clinical questions that can potentially affect clinical decision making—exactly the type of data whose scientific value should be maximised for public benefit. Finally, some information of interest was not asked in order to reduce response burden, including questions about whether funders have explicitly prohibited data sharing and experiences negotiating data ownership among funders, trialists, and secondary users. Data ownership remains a critical obstacle to data sharing efforts,<sup>32 34</sup> particularly for industry funded research, with sponsors often explicitly retaining ownership rights.<sup>41 45 46</sup>

## Conclusions

We found strong support for sharing clinical trial data among corresponding authors of recently published trials in high impact general medical journals who responded to our survey. However, practical concerns must be addressed if the efforts of clinical trial funders, journals, and secondary users of clinical trial data to promote data sharing are to succeed. The clinical trialist community not only has to cooperate with these efforts, but must trust that data sharing is in the best interests of the public and science.

Contributors: VR, CPG, IH, SJ, and JSR were responsible for the conception and design of this work. VR, KD, and JSR were responsible for acquisition of data. VR and JSR drafted the manuscript. KD and KMS conducted the statistical analysis. JSR provided supervision. All authors participated in the analysis and interpretation of the data and critically revised the manuscript for important intellectual content.

Data access and responsibility: All authors had full access to all the data in the study, and JSR takes responsibility for the integrity of the data and the accuracy of the data analysis.

Funding: This study was not supported by any external grants or funds. VR received support from the Yale University School of Medicine Medical Student Research Fellowship. The five \$100 Amazon gift certificates were paid for from JSR's institutional discretionary funds. HMK and JSR receive support from the Centers of Medicare and Medicaid Services (CMS) to develop and maintain hospital performance measures that are used for public reporting. HMK is supported by a National Heart Lung Blood Institute Cardiovascular Outcomes Center Award (1U01HL105270-02). JSR is supported by the National Institute on Aging (K08 AG032886) and by the American Federation for Aging Research through the Paul B Beeson Career Development Award Program.

Competing interests: All authors have completed the ICMJE uniform disclosure form at [www.icmje.org/coi\\_disclosure.pdf](http://www.icmje.org/coi_disclosure.pdf) (available on request from the corresponding author) and declare: no support from any organisation for the submitted work; CPG, HMK, and JSR receive research support from Medtronic to develop methods to promote data sharing; CPG and JSR are on a scientific advisory board for FAIR Health, a not-for-profit organisation with the mission to achieve fairness and transparency in healthcare reimbursement; HMK chairs a scientific advisory board for UnitedHealthcare, a health insurance company; SJ is a paid member of a data monitoring committee for Genzyme/Sanofi; no other relationships or activities that could appear to have influenced the submitted work.

Data sharing: Requests for statistical code and dataset can be made to the corresponding author at [joseph.ross@yale.edu](mailto:joseph.ross@yale.edu). The dataset will be made available via a publicly accessible repository on publication.

- 1 Gotzsche PC. Why we need easy access to all data from all clinical trials and how to accomplish it. *Trials* 2011;12:249.
- 2 National Institutes of Health (NIH). Final NIH statement on sharing research data. 2003. <http://grants.nih.gov/grants/guide/notice-files/NOT-OD-03-032.html> (accessed 6 Oct 2012).
- 3 Medical Research Council (MRC). MRC policy on research data-sharing. [www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/datasharing/Policy/index.htm](http://www.mrc.ac.uk/Ourresearch/Ethicsresearchguidance/datasharing/Policy/index.htm) (accessed 6 Oct 2012).
- 4 Bill and Melinda Gates Foundation. Our commitment to sharing information. [www.gatesfoundation.org/global-health/Pages/our-commitment.aspx](http://www.gatesfoundation.org/global-health/Pages/our-commitment.aspx) (accessed 6 Oct 2012).
- 5 Laine C, Goodman SN, Griswold ME, Sox HC. Reproducible research: moving toward research the public can really trust. *Ann Intern Med* 2007;146:450-3.
- 6 Groves T. BMJ policy on data sharing. *BMJ* 2010;340:c564.
- 7 Alsheikh-Ali AA, Qureshi W, Al-Mallah MH, Ioannidis JP. Public availability of published research data in high-impact journals. *PLoS One* 2011;6:e24357.
- 8 Nature Publishing Group. Authors and references, policies: availability of data and materials. [www.nature.com/authors/policies/availability.html](http://www.nature.com/authors/policies/availability.html) (accessed 6 Oct 2012).
- 9 Trials. About Trials: editorial policies, data and materials release. [www.trialsjournal.com/about#trials](http://www.trialsjournal.com/about#trials) (accessed 6 Oct 2012).
- 10 Altman DG, Hrynaskiewicz I, Furberg CD, Grimshaw JM, Rothwell PM. Five years of Trials. *Trials* 2011;12:248.
- 11 Cochrane Collaboration. Press release: Cochrane Collaboration statement on access to clinical trial data. 2011. [www.cochrane.org/features/clinical-trials-statement-press-release](http://www.cochrane.org/features/clinical-trials-statement-press-release) (accessed 6 Oct 2012).
- 12 Ross JS, Gross CP, Krumholz HM. Promoting transparency in pharmaceutical industry-sponsored research. *Am J Public Health* 2012;102:72-80.
- 13 Ross JS, Lehman R, Gross CP. The importance of clinical trial data sharing: toward more open science. *Circ Cardiovasc Qual Outcomes* 2012;5:238-40.
- 14 Gotzsche PC, Jorgensen AW. Opening up data at the European Medicines Agency. *BMJ* 2011;342:d2686.
- 15 Chan AW. Access to clinical trial data. *BMJ* 2011;342:d80.
- 16 Godlee F, Loder E. Missing clinical trial data: setting the record straight. *BMJ* 2010;341:c5641.
- 17 Wieseler B, McGauran N, Kaiser T. Still waiting for functional EU Clinical Trials Register. *BMJ* 2011;342:d3834.
- 18 Hampton T. European drug agency works to improve transparency, but skepticism remains. *JAMA* 2012;308:850-1.
- 19 Krumholz HM, Ross JS. A model for dissemination and independent analysis of industry data. *JAMA* 2011;306:1593-4.
- 20 Loder E. Liberating clinical trial data. 2012. <http://blogs.bmj.com/bmj/2012/10/09/elizabeth-loder-liberating-clinical-trial-data/> (accessed 15 Oct 2012).
- 21 Thomas K. Glaxo opens door to data on research. *New York Times* 2012 Oct 11.
- 22 Krumholz HM. Open science and data sharing in clinical research: basing informed decisions on the totality of the evidence. *Circ Cardiovasc Qual Outcomes* 2012;5:141-2.
- 23 Hrynaskiewicz I, Altman DG. Towards agreement on best practice for publishing raw clinical trial data. *Trials* 2009;10:17.
- 24 Rodwin MA, Abramson JD. Clinical trial data as a public good. *JAMA* 2012;308:871-2.
- 25 Piwowar HA, Day RS, Fridsma DB. Sharing detailed research data is associated with increased citation rate. *PLoS One* 2007;2:e308.
- 26 Kirwan JR. Making original data from clinical studies available for alternative analysis. *J Rheumatology* 1997;24:822-5.
- 27 Savage CJ, Vickers AJ. Empirical study of data sharing by authors publishing in PLoS journals. *PLoS One* 2009;4:e7078.
- 28 Reidpath DD, Allotey PA. Data sharing in medical research: an empirical investigation. *Bioethics* 2001;15:125-34.
- 29 Blumenthal D, Campbell EG, Anderson MS, Causino N, Louis KS. Withholding research results in academic life science. Evidence from a national survey of faculty. *JAMA* 1997;277:1224-8.
- 30 Blumenthal D, Campbell EG, Gokhale M, Yucel R, Clarridge B, Hilgartner S, et al. Data withholding in genetics and the other life sciences: prevalences and predictors. *Acad Med* 2006;81:137-45.
- 31 Campbell EG, Clarridge BR, Gokhale M, Birenbaum L, Hilgartner S, Holtzman NA, et al. Data withholding in academic genetics: evidence from a national survey. *JAMA* 2002;287:473-80.
- 32 Vickers AJ. Whose data set is it anyway? Sharing raw data from randomized trials. *Trials* 2006;7:15.
- 33 Walport M, Brest P. Sharing research data to improve public health. *Lancet* 2011;377:537-9.
- 34 Hrynaskiewicz I, Norton ML, Vickers AJ, Altman DG. Preparing raw clinical data for publication: guidance for journal editors, authors, and peer reviewers. *BMJ* 2010;340:c181.
- 35 Fowler FJ. *Survey research methods*. Sage, 1993.
- 36 Sudman S, Bradburn NM. *Asking questions: a practical guide to questionnaire design*. Jossey-Bass, 1982.
- 37 Asch DA, Jedziewski MK, Christakis NA. Response rates to mail surveys published in medical journals. *J Clin Epidemiol* 1997;50:1129-36.
- 38 Keyhani S, Federman A. Doctors on coverage—physicians' views on a new public insurance option and Medicare expansion. *N Engl J Med* 2009;361:e24.
- 39 Shanafelt TD, Boone S, Tan L, Dyrbye LN, Sotile W, Satele D, et al. Burnout and satisfaction with work-life balance among US physicians relative to the general US population. *Arch Intern Med* 2012 Aug 20:1-9.
- 40 Haahr MT, Hrobjartsson A. Who is blinded in randomized clinical trials? A study of 200 trials and a survey of authors. *Clin Trials* 2006;3:360-5.
- 41 Lundh A, Krogsboll LT, Gotzsche PC. Access to data in industry-sponsored trials. *Lancet* 2011;378:1995-6.
- 42 Rochon PA, Sekeres M, Hoey J, Lexchin J, Ferris LE, Moher D, et al. Investigator experiences with financial conflicts of interest in clinical trials. *Trials* 2011;12:9.
- 43 Krumholz HM. Documenting the methods history: would it improve the interpretability of studies? *Circ Cardiovasc Qual Outcomes* 2012;5:418-9.
- 44 Piwowar HA, Bechik MJ, Bilofsky H, Crowley RS. Towards a data sharing culture: recommendations for leadership from academic health centers. *PLoS Med* 2008;5:e183.
- 45 Gotzsche PC, Hrobjartsson A, Johansen HK, Haahr MT, Altman DG, Chan AW. Constraints on publication rights in industry-initiated clinical trials. *JAMA* 2006;295:1645-6.
- 46 Lundh A, Krogsboll LT, Gotzsche PC. Sponsors inverted question mark participation in conduct and reporting of industry trials: a descriptive study. *Trials* 2012;13:146.

Accepted: 03 November 2012

**What is already known on this topic**

Data sharing policies are increasingly promoted to improve access to clinical trial data to inform evidence based practice  
Little is known about support for these policies among clinical trialists

**What this study adds**

Among corresponding authors of recently published trials in high impact general medical journals who responded to our survey, about three quarters supported initiatives for sharing clinical trial data  
Respondents reported a willingness to share data, along with several practical concerns related to appropriate data use, investigator or funder interests, and protection of research subjects

Cite this as: *BMJ* 2012;345:e7570

This is an open-access article distributed under the terms of the Creative Commons Attribution Non-commercial License, which permits use, distribution, and reproduction in

any medium, provided the original work is properly cited, the use is non commercial and is otherwise in compliance with the license. See: <http://creativecommons.org/licenses/by-nc/2.0/> and <http://creativecommons.org/licenses/by-nc/2.0/legalcode>.

## Tables

**Table 1 | Comparison of characteristics of survey respondents and non-respondents**

	Respondents (n=317)	Non-respondents (n=366)	P value
Corresponding author characteristic			
Primary employer:			
Medical school or hospital	278 (88)	319 (87)	0.72
Government	19 (6)	18 (5)	
Private industry	4 (1)	8 (2)	
Other	16 (5)	21 (6)	
Location:			
US or Canada	167 (53)	172 (47)	0.11
Western Europe	113 (36)	159 (43)	
Other	37 (12)	35 (10)	
Clinical trial characteristic			
Funding agency:			
Government	120 (38)	87 (24)	0.001
Private industry	71 (22)	107 (29)	
Non-profit organisation or other	45 (14)	54 (15)	
Mixed (two or more funders)	81 (26)	118 (32)	
Median (interquartile range) enrolment size	524 (234–1700)	502 (185–1585)	0.17
Journal where trial was published:			
<i>New England Journal of Medicine</i>	113 (36)	137 (37)	0.30
<i>Lancet</i>	70 (22)	98 (27)	
<i>BMJ</i>	53 (17)	51 (14)	
<i>JAMA</i>	43 (14)	51 (14)	
<i>Annals of Internal Medicine</i>	24 (8)	21 (6)	
<i>PLoS Medicine</i>	14 (4)	8 (2)	

Table 2| Sociodemographic and professional characteristics of survey respondents (n=317)

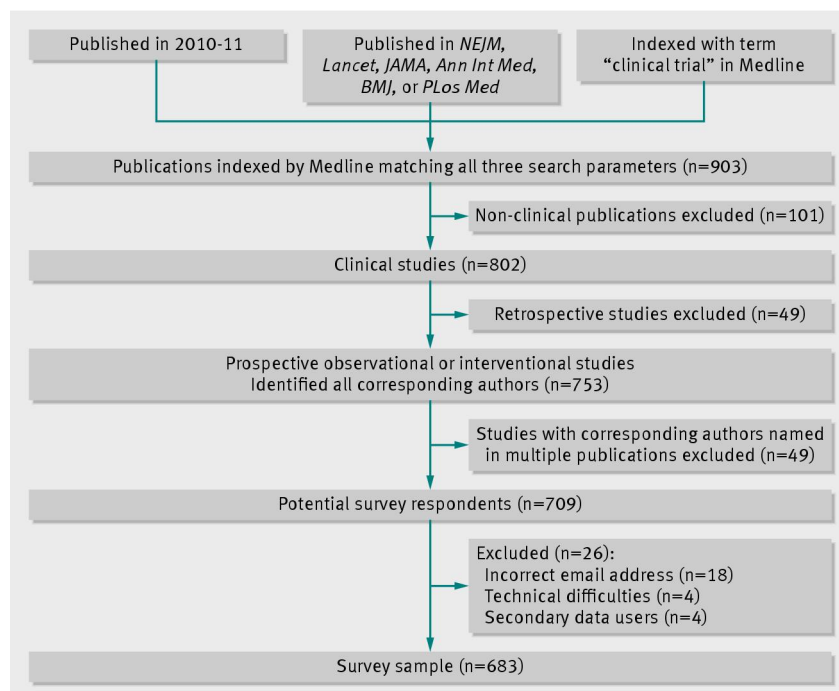
Characteristic	No (%) of respondents
Age (years):	
≤49	126 (40)
50–64	159 (50)
≥65	31 (10)
Sex:	
Men	243 (77)
Women	73 (23)
Location of scientific training:	
US or Canada	149 (47)
Western Europe	132 (42)
Elsewhere	33 (11)
Years since completion of highest degree:	
≤9	46 (15)
10–24	155 (49)
≥25	116 (37)
Primary employer:	
Non-academic institution	53 (17)
Academic institution:	264 (83)
Rank of assistant professor or below*	43 (16)
Rank of associate professor*	47 (18)
Rank of full professor*	174 (66)
Employment effort devoted to research (%):	
≤49	104 (33)
50–74	104 (33)
≥75	109 (34)
<b>Research activity in past 3 years</b>	
No of publications:	
1–10	71 (22)
11–25	117 (37)
≥25	129 (41)
No of grants or contracts:	
0	33 (10)
1–3	138 (44)
≥4	146 (46)
Direct funding costs from grants or contracts (\$million):	
0	33 (10)
<1	120 (38)
≥1	164 (52)
External funding source of grants or contracts†:	
Government	211 (74)
Non-profit organisation	148 (52)
Private industry	149 (52)

\*Rank provided only for the 262 respondents with an academic institution as their primary employer.

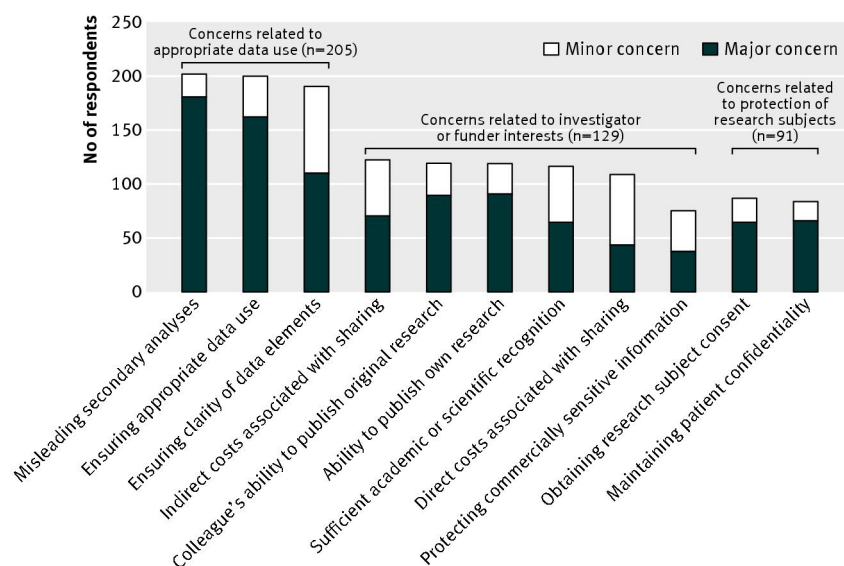
†External funding source of grants or contracts provided only for the 284 respondents who received a grant or contract in the past 3 years (n=284). Funding sources are not exclusive (funding from more than one source could be received).



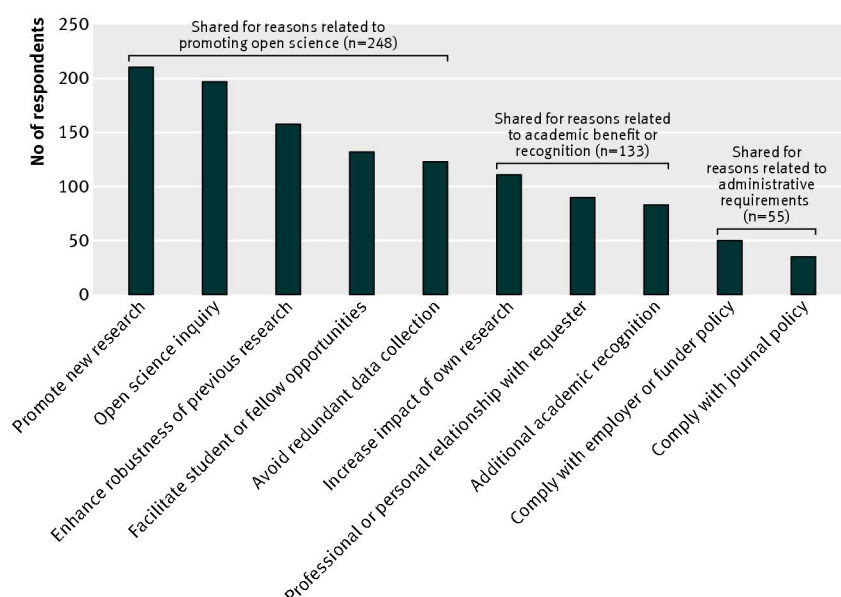
## Figures



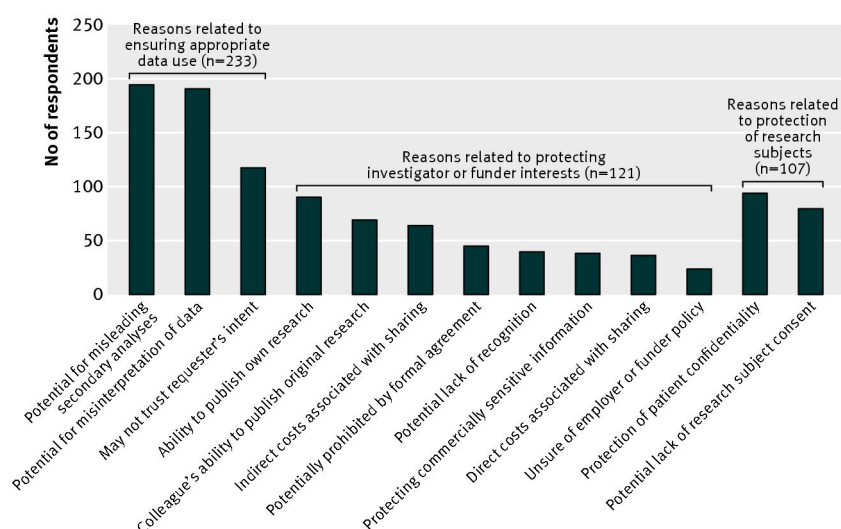
**Fig 1** Flow chart showing identification and selection of potential survey respondents



**Fig 2** Major and minor concerns with data sharing through repositories described by 317 survey respondents, grouped by overarching concern related to appropriate data use, investigator or funder interests, and protection of research subjects. (The sample number for each overarching category indicates the number of respondents who initially selected that category and were then given the opportunity to select more detailed concerns from among multiple choice responses)



**Fig 3** Reasons for granting individual requests to share data described by 317 survey respondents, grouped by overarching reason related to promoting open science, enhancing academic benefit or recognition, and satisfying administrative requirements. (The sample number for each overarching category indicates the number of respondents who initially selected that category and were then given the opportunity to select more detailed reasons from among multiple choice responses)



**Fig 4** Reasons for denying individual requests to share data described by 317 survey respondents, grouped by overarching reason related to ensuring appropriate data use, protecting investigator or funder interests, and protecting research subjects. (The sample number for each overarching category indicates the number of respondents who initially selected that category and were then given the opportunity to select more detailed reasons from among multiple choice responses)