



Published in final edited form as:

Clin Trials. 2016 April ; 13(2): 149–160. doi:10.1177/1740774515602868.

Risks of phase I research with healthy participants: A systematic review

Rebecca A Johnson¹, Annette Rid², Ezekiel Emanuel³, and David Wendler^{4,*}

¹Department of Sociology, Princeton University, Princeton, NJ, USA

²Department of Social Science, Health & Medicine, King's College London, London, United Kingdom

³Department of Medical Ethics and Health Policy, University of Pennsylvania, Philadelphia PA, USA

⁴Department of Bioethics, NIH Clinical Center, Bethesda MD, USA

Abstract

Background/Aims—Tragedies suggest that phase I trials in healthy participants may be highly risky. This possibility raises concern that phase I trials may exploit healthy participants to develop new therapies, making the translation of scientific discoveries ethically worrisome. Yet, few systematic data evaluate this concern. The present paper systematically reviews the risks of published phase I trials in healthy participants and evaluates trial features associated with increased risks.

Methods—Data on adverse events and trial characteristics were extracted from all phase I trials published in PubMed, Embase, Cochrane, Scopus, and PsycINFO (1 January 2008 through 1 October 2012). Inclusion criteria were phase I studies that enrolled healthy participants of any age, provided quantitative adverse event (AE) data, and documented the number of participants enrolled. Exclusion criteria included 1) AE data not in English, 2) a “challenge” study in which participants were administered a pathogen, and 3) no quantitative information about serious AE's. Data on the incidence of adverse events, duration of AE monitoring, trial agent tested, participant demographics, and trial location were extracted.

Results—In 475 trials enrolling 27,185 participants, there was a median of zero serious adverse events (interquartile range, 0-0) and a median of zero severe adverse events (interquartile range, 0-0) per 1000 treatment group participants/day of monitoring. The rate of mild and moderate

*To whom correspondence should be addressed: David Wendler, Ph.D., Department of Bioethics, NIH Clinical Center, 10/1C118, Bethesda, MD 20892, USA, Phone: 301.496.2429, Fax: 301.496.0760, dwendler@mail.nih.gov.

Ethical approval: No ethical approval was required for the present systematic review.

Contributions: Conceived and designed the experiments (A.R., D.W., R.J., Z.E.); analyzed the data (R.J.); wrote the first draft of the manuscript (A.R., R.J.); contributed to the writing of the manuscript (A.R., D.W., R.J., Z.E.); agree with manuscript results and conclusions (A.R., D.W., R.J., Z.E.).

Competing interests: This work was completed as part of the authors' (R.J. and D.W.) official duties as employees of the U.S. NIH Clinical Center. The NIH had no role in the analysis, writing of the manuscript, or the decision to submit it for publication.

Disclaimer: The views expressed are the authors' own. They do not represent the position or policy of the NIH, DHHS, or US government.

adverse events was a median of 1147.19 per 1000 participants (interquartile range, 651.52 – 1730.9) and 46.07 per 1000 participants/AE monitoring day (interquartile range, 17.80 – 77.19).

Conclusions—We conclude that phase I trials do cause mild and moderate harms but pose low risks of severe harm. To ensure that this conclusion also applies to unpublished trials, it is important to increase trial transparency.

Keywords

Ethics; phase I clinical trials; healthy volunteers; data transparency; research design

Introduction

Phase I studies with healthy participants are a vital step in translating scientific discoveries into clinical therapies. Yet some worry that phase I studies expose healthy participants to high risks. First, unlike phase I oncology trials that enroll persons with cancer or later phase disease studies that enroll patients, phase I studies in healthy participants offer no potential for clinical benefit, raising concerns about exploitation.^{1,2} Second, payments may lead healthy participants in phase I studies to underestimate the risks they face.^{3,4} Third, payments may result in some individuals regarding repeated research participation as a full-time occupation that might pose high risks and offers few rewards.^{5,6} Fourth, some express concern that phase I trials enroll low-income participants who endure high trial risks but have limited access to the medical products they help develop.^{6–9}

Momentous tragedies, such as the TeGenero case in which six healthy participants experienced life-threatening reactions to a monoclonal antibody or the death of a healthy volunteer in a John Hopkins asthma study, seem to support the view that phase I trials are very risky.^{10,11} Yet there are few systematic data on the harms experienced by healthy participants in phase I research. This lack of systematic data raises the need for reviews “documenting the prevalence, nature, and severity of risks among healthy volunteers” in phase I studies to assess their ethical appropriateness.⁵

Extant studies on the risks of phase I studies with healthy participants suffer from important shortcomings. First, the data come from trials conducted in a single country or at a single clinical research unit.^{12–16} Second, most enroll a narrow ethnic or national population such as Japanese men or French medical students.^{12,13,17} Third, existing studies often rely on surveys of investigators, which may lead to recollection biases.^{12,18} Fourth, most reviews include fewer than 10,000 participants,^{14,16–18} thus undermining their ability to detect rare but serious harms. A recent review also addresses these shortcomings, using published and unpublished data from Pfizer-sponsored trials conducted at trial centers located in different continents to assess the risks of phase I trials for healthy participants.¹⁹ The present review has the advantage of examining data from a range of trial sponsors, while the other review has the advantage over the present paper of including both published and unpublished trial data. The two reviews complement each other by each providing substantial data improvements over past reviews, serving as important first steps in building a more robust body of research on the risks of phase I trials.

Methods

We followed the Preferred Reporting Items for Systematic Reviews and Meta-Analysis (PRISMA) guidelines in reporting the results of our review.

Data sources and searches

The present paper searched the following databases for dates of coverage ranging from 1 January, 2008 through 1 October, 2012: PUBMED, EMBASE, COCHRANE CENTRAL REGISTER OF CONTROLLED TRIALS, SCOPUS, and PSYCINFO using terms describing the phase of trial, health of participants, and AE's (Figure S1). The most rare adverse events in past research were one drug-related, clinically serious adverse event per 5,000 participants.¹³ Therefore, the almost five year coverage period was intended to capture a sufficient number of trials to detect rare but serious harms.

Criteria for selecting studies

Studies were included in the analysis if they: 1) were described as phase I; 2) enrolled healthy participants of any age, which meant trials enrolling not only adult but also pediatric and infant participants; 3) provided quantitative serious AE data; and 4) detailed the number of enrolled participants. Exclusion criteria were: 1) AE data not in English (we included articles where the rest of the text was in a non-English language as long as AE information was reported in English), 2) a "challenge" study in which participants were administered a pathogen, and 3) no quantitative information about serious AE's. Challenge studies were excluded because we were interested in phase I trials that test the toxicity of an active investigational agent, rather than trials where adverse events stem from both the agent and the challenge period infection.

Study selection process

Two of the authors (R.J., D.W.) first independently performed a title screen on the 1,523 identified articles and then an abstract screen, classifying articles as ineligible, unclear eligibility, possibly eligible, or probably eligible. Any disagreements were resolved by consensus, after which ineligible articles were excluded. One author (R.J.) read each of the 567 articles and excluded 174 articles that did not meet the inclusion/exclusion criteria, leaving 393 eligible articles (Figure 1, Figure S2). The primary reasons for exclusion were that the trial did not provide *any* adverse event information or that the trial did not specify the number of serious AE's, instead only stating that serious adverse events were "rare." Multiple phase I trials in the same published article were coded separately, yielding 475 trials for analysis.

Data extraction

One author (R.J.) extracted the data on trial features and adverse events depicted in Tables 1–3. In addition, one physician (A.R) coded each trial's target disease area using clinicaltrials.gov categories of conditions.²⁰ Only some of the eligible trials explicitly indicated the system used to classify AE severity, including the National Cancer Institute (NCI's) Common Terminology Criteria for Adverse Events (CTCAE) or the Food and Drug Administration (FDA's) COSTART.^{21,22} Because the trials did not provide sufficient detail

to recode the findings on a common scale, our analysis reports the AE severity, type, and relatedness used in the published articles. Trials did generally adhere to the distinction between *severe* versus *serious* adverse events.²³

To assess the accuracy of the data extraction, a research assistant independently extracted data from a random sample of 27 trials (5.7%) and compared the results with those of the primary extractor. The primary extractor's error rate of 1.8% (19 errors out of 1,066 data points) was deemed acceptable, especially since the errors primarily concerned aspects of a trial, such as inpatient versus outpatient setting (7 errors), that were not used in final analyses. There was only one error in extracting adverse events out of a total of 162 AE data points, the main variable of interest.

Data quality assessment

To confirm that published articles offer the best publicly available source of phase I safety data, we searched clinicaltrials.gov (keywords: *normal*; *healthy*) for phase I trials in healthy participants that were completed during 01/01/2006 to 10/01/2010. This window allows for a roughly two-year lag between completion of the trial and publication of the trial results.²⁴ We found 3,851 registered phase I trials, suggesting that our sample reflects fewer than 12% of phase I trials with healthy participants conducted during a 5-year period.²⁵ However, only 380 of these 3,851 registered trials (9.9%) reported study results on clinicaltrials.gov. Furthermore, the published trials offered more adverse event details than clinicaltrials.gov results.

We also compared data quality in our sample to New Drug Application (NDA) review documents available through the Food and Drug Administration's (FDA) database Drugs@FDA for specific drugs and disease areas.²⁶ However, the FDA does not guarantee that phase I results are posted for all approved therapies, review documents may not contain safety outcomes for each phase I trial of an approved therapy, and the database does not include information on phase I studies of drugs for which U.S. approval was never sought.²⁷⁻²⁹ Therefore, the present paper uses the most comprehensive, publicly available data source.

Data synthesis and analysis

We used SPSS Version 21.0 (SPSS, Chicago, Illinois) to perform descriptive analyses. We then used STATA Version 13.1 (STATA, College Station, Texas) to fit the Poisson regression model. For this model, the dependent variable was the number of severe adverse events in the treatment group and we used STATA's exposure command to adjust this count for a trial's number of treatment group participants at risk for an event by taking the natural log of this variable. Predictors were chosen based trial features that others have argued predict higher risk: income level of trial host country,³⁰ trials conducted in CRO's as opposed to academic sites,^{31,32} agent type (e.g. biologic v. small molecule),³³ sponsorship source,³⁴ mean BMI of participants,¹⁷ percentage of female participants,¹⁷ and approval status of agent under study. Factor variables were recoded into 1 = yes; 0 = no dummy indicators, with one category left out of the regression specification.

Results

Trial characteristics

A total of 475 trials, enrolling 27,185 participants, satisfied the inclusion/exclusion criteria (Figure 1, Table 1). The median study enrollment was 30 participants (interquartile range, 18–51 participants).

Over half of the trials tested small, chemically manufactured molecule agents (251 trials). A total of 331 trials (69.7%) tested agents that were not approved by the FDA, EMEA, or other regulatory agency in the trial country. Conversely, 94 trials (19.8%) tested agents that were approved at the time of study, with approximately one third of these testing interactions between two approved therapies. Approval status was coded based on the sponsor's reporting, so the classification should reflect the agent's status at the time the trial was conducted rather than at the time of publication. Most trials reported a target disease area despite enrolling healthy participants, likely due to the fact that Investigational New Drug applications required to begin phase I testing inquire about a therapy's proposed disease indication.³⁵ Viral diseases were the most common disease under study (14.3% of the sample). The median duration of AE monitoring was 16 days (interquartile range, 7–42 days). This duration reflects AE monitoring both *during* drug administration/exposure and monitoring that occurs in the days *after* this administration/exposure.

Pharmaceutical, biotechnology, or medical device companies exclusively funded the majority of the trials (60.8%); government funded the next highest proportion: 76 trials (16.0%). Contract research organizations were the most frequent trial site (30.3%), although 161 trials (33.9%) did not report where the trial was conducted and did not have a clinicaltrials.org or EudraCT registration with this information.

North America hosted 169 trials (35.6%), with Europe next with 131 trials (27.6%). Using the World Bank classification of countries based on 2010 GDP data (the midpoint year of our sample),³⁶ 70.4% of the eligible trials took place in high-income countries (e.g. U.S., France), 12.6% took place in upper-middle-income countries (e.g. Romania, China), 2.3% took place in lower-middle income countries (e.g. Ukraine, India), 2.3% took place in low-income countries (e.g. Kenya, Tanzania), and 12.4% did not indicate the trial country.

Participant characteristics

Overall, 91.9% of participants received an active agent and 10.8% received a placebo, with some participants taking part in both an active arm and placebo arm as part of a crossover study (Table 2). There were more male (49.9%) than female participants (39.4%), with 10.7% participating in trials that did not provide a sex breakdown. The most frequent racial/ethnic group was Caucasian (49.8%), representing a significantly more ethnically diverse sample than past reviews that often feature solely Caucasian or solely Asian ethnicities. The mean age was 31.2 years (SD = 10.0), mean weight 71.2 kg (SD = 16.2), and mean BMI was 24.8 (SD = 4.0). These means were calculated by taking the overall, unweighted mean of the trial-level means, rather than a mean weighted by trial size.

Adverse events

All of the 475 eligible trials reported the rate of serious adverse events (SAE) and deaths, since we specified numerical SAE data as an inclusion criteria. There were a total of 284 serious AE's (0.5% of the total AE's) in the 27,185 participants (Table 3). Only 15 of the serious AE's were deemed possibly related to the investigational agent, typically by the investigator or trial sponsor; the remaining 269 were deemed unrelated to the investigational agent. There were 5 deaths among the 27,185 trial participants, such as a traffic accident and cocaine overdose, all of which were classified by the investigator or trial sponsor as unrelated to the investigational agent. Of the 52,798 adverse events (Table 3), 3.7% were severe and 95.8% were mild or moderate, with the numbers adding up to less than 100% because some trials reported the number of non-serious AEs but not a breakdown of their severity levels. With 27,185 participants and zero deaths, assuming a binomial distribution of adverse events and using the Clopper Pearson exact method of calculating binomial proportion confidence intervals, the study excludes a death rate above 0.14 deaths per 1,000 participants at the 95% confidence level. With 15 serious adverse events and 27,185 participants, the study excludes a serious adverse event rate above 0.91 serious adverse events per 1,000 participants.

Severe adverse events

A total of 48 trials enrolling 2,884 participants 1) had both a treatment and placebo group (and no participants in both groups as in a crossover design), 2) reported the sample size and the number of mild-severe AE's for each group, 3) separated events into mild/moderate versus severe categories, and 4) reported duration of adverse event monitoring, allowing us to compare AE incidence rates between the two groups (Table 3). Since a large number of trials had zero severe events, thus violating parametric assumptions, we performed a Wilcoxon signed-rank test that examined the magnitude and sign of the difference between severe AE incidence in the treatment versus placebo arms of a trial, treating each trial as an observation. This revealed that the treatment groups in trials displayed a significantly higher incidence of severe adverse events per 1000 participants (median = 0, interquartile range, 0 – 12.96) than the placebo groups (median = 0, interquartile range, 0-0) ($Z = -2.471$, $p = 0.013$). While the test discards all trial pairs where there is zero difference between the treatment and placebo groups, thus discarding trials where there were zero severe adverse events in each group, this works against the test showing significant differences in the incidence rates and makes it more noteworthy that we found a significant result.

The median number of severe adverse events in trial treatment groups was zero, but there could be a small fraction of trials on the upper end of the risk distribution where participants face high risks of harm. Among the trials that have some severe adverse events rather than the median of zero, is there a high incidence rate of severe events? To address this question, we re-examined the median and ranges to only look at trials with more than zero severe adverse events. The median rate of severe adverse events in these trials was 111.92 events/1000 treatment group participants (interquartile range, 63.54–189.28), which, when adjusted for days of AE monitoring, falls to a median of 1.66 events/1000 participants/monitoring day (interquartile range, 0.46–4.15).

Analyzing mild and moderate adverse events using a Wilcoxon signed-rank test, the treatment group had a significantly higher median incidence of events per 1000 participants (median = 1147.19 interquartile range, 651.52–1730.90) than the placebo group (median = 701.39, interquartile range, 184.52 – 1175.00)($Z = -4.646$, $p = 0.000$).

Trial features associated with increased risk

We operationalized trial risk as incidence of severe events/participant in the treatment group. The data was highly skewed towards zero, with 67% of trials reporting zero severe adverse events, so we determined that a Poisson regression model would best account for this large proportion of zero outcomes. We used a dependent variable of the number of severe adverse events in the treatment group and included an adjustment for the number of treatment group participants to control for trials that had a larger number of participants who are “at risk” for a severe event. In addition, since there were several predictors that others have argued contribute to trial risk—testing a vaccine or biologic associated with higher risk than small molecules; pharmaceutical funding associated with higher risk compared to government or academic funding; trials hosted in low or middle-income countries facing higher risk than trials hosted by higher-income countries; trials where participants have a lower mean BMI experience higher rates of adverse events; trials testing previously unapproved agents rather than drug-drug interactions between approved agents facing higher risk—we fit the model using a forward stepwise regression with a p value to enter of 0.05, and a p value to remove of 0.10.

The model as a whole was statistically significant ($p < 0.001$). Overall, trials that tested vaccines had significantly higher incidences of severe adverse events per treatment group participant compared to all other agent types (incidence rate ratio = 6.21, 95% CI = 3.71, 10.37). The remaining results are summarized in Table 4, and illustrate that a large number of variables were excluded during the stepwise regression because, when other significant predictors of risk were included, they neither significantly increased a trial’s risk level nor significantly decreased its risk level. To examine the sensitivity of these results to our choice of stepwise modeling, we specified a Poisson regression without using a stepwise procedure. The results are substantively equivalent: vaccines remain a significant predictor of higher risk at the $p < 0.001$ level; small molecules remain a significant predictor of lower risk at the $p < 0.001$ level; funded by government, nonprofit or academic funding is a significant predictor of lower risk at the $p < 0.001$ level; hosted by academic or nonprofit trial site is a significant predictor of higher risk at the $p < 0.01$ level. Conducted in a low or middle income country is no longer a significant predictor of lower risk. This suggests that while some hypothesized features of a trial significantly increase its risk—for example, it testing a vaccine—others are not significant contributors to risk.

Discussion

This systematic review of 475 trials published from 2008 to 2012, involving 27,185 participants, finds that healthy participants in phase I trials do not experience high rates of significant harm. Participants in over 98% of the trials experienced no drug-related serious adverse events. There were 5 deaths, none of which were deemed related to the study drug.

Participants in trial treatment groups faced higher mild and moderate adverse events than those in the placebo groups, yet the median rate of *severe* adverse events per treatment group was zero. In sum, mild harms were prevalent across trials, while serious harms were confined to a small subset of trials. In this small subset, rates of adverse events do not support the premise that risks of phase I trials are extremely high, with a median of 1.66 severe adverse events/1000 participants/day of AE monitoring. Though some healthy participants have referred to phase I research as a “mild torture economy,”³⁷ our review suggests that the risks of phase I research in healthy participants are not dramatically high. Yet the importance of ensuring that drug translation does not exploit healthy participants suggests that further research should address possible publication biases in safety outcomes.

One concern expressed regarding phase I trials involves possible “off-shoring” to lower-income countries with more vulnerable participants and less stringent review and safety monitoring.³⁰ However, in the present data, the majority of trials (70.4%) took place in high-income countries and there was no significant increase in risk for trials in lower or middle-income countries compared to trials in high-income countries. While this suggests that trials appear to remain in higher-income countries, we lacked participant-level demographic data that could tell us whether trials within these countries disproportionately enroll vulnerable populations such as those in poverty. Similarly, it has been argued that studies conducted in contract research organization or pharmaceutical-owned clinics may be riskier than studies conducted in academic centers or hospitals.^{31,32} This trial feature also was not associated with increased risk after controlling for the agent type tested by the trial.

Our findings about the relative frequency of mild versus serious harm are in line with results of past reviews of phase I research. The median of mild adverse events in our sample (38.37/1000 treatment group participants/day) is lower than past estimates that range from 76.7 events/1000 participants/day to 124.8 events/1000 participants/day.^{14,17} The overall rate of related and unrelated serious adverse events per 100 participants/day (0.025) (as opposed to the median across trials, which equals zero) is within the range of past research, which reports rates ranging from 0.020 to 0.074.^{14,17} This comparison with other reviews suggests that our results are roughly in line with the findings of past reviews of phase I research. The study also improves upon limitations of previous reviews, such as small sample sizes inadequately powered to detect rare but serious harms, data confined to a single country or single clinical research unit, homogeneous trial populations (e.g. Japanese men), and no statistical analysis of features associated with an increased risk of harm.

Limitations

This review has several limitations. First, the analyzed data are limited to published studies, creating three shortcomings. First, sponsors and investigators may only publish studies with sufficiently low rates of adverse events to progress to later stages of testing or to receive regulatory approval.^{38,39} Compounds tested in phase I trials have roughly a 70% chance of progressing to the next phase of testing, but only a 19% chance of eventual approval.⁴⁰ And a general analysis of publication rates for drugs that fail to reach the licensing stage of drug development suggests that fewer than 50% of these “stalled” drugs have a full journal publication for any of the drug’s trials.⁴¹ Therefore, if sponsors only publish phase I results

from approved compounds, the resulting sample of phase I trials would not represent all trials conducted. This is especially the case since phase I trials may experience especially low rates of publication: one study of French clinical trials shows a 17% publication rate for phase I trials compared to an average of 43% for the other phases.⁴² Second, for the 19% of compounds that are approved, sponsors may withhold publication of unfavorable results. Third, for trials that are published, the published article may present the trial data in a way that minimizes harms. These data reporting issues may mean our sample underreports the harms of phase I trials. While these publication biases are important limitations, the present review makes substantial data advances over past phase I reviews of small, homogeneous participant populations.

As our data quality assessment in the methods section underscores, other sources of publicly available trial data—clinicaltrials.gov, FDA New Drug Application review documents—contain both similar biases towards safer trials and less detailed adverse event and trial characteristics data than the present sample of published articles. These biases in current, publicly available trial data highlight the importance of initiatives, such as the AllTrials campaign, to systematically collect the outcomes of all clinical trials for both approved and non-approved therapies.^{43–47} Our results highlight that these campaigns should focus on both improving *whether* phase I results are published and the *quality* of data reporting, as a non-negligible proportion of the trials we reviewed lacked basic information such as mean BMI and the racial/ethnic breakdown of participants.

Second, though we used a comprehensive set of search terms that incorporated both the phase of the trial and the health status of the participants, it is possible that our search missed some published phase I trials—for instance, those that did not specify the trial phase in the title or in the full text. We tried to minimize this limitation through strategies such as allowing trials to qualify by either mentioning enrollment of healthy participants or mentioning a vaccine trial where the participants are healthy but may be less likely to be referred to that way in the trial title (Figure S1).

Third, there was not sufficient information to score the severity of the AEs from the different studies on a common scale and trials varied in how they graded outcomes; for instance, few investigators used uniform systems like the Food and Drug Administration (FDA's) Coding Symbols for a Thesaurus of Adverse Reaction Terms (COSTART). This limitation affects the interpretation of our regression findings on features of a trial that predict higher risk, since sponsors might have used different thresholds for what counts as a severe adverse event. For instance, some vaccine trials reported vaccine site irritation above a certain circumference as a severe adverse event, a classification that may not have occurred in other trial types. Future trial transparency efforts should emphasize adoption of a uniform severity rating system. Fourth, qualitative data show that participants may find the logistical elements of study participation—fixed meals, confinement to a study unit—more unpleasant than the side effects of trial agents.^{37,48} Adverse events reporting also focuses more on harms associated with the investigational agent rather than harms and burdens associated with study procedures such as biopsies, venipuncture, and other procedures. Therefore, certain harms salient to participants are not documented in reviews such as the present study that analyze clinical adverse events largely stemming from the investigational

drug. However, because the risks stemming from drug exposure, rather than the risks from study procedures, are thought to be a risk particularly high for phase I trials, the present review focused on the trial component that provokes the most ethical concern. Fifth, we relied on the sponsors', investigators', and safety monitors' assessments of the relatedness of serious adverse events. These assessments classified the majority of serious adverse events as unrelated to the investigative agents. The present study did not have sufficient detail about the adverse events and surrounding circumstances to make independent relatedness judgments. Sixth, the lack of detailed participant data, such as income data, information on repeat trial participation, or insurance status, prevented us from addressing concerns about trials recruiting low-income persons who participate in repeated phase I studies.

Conclusion

Phase I research in healthy participants raises important ethical concerns about exploiting healthy participants by exposing them to high risks that then benefit medical patients. Our review of published phase I trials shows the rarity of severe harm. In particular, though participants face a higher likelihood of mild harms than they might face in daily life, participants face a low risk of severe harm. In arriving at these results, the present review makes important advances over past reviews of phase I research, which are outdated, cover small samples, analyze single study unit or single country data, and do not examine features of a trial that predict higher risk. However, the discrepancy between our findings and the perception of phase I research as highly risky highlights two points: first, some critics of phase I trials may have overstated the risk the trials pose to the average participant. Second, given limitations of published data sources, the present review should serve as a first step in trial transparency efforts that can help measure the risks of unpublished trials. The need for this robust literature underscores the importance of current trial data transparency initiatives.^{43–47} It also suggests that the initiatives should expand their mandate from making data *available* to ensuring the comparability of adverse events across trials, providing more detailed participant characteristics, and other features crucial for accurately assessing the risks that phase I research poses to healthy participants.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

The authors thank Remy Brim, PhD, Marion Danis, MD, Dennis Dixon, PhD, Andrea Dockry, BS, Susan Hilsenbeck, PhD, Steve Pearson, MD, Sarah Solomon, Karen Smith, MLS, Robert Wesley, PhD, and Amina White, MD for their valuable input on aspects of this manuscript. Rebecca Johnson had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

References

1. Shamoo AE, Resnik DB. Strategies to minimize risks and exploitation in phase one trials on healthy subjects. *Am J Bioeth.* 2006; 6:W1–W13. [PubMed: 16754430]
2. Lemmens T, Elliott C. Justice for the professional guinea pig. *Am J Bioeth.* 2001; 1:51–53. [PubMed: 11951892]

3. McNeill P. Paying people to participate in research: Why not? *Bioethics*. 1997; 11:390–396. [PubMed: 11655118]
4. Steinbrook R. Compensation for injured research subjects. *N Engl J Med*. 2006; 354:1871–1873. [PubMed: 16672697]
5. Kass NE, Myers R, Fuchs EJ, et al. Balancing justice and autonomy in clinical research with healthy volunteers. *Clin Pharmacol Ther*. 2007; 82:219–227. [PubMed: 17410122]
6. Elliott C, Abadie R. Exploiting a research underclass in phase 1 clinical trials. *N Engl J Med*. 2008; 358:2316–2317. [PubMed: 18509119]
7. Fisher, J. *Medical research for hire: The political economy of pharmaceutical clinical trials*. New Brunswick, NJ: Rutgers University Press; 2008.
8. Harris J. Scientific research is a moral duty. *J Med Ethics*. 2005; 31:242–248. [PubMed: 15800367]
9. Rhodes R. Rethinking research ethics. *Am J Bioeth*. 2005; 5:7–28. [PubMed: 16036651]
10. Sheridan C. TeGenero fiasco prompts regulatory rethink. *Nat Biotechnol*. 2006; 24:475–476. [PubMed: 16680111]
11. Savulescu J, Spriggs M. The hexamethonium asthma study and the death of a normal volunteer in research. *J Med Ethics*. 2002; 28:3–4. [PubMed: 11834748]
12. Kumagai Y, Fukazawa I, Momma T, et al. Piii-48. *Clin Pharmacol Ther*. 2006; 79:P71.
13. Sibille M, Donazzolo Y, Lecoz F, et al. After the London tragedy, is it still possible to consider Phase I is safe? *Br J Clin Pharmacol*. 2006; 62:502–503. [PubMed: 16817849]
14. Lutfullin A, Kuhlmann J, Wensing G. Adverse events in volunteers participating in phase I clinical trials: a single-center five-year survey in 1,559 subjects. *Int J Clin Pharmacol Ther*. 2005; 43:217–226. [PubMed: 15906587]
15. Royle J, Snell E. Medical research on normal volunteers [letter]. *Br J Clin Pharmacol*. 1986; 21:548–549. [PubMed: 3718813]
16. Arnold J. Incidence of injury during clinical pharmacology research and indemnification of injured research subjects at Quincy research center. In: *Compensating for research injuries: the ethical and legal implications of programs to redress injured subjects*. Append Pres Comm Study Ethical Probl Med Biomed Behav Res. 1982:275–302. U.S. Government Printing Office, Stock No. 040-000-00456-4.
17. Sibille M, Deigat N, Janin A, et al. Adverse events in phase-I studies: a report in 1015 healthy volunteers. *Eur J Clin Pharmacol*. 1998; 54:13–20. [PubMed: 9591924]
18. Orme M, Harry J, Routledge P, et al. Healthy volunteer studies in Great Britain: the results of a survey into 12 months activity in this field. *Br J Clin Pharmacol*. 1989; 27:125–133. [PubMed: 2713210]
19. Emanuel EJ, Bedarida G, Macci K, et al. Quantifying the risks of non-oncology phase I research in healthy volunteers: meta-analysis of phase I studies. *BMJ*. 2015; 350:h3271. [PubMed: 26115663]
20. [accessed 10 March 2013] National Institutes of Health clinicaltrials.gov. Studies by topic. http://clinicaltrials.gov/ct2/search/browse?brwse=cond_cat
21. Food and Drug Administration. *Coding symbols for thesaurus of adverse reaction terms*. Rockville, MD: 1995.
22. National Cancer Institute. *Common Terminology Criteria for Adverse Events (CTCAE)*. Bethesda, MD: U.S. Department of Health and Human Services; 2009.
23. *Clinical Safety Data Management: Definitions and standards for expedited reporting E2A*. Geneva: 1994.
24. Ross JS, Mulvey GK, Hines EM, et al. Trial publication after registration in ClinicalTrials.gov: a cross-sectional analysis. *PLoS Med*. 2009; 6:e1000144. [PubMed: 19901971]
25. Wood AJJ. Progress and deficiencies in the registration of clinical trials. *N Engl J Med*. 2009; 360:824–830. [PubMed: 19228628]
26. Food and Drug Administration. [accessed 11 January 2014] *Drugs@FDA: FDA approved drug products*. <http://www.accessdata.fda.gov/scripts/cder/drugsatfda/>
27. Turner EH. A taxpayer-funded clinical trials registry and results database. *PLoS Med*. 2004; 1:e60. [PubMed: 15562322]

28. O'Connor AB. The need for improved access to FDA reviews. *JAMA*. 2009; 302:191–193. [PubMed: 19584349]
29. Zarin DA, Ide NC, Tse T, et al. Issues in the registration of clinical trials. *JAMA*. 2007; 297:2112–2120. [PubMed: 17507347]
30. Petryna, A. When experiments travel: clinical trials and the global search for human subjects. Princeton, NJ: Princeton University Press; 2009.
31. Foster C, Malik AY. The elephant in the (board) room: the role of contract research organizations in international clinical research. *Am J Bioeth*. 2012; 12:49–50. [PubMed: 23072682]
32. Shuchman M. Commercializing clinical trials--risks and benefits of the CRO boom. *N Engl J Med*. 2007; 357:1365–1368. [PubMed: 17914036]
33. Kenter MJH, Cohen AF. Establishing risk of human experimentation with drugs: lessons from TGN1412. *Lancet*. 2006; 368:1387–1391. [PubMed: 17046471]
34. Rehnquist, J. The globalization of clinical trials: a growing challenge in protecting subjects. Boston, MA: Office of the Inspector General: U.S. Department of Health and Human Services; 2001.
35. Food and Drug Administration. Investigational new drug application. Rockville, MD: 2010.
36. World Bank. Data: how we classify countries. Washington, DC: 2013.
37. Abadie, R. The professional guinea pig: big pharma and the risky world of human subjects. Durham, NC: Duke University Press; 2010.
38. Rising K, Bacchetti P, Bero L. Reporting bias in drug trials submitted to the Food and Drug Administration: review of publication and presentation. *PLoS Med*. 2008; 5:e217. [PubMed: 19067477]
39. Bourgeois FT, Murthy S, Mandl KD. Outcome reporting among drug trials registered in ClinicalTrials.gov. *Ann Intern Med*. 2010; 153:158–166. [PubMed: 20679560]
40. DiMasi JA, Feldman L, Seckler A, et al. Trends in risks associated with new drug development: success rates for investigational drugs. *Clin Pharmacol Ther*. 2010; 87:272–277. [PubMed: 20130567]
41. Hakala A, Kimmelman J, Carlisle B, et al. Accessibility of trial reports for drugs stalling in development: a systematic assessment of registered trials. *BMJ*. 2015; 350:h1116. [PubMed: 25752879]
42. Decullier E, Chan AW, Chapuis F. Inadequate dissemination of phase I trials: a retrospective cohort study. *PLoS Med*. 2009; 6:e1000034. [PubMed: 19226185]
43. Lemmens T, Telfer C. Access to information and the right to health: the human rights case for clinical trials transparency. *Am J Law Med*. 2012; 38:63–112. [PubMed: 22497094]
44. Chalmers I, Glasziou P, Godlee F. All trials must be registered and the results published. *BMJ*. 2013; 346:f105. [PubMed: 23303893]
45. GlaxoSmith Kline. [accessed 12 December 2014] GSK announces support for AllTrials campaign for clinical data transparency. 2013. <http://www.gsk.com/en-gb/media/press-releases/2013/GSK-announces-support-for-AllTrials-campaign-for-clinical-data-transparency/>
46. IdeaPoint Inc. [accessed 1 January 2015] IdeaPoint, Inc. announces launch of multi-sponsor portal providing researchers opportunity to request access to clinical data from five global study sponsors. 2014. <http://www.prnewswire.com/news-releases/ideapoint-inc-announces-launch-of-multi-sponsor-portal-providing-researchers-opportunity-to-request-access-to-clinical-data-from-five-global-study-sponsors-238410261.html>
47. European Medicines Agency. [accessed 1 January 2015] European Medicines Agency to push ahead in 2014 towards publication and access to clinical trial data. 2013. http://www.ema.europa.eu/ema/index.jsp?curl=pages/news_and_events/news/2013/12/news_detail_001991.jsp
48. Ondrusek, N. Making participation work: a grounded theory describing participation in phase I drug trials from the perspective of the healthy subject. PhD Thesis. Canada: University of Toronto; 2010.

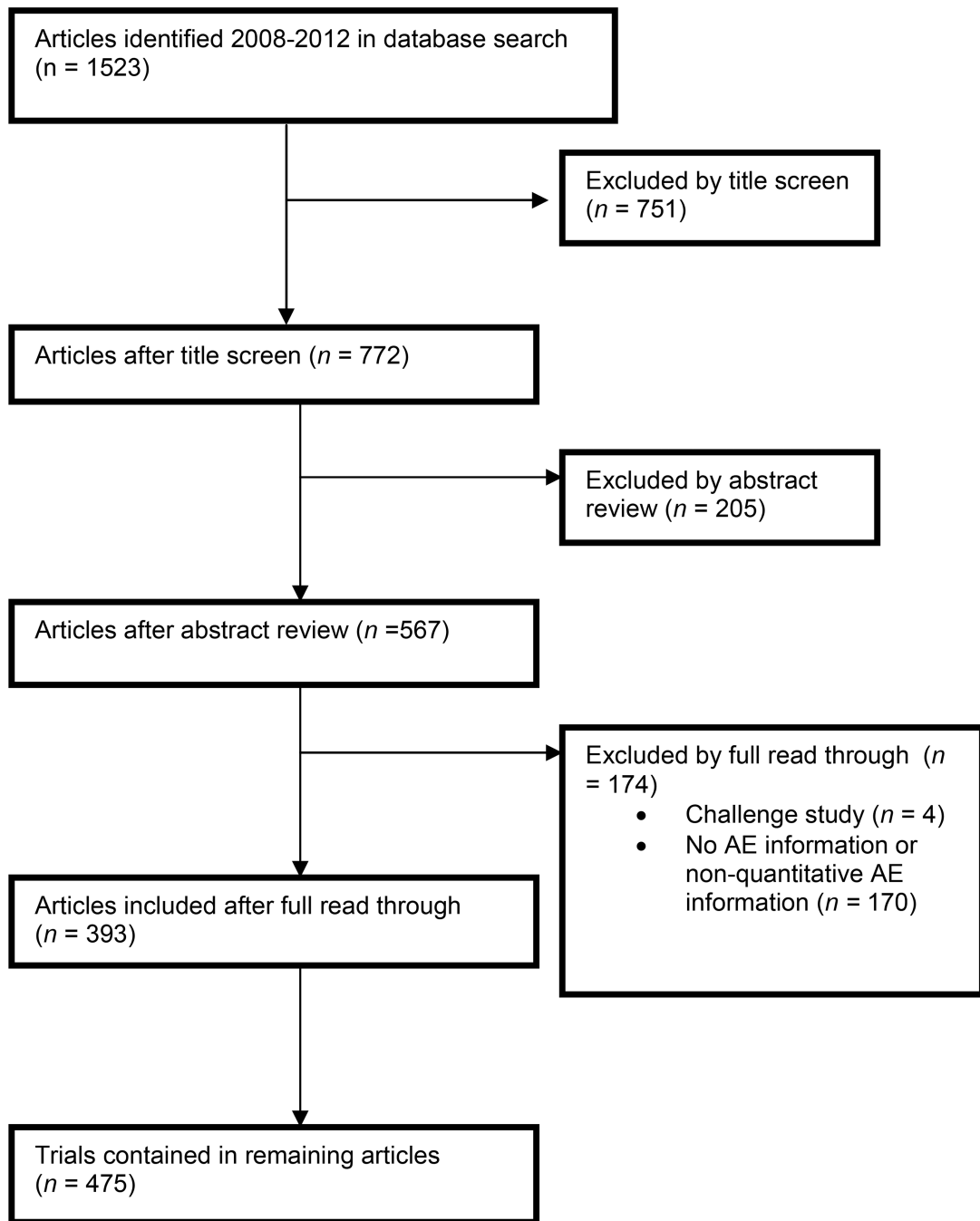


Figure 1.
Summary of literature search and selection

Table 1

Trial characteristics

N = 475 trials		N (%)
Investigational agent under study		
Non-antibiotic small molecule (e.g. Abiretone acetate, an orally administered enzyme inhibitor)		251 (52.8)
Vaccine (e.g. Inactivated whole-virus Ross River virus vaccine)		124 (26.1)
Biologic (e.g. TB-403, a humanized recombinant immunoglobulin (Ig)G1 monoclonal antibody)		40 (8.4)
Antibiotic (e.g. Telavancin, a lipoglycopeptide that inhibits cell wall synthesis in Gram-positive bacteria)		16 (3.4)
Other or unclear		44 (9.3)
Approval status of investigational agent		
Not approved		331 (69.7)
Previously approved, now testing:		94 (19.8)
	Drug-drug interactions	31 (6.5)
	Different formulation or addition of adjuvant	23 (4.8)
	Substantially different dosing	6 (1.3)
	Different population (e.g. infants; the elderly)	8 (1.7)
	Different disease indication	2 (0.4)
	Other	24 (5.1)
Mixed (e.g. one drug approved; other not approved)		29 (6.1)
Not approved but already widely used*		5 (1.0)
Other/unclear		16 (3.4)
Disease area under study[†]		
Viral diseases		68 (14.3)
Immune system diseases		62 (13.1)
Heart and blood diseases		51 (10.7)
Bacterial and fungal diseases		50 (10.5)
Nervous system diseases		35 (7.4)
Symptoms and general pathology		24 (5.1)
Parasitic diseases		22 (4.6)
Gland and hormone related diseases		21 (4.4)
Cancers and other neoplasms		21 (4.4)
Urinary tract, sexual organs, and pregnancy conditions		19 (4.0)
Digestive systems diseases		17 (3.6)
Respiratory tract (lung and bronchial) diseases		17 (3.6)
Blood and lymph conditions		14 (2.9)
Behaviors and mental disorders		11 (2.3)
Muscle, bone, and cartilage diseases		11 (2.3)
Eye diseases		9 (1.9)

N = 475 trials	N (%)
Nutritional and metabolic diseases	4 (0.8)
Skin and connective tissue disorders	3 (0.6)
Ear, nose, and throat diseases	2 (0.4)
Wounds and injuries	1 (0.2)
Not classifiable	20 (4.2)
Funding source	
Pharmaceutical/biotech/medical device industry	289 (60.8)
Government	76 (16.0)
Government + pharmaceutical/biotech/medical device industry	24 (5.1)
Non-profit/NGO	15 (3.2)
Academic + government	4 (0.8)
Academic	4 (0.8)
Academic + pharmaceutical/ biotech/medical device industry	1 (0.2)
Other/unclear	62 (13.1)
Trial unit type	
Contract research organization (CRO)	144 (30.3)
Hospital	53 (11.2)
Academic or nonprofit research unit [‡]	49 (10.3)
Pharmaceutical company owned research unit	30 (6.3)
Government or army-owned testing site	10 (2.1)
Mixed [§]	28 (5.9)
Other or unclear	161 (33.9)
Location	
North America	169 (35.6)
Europe	131 (27.6)
Asia	83 (17.5)
Africa	14 (2.9)
Australia	10 (2.1)
Multiple continents	7 (1.5)
Central or South America	4 (0.8)
Not given	57 (12.0)

* Article explicitly states agent has been in use for many years (e.g. piperazine phosphate tested before EMEA approval)

[†]Total percentage > 100 because some trials tested drugs aimed at multiple target disease areas.

[‡]Not otherwise specified as hospital.

[§]Multiple units falling under multiple categories.

Table 2

Participant characteristics (total n=27185)

	N (%)	Observed N	Missing N
Study arm *			
Treatment arm	24988 (91.9)		
Placebo arm	2947 (10.8)	27026 (99.4)	159 (0.6)
Sex			
Male	13573 (49.9)		
Female	10696 (39.4)	24269 (89.3)	2916 (10.7)
Race/ethnicity			
Caucasian	13550 (49.8)		
Asian	2877 (10.6)		
Black/African-American/African	2821 (10.4)		
Hispanic/Latino	1063 (3.9)		
Other	739 (2.7)	21050 (77.4)	6135 (22.6)
Mean age	31.2 (SD = 10.0)	20925 (77.0)	6260 (23.0)
Mean weight	71.2 (SD = 16.2)	8782 (32.3)	18403 (67.7)
Mean BMI	24.8 (SD = 4.0)	7079 (26.04)	20106 (74.0)

* Study arms total to >100% and >27,185 participants because of crossover studies that have the same participant crossover to both the treatment and placebo group.

Table 3

Absolute number and median incidence of AE's. Total AE = 52798. Total participants for each severity level of AE varies due to irregular reporting. A total of $n = 9747$ participants experienced no AE.

Adverse event (AE) severity	Number of AE's (% of total AE's)*	Median incidence of AE's across all agent types (interquartile range)		Median incidence of AE's for vaccines alone (interquartile range)		Examples of AE's
		N per 1000 participants	N per 1000 participants per day of AE monitoring [†]	N per 1000 participants	N per 1000 participants per day of AE monitoring	
Serious (includes death)						
Possibly related	15 (0.03)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	Overdose of study medication; miscarriage; severe malaise
Unrelated	269 (0.5)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	Alcohol withdrawal; acute bronchitis; gastric ulcer
Treatment arm	N/G [‡]	N/G	N/G	N/G	N/G	--
Placebo arm	N/G	N/G	N/G	N/G	N/G	--
Death						
Possibly related	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	None
Unrelated	5 (0.01)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	Ischemic stroke; post-operative embolism; cocaine overdose; traffic accident; suicide in person with history of depression
Treatment arm [§]	4 (0.008)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	Ischemic stroke; post-operative embolism; traffic accident; suicide in person with history of

Adverse event (AE) severity	Number of AE's (% of total AE's)*	Median incidence of AE's across all agent types (interquartile range)		Median incidence of AE's for vaccines alone (interquartile range)		Examples of AE's
		N per 1000 participants	N per 1000 participants per day of AE monitoring [†]	N per 1000 participants	N per 1000 participants per day of AE monitoring	
Placebo arm ⁴	1 (0.002)	0 (0, 0)	0 (0, 0)	0 (0, 0)	0 (0, 0)	depression Cocaine overdose
Severe						
Total	1944 (3.7)					Fever; nasopharyngitis
Treatment arm		0.00 (0.00, 12.96)	0.00 (0.00, 0.25)	82.75 (1.14, 153.57)	0.41 (0.00, 1.61)	Vasovagal syncope; vomiting
Placebo arm		0.00 (0.00, 0.00)	0.00 (0.00, 0.00)	0.00 (0.00, 25.00)	0.00 (0.00, 0.68)	Anemia; headache; croup
Mild or Moderate						
Total*	50570 (95.8)					Diarrhea; nausea; headache
Treatment arm		1147.19 (651.52, 1730.90)	46.07 (17.80, 77.19)	2473.21 (1539.97, 5892.26)	42.83 (16.47, 100.79)	Diarrhea; nausea
Placebo arm		701.39 (184.52, 1175.00)	22.60 (5.49, 51.25)	1583.33 (1041.67, 3175.00)	22.41 (10.23, 42.01)	Injection site pain; common cold
Mild versus moderate						
Mild [†]	21773	916.67 (333.33, 2307.69)	38.37 (18.11, 149.57)	2501.67 (1040.57, 4461.25)	40.82 (14.35, 146.62)	Vomiting; fatigue; pain
Moderate	702	55.56 (0.00, 333.33)	2.62 (0.00, 11.02)	500.00 (180.79, 932.39)	5.29 (2.04, 34.02)	Diarrhea; nausea; headache

* Includes all trials reporting the number of serious AE's and/or the number of mild-severe AE's, even if trial did not indicate the breakdown of treatment versus placebo group AE's or treatment versus placebo group participants for mild-severe events.

[†] 349 trials with n = 22635 participants reported the duration of AE monitoring

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript

[¶] N/G = data not given in the article

[§] For $n = 159$ participants, treatment versus placebo group assignment was unclear

^{*} This total derives from both trials that separated events into mild versus moderate and trials that reported the total number of mild/moderate AE's but did not separate the two severity levels. Therefore, it is greater than the sum of mild and moderate AE's reported below, which only include trials that separate these AE's by mild versus moderate severity level.

[†] For the $n = 201$ trials that separate mild from moderate AE's

Table 4

Features of a trial that predict increased risk

	Severe adverse event incidence rate ratio (95% CI)	P
Characteristic		
Vaccine agent type (reference category: other/unclear agent type)	6.21 (3.71, 10.37)	<0.001
Small molecule agent type (reference category: other/unclear agent type)	0.33 (0.19, 0.57)	<0.001
Funded by government, nonprofit, or academic source (reference category: funded by pharmaceutical or biotech company)	0.08 (0.03, 0.20)	<0.001
Hosted by academic or nonprofit trial site (reference category: hosted at non-university and non-government hospital)	5.06 (2.25, 11.40)	<0.001
Conducted in low or middle income country (reference category: conducted in high-income country)	0.32 (0.18, 0.56)	<0.001

Variables excluded during forward stepwise regression: biologic agent type (reference category: other/unclear agent type); antibiotic agent type (reference category: other/unclear agent type); held at contract research organization or pharmaceutical owned site (reference category: non-university and non-government hospital); held at government or army-owned site (reference category: non-university and non-government hospital); not previously approved (reference category: previously approved); mean BMI; percentage of female participants