Published in final edited form as:

J Empir Res Hum Res Ethics. 2011 March ; 6(1): 21–24. doi:10.1525/jer.2011.6.1.21.

A Delphi Study to Establish Important Aspects of Ethics Review

Michael E. Geisser, Kevin N. Alschuler, and Raymond Hutchinson

University of Michigan

Abstract

Little research has been done to examine the cost-effectiveness of REC review, or the components of review that make the greatest contributions to the protection of human subjects. We describe a process used to obtain consensus on the important categories and outcomes of REC review using the Delphi method and an array of stakeholders in a limited domain of research (biomedical). Study participants recruited from the University of Michigan Medical School's RECs, REC council, and principal investigators identified the following six categories as being the most important aspects of REC review: (1) A favorable risk/benefit ratio; (2) minimization of risk to subjects; (3) clarity of consent; (4) protection of vulnerable populations; (5) protection of privacy and confidentiality; and (6) review time. We believe that this kind of information can be used to assist in the development of a metric to assess the effectiveness and efficiency of REC review in the various research domains.

Keywords

REC; IRB; ethical review; REC review time; Delphi method

As noted in the article by Abbott and Grady in this issue, considerable research is needed to examine the structure and organization, processes, and outcomes of research ethics committees (RECs). We provide a brief overview of how the Delphi technique was used at our institution to develop consensus regarding the important aspects of REC review, as well as operational definitions of each of the important elements identified. Specific details regarding the methods and results are available in a supplemental online document at dx.doi.org/10.1525/jer.2011.6.1.3-19.

Any examination of concerns about RECs, such as Abbott and Grady review, focuses on a wide range of problems over a broad spectrum of institutions, and typically is based on an assessment by a limited set of stakeholders such as researchers. We suggest that this may be an inadequate approach to developing a metric of REC performance adequacy for one's particular institution. Any given institution faces a limited range of REC problems relative to all that have been catalogued. For example, problems due to an REC's unfamiliarity with relatively new qualitative methods such as photo-voice, or the methodology of oral historians or urban ethnographers, are unlikely to rise to a level of importance in medical school RECs; and problems of poor risk/benefit assessment are unlikely to rise to a level of importance in social or behavioral RECs. Conversely, the performance of RECs based on the assessments of investigators alone is likely to focus only on aspects of REC performance that researchers find inconvenient, while surveys of REC performance based on the assessment of REC members alone are likely to focus on factors that investigators do not

take seriously. We suggest that for large institutions, such as the University of Michigan Medical School which has many RECs, a better approach to discovering and characterizing the important roles of the REC is to survey a broad array of the stakeholders across the many biomedical RECs in that particular setting. Their collective judgment about that medical school's RECs' most important roles is probably less biased and more useful than some of the currently available assessments. In turn, such an assessment, based on a robust sample of RECs and stakeholders, may prove generalizable and therefore useful to similar other institutions, including much smaller medical schools that have just one or two RECs, and hence a less robust sample of stakeholders.

The Delphi technique (see Hsu & Bradford, 2007, and Powell, 2003, for reviews), originally developed by the Rand Corporation, is a widely accepted method used to obtain consensus from a group of experts on a particular topic. The technique is a group communication process designed to examine and discuss a specific topic in depth. The technique is conducted in stages, with initial stages being more open-ended. In subsequent stages, participants are given feedback on the prior results, and are asked to reconsider their responses given the information obtained in prior rounds. Consensus on a topic is generally achieved by having participants provide ratings, and examining group responses. Although there is no set number of iterations recommended for the technique, consensus can be generally reached in three rounds. The Delphi technique is frequently used to develop clinical guidelines for treating certain illnesses or injuries, as well as developing recommendations for important assessment domains that should be measured in clinical trials of treatments for various diseases.

In the present study, we sought to obtain data on the important aspects of REC review, and operational definitions of these key concepts, through polling a sample of institutional researchers, regulatory officials, ethicists, and REC members at the University of Michigan Medical School. A Delphi method was utilized to identify and rank the importance of each aspect of REC review, as well as to provide input into and refine operational definitions of each of these factors. Ethical aspects of review, as well as review time, were included as factors examined in the survey, given the charge in the regulations to assure human subjects protections and to promote clinical research by conducting efficient and timely review.

Method

Participants

Participants were recruited from the University of Michigan's REC, REC council, and bioethics committees, and principal investigators (PIs) who were among the top 20% with active research protocols within the medical school.

Procedures

This study was approved by the University of Michigan Institutional Review Board. Participants were fully informed of the procedures in a cover letter sent with the questionnaire. A waiver of written documentation of consent was obtained for this research.

Briefly, a Delphi study design was used to rank the importance of different factors in REC review and reach consensus on operational definitions for the measurement of each construct. The Delphi exercise included three rounds of questionnaires distributed to participants by e-mail. The researchers initially selected some items for inclusion in the first questionnaire based on the components presented by Taylor (2007), the Common Rule, and review time items. For Round 1, participants were asked to assign a point value to each item to indicate the level of importance of that item to REC review, such that the rankings of all items in the questionnaire totaled 100 points. Participants were also allowed to include their

own items, and rank them as well. A similar format was used to have respondents rate outcomes representative of effective REC review, which served as the basis for developing operational definitions for the review items.

Round 2 consisted of having participants re-rank the top six responses for items and outcomes for Round 1. Also, in this round, we attempted to match items to outcomes, and asked respondents to indicate whether they agreed or disagreed with the pairings.

Finally, the Round 3 questionnaire contained the six highest-ranked categories of regulatory review and associated operational definitions. These categories and operational definitions were refined based on the Round 2 responses. Participants were asked to indicate their agreement with the definition and were provided space to suggest improvements to the definition if they disagreed.

Results

Twenty-nine out of 71 persons invited to participate in the study responded to the first questionnaire. In the first round, a favorable risk/benefit ratio was rated as the most important aspect of REC review. The two most highly rated outcomes in terms of importance were the length of the review and the number of serious adverse events that occur in a study. In this round, REC council members rated a favorable risk/benefit ratio higher compared to REC members and principal investigators, and principal investigators weighted prompt and clear communication to the study team higher compared to REC members. Six review categories were selected for Round 2: (1) a favorable risk/benefit ratio; (2) risks to subjects are minimized; (3) the consent process is clear; (4) there are adequate protections for vulnerable populations; (5) there are appropriate privacy protections; and (6) the review time is low. Ratings from this round ($n = 19$) revealed three categories that were highly rated: risk/benefit ratio, minimization of risk, and clarity of consent. Additionally, participants generally agreed with the outcomes associated with each review category, and REC members rated subject understanding of the study higher when compared to PIs. Some participants did add additional categories and outcomes, but none of these were mentioned by other participants, so additional categories were not included in later rounds.

In Round 3, 20 participants (65% of sample that agreed to participate in Round 1) responded to the questionnaire. Consensus was reached on operational definitions of five of the six categories of REC review. The only definition to not receive overwhelming support was the item on review time. The participants who disagreed with that definition suggested that investigator satisfaction with review be removed from the definition.

Discussion

Abbott and Grady indicate that while it is generally recognized that there is a need to evaluate REC effectiveness, no study in their literature review included an assessment of the quality of REC review. In addition, they indicate that additional research is needed to understand how RECs work, the issues they find important, what constitutes a quality REC review, and whether REC review is effective in terms of protecting human research participants. We describe how the Delphi method was employed at our institution to reach consensus on the important aspects of REC review, and operational definitions of each of these dimensions. We plan to use this information to inform the development of a metric to assess the effectiveness of REC review in the University of Michigan Medical School. It remains to be seen whether the data generated from this study can be used to develop a reliable and valid measure of the quality of REC review in this and other related institutions.

The present study revealed important information regarding opinions on the key elements of REC review, as well as operational definitions of each of these categories. Specifically, the sample reached a general consensus that the most important elements of REC review are assuring that the study has a favorable risk/benefit ratio, risks to subjects are minimized, the consent process is clear, there are adequate protections for vulnerable populations, there are appropriate privacy protections, and review time is appropriate. These findings are similar to the two prior studies that examined the important aspects of ethical review (Advisory Council on Human Radiation Experiments, 1995; Meslin et al., 1994), with the exception, of course, that review time is not an ethical consideration element. In addition, the sample was able to reach consensus on operational definitions of each of these elements, with the exception of review time. With regard to this latter element, there was disagreement as to when the review was considered complete (i.e., when contingent approval was given by the REC versus final approval granted after the contingencies were met).

We believe that the Delphi method is one research tool that can be used to examine certain aspects of the key research questions raised above by Abbott and Grady. Our panel of “experts” was limited to members of our institution, and to persons and investigators associated with our biomedical REC. Especially in view of our findings, it is likely that RECs with a different content focus (e.g., social science), or participants from other kinds of institutions, would have provided different responses to the surveys.

Typically, prior studies that have examined the important aspects of REC review have recruited only REC members as subjects (Meslin et al., 1994). We believe that a strength of our study is the heterogeneity of the sample across stakeholders in the IRB process, which potentially increases the generalizability of the findings. However, the response rate to the Round 1 questionnaire in the current study was low (29 of 71), and we had greater difficulty recruiting PIs compared to other kinds of participants. Perhaps, REC members and other persons who work in the areas of research regulation and/or ethics may have more of a personal vested interest in participating in this type of research. It is possible that recruitment in general could have been improved by using financial incentives, as no incentives were provided for participation in the current study.

Other limitations of the study may include the sample. However, the goal of Delphi studies is to not necessarily obtain a representative sample in a statistical sense, but to have the panel be representative of the expert opinion sought. One review (Ludwig, 1997) suggests that most Delphi studies should recruit between 15–20 participants, and the present study obtained responses from this number of participants or more in each of the three rounds. Also, it should be noted that the categories produced in this study may not exhaust all of the factors deemed important in terms of REC review of biomedical research.

The factors related to review time and ethical consideration, while both deemed important, represent two quite different dimensions of how REC review could be evaluated. Thus, quantitative or qualitative evaluation of these dimensions should be considered separately, as scores or outcomes on these different dimensions likely have very different meanings or implications. For example, a poor score on the ethical review categories might suggest that a particular protocol might not be approvable or have significant ethical weaknesses, while an evaluation of review time may have no relationship to the ethical quality of a study (although a study of poor ethical quality might undergo several revisions before it reaches an approvable state, thus increasing the time of review).

In conclusion, this study represents a first step in developing a measure of the quality of REC review at our institution. This study provides insight into the important dimensions of REC review from the viewpoint of a heterogeneous sample of stakeholders in a specific

biomedical setting. In addition, consensus was reached on operational definitions of the various dimensions that were deemed important, with the exception of review time. We believe that the results will be useful in assisting in the development of an index of the quality of REC review at our institution, which could then be used to assess REC performance and examine how changes to the REC process impact review quality as well as speed of review.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

This study was supported by the University of Michigan's Clinical and Translational Science Award (CTSA grant # UL1RR024986).

References

- Advisory Committee on Human Radiation Experiments. Final report: Advisory Committee on Human Radiation Experiments (Publication No. 061-000-00-848-9). Pittsburgh, PA: U.S. Government Printing Office; 1995.
- Hsu, C.; Sandford, BA. The Delphi technique: Making sense of consensus. Practical Assessment Research & Evaluation. 2007. Retrieved from <http://pareonline.net/getvn.asp?v=12&n=10>
- Ludwig, B. Predicting the future: Have you considered using the Delphi methodology?; Journal of Extension. 1997. p. 1-4. Retrieved from <http://www.joe.org/joe/1997october/tt2.html>
- Meslin EM, Lavery JV, Sutherland HJ, Till JE. Judging the ethical merit of clinical trials: What criteria do research ethics board members use? REC: A Review of Human Subjects Research. 1994; 16(4): 6–10.
- Powell, C. The Delphi technique: Myths and realities; Journal of Advances Nursing. 2003. p. 376-382. Retrieved from http://rachel.org/files/document/The_Delphi_Technique_Myths_and_Realities.pdf
- Taylor HA. Moving beyond compliance: Measuring ethical quality to enhance the oversight of human subjects research. IRB: Ethics & Human Research. 2007; 29(5):9–14. [PubMed: 17969778]

Biographies

Michael E. Geisser is Professor of Physical Medicine and Rehabilitation and Co-Chair of the Medical School Institutional Review Board, University of Michigan, Ann Arbor, MI.

Kevin N. Alschuler is a Fellow in the Department of Rehabilitation Medicine, University of Washington, Seattle, WA.

Raymond Hutchinson is Professor of Pediatrics and Communicable Diseases and Associate Dean for Regulatory Affairs, University of Michigan Medical School, Ann Arbor, MI.