U.S. Department of Veterans Affairs

Public Access Author manuscript

J Empir Res Hum Res Ethics. Author manuscript; available in PMC 2016 December 01.

Published in final edited form as:

J Empir Res Hum Res Ethics. 2015 December ; 10(5): 460–469. doi:10.1177/1556264615612195.

Using the IRB Researcher Assessment Tool to Guide Quality Improvement

Daniel E. Hall^{a,b}, Barbara H. Hanusa^a, Bruce S. Ling^{a,b}, Roslyn A. Stone^{a,b}, Galen E. Switzer^{a,b}, Michael J. Fine^{a,b}, and Robert M. Arnold^b

^aVA Pittsburgh Healthcare System

^bUniversity of Pittsburgh

Abstract

Institutional Review Boards (IRBs) are intended to protect those who participate in research. However, because there is no established measure of IRB quality, it is unclear if these committees achieve their goal. The IRB Researcher Assessment Tool is a previously validated, internally normed, proxy measure of IRB quality that assesses 45 distinct IRB activities and functions. We administered this instrument to a sample of investigators and IRB members at a large urban VA Medical Center. We describe a systematic approach to analyze and interpret survey responses that can identify the IRB activities and functions most in need of quality improvement. The proposed approach to empirical data analysis and presentation could inform local initiatives to improve the quality of IRB review.

Keywords

Research Ethics Committee; Institutional Review Board; Quality Improvement; Engineering

Since 1974, the Code of Federal Regulations has required Institutional Review Board (IRB) evaluation of all human subjects research conducted in the United States of America. The purpose of IRB review is to protect research participants. The annual institutional cost of this review is substantial, ranging from approximately \$100,000 for low volume IRBs to more than \$1,000,000 for high volume IRBs (Wagner, Bhandari, Chadwick, & Nelson, 2003). However, it is not clear whether these committees actually improve the protection of human research participants because there is no established measure to assess the quality of the IRB review process (Coleman & Bouesseau, 2008).

Methodological and conceptual barriers limit the development of direct measures of IRB protections (Taylor 2007). For example, no system exists currently to aggregate the adverse event reports across studies, sites, and IRBs. Furthermore, the rarity of critical failures, such as the death of Jesse Gelsinger from experimental gene transfer (Fiscus, 2001), makes it difficult to detect the impact of regulations intended to mitigate risks of such catastrophic events. Existing attempts to measure IRB quality have focused on proxy measures, such as

administrative compliance (Tsan, Smith, & Gao, 2010) or self-reports (Keith-Spiegel & Tabachnick, 2006).

The IRB Researcher Assessment Tool (IRB-RAT) is a self-report measure of IRB quality (Keith-Spiegel & Tabachnick, 2006) that consists of 45 statements (“items”) that describe a variety of IRB activities and functions (Table 1). For each item, respondents use a 7-point Likert scale to indicate how well the statement describes their “ideal” IRB as well as their “actual” IRB (1=definitely does not describe; 2=does not describe; 3=only slightly describes; 4=describes somewhat; 5=describes well; 6=describes very well; 7= describes extremely well). The IRB-RAT functions as a self-report measure of IRB performance that is internally normed to each respondent’s standard of ideal quality for each activity or function.

The IRB-RAT was validated initially in a sample of 886 behavioral scientists and biomedical researchers, providing an initial, ideal rating of IRB functions (Keith-Spiegel & Koocher, 2005; Keith-Spiegel & Tabachnick, 2006). A subsequent administration of the IRB-RAT to 115 investigators, research coordinators and IRB committee members (Reeser, Austin, Jaros, Mukesh, & McCarty, 2008) demonstrated that the ratings of the IRB activities differed according to respondent role (e.g., investigator, IRB committee member). Because it is unclear how to make the results from the IRB-RAT actionable, this instrument has not yet been used to guide IRB quality improvement. We describe a systematic approach for analyzing responses to the IRB-RAT that can inform processes of quality improvement by identifying those IRB activities and functions most in need of improvement, defined as those items with the greatest discrepancy between the ratings of the actual and ideal IRBs as well as comparatively high ratings of the ideal IRB.

Methods

We designed an anonymous electronic survey using Survey Monkey,TM to assess attitudes and opinions about the IRB of one large VA Medical Center using the IRB-RAT. We emailed the survey to all principal investigators and project coordinators listed on the IRB’s portfolio of active protocols during the month of April 2010. We also sent the survey to all members and staff of the IRB. The survey was open for 10 days, with reminder emails sent to non-responders on days 3, 7 and 10. All responses were recorded anonymously. These methods were reviewed by the IRB of the VA Pittsburgh Healthcare System and determined to be exempt from IRB oversight.

We downloaded survey responses into SPSS (IBM Corp. Released 2012, IBM SPSS Statistics for Windows, Version 21.0, Armonk, NY), assessed the data quality, and summarized response rates by respondent type (i.e., investigator/project coordinator or IRB member/staff). For each IRB-RAT item, we computed the sample averages of ratings for the ideal and actual IRB as well as the average difference between the actual and ideal IRB ratings. We assessed concordance with the national validation sample using Pearson correlation (r). We then constructed bivariate scatter plots for each respondent type to highlight associations between the item-specific ideal ratings and discordant ratings of the ideal and actual IRBs. Reference lines on these plots were estimated from a linear mixed

model fit using Stata version 13. (See appendix [online digital content] for model details and Stata code).

Results

In April 2011, we emailed questionnaires to 178 principal investigators/project coordinators and 28 IRB members/staff (Table 2). The 98 individuals who replied initially included 31 who chose not to participate after reading an explanation of the survey. Of the 67 respondents who initiated the survey (32.5% of those sampled), 47 (70%) completed the entire survey; responses to some RAT items were missing. Overall, about 10% of the IRB RAT item responses were missing, mostly items at the end of the survey regarding actual IRB performance. The estimated median time to complete the IRB-RAT was about 13 minutes.

The average ratings for each of the 45 IRB-RAT items are ordered by their ranking in the national validation sample for the ideal IRB (Table 1). The VA sample ratings of the ideal IRB by investigators/project coordinators and IRB members/staff correlate reasonably well with the national validation sample ($r = 0.75$ and 0.68 , respectively, $p < 0.001$ for each). The average difference between the actual and ideal IRB was calculated from the paired data where both ratings were available. Due to missing data, slight discrepancies are reported in Table 1 between the item-specific average differences and the average of the respondent-specific differences between ratings for the actual and ideal IRBs on each item.

The average ratings for each of the 45 IRB-RAT items are plotted in Figure 1 by IRB type and respondent type. As expected, both respondent types rated the actual IRB lower than the ideal IRB. Based on the linear mixed model, the estimated mean item ratings for the ideal IRB were similar for both investigators/project coordinators and IRB members/staff ($p = 0.19$). However, investigators/project coordinators rated their actual IRB lower than did the IRB members/staff (estimated mean difference = -0.54 , 95% CI: -0.92 to -0.15 ; $p < 0.01$). In addition, the discrepancy between the actual and ideal IRBs for investigators/project coordinators (-1.50) was nearly twice that of IRB members/staff (-0.79 , 95% CI for difference: -0.83 to -0.59 , $p < 0.001$).

Figure 2 provides two additional nuances. First, it includes a graphical representation of the error around each point estimate and these standard errors are not presented in Table 1. Second, Figure 2 demonstrates the spatial relationship between the items according to the 2-dimensional scoring paradigm, and it is this spatial relationship that provides the key to our interpretation. The figure shows the cross-classification of the average difference between the actual and ideal ratings for each item by the average ratings of the ideal IRB for IRB members/staff (Figure 2a) and principal investigators/project coordinators (Figure 2b). Quadrants are defined by the estimated mean ratings of the ideal IRB and the estimated mean differences across all 45 items for the respective respondent groups and dotted lines represent the corresponding 95% confidence intervals, based on the linear mixed model.

From left to right in Figure 2, the 45 IRB-RAT items in the x-axis are ordered by increasing “ideal” rating, with item 7 rated the highest by IRB members/staff (Figure 2a), and items 2

and 24 rated the highest by principal investigators/study coordinators (Figure 2b). From top to bottom, items are plotted according to the magnitude of the difference between the actual and ideal IRB ratings such that items with the greatest differences are plotted at the bottom of each graph, i.e., item 33 for IRB members/staff (Figure 2a) and item 30 for investigators/project coordinators (Figure 2b).

Items above the upper confidence limit for the mean difference between the actual and ideal ratings of the IRB denote those items with relatively “good performance”; they are highlighted blue in both Figure 2 and Table 1. Items below the lower confidence interval (red) denote relatively “poor performance”, where the discordance between the actual and ideal IRB ratings is greatest. Items above the confidence limit for the ideal IRB (yellow) are most reflective of the ideal IRB and indicate the items of greatest importance.

Items highlighted orange have both relatively poor performance and high importance, thus constituting the items with the greatest need for quality improvement. Investigators/project coordinators rate 7 items in this region (Figure 2b, items 1, 2, 6, 8, 9, 10, 21, 24 and 30), including timeliness of review (item 1), absence of bias (item 21) and timely action for regulatory violation (item 30). By contrast, IRB members/staff do not rate any items as having a relatively greater need for quality improvement (Figure 2a). However, the halos around each item reflect uncertainty around the point estimates (radius= 1 standard error), and the items whose halos extend into the orange region of Figure 2a are those closest to our criteria for items in greatest need for quality improvement. Reference to Table 1 confirms these as items 1, 3, 6 and 9 for IRB members/staff; these items also were identified by the investigators/project coordinators (Figure 2b, orange region).

Items highlighted green have both relatively good performance and high importance. Protection of human participants (item 7, Figures 2a and 2b), knowledgeable IRB membership (item 5, Figure 2b) and management of conflicted interests (item 22, Figure 2a) are each rated in this way by either or both respondent types.

Many of the items with good performance (blue) are located below the lower confidence interval of the mean rating of the ideal IRB, indicating that the good performance is found in areas that are comparatively less descriptive of the ideal IRB (i.e., less important). In fact the single highest rated item was also the least important (Figures 2a and 2b, item 44), suggesting that although the IRB was good at making editorial suggestions, those suggestions were not particularly valued by either the investigators/project coordinators or the IRB members/staff. Finally, when comparing the sets of good and poor performing items for both respondent types, there is agreement between IRB members/staff and investigators/project coordinators that the highest performing items are 39 and 40, the lowest performing items are 4, 21 and 27, and that the items in greatest need of improvement are 1, 2, 6 and 9.

Discussion

This report describes a novel strategy for analyzing the IRB-RAT to inform processes of IRB quality improvement. Stakeholders in an IRB review process rated their IRB's performance across 45 activities and functions. By analyzing their relative ratings, we

identified a limited set of activities and functions most in need of quality improvement as well as those activities and functions with relatively good performance. We suggest that these empirical data could inform local processes of quality improvement and provide the evidence necessary to catalyze local change.

For example, our data demonstrate that this IRB is doing a comparatively good job protecting human subjects (item 7) and monitoring projects in line with federal regulations (item 39). However, there is similar consensus that improvement is most needed for a limited set of activities and functions that include the timeliness of review (item 1) along with perceptions among investigators/project coordinators that the IRB is biased (items 2 and 21), requires more resources to carry out its functions efficiently (item 24), and does not always conduct its evaluations conscientiously (item 6) or explain its determinations completely (item 9).

Armed with these data, stakeholders in the IRB review process might propose initiatives to improve the perception of the IRB and the timeliness of review. Policies could be drafted to mitigate bias against particular research techniques, and new members with expertise in such areas could be recruited. Those responsible for leading IRB meetings and drafting correspondence with investigators could formalize ways to improve the clarity of communication and the conscientiousness of the reviews. The IRB-RAT could then be administered at regular intervals to measure the impact of these initiatives.

Local initiatives for quality improvement might be further informed if the IRB-RAT were administered to a representative sample of IRBs and investigators across a range of disciplines, methodologies, centers and regions. Findings from such a representative sample would provide a comprehensive picture that could place local initiatives for quality improvement within the broader context of IRB performance.

We are not aware of any other actionable measure of IRB quality that could guide processes of quality improvement. While the IRB-RAT remains only a proxy measure of protections that IRBs are intended to provide, it is both comprehensive in scope and flexible to administer with a relatively light respondent burden at approximately 13 minutes. Furthermore, if stakeholders in the IRB review process perceive concerns not covered by the 45 items currently in the IRB-RAT, new items pertinent to those specific concerns could be drafted easily. Administration of the new items could probe the relative importance of emerging concerns within the context of the other activities and functions that constitute the IRB's work.

We also found that investigators/project coordinators rated the IRB-RAT differently than did IRB members/staff, essentially doubling the size of the discrepancy between the actual and ideal IRBs. This is consistent with other data that showed that IRB members rated the IRB higher than investigators and research coordinators (Chenneville et al., 2014), and that investigators *who had served on an IRB* rated the actual IRB higher than research coordinators (Reeser et al., 2008).

These differences between respondent types may reflect a tendency for IRB members/staff to "cut themselves some slack" or for investigators/project coordinators to rate regulatory

oversight harshly. However, because the patterns of response change along with the magnitude of the ratings, we suggest the discrepancies between respondent types reflect substantive differences in perspective and value that could inform quality improvement. For example, if investigators are considered the end users and customers of IRBs, their concerns should be the primary focus of quality improvement. Alternatively, if IRB members are thought to have a broader perspective on the purposes of research oversight, their concerns may reflect important issues for quality improvement obscured by investigators' narrower focus on the efficient conduct of research. Further research is needed to further elucidate the cause of the consistently observed differences between respondent types.

Our specific findings about the IRB activities and functions with relatively good and poor performance are limited to the single site studied. They are further limited by a relatively low sample size and response rate. However, the analytical approach generalizes to other research centers with IRBs and could be applied by other investigators interested in improving IRB quality and performance. We used the 95% confidence interval to denote those items with relatively good or poor performance, but the standard error of each item's point estimates raise questions regarding the classification of borderline cases (e.g., items 7, 8 and 22 in Figure 2b or items 1, 3, 6, 7 and 9 in Figure 2a). However, stakeholders in the process of quality improvement can use their discretion when considering these borderline cases.

In conclusion, this approach to analyzing and interpreting data from the IRB-RAT provides an intuitive strategy for identifying the IRB activities and functions most in need of quality improvement. This information could focus local initiatives to monitor and improve the quality of IRB review.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgements

The authors would like to thank Ulrike Feske, PhD for her helpful comments on the manuscript. This research was supported by the US Department of Veterans Affairs, Veterans Health Administration, Office of Research and Development, Health Services Research and Development (CDA 08-281 and SDR 11-399-1). The funder did not participate directly in the collection, analysis or interpretation of data. The views expressed in this article are those of the authors and do not necessarily reflect the position or policy of the Department of Veterans Affairs or the United States government.

Biographies

Daniel Hall, MD, MDiv, MHSc, is a Core Investigator at the VAPHS Center for Health Equity Research and Promotion, Associate Professor of Surgery at the University of Pittsburgh, and member of the VAPHS IRB. He was funded by VA HSR&D to improve the quality and efficiency of IRB review throughout the VA network. He conceived the project, managed its execution and drafted the manuscript.

Barbara Hanusa, PhD, is a consulting statistician at the VAPHS Mental Illness Research, Education and Clinical Center (MIRECC). She has extensive experience analyzing

psychometric data. She was responsible for the data management, conducted the statistical analyses, and contributed to the drafting and critical review of the manuscript.

Bruce Ling, MD, MPH is a Core Investigator at the VAPHS Center for Health Equity Research and Promotion, Assistant Professor of Medicine at the University of Pittsburgh, and Chair of the VAPHS IRB. Based on his IRB experience he contributed to the design of the project, the interpretation of data and the drafting of the manuscript.

Roslyn Stone, PhD, is a Professor of Biostatistics at the University of Pittsburgh Graduate School of Public Health and a senior statistician in the Biostatistics and Informatics Core of the VAPHS Center for Health Equity Research and Promotion. She has extensive experience analyzing health services research data. She participated in the conceptualization and conduct of the analysis as well as drafting and critical review of the manuscript.

Galen Switzer, PhD is a Professor of Medicine and Psychiatry at the University of Pittsburgh with expertise in psychometrics and survey methodology. He is Co-chief of the Measurement Core at the VAPHS Center for Health Equity Research and Promotion. He assisted with the survey design, implementation and interpretation.

Michael Fine, MD is the Director of the VAPHS Center for Health Equity Research and Promotion and Professor of Medicine at the University of Pittsburgh. He has extensive expertise in health services research methodology. He contributed to the design of the project, interpretation of data, and editing of the final manuscript.

Robert Arnold, MD is a Co-Investigator on a recently completed NIH sponsored study of 20 IRBs at 10 of the largest university medical centers. He is Professor of Medicine at the University of Pittsburgh with extensive experience in research and clinical ethics. He contributed to the design of the project, the interpretation of the data and critical review of the manuscript

References

- Chenneville T, Menezes L, Bylsma LM, Mann A, Kosambiya J, Baxi R. Assessing Institutional Ethics Committees in India Using the IRB-RAT. *Journal of Empirical Research on Human Research Ethics*. 2014; 9(4):50–59. [PubMed: 25747296]
- Coleman CH, Bouesseau M-C. How do we know that research ethics committees are really working? The neglected role of outcomes assessment in research ethics review. *BMC Medical Ethics*. 2008; 9:6. [PubMed: 18373857]
- Fiscus PW. Protecting human research subjects. *J Biolaw Bus*. 2001; 4(4):18–22. [PubMed: 12751493]
- Keith-Spiegel P, Koocher GP. Institutional Review Board Researcher Assessment Tool (IRB-RAT) Users Guide. 2005 Retrieved from [www.ethicsresearch.com](http://www.ethicsresearch.com/images/IRB_RAT_User_s_Guide.revised_3-11-05.pdf) website: http://www.ethicsresearch.com/images/IRB_RAT_User_s_Guide.revised_3-11-05.pdf.
- Keith-Spiegel P, Tabachnick B. What scientists want from their research ethics committee. *Journal of Empirical Research on Human Research Ethics*. 2006; 1(1):67–82. [PubMed: 19385866]
- Reeser JC, Austin DM, Jaros LM, Mukesh BN, McCarty CA. Investigating Perceived Institutional Review Board Quality and Function Using the IRB Researcher Assessment Tool. *Journal of Empirical Research on Human Research Ethics*. 2008; 3(1):25–34. [PubMed: 19385780]
- Taylor HA. Moving beyond compliance: measuring ethical quality to enhance the oversight of human subjects research. *IRB: Ethics & Human Research*. 2007; 29:9–14. [PubMed: 17969778]

- Tsan M-F, Smith K, Gao B. Assessing the quality of human research protection programs: the experience at the Department of Veterans Affairs. *Irb: a Review of Human Subjects Research*. 2010; 32(4):16–19. [PubMed: 20853799]
- Wagner TH, Bhandari A, Chadwick GL, Nelson DK. The cost of operating institutional review boards (IRBs). *Academic Medicine*. 2003; 78(6):638–644. [PubMed: 12805049]

Best Practices

Quality improvement of IRB review depends on reliable and repeated assessment of quality. Although it remains only a self-report measure, we suggest that the IRBRAT is a useful tool to begin longitudinal quality measurement, and the analysis strategy described here provides actionable results that can inform initiatives for quality improvement.

Research Agenda

In addition to administering the IRB-RAT to representative samples of investigators and IRB members, future research could focus on how performance on the IRBRAT changes after implementing initiatives aimed at addressing some of the activities and functions identified as needing improvement.

Educational Implications

Quality improvement requires investment from all stakeholders in a system, including the IRB. Performance metrics should be shared with the stakeholders to clarify how they will be measured. Using the IRB-RAT will inform both investigators and IRB members about the range of activities and functions that comprise IRB quality, perhaps suggesting ways to improve, edit or expand the existing 45 question items. As change is measured over time, these findings should be communicated back to the stakeholders, hopefully demonstrating that their system is responsive to needed change.

Figure 1.

Item-specific average ratings of the ideal and actual IRB for IRB members/staff and investigators/project coordinators.

IRB member/staff ratings are plotted in blue circles. Investigator/project coordinator ratings are plotted in red squares. Ratings of the ideal IRB are plotted with solid circles and squares. Ratings of the actual IRB performance are plotted with open circles and squares. Rating are on a scale of 1 (“definitely does not describe”) to 7 (“describes extremely well”). Items are ordered according to the IRB members’ rating of the ideal IRB. Items are identified by item numbers in Table 1.

Figure 2.

Differences between actual and ideal IRB ratings according to the ideal IRB rating for both IRB members/staff (2a) and Investigators/project coordinators (2b).

Note: Each point plots the average ideal rating vs. the average difference between the actual and ideal IRB rating for a specific item from the IRB-RAT. Item numbers follow the key found in Table 1. Quadrants are defined by the estimated mean ratings of the ideal IRB and the estimated mean differences across all 45 items for the respective respondent groups and dotted lines represent the corresponding 95% confidence intervals, based on a linear mixed

model. The diameter of the circle around each point estimate represents the standard error of the average difference between the actual and ideal IRB ratings. Items highlighted blue and red are those with relatively good and poor performance, respectively. Items highlighted yellow are relatively more important, items highlighted orange are most in need of quality improvement, and those in green are the most important items with relatively good performance.

Table 1

Mean ratings of the 45 items from the IRB Researcher Assessment Tool for ideal IRBs from the national validation sample and investigators/project coordinators and IRB members/staff from the current sample as well as actual ratings of the IRB from the current sample.

	Item text	Ideal IRB			Actual IRB			Actual – Ideal IRB
		National Validation Sample	Investigator/Project Coordinator	IRB Member/Staff	Investigator/Project Coordinator	IRB Member/Staff	Investigator/Project Coordinator	
1	An IRB that reviews protocols in a timely fashion	6.43	6.62	6.36	4.45	5.23	-2.15	-1.15
2	An IRB whose members do not allow personal biases to affect their evaluation of protocols	6.17	6.68	6.36	4.86	5.50	-1.82	-0.92
3	An IRB that does a good job of upholding participants' rights while, at the same time, facilitating the conduct of research	6.10	6.37	6.36	4.82	5.33	-1.43	-1.17
4	An IRB that does not use its power to suppress research that is otherwise methodologically sound and in compliance with federal policy whenever it perceives potential criticism from outside the scientific community.	6.08	6.29	5.83	4.44	4.80	-1.89	-1.20
5	An IRB with members who are very knowledgeable about IRB procedures and federal policy	6.01	6.64	6.29	5.32	5.67	-1.34	-0.83
6	An IRB that conducts a conscientious and complete review of protocols	5.86	6.58	6.36	4.78	5.42	-1.84	-1.08
7	An IRB that views protection of human participants as its primary function	5.80	6.49	6.57	5.90	6.33	-0.52	-0.33
8	An IRB that responds in a timely manner to investigators' inquiries about its processes and decisions	5.80	6.50	6.00	4.74	5.25	-1.74	-0.83
9	An IRB that gives a complete explanation for any required changes to or disapprovals of protocols	5.73	6.60	6.36	4.47	5.33	-2.12	-1.08
10	An IRB that is willing to work with investigators to find mutually satisfying solutions whenever disagreements exist	5.71	6.55	6.07	4.80	5.08	-1.75	-1.17
11	An IRB whose members fully understand and act within the scope of their function	5.67	6.43	6.29	5.12	5.69	-1.31	-0.62
12	An IRB that includes a complete rationale when it denies or mandates changes in a protocol based on criteria that are more stringent than or different from federal research policy (i.e., application of "local standards")	5.59	6.25	6.33	5.09	5.83	-1.00	-0.50
13	An IRB that views its role as being an investigator's ally rather than as being a hurdle to clear	5.57	6.37	6.21	4.85	5.33	-1.42	-0.92
14	An IRB that conducts a conscientious, informed analysis of potential benefits weighed against potential risks before making decisions	5.54	6.06	6.21	5.14	5.85	-0.93	-0.38
15	An IRB that takes timely and appropriate action whenever scientific misconduct is alleged	5.52	6.42	6.29	5.37	5.83	-1.00	-0.50

	Item text	Ideal IRB			Actual IRB			Actual – Ideal IRB	
		National Validation Sample	Investigator/Project Coordinator	IRB Member/Staff	Investigator/Project Coordinator	IRB Member/Staff	Investigator/Project Coordinator	IRB Member/Staff	
16	An IRB that is open to reversing its earlier decisions (i.e., willing to carefully listen to investigators' appeals)	5.52	6.36	6.21	4.00	5.58	-2.37	-0.75	
17	An IRB that invites investigators to present their position whenever a question or concern about a research protocol arises	5.51	6.34	5.86	4.38	5.08	-1.98	-0.92	
18	An IRB that maintains complete and accurate records	5.50	6.62	6.36	5.14	5.92	-1.48	-0.50	
19	An IRB that can competently distinguish exempt from nonexempt research	5.48	6.41	6.14	4.00	5.17	-2.46	-1.00	
20	An IRB that treats investigators with respect	5.45	6.23	6.21	5.08	5.33	-1.14	-1.00	
21	An IRB that holds no preconceived biases against particular research topics	5.45	6.56	6.14	4.04	5.00	-2.44	-1.25	
22	An IRB that requires members to abstain from evaluating protocols whenever a real or apparent conflict of interest arises	5.44	6.48	6.43	4.80	5.85	-1.67	-0.54	
23	An IRB that holds no preconceived biases against particular research techniques	5.43	6.35	6.21	5.24	5.83	-1.00	-0.50	
24	An IRB that is allocated sufficient resources to carry out functions efficiently and thoroughly	5.38	6.67	6.21	4.59	5.77	-2.04	-0.46	
25	An IRB that acknowledges full responsibility for its errors or delays in processing protocols and attempts to correct them as expeditiously as possible	5.33	6.40	6.08	4.97	5.42	-1.42	-0.91	
26	An IRB that offers investigators information to improve the chances of gaining IRB approval	5.31	6.43	5.93	5.23	5.75	-1.15	-0.33	
27	An IRB that recognizes when it lacks sufficient expertise to evaluate a protocol and seeks an outside evaluator	5.28	6.23	6.21	4.29	5.00	-1.94	-1.50	
28	An IRB that is open to innovative approaches to conducting research	5.28	6.32	5.79	4.69	5.23	-1.61	-0.62	
29	An IRB that applies appropriately flexible standards regarding voluntary and informed consent requirements (e.g. required wording is less demanding for minimal risk research using competent adult participants)	5.23	6.19	5.71	4.22	5.00	-1.94	-0.83	
30	An IRB that takes timely action when an investigator has violated the specifications of its rulings	5.22	6.53	6.21	3.71	5.25	-2.86	-1.08	
31	An IRB that ensures that at least one member is knowledgeable about the content domain and discipline of submitted protocols	5.13	6.35	6.21	5.16	5.85	-1.16	-0.54	
32	An IRB composed of members who arrive at meetings well prepared.	5.07	6.40	6.07	5.24	5.50	-1.06	-0.67	
33	An IRB that shows considerable evidence that the advancement of science is part of its mission	4.82	6.32	6.21	4.59	4.62	-1.76	-1.62	

	Item text	Ideal IRB			Actual IRB			Actual – Ideal IRB
		National Validation Sample	Investigator/Project Coordinator	IRB Member/Staff	Investigator/Project Coordinator	IRB Member/Staff	Investigator/Project Coordinator	
34	An IRB that requires that its Chair be an experienced investigator	4.75	6.33	6.00	5.08	5.58	-1.17	-0.67
35	An IRB that is open and pleasant in its interactions with investigators	4.72	6.31	5.86	4.39	5.08	-2.15	-0.92
36	An IRB whose Research Compliance Officer (or staff member in charge of IRB functions) has a background in conducting research	4.68	6.37	5.86	4.81	5.58	-1.46	-0.42
37	An IRB that is empathetic with the difficulties that can present themselves during the design or conduct of research	4.66	6.38	6.29	5.13	5.33	-1.16	-1.00
38	An IRB that is composed primarily of highly competent investigators	4.46	6.35	5.57	5.12	5.08	-1.22	-0.58
39	An IRB that monitors the progress of each approved research project in line with federal policy	4.39	6.40	6.14	5.50	6.17	-0.92	-0.17
40	An IRB that provides a comprehensive training program for its new members	4.34	5.74	5.71	4.78	5.75	-0.94	0.00
41	An IRB that offers consultation during the development of research protocols or grant applications	4.30	6.37	5.50	5.00	5.15	-1.35	-0.31
42	An IRB that has a diverse membership (i.e., includes women, minorities and both junior and senior members of the institution)	4.07	5.83	6.00	5.35	5.46	-0.65	-0.54
43	An IRB that offers investigators opportunities to be educated about federal research policy	4.03	6.18	6.07	5.54	5.92	-0.53	-0.33
44	An IRB that offers editorial suggestions regarding consent documents and protocols (e.g., typos, grammar, clarity)	3.20	5.60	5.14	5.22	5.62	-0.41	0.23
45	An IRB that is composed of more than one public member	2.68	5.70	5.64	4.62	5.50	-1.20	-0.17
	Overall Mean		6.34	6.09	4.85	5.38	-1.50	-0.71
	95% Confidence Interval		6.17, 6.52	5.75, 6.43	4.67, 5.02	5.04, 5.73	-1.56, -1.43	-0.83, -0.59

Note: Items are identified by the rank order of the average ratings of the ideal IRB from the validation sample of 886 behavioral scientists (Keith-Spiegel & Koocher, 2005). Item 1 had the highest average rating in the validation sample; item 45 had the lowest average rating. Average ratings for the actual and ideal IRBs are shown for our sample of investigators/project coordinators and IRB members/staff. Due to missing data, the differences between the item-specific averages are not identical to the averages of the respondent-specific differences between ratings for the actual and ideal IRBs on each item; to aid interpretation of Figure 2, we report the average differences calculated from paired data where both an actual and ideal ratings were available. The estimated overall mean ratings and 95% confidence intervals are estimated from a linear mixed model. Items highlighted blue and red are those with relatively good and poor performance, respectively (see text and Figure 2). Items highlighted yellow are relatively more important (see text and Figure 2). Items highlighted orange are most in need of quality improvement, and those in green are the most important items with relatively good performance.

Table 2

Summary of survey response patterns for IRB members/staff and investigators/project coordinators

Status	IRB Members/ Staff (N=28)	Investigators/ Project Coordinators (N=178)	Total (N=206)
No Response	9	99	108
Initial Response	19	79	98
Opted Out	5	26	31
Opted In	14	53	67
Response Rate	50.0%	29.8%	32.5%

Note: Initial response indicates those participants who followed the hyperlink from the invitation email to the Survey Monkey site. Those who opted out viewed only the initial page explaining the survey and requesting participation, but completed no questions. Those who opted in completed at least 1 survey item.