Investigating Perceived Institutional Review Board Quality and Function Using the IRB Researcher Assessment Tool

Jonathan C. Reeser, Diane M. Austin, Linda M. Jaros, Bickol N. Mukesh, and Catherine A. McCarty
Marshfield Clinic Research Foundation

ABSTRACT: **THE INSTITUTIONAL REVIEW BOARD-RESEARCHER ASSESSMENT TOOL** (IRB-RAT) was designed to assess the relative importance of various factors to the effective functioning of IRBs. We employed the IRB-RAT to gain insight into the ways in which our IRB is perceived to be deficient by those who routinely interact with our Office of Research Integrity and Protections. Respondents ranked qualities thought to be characteristic of an "ideal" IRB and then compared our IRB to that internal standard. We observed that the rate of study participation varied by role. The composite relative ranking of the 45 items that comprise the IRB-RAT differed significantly from the rank order reported by Keith-Spiegel *et al*. Our data furthermore suggest that role influences scoring of the IRB-RAT (*e.g.*, investigators awarded our IRB significantly higher scores in several areas than did research coordinators). Additional research is warranted to determine if the observed role-dependent differences in the perceived quality of our IRB simply reflect the local research culture or if they are indicative of a more fundamental and generalizable difference in outlook between investigators and research coordinators.

KEY WORDS: Institutional Review Board, IRB Researcher Assessment Tool, benchmarking

Received: April 6, 2007; revised: December 17, 2007

OUR CLINIC IS A LARGE NOT-FOR-PROFIT, multispecialty medical practice that serves the western, central and northern regions of our state and houses the state's largest private research foundation. The research foundation supports approximately 400 active clinical and non-clinical research protocols, and the Office of Research Integrity and Protections maintains an institutional review board (IRB) consisting of two committees which together review over 500 items (e.g., new protocols, continuing reviews, reports of unanticipated problems, etc.) per quarter. Despite performance benchmarks that compare favorably with other IRBs (mean time to review for a new protocol receiving full board review = 13 days, mean time to expedited review = 2.7 days), we routinely receive complaints regarding various aspects of our human subjects protection program. In an effort to better understand the ways in which our IRB is perceived as being deficient, we surveyed those individuals who routinely interact with our Office of Research Integrity and Protections regarding their perceptions of our boards' strengths and weaknesses.

We employed the IRB-Researcher Assessment Tool (IRB-RAT) as our data collection instrument. Developed by Keith-Spiegel, Koocher, & Tabachnick (2006), the IRB-RAT consists of a 45-item questionnaire designed to probe beliefs regarding the relative importance of eight different *a priori* factor domains to the function of IRBs: (1) procedural justice, (2) interactional justice, (3) absence of bias, (4) pro-science sensitivity and commitment, (5) competence, (6) IRB outreach, (7) IRB formal functioning, structure and composition, and (8) upholding the rights of human research participants. Keith-Spiegel et al. (2006) used this instrument to determine which elements represented the most important characteristics of an "ideal" IRB for 886 biomedical and social and behavioral scientists from throughout the United States. While social/behavioral scientists outnumbered biomedical scientists by a 2:1 ratio, few compelling differences were detected between the two types of investigators. In addition, no significant differences were uncovered between investigators with and those without a history of service to an IRB. The authors conclude that "the ideal ethics committee appears to be a just body that employs fair procedures, treats investigators with respect, and accords them the opportunity to have a voice when disagreements arise."

To our surprise, and in contrast to the work of Keith-Spiegel et al., our investigation revealed substantial differences in the scores awarded to the various IRB-RAT

Journal of Empirical Research on Human Research Ethics, pp. 25–34. print issn 1556-2646, online issn 1556-2654.
DOI: 10.1525/jer.2008.3.1.25

items by different groups of respondents. Specifically, we detected differences between research coordinators, investigators and IRB members, and also between investigators with and without a history of service to an IRB. We believe these novel observations have considerable relevance to and potential importance for future research into the IRB-RAT based assessment of IRB quality and function.

Methods

Following approval of the research protocol by our IRB (via expedited review), we mailed a total of 349 IRB-RAT instruments to Clinic and Foundation personnel who were identified as having reason to interact with the foundation's IRB, including investigators (both physician researchers and research scientists), research coordinators, and IRB members. Of the distributed surveys, 60 (17.2%) were sent to research coordinators, 271 (77.7%) were sent to investigators, and 18 (5.2%) were sent to "IRB members only" (i.e., IRB members who had no other research role). Forty-five (16.5%) of the 271 investigators had served on our institution's IRB at some point. Two electronic reminders were sent out at intervals to promote participation. Participation was entirely voluntary, and no incentives or remuneration were offered to entice participation. Investigators remained blinded to the identity of the respondents.

The 45-item IRB-RAT instrument was designed to assess opinions regarding the relative importance of various qualities to the functioning of an IRB (Keith-Spiegel et al., 2006). We instructed participants to first provide responses that reflected their perception of the relative importance of each of the 45 items to creating an IRB that would allow them to do their "best work." In essence, we wanted to know what elements they felt were most characteristic of and important to an *ideal IRB*. For each item, the respondent was asked to provide an integer score ranging from 1 (not important) to 7 (absolutely essential). Each item on the IRB-RAT was then administered a second time, with instructions to assign a score indicative of the extent to which the item described the *actual performance* of *our IRB*. Scoring options again ranged from 1 (not at all descriptive) to 7 (highly descriptive), with the scores serving to compare the quality and function of our IRB with the respondent's just-defined ideal internal standard for each of the 45 items. Respondents were permitted to view their initial response regarding the ideal IRB when formulating their numerical performance score for our IRB on a given item.

The responses to the 45 individual items were ranked according to their mean, and the Wilcoxon signed ranks test was employed to compare the rank order of scores with the rankings previously published by Keith-Spiegel et al. (2006). The Pearson correlation coefficient was used to assess the correlation between items grouped into the eight *a priori* domains. Because the response variable was negatively skewed, the data were initially converted to positively skewed data by subtracting all the values for a variable from (1 + |maximum value for the variable|). Subsequently, we performed a log transformation on the positively skewed data prior to conducting any statistical comparisons. The differences in IRB domain scores between various subgroups (e.g., gender, research role, and area of research) were assessed using multiple analyses of variance (MANOVA). Due to the multiple comparisons performed on these data, we adopted a P-value of <0.01 as an indication of statistical significance.

Results

One hundred fifteen IRB-RAT study instruments were returned, representing a 33% overall response rate. Sixty-two (54%) of the respondents were investigators, 26 (23%) were research coordinators, 18 (16%) were "IRB members only," and the remaining 9 (8%) were "others" (*e.g.*, managers, research assistants, and research nurses). Of the 62 investigators, 29 (47%) had ever served on the IRB. The majority of respondents (65%) were involved in clinical trials, while 23% of participants were engaged in non-clinical research. Twenty-four percent of the respondents reported involvement in genetic research. There was an even distribution between male (48%) and female (49%) respondents. However, the response rate did vary tremendously by role: only 23% (62/271) of the investigators who had been invited to participate returned completed instruments, compared to 43% (26/60) of the research coordinators. Of the 45 investigators with a history of IRB service, 29 (64%) responded, while fully 100% (18/18) of the "IRB members only" returned completed instruments.

Table 1 displays the 45 different items that comprise the IRB-RAT, with relative rankings (and mean scores) listed according to the following criteria:

- The item's perceived importance to an ideal IRB, as defined by the participants in our study ("Ideal IRB")
- The degree to which our IRB could be characterized by the descriptor, as judged by study respondents ("Our IRB")
- The rank order (and mean score) of IRB-RAT items as reported by Keith-Spiegel et al. (2006) ("National Data")

TABLE 1. Rank of (and mean score awarded by all study participants to) Each IRB-RAT Item.

IRB-RAT Item (grouped by domain)	Ideal IRB	Our IRB	National Data
Procedural Justice			
• Reviews protocols in a timely fashion	4 (6.52)	30 (5.54)	1 (6.43)
• Conducts a conscientious and complete review of protocols	3 (6.53)	7 (6.12)	6 (5.86)
• Gives a complete rationale for any required changes to or disapprovals of protocols	8 (6.47)	24 (5.61)	9 (5.73)
• 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	19 (6.28)	31 (5.51)	12 (5.59)
• Open to reversing its earlier decisions	22 (6.25)	35 (5.35)	16 (5.52)
• Invites investigators to present their position whenever a question or concern arises	20 (6.27)	19 (5.71)	17 (5.51)
• Recognizes when it lacks sufficient expertise to evaluate a protocol and seeks an outside experts	35 (6)	34 (5.43)	27 (5.28)
Interactional Justice			
• Responds in a timely manner to investigators' inquiries about its processes and decisions	17 (6.32)	22 (5.64)	8 (5.8)
• Works with investigators to find mutually satisfying solutions whenever disagreements exist	26 (6.19)	40 (5.3)	10 (5.71)
• Treats investigators with respect	5 (6.5)	9 (6.04)	20 (5.45)
• Acknowledges full responsibility for its errors or delays in processing protocols and attempts to correct them as expeditiously as possible	21 (6.25)	37 (5.32)	25 (5.33)
• Open and pleasant in its interactions with investigators	25 (6.21)	21 (5.67)	35 (4.72)
Absence of Bias			
• Members who do not allow their personal biases to affect their evaluation of protocols	10 (6.44)	33 (5.44)	2 (6.17)
• Hold no preconceived biases against particular research topics	31 (6.14)	20 (5.67)	21 (5.45)
• Requires members to abstain themselves from evaluating protocols whenever a real or apparent conflict-of-interest arises	6 (6.49)	5 (6.16)	22 (5.44)
• Members hold no preconceived biases against particular research techniques	30 (6.17)	16 (5.78)	23 (5.43)
• Open to innovative approaches to conducting research	28 (6.18)	41 (5.29)	28 (5.28)
Pro-Science Sensitivity			
• Does a good job of upholding participants' rights while facilitating the conduct of research	7 (6.47)	10 (6.04)	3 (6.1)
• Does not use its power to suppress research that is otherwise methodologically sound and in compliance with federal policy whenever it perceives criticism outside the scientific community	16 (6.33)	18 (5.77)	4 (6.08)
• Views itself as an investigator's ally rather than a hurdle to clear	27 (6.19)	44 (5.12)	13 (5.57)
• Shows considerable evidence that the advancement of science is part of its mission	33 (6.04)	28 (5.58)	33 (4.82)
• Shows empathy with the difficulties that can present themselves during the conduct of research	38 (5.85)	38 (5.32)	37 (4.66)
IRB Competence			
• Members are very knowledgeable about IRB procedures and federal policy	9 (6.46)	11 (6.03)	5 (6.01)
• Conducts a conscientious informed analysis of potential benefits weighed against potential risks before making decisions	11 (6.44)	13 (6.01)	14 (5.54)
• Competently distinguishes exempt from nonexempt research	18 (6.29)	14 (5.98)	19 (5.48)
• Ensures that at least one member who is knowledgeable about the content of submitted protocols	34 (6.01)	26 (5.59)	31 (5.13)
• Members who arrive at meetings well-prepared	23 (6.22)	25 (5.6)	32 (5.07)

(*continued*)

TABLE 1. (*Continued*)

IRB-RAT Item (grouped by domain)	Ideal IRB	Our IRB	National Data
• Research Compliance Officer who has a research background	37 (5.87)	43 (5.15)	36 (4.68)
• Composed primarily of members regarded as highly competent investigators	45 (4.89)	45 (4.77)	38 (4.46)
• Provides a comprehensive training program for its new members	29 (6.18)	29 (5.55)	40 (4.34)
IRB Outreach			
• Offers investigators information to improve the chances of gaining IRB approval	36 (5.93)	42 (5.19)	26 (5.31)
• Offers consultation during the development of research protocols or grant applications	41 (5.66)	36 (5.34)	41 (4.3)
• Offers investigators opportunities to be educated about federal research policy	42 (5.63)	23 (5.64)	43 (4.03)
• Offers editorial suggestions regarding consent documents and protocols	43 (5.23)	39 (5.3)	44 (3.2)
IRB Formal Functioning, Structure, and Composition			
• Members fully understand and act within the scope of their function	15 (6.37)	27 (5.58)	11 (5.67)
• Maintains accurate records	1 (6.75)	2 (6.36)	18 (5.5)
• Allocated sufficient resources to carry out its functions	24 (6.21)	17 (5.77)	24 (5.38)
• Requires that its Chair be an experienced investigator	39 (5.72)	15 (5.88)	34 (4.75)
• Monitors the progress of each approved research project in line with federal policy	13 (6.4)	1 (6.45)	39 (4.39)
• Has diverse membership	40 (5.7)	8 (6.11)	42 (4.07)
• Composed of more than one public member	44 (5.22)	12 (6.02)	45 (2.68)
Upholding the Rights of Human Participants			
• Views protection of human participants as its primary function	14 (6.39)	3 (6.32)	7 (5.8)
• Takes timely and appropriate action whenever scientific misconduct is alleged	2 (6.57)	4 (6.23)	15 (5.52)
• Applies appropriately flexible standards regarding voluntary and informed consent requirements	32 (6.13)	32 (5.48)	29 (5.23)
• Takes timely action when an investigator has violated its decisions	12 (6.41)	6 (6.14)	30 (5.22)

As mentioned, Keith-Spiegel et al. clustered the 45 items into eight domains. Factor analysis revealed these eight domains all correlated strongly with one another, suggesting a high internal validity to the IRB-RAT instrument (data not shown). Meaningful factor analysis of the individual items was not possible due to the small size of our sample.

We found that the rank order of the 45 mean Ideal IRB item scores provided by the study respondents differed significantly (Wilcoxon signed ranks test, $P < 0.001$) from the ranking of the IRB-RAT items reported by Keith-Spiegel et al. (2006) ("National Data" in Table 1). We also observed that the mean score for any given Ideal IRB item in our study was consistently greater than the score reported for that item in Keith-Spiegel et al. (2006).

The mean scores awarded to our IRB for each of the eight factor domains are presented in Table 2. MANOVA analysis failed to reveal any significant scoring difference in mean factor domain scores between male and female respondents, or between those respondents participating in clinical versus non-clinical or genetic versus non-genetic research. However, we did find that the performance scores awarded to our IRB by investigators and "IRB members only" were consistently greater than the scores provided to our IRB by research coordinators. These differences reached statistical significance for several of the *a priori* domains, including procedural justice, interactional justice, pro-science sensitivity, and IRB outreach. The most significant differences were noted between investigators without IRB experience and research coordinators in the domains of interactional justice and IRB outreach ($P < 0.001$). A comparison of the individual IRB-RAT items included in these domains revealed that research coordinators awarded significantly lower scores than investigators on 4 of the 5 RAT items in the interactional justice domain, and on 1 of the 4 items in the IRB outreach domain (Table 3). Further inspection of Table 2 reveals that for 7 of the 8 factor domains, the highest scores were awarded to our IRB by "IRB members only."

TABLE 2. Comparison, by Role, of Mean Factor Domain Scores Awarded to Our IRB and the Ideal IRB.

Domain	Research Coordinators (*N* = 26)	Investigators with IRB Experience (*N* = 29)	Investigators without IRB Experience (*N* = 33)	IRB Members Only (*N* = 18)
Procedural justice	5.18 ± 0.90 (6.45 ± 0.65)	5.51 ± 1.34*[b] (6.48 ± 0.57)	5.89 ± 0.91*[a] (5.99 ± 0.97)	5.92 ± 0.74 (6.28 ± 0.75)
Interactional justice	4.65 ± 1.26 (6.50 ± 0.82)	5.65 ± 1.42*[a] (6.34 ± 0.64)	6.02 ± 0.88**[a] (6.10 ± 0.80)	5.97 ± 0.71*[a] (6.16 ± 0.80)
Absence of bias	4.99 ± 1.38 (6.25 ± 0.97)	5.52 ± 1.39 (6.28 ± 0.77)	5.49 ± 1.37 (6.21 ± 0.71)	6.09 ± 0.90 (6.28 ± 1.05)
Pro-science sensitivity	5.09 ± 1.32 (6.52 ± 0.65)	5.40 ± 1.42 (6.22 ± 0.84)	5.70 ± 0.87*[a] (5.98 ± 0.87)	5.99 ± 0.91**[a] (5.85 ± 1.06)
IRB competence	5.07 ± 1.01 (6.23 ± 0.71)	5.30 ± 1.09 (5.92 ± 0.80)	5.85 ± 0.82 (5.91 ± 0.63)	5.85 ± 0.73 (6.03 ± 0.75)
IRB outreach	4.78 ± 1.20 (6.07 ± 0.95)	5.44 ± 1.05*[a] (5.51 ± 1.11)	5.54 ± 1.32**[a] (5.37 ± 0.98)	5.59 ± 0.86 (5.48 ± 0.85)
Formal functioning	5.89 ± 1.12 (6.18 ± 0.82)	5.90 ± 0.96 (6.06 ± 0.68)	6.08 ± 0.72 (5.68 ± 0.71)	6.15 ± 0.80 (6.27 ± 0.76)
Upholding human rights	5.48 ± 1.27 (6.31 ± 0.68)	6.17 ± 0.82 (6.43 ± 0.59)	5.92 ± 1.03 (6.25 ± 0.62)	6.19 ± 0.78 (6.43 ± 0.79)

Data represents mean ± standard deviation.
*$P < 0.01$, **$P < 0.001$ (statistical comparisons conducted on Our IRB data only).
[a]Compared to research coordinators.
[b]Compared to investigators without IRB experience.

By comparison, Ideal IRB scoring by domain was somewhat less predictable. The shaded cells in Table 2 contain the mean scores for each "Ideal IRB" factor domain, and inspection reveals that the highest scores were scattered across three different roles. However, in none of the 8 domains did "Investigators without IRB experience" generate the highest mean score. In fact, investigators without IRB experience awarded the lowest mean Ideal IRB score for 7 of the 8 domains.

Table 4 compares the relative ranking by role of the items identified by Keith-Spiegel et al. (2006) as the 10 most important qualities of an ideal IRB. This comparison reveals that of the 10 top-ranked items in the "National Data," responding research coordinators ranked 7 in their top 10, while participating investigators ranked only 5 of the items in their top 10. Furthermore, the composite mean rank of these 10 items among research coordinators was 10.3, suggesting

TABLE 3. Grading the Actual Performance of Our IRB: Significant IRB-RAT Scoring Differences Between Research Coordinators and Investigators with IRB Experience.

IRB-RAT Item	Research Coordinators Mean ± SD	Investigators with IRB Experience Mean ± SD	*P* Value
Interactional Justice			
• Works with investigators to find mutually satisfying solutions whenever disagreement exists	4.4 ± 1.5	5.4 ± 1.4	0.01
• Responds in a timely manner to investigators' inquiries about its processes and decisions	4.6 ± 1.5	5.8 ± 1.4	0.007
• Acknowledges full responsibility for its errors or delays in processing protocols and attempts to correct them as expeditiously as possible	4.1 ± 1.6	5.5 ± 1.6	0.002
• Open and pleasant in its interactions with investigators	4.5 ± 1.7	5.8 ± 1.5	0.005
IRB Outreach			
• Offer editorial suggestions regarding consent documents and protocols	4.7 ± 1.4	5.6 ± 1.3	0.009

TABLE 4. Rank (and mean score) Awarded by Participating Investigators (PI), Investigators with IRB Experience, and Research Coordinators (RC) to the 10 IRB-RAT Items Ranked Highest in Keith-Spiegel *et al.* (National Data).

		Ideal IRB			Our IRB		
IRB-RAT Items	National Data	All PI	PI with IRB Experience	RC	All PI	PI with IRB Experience	RC
• Reviews protocols in a timely fashion	1 (6.43)	4 (6.48)	5 (6.69)	6 (6.64)	29 (5.53)	30 (5.45)	30 (5.00)
• Members who do not allow their personal biases to affect their evaluation of protocols	2 (6.17)	6 (6.43)	11 (6.57)	8 (6.60)	34 (5.40)	35 (5.32)	19 (5.54)
• Does a good job of upholding participants' rights while, at the same time, facilitating the conduct of research	3 (6.10)	10 (6.39)	6 (6.66)	1 (6.81)	12 (5.96)	11 (5.97)	6 (6.00)
• Does not use its power to suppress research that is otherwise methodologically sound and in compliance with federal policy whenever it perceives criticism outside the scientific community	4 (6.08)	7 (6.42)	4 (6.70)	33 (6.21)	19 (5.73)	19 (5.78)	23 (5.31)
• Members are very knowledgeable about IRB procedures and federal policy	5 (6.01)	5 (6.47)	14 (6.48)	9 (6.60)	6 (6.13)	9 (6.04)	11 (5.84)
• Conducts a conscientious and complete review of protocols	6 (5.86)	12 (6.35)	8 (6.62)	5 (6.68)	7 (6.11)	7 (6.07)	5 (6.05)
• Views protection of human participants as its primary function	7 (5.80)	16 (6.34)	21 (6.31)	13 (6.50)	4 (6.35)	5 (6.34)	8 (5.95)
• Responds in a timely manner to investigators' inquiries about its processes and decisions	8 (5.80)	18 (6.28)	18 (6.39)	15 (6.46)	13 (5.94)	16 (5.82)	40 (4.60)
• Gives a complete rationale for any required changes to or disapprovals of protocols	9 (5.73)	14 (6.35)	9 (6.59)	3 (6.76)	20 (5.69)	22 (5.64)	31 (4.86)
• Works with investigators to find mutually satisfying solutions whenever disagreements exist	10 (5.71)	30 (6.03)	26 (6.14)	10 (6.56)	32 (5.44)	33 (5.43)	43 (4.39)
Mean rank	5.5	12.2	12.2	10.3	17.6	18.7	21.6

that the research coordinators' opinion of the relative importance these RAT items was closer to the national standard (mean rank = 5.5) than was that of the investigators (mean rank = 12.2).

Discussion

Despite the Federal mandate to oversee research involving human subjects, IRBs are often lightning rods for poorly veiled criticism and complaints, seemingly from all sectors of the research enterprise (Burke, 2005; Hohmann, 2005; IRB Advisor, 2006a; 2006b; Levine, 2006; Burris & Moss, 2006; Edgar & Rothman, 1995; Goldman & Katz, 1982; Sansone & McDonald, 2004; Lux, Edwards, & Osborne, 2000; Shah, Whittle, Wilfond, Gensler, & Wendler, 2004). Our IRB is not immune to such criticism, and so, in an effort to better understand the ways in which we are perceived as being deficient, we decided to survey the individuals who commonly interact with our IRB. The study yielded unexpected results, suggesting that an individual's role in the research enterprise significantly influences his or her perception of the quality of our IRB.

While we acknowledge that our data may be limited by the preliminary nature of our inquiry, we nevertheless believe that our observations have potentially important practical implications regarding the assessment of IRB quality and function. Although our sample size is small and (at 34%) our overall response rate is lower than what has been reported as typical for mailed surveys published in medical journals (Asch, Jedrziewski, & Christakis, 1997), we feel that our response rate does not preclude meaningful analysis of the data, particularly when compared to the 38.8% response rate obtained by

Keith-Spiegel et al. (2006). Furthermore, the comparatively lower response rate we observed from investigators collectively is consistent with the pattern of behavior reported by Asch and colleagues. Thus, although lower response rates do suggest the possibility of response bias, our response rates are in fact not too dissimilar from previously published studies in the IRB literature. That stated, we submit that the tendency for investigators (as a group) to generally fail to reply to surveys should prompt researchers to analyze survey-based data collected during IRB self-assessment research with appropriate caution so as to avoid over-interpretation.

The data we collected during this investigation provided us with a benchmark of the perceived performance of our IRB. Upon examining the collected data (Table 1), we were generally encouraged by the fact that of the 10 items felt to be *most* desirable in an *ideal* IRB (as defined by our respondents), 6 were ranked among the top 10 strengths of our IRB (i.e., those areas in which our IRB scored highest in terms of actual performance). Those items felt to warrant inclusion in the Ideal IRB top 10 but which did not appear in the top 10 list of Our IRB's strengths will serve as focal-points for future process-improvement initiatives.

Despite the fact that 6 of the IRB-RAT items shared a ranking among the 10 most important elements of an ideal IRB as defined both by our study participants and by the national sampling of researchers reported in Keith-Spiegel et al. (Table 1), further analysis revealed that the rank order of mean scores awarded to the Ideal IRB by survey respondents differed significantly from the rank order of the National Data reported by Keith-Spiegel et al. (2006) (Wilcoxon signed ranks test, $P < 0.001$). Possible explanations for this difference include:

- Keith-Spiegel et al. sampled researchers almost exclusively. In contrast, our study population was much less homogeneous and included research coordinators and IRB members, as well as investigators;
- Social and behavioral science researchers comprised 64% of the Keith-Spiegel sample, while biomedical researchers comprised 30%, whereas the investigators at our institution are almost exclusively biomedical in orientation;
- Keith-Spiegel et al. primarily sampled academic centers from across the country, while our study population came from a single private, not-for-profit rural medical center.

Keith-Spiegel et al. (2006) detected few significant differences between the two subsets of investigators in their study. Biomedical researchers gave higher ratings for formalities, but otherwise no significant differences were noted. If we therefore assume that the "National Data" in Table 1 is indeed reflective of the biomedical researcher cohort within the larger study population, then the analysis of comparative RAT item ranking based upon role offered in Table 4 implies that some differences do appear to exist between the investigators who participated in this study and those surveyed by Keith-Spiegel et al. The relative importance of the environment (i.e., academic medical center versus a large rural multispecialty clinic) in explaining these findings, and indeed a better overall accounting of which of the proposed explanations serves to account for the observed differences, must await further investigation, possibly through hypotheses-driven collaborative research projects.

Our data also clearly suggest that role-dependent differences exist not only in the relative importance afforded to various factors in the assembly of an ideal IRB, but in the perception of how our IRB actually functions as well. We interpreted the scores awarded to Our IRB to reflect the extent to which respondents thought that our IRB had achieved the ideal IRB standard. The most compelling observation in this regard is that research coordinators awarded Our IRB lower scores, on average, than did investigators or "IRB members only" for all eight factor domains of the IRB-RAT (Table 2). More specifically, research coordinators awarded our IRB a mean score that was significantly lower than that awarded by investigators with IRB experience for 2 of the 8 factor domains, while the difference between the coordinators' scores and those of the investigators without IRB experience reached significance in 4 domains. Significant differences were found to exist between research coordinators and "IRB members only" in 2 of the 8 domains.

What accounts for the significant difference in the perceived performance of our IRB by research coordinators and others? Although the majority of the research coordinators were females, MANOVA did not reveal a consistent correlation between gender and scoring. One plausible explanation is that the lower scores offered by the research coordinators simply reflect the practical work environment of, and the demanding job-related expectations placed on, the research coordinators at our institution. Among other tasks, our research coordinators typically serve to buffer the investigators from interaction with the IRB. They therefore must satisfy the demands of both the investigators and the IRB (to say nothing of the subject and the study sponsor), and there may well be a component of frustration reflected in the lower scores offered by the research coordinators.

In this regard, we think it potentially worrisome that the lowest mean domain score awarded to Our IRB was

in the realm of "interactional justice." We find this concerning because it is conceivable that this score might reflect a perceived lack of respect for research coordinators at our institution. Of course, there are numerous factors that influence the scoring of these IRB-RAT items, including (but not limited to) the way in which research coordinators are employed at our institution. It would be interesting to see whether similar findings would be observed at different institutions with different organizational structures and cultures. Further research into the factors which might account for the research coordinators' comparatively harsher assessment of our IRB appears to be warranted.

Another way to approach the same issue is to inquire why investigators should have awarded generally higher scores to Our IRB than did research coordinators. An analysis of the investigators' scoring of the IRB-RAT prompted us to compare those with a history of IRB service to those without a comparable history. In contrast to the findings of Keith-Spiegel et al., we found that IRB service may influence the way in which investigators score the IRB-RAT. Specifically, investigators with IRB experience tended to award somewhat lower scores to Our IRB than did their colleagues without a history of IRB service (Table 2). Furthermore, investigators without IRB experience provided the lowest mean scores in 7 of the 8 Ideal IRB domains. These findings suggest that familiarity with the actual process of IRB review and the regulations governing human subjects protections may have multiple consequences. On one level, such experience may influence investigators' opinions regarding the overall importance of the IRB to the research enterprise (i.e., investigators with IRB experience may value the potential contributions of an ideal research ethics board to the pursuit of science more highly than those without IRB experience, thereby accounting for the generally higher Ideal IRB scores). On another level, due to their familiarity with IRB protocol, those investigators with IRB experience may tend to be more critical of the actual functioning of our local Board, thereby accounting for the generally lower Our IRB scores among investigators with IRB experience.

Conversely, "IRB members only" awarded the highest mean score to Our IRB for 7 of the 8 domains (Table 2). Based upon the argument outlined above (that IRB experience prompts individuals to hold the IRB to a higher standard), it might have been expected for this group to assess our IRB more critically. That they did not suggests a degree of member-related bias in the evaluation of our IRB by this subgroup. The underlying motivation for this apparent bias may be a sense of loyalty to the Board, or it may reflect a simple tendency to avoid negative self-assessment. It should be remembered that respondents in this category have no other role within the Foundation besides service to the IRB, and therefore there may be other unique factors that influence their responses. Whatever the etiology, the observation serves to reinforce the point that an individual's role in the research process seems to influence his or her response to the IRB-RAT, a conclusion that has practical implications for the conduct of future research into IRB quality using the IRB-RAT. Specifically, our data imply that the outcome of such studies will depend in part on research role-related demographic factors within the population surveyed.

Burke (2005) wrote "the IRB as a social entity takes on a distinct culture that is influenced by the culture of the institution it serves." We believe that our results support this statement, insofar as the data indicate our study participants responded to the IRB-RAT instrument in a manner that is substantively different than the national sampling of investigators described in Keith-Spiegel et al. (2006). We propose that cultural and environmental factors (as yet largely undefined) influence the interactions between an IRB and those it serves. We furthermore suggest that role within the research enterprise represents one of those factors.

In our opinion, it seems likely that institutions and their employees develop unique expectations of their ethics boards, and that one's set of expectations of the local IRB may be colored not only by the institutional culture and organizational structure but by the role one has in the conduct of human subjects research. While the results of this preliminary investigation are intriguing, they need to be corroborated by similar undertakings at research institutions that have different organizational structures and cultures. In this way, it should be possible to determine the specific factors that are associated with various beliefs and/or systematic biases regarding IRB quality and function. Until such time as sufficient work has been performed to permit a more detailed and nuanced understanding of this topic, we would recommend that research on the functioning of IRBs (particularly research utilizing the IRB-RAT) account for institutional and role dependent variables that may influence the outcome of the investigation (including such factors as role dependent participation rates).

Best Practices

To our knowledge, this is the first published study to describe the use of the IRB-RAT as an IRB self assessment tool. From a practical standpoint, we found the instrument useful and informative, and we therefore

believe the instrument may well provide IRBs with a useful means of establishing baseline benchmark data which can then serve as the basis for future comparison (for example, pre- and post-accreditation). However, we believe this preliminary investigation has revealed that the items that comprise the IRB-RAT may be sensitive to the respondent's role in the research enterprise, which may compromise the instrument's utility in comparing and contrasting IRBs from different institutions. We would therefore recommend that future users of the IRB-RAT take care to account for the role of their respondents, as it appears that significant scoring differences may be overlooked if one simply considers the data in aggregate.

Research Agenda

In this study, we observed a distinct difference in responses to the IRB-RAT items provided by research coordinators in comparison to those provided by investigators. The basis for this striking finding remains open to speculation. A multi-center collaborative research project would permit an improved understanding of the factors influencing this role-dependent perceptual difference. Such a project would by necessity include a variety of different institutional research paradigms in order to examine the effect of institutional structure and culture on the responses provided. For example, it would seem appropriate to compare the perceptions regarding IRB function (as reflected by the IRB-RAT) from a number of different institutions, including those which are highly dependent on research coordinators in their clinical research activities versus those which are not, those which depend heavily on central IRBs (commercial or non-commercial) versus those who do not, and those with accredited IRBs versus those with unaccredited boards. Of course other variables of significance may be considered and analyzed for their effect on the IRB assessment process, including (but not limited to) IRB composition, size, benchmarks of performance, etc. The causes for role dependent variation in response rates may also become more evident from such future research. *N.B.* We would be willing to share our data with other research groups interested in conducting research into the nature of this observation.

Educational Implications

The success of a given clinical investigation depends on the integrated contributions of many individuals serving in different capacities. Our research has revealed that selected members of the clinical research team at our institution (i.e. IRB members, investigators and research coordinators) value the various functions and competencies of our IRB differently. The reasons for these differences and their practical impact on the research process have yet to be determined, but it seems plausible to speculate that role-dependent attitudes and biases may influence the success of research. Those involved in the conduct and oversight of human subjects' research should remain sensitive to the possibility that these role-dependent differences exist and that they may influence the conduct of research in significant ways. Exploring these differences through workshops designed to understand and contrast roles and associated attitudes would enable both research administrators and those invested in the research to work together to develop educational materials and support programs designed to optimize team performance.

Acknowledgments

The authors are grateful to Linda Weis and Alice Stargardt for their professional assistance in preparing the manuscript for publication. We thank Nancy Hathaway, J.D., for critically reviewing the manuscript prior to its submission.

Author Note

Address correspondence to: Jonathan C. Reeser, Office of Research Integrity and Protections, Marshfield Clinic Research Foundation, 1000 North Oak Avenue, Marshfield, WI 54449; Phone 715-389-3162; E-MAIL: reeser.jonathan@marshfieldclinic.org.

Authors' Biographical Sketches

Jonathan C. Reeser, M.D., Ph.D., is a Physical Medicine and Rehabilitation specialist at Marshfield Clinic in Marshfield, Wisconsin. He chairs the Marshfield Clinic Research Foundation's (MCRF) Institutional Review Board.

Diane M. Austin, CRA, CIP, is Director of the Office of Research Compliance at the Marshfield Clinic Research Foundation. Ms. Austin has been involved in the administration of MCRF's human subject protection program for 13 years.

Linda M. Jaros, CIP, is the Institutional Review Board Administrator for the Marshfield Clinic

Research Foundation. She also serves as a member of the MCRF IRB.

Bickol N. Mukesh, Ph.D., is a biostatistician in the Biomedical Informatics Research Center at the Marshfield Clinic Research Foundation. Dr. Mukesh also serves as a member of the MCRF IRB.

Cathy A. McCarty, Ph.D., is Interim Director of the Center for Human Genetics at the Marshfield Clinic Research Foundation and Principal Investigator of the Personalized Medicine Research Project (the largest population-based DNA biobank in the United States). Dr. McCarty also serves as a member of the MCRF IRB.

References

Asch, D. A., Jedrziewski, M. K., & Christakis, N. A. (1997). Response rates to mail surveys published in medical journals. *Journal of Clinical Epidemiology, 50*(10), 1129–1136.

Burke, G. S. (2005). Looking into the institutional review board: Observations from both sides of the table. *The Journal of Nutrition, 135*(4), 921–924.

Burris, S. & Moss, K. (2006). U.S. health researchers review their ethics review boards: A qualitative study. *Journal of Empirical Research on Human Research Ethics, 1*(2), 39–58.

Edgar, H. & Rothman, D. J. (1995). The institutional review board and beyond: future challenges to the ethics of human experimentation. *The Milbank Quarterly, 73*(4), 489–506.

Goldman, J. & Katz, M. D. (1982). Inconsistency and institutional review boards. *The Journal of the American Medical Association, 248*(2), 197–202.

Hohmann, E. with response by Woodson, J. (2005). Inefficient, arbitrary, inconsistent: A frank look at how some investigators view IRBs and a few suggestions for improvement. *Protecting Human Subjects, 12*, 12–14. Retrieved March 14, 2007 from http://www.science.doe.gov/ober/ humsubj/Summer05.pdf.

Fairness and common sense can ease tensions. (2006a). *IRB Advisor, 6*(8), 89–91.

How IRBs can improve investigators' opinion. (2006b). *IRB Advisor, 6*(8), 91–93.

Keith-Spiegel, P., Koocher, G. & Tabachnick, B. (2006). What scientists want from their research ethics committee. *Journal of Empirical Research on Human Research Ethics, 1*(1), 67–82.

Levine, R. J. (2006). Empirical research to evaluate ethics committees' burdensome and perhaps unproductive policies and practices: A proposal. *Journal of Empirical Research on Human Research Ethics, 1*(3), 1–4.

Lux, A. L., Edwards, S. W., & Osborne, J. P. (2000). Responses of local research ethics committees to a study with approval from a multicentre research ethics committee. *British Medical Journal, 320*(7243), 1182–1183.

Sansone, R. A. & McDonald, S. (2004). Quality improvement within institutional review boards. *Quality Assurance Journal, 8*(4), 258–261.

Shah, S., Whittle, A., Wilfond, B., Gensler, G., & Wendler, D. (2004). How do institutional review boards apply the federal risk and benefit standards for pediatric research? *The Journal of the American Medical Association, 291*(4), 476–482.