



Published in final edited form as:

J Empir Res Hum Res Ethics. 2011 March ; 6(1): 3–19. doi:10.1525/jer.2011.6.1.3.

A Systematic Review of the Empirical Literature Evaluating IRBs: What We Know and What We Still Need to Learn

Lura Abbott and Christine Grady

National Institutes of Health

Abstract

institutional review boards (IRBs) are integral to the U.S. system of protection of human research participants. Evaluation of IRBs, although difficult, is essential. To date, no systematic review of IRB studies has been published. We conducted a systematic review of empirical studies of U.S. IRBs to determine what is known about the function of IRBs and to identify gaps in knowledge. A structured search in PubMed identified forty-three empirical studies evaluating U.S. IRBs. Studies were included if they reported an empirical investigation of the structure, process, outcomes, effectiveness, or variation of U.S. IRBs. The authors reviewed each study to extract information about study objectives, sample and methods, study results, and conclusions. Empirical evidence collected in forty-three published studies shows that for review of a wide range of types of research, U.S. IRBs differ in their application of the federal regulations, in the time they take to review studies, and in the decisions made. Existing studies show evidence of variation in multicenter review, inconsistent or ambiguous interpretation of the federal regulations, and inefficiencies in review. Despite recognition of a need to evaluate effectiveness of IRB review, no identified published study included an evaluation of IRB effectiveness. Multiple studies evaluating the structure, process, and outcome of IRB review in the United States have documented inconsistencies and inefficiencies. Efforts should be made to address these concerns. Additional research is needed to understand how IRBs accomplish their objectives, what issues they find important, what quality IRB review is, and how effective IRBs are at protecting human research participants.

Keywords

institutional review boards; IRBs; evaluation; empirical data

The IRB process is too important not to undergo periodic evaluation. Evaluations can help an IRB to determine whether it is effectively protecting human subjects, whether it is operating efficiently, and whether it has adequate authority (Office of the Inspector General, 1998a, p. 20).

Institutional review boards (IRBs), a fixture in the U.S. research firmament (McCarthy, 1996), review most research involving human participants before it is initiated and at least annually until it is complete. IRBs review research proposals to assure they adhere to federal regulations (Department of Health and Human Services; Federal Drug Administration), include adequate protections of study participants' rights and welfare, and are ethically sound. But, little is known about how well IRBs accomplish these goals.

© 2011 by joan sieber

Address correspondence to: Christine Grady, Department of Bioethics, NIH Clinical Center, Building 10/1C118, Bethesda, MD 20892. Phone: 301-496-2429; cgrady@nih.gov..

In recent years, investigators and others have expressed dissatisfaction with the IRB system, criticizing it as dys-functional (Fost & Levine, 2007), overburdened (Burman et al., 2001; Office of the Inspector General, 1998b), and overreaching (Gunsalus et al., 2006). Clinical investigators complain that the IRB review process is inefficient and delays their research for what seem like minor modifications (Whitney et al., 2008). Research sponsors object that IRB review is time consuming, leading to delays that can significantly increase the costs of research. The public primarily hears about problems and hence fears that research might be unsafe and existing protections ineffective (Lemonick, Goldstein, & Park, 2002). The current IRB system has also been described as outdated and inappropriate for the scope and type of research being conducted in the twenty-first century (Maschke, 2008).

Proposals to reform and improve the IRB system abound, including proposals to centralize, regionalize, or consolidate IRBs, strengthen and demystify federal oversight, infuse more support and resources into the system, augment IRB member training, require credentialing of IRB professionals, mandate independent accreditation, educate the public, and continue to investigate “alternative” models of review (IOM, 2002; NIH, 2006; Steinbrook, 2002). A common and persistent call for data on IRB quality and a method for monitoring IRB effectiveness is found in various proposals for change.

IRB evaluation has occurred through sporadic auditing or for-cause investigation by research institutions and regulatory agencies such as the U.S. FDA and the DHHS Office of Human Research Protections (OHRP). Events such as the deaths of Jesse Gelsinger at the University of Pennsylvania and a young female employee at Johns Hopkins University, and several high-profile OHRP-initiated suspensions in the last decade, stunned the scientific community, escalated concerns about study oversight, and focused attention on the overburden, inefficiencies, inherent conflicts of interest, inconsistencies, and lack of data about the function and effectiveness of IRBs (Emanuel et al., 2004).

In 2001, the Association for the Accreditation of Human Research Protection Programs, Inc. (AAHRPP) established a process for accreditation of Human Research Protection Programs described as voluntary, peer driven, and educational. Among other things, AAHRPP determines whether an institution has appropriate arrangements for prospective independent scientific and ethical review and whether the IRB is satisfying federal regulations for review of human subjects research. Accreditation encourages the development of standardized policies and procedures that might lead to improvements in IRB function; however, the focus is primarily on structure and process.

We systematically reviewed available empirical studies evaluating IRBs in order to determine what is known about how well IRBs function and what has been overlooked.

Methods

The goal of this systematic review is to comprehensively describe empirical data on evaluation of IRBs in the United States (Strech, Synofik, & Marckmann, 2008). A PubMed search was used to identify empirical studies that evaluated IRBs. The final search strategy was limited to English language only and used a combination of Medical Subject Heading (MeSH) terms including ethics committees, research; institutional review boards; and evaluation, program evaluation, outcome assessment, observation, questionnaires, and a few additional subheadings (see Figure 1). A total of 1,576 English language publications were identified. Personal experiences or anecdotes, literature reviews, and discussions of whether certain types of activity, such as quality improvement, are appropriate for IRB review were excluded. A total of 111 publications reported empirical findings from studies about IRBs or research ethics committees, using either qualitative or quantitative methods. Of these, fifty-

nine were excluded because they studied review committees outside of the United States. Forty-three of the remaining fifty-two were included in our analysis. The other studies were analyses of policies or studies of the views of investigators.

Both authors read and categorized each published study into one of the following groups for analysis: (1) studies of IRB structure involving a general description of IRB volume, characteristics of IRB members, or costs associated with IRB review (Table 1); (2) process studies examining the extent to which federal regulations are implemented by the IRB (Table 2); (3) studies documenting variation in the process or outcome of IRB review in multicenter research (Table 3); and (4) outcome studies including those that examine IRB decisions and the results of IRB deliberations (Table 4). Each table contains information extracted from the published articles about study objectives, methods, results, and conclusions.

Results: What Do We Know About IRBs?

Sample Characteristics

Our search identified forty-three empirical studies of various aspects of U.S. IRB structure, process, outcome, or variation in process or outcomes among different IRBs reviewing multicenter studies. A broad array of study methodologies was used, including surveys; analysis of written documents such as IRB minutes, stipulations, and resource utilization data; interviews with IRB members, administrators, or investigators; and site visits. Sample size ranged from analysis of a single IRB (Grodin, Zaharoff, & Kaminow, 1986) to 491 IRBs (Bell, 1998), and survey responses ranged from ten in one study (McClure et al., 2007) to almost three thousand (2,989) faculty members who serve on IRBs (Campbell et al., 2003)

STUDIES OF IRB STRUCTURE—IRB structure studies ($n = 16$) included evaluation of IRB membership characteristics, IRB costs, the volume of studies reviewed, and the experience of nonaffiliated, nonscientific, and nurse IRB members (see Table 1). In several studies, it was reported that IRB members are predominantly white, well-educated, male investigators (Bell, 1998; Campbell et al., 2003; DeVries & Forsberg, 2002; Hayes, Hayes, & Dykstra, 1995). Financial relationships with industry and financial conflicts of interest were found to be common among medical school faculty members who serve on IRBs as well as among members of academic medical center IRBs (Campbell et al., 2003; Campbell et al., 2006). Investigators in one study cited lack of necessary administrative support to monitor the volume of ongoing research (DeVries & Forsberg, 2002). In two studies investigators described a need for role clarification and additional training for the nonscientist and nonaffiliated IRB members (Anderson, 2006; Sengupta & Lo, 2003). In another study, authors recommended adding more nurse members after comparing the attitudes and perceived influence of physician, nurse, unaffiliated, and other members (Rothstein & Phuong, 2007). Studies of the costs of IRB operations showed that although IRB costs are highly variable, many IRBs are economically inefficient with large economies of scale favoring high-volume IRBs (Byrne et al., 2006; Sobolski, Flores, & Emanuel, 2007; Wagner et al., 2003; Wagner, Cruz, & Chadwick, 2004). Importantly, data demonstrated variation in the overall operating and per protocol costs of IRBs of similar size, suggesting that other factors contribute to differences in efficiency. The authors of one study estimated that centralizing review of multisite protocols could result in a 10–35% cost savings (Sobolski et al., 2007), although it was noted that few IRBs track whether the studies they review are multisite.

The need for standards for measuring IRB quality and efficiency was recognized early in the evolution of IRBs, but never realized. For example, in the 1970s the U.S. National

Commission recruited Bradford Gray and colleagues to evaluate IRBs and assess their effectiveness. After an intensive review of a single IRB (Gray, 1975) and later a larger survey of IRB members and investigators from 421 institutions (Cooke, Tannenbaum, & Gray, 1977), they concluded that lack of performance evaluation with objective measures hindered assessment of IRB impact and that serious efforts to understand and improve IRB performance were needed.

Concern in the 1990s about the federal oversight process and its effectiveness in protecting human subjects from harm led to two federal reviews, one by the U.S. Government Accounting Office (GAO, 1996) and another by the Inspector General's (IG) Office of the Department of Health and Human Services (Office of the Inspector General, 1998a, 1998b, 1998c). Both warned that IRBs are under pressure because of increasing workload without adequate support or resources. The GAO concluded that current oversight mechanisms were preventing major abuses, but that measures of IRB performance were lacking. The IG, in contrast, concluded that IRBs' ability to safeguard the rights and welfare of human research subjects was seriously strained because of the high volume of studies and pervasive conflicts of interest.

In sum, the studies in this section recommended increasing the diversity of IRB membership and enhancing their training, managing IRB member conflicts of interest, decreasing IRB costs by consolidating and centralizing IRB review, and developing performance measures for IRBs.

STUDIES OF IRB PROCESS—Five studies evaluated a particular aspect of the IRB review process (see Table 2). In one qualitative study, IRB members were interviewed about the process of reviewing emergency research with a consent exemption. The authors concluded that although IRB members found the community consultation requirement vague and difficult to implement, current regulations, if properly adhered to, would adequately protect subjects (McClure et al., 2007). In a survey about pediatric assent requirements, half of 188 IRB chairs said they used a standard age at which they required assent, while the other half relied on investigators' judgment (Whittle et al., 2004). An analysis of FDA warning letters to hospital, university, and commercial IRBs included citations for process failures, such as failure to have or follow adequate written procedures for research review; to prepare and maintain adequate documentation; and to conduct adequate continuing review (Bramstedt & Kassimatis, 2004). A review of OHRP compliance oversight determination letters found that citations of non-compliance were primarily for deficiencies in following the regulations for initial IRB review and deficiencies in consent documents (Weil et al., 2010). A study commissioned by the President's Commission (President's Commission, 1983) concluded that IRB processes could feasibly be evaluated through peer site visits. The studies in this group recommended clearer guidance and training for IRB members, more consistent application of federal regulations, and methods for evaluation of IRB performance.

STUDIES OF MULTICENTER VARIATION IN PROCESS AND OUTCOME—A subset of published empirical studies of IRBs (n = 16) evaluated variation in the processes and/or outcomes of review by different IRBs for multicenter studies (see Table 3). In each of these studies, the same research proposal was submitted to IRBs at multiple sites, and data were collected that demonstrate considerable variation among them in the type of review required (Dziak et al., 2005; Larson et al., 2004; McWilliams et al., 2003); the time it took to review the proposed research (Clark et al., 2006; Dziak et al., 2005; Green, Lowery, & Wyszewianski, 2006; Greene, Geiger, & Harris, 2006; Helfand et al., 2009; Hirshon et al., 2002; Larson et al., 2004; Mansbach et al., 2007; Stair et al., 2001); the designation of risk level (Mansbach et al., 2007; McWilliams et al., 2003); acceptable methods for recruitment

of subjects (Clark et al., 2006; Silverman, Hull, & Sugarman, 2001); the number and type of IRB concerns expressed or changes required (Clark et al., 2006; Green et al., 2006; S. Greene et al., 2006; Greene & Geiger, 2006; Helfand et al., 2009; Stair et al., 2001; Stark, Tyson, & Hibberd, 2009; Sherwood et al., 2006); and the IRB determination (Stair et al., 2001; Stark et al., 2009). Variation in informed consent practices and inadequacies in regulatory completeness of the consent forms were documented in one study (Silverman et al., 2001), and variation in cost in another (Vick et al., 2005). Further, study investigators argued that inconsistencies can have negative consequences for investigators, for research participants, and for the integrity of the science. All studies in this section recommended the development of a process for multicenter review to increase efficiency and reduce variation.

STUDIES OF IRB OUTCOME—Particular outcomes beyond multicenter variation in review practices and outcomes were evaluated in six studies, and again considerable variation among IRBs was shown (see Table 4). In one study, wide differences were noted in IRB decisions about compensation and assent in pediatric research (Kimberly et al., 2006). In another study in which IRB chairs rated the risk level of certain pediatric research procedures, only a single blood draw was rated as minimal risk by the majority (Shah et al., 2004). Burman and colleagues found that for a centrally approved TB protocol, most IRBs made no changes in the protocol, but requested consent form changes that made the forms longer and more complex (Burman et al., 2003). Another study documented variation from existing guideline recommendations in consent forms for research with stored samples (White & Gamm, 2002). A survey of multiple IRBs showed the most common reason for rejection of studies was problems with informed consent (Jones et al., 1996). A final study in this group showed that the majority of proposals reviewed by one IRB over twelve years were approved with stipulations; no proposals were rejected (Grodin et al., 1986). The studies in this section showed that IRBs have different views about the level of risk of common procedures, that IRB-requested changes make consent forms longer and more complex, and that although no studies were rejected in a study of one IRB, in another study of multiple IRBs the majority of rejections were related to the consent document.

Discussion

This systematic review of empirical studies of IRBs provides valuable information about what we know about IRBs and points to critical gaps in our knowledge. Considerable effort has been expended over several decades to evaluate and document practices, inconsistencies, and variation in the structures, processes, and outcomes of IRB review. These data from forty-three published studies show that U.S. IRBs differ in their application of the same set of regulations and are somewhat inconsistent in their judgments.

Data presented in this review provide support to commentators who have complained about inconsistencies, delays, inefficiencies, “redundant reviews in multi-site trials, and needless tinkering with consent forms” (NIH, 2006) for many kinds of studies reviewed by U.S. IRBs. Overwhelmingly, these data show that IRB practices and decisions, including determinations of which studies require full or expedited review, whether the level of risk was minimal or greater than minimal, and practices related to recruitment, vary from IRB to IRB, often without a clear justification for the variation.

Of note, investigators of several of these studies suggest that changes required by individual IRBs can sometimes jeopardize the scientific integrity of multicenter studies and national studies. Data also suggest that difficulty and delays with the local IRB approval process sometimes result in sites or investigators choosing not to participate in research (Mansbach et al., 2007). The extent to which IRB review inhibits important and ethically appropriate research is not known, but potentially troublesome.

Although some data suggest that IRBs focus on consent documents, little is documented about other issues that IRBs consider important to their decisions or about the substance of IRB deliberations. Although investigators of reviewed studies and other commentators have repeatedly called for measures of IRB quality and effectiveness, we could not identify one study that evaluated the effect that IRB review has on the protection of human subjects.

Investigators of reviewed studies document substantial inefficiencies in existing structures for IRB review, and many argue that centralizing review for multicenter studies would not only be more efficient, but better for the integrity of the science without jeopardizing protection of human subjects. The National Cancer Institute, the Veterans' Administration, and others have important central IRB review initiatives for multisite studies, and such initiatives are likely to expand (NCI Central IRB Initiative; VA Central Institutional Review Board). However, some are reluctant to use central review principally because of belief in the value of local input and concerns about local institutional liability (Loh & Meyer, 2004). Possible positive aspects of local review might include knowledge of the local participant populations and the local investigators, and opportunities for mentoring junior investigators and research staff (NIH, 2006). Research documenting the advantages of local review over central review in protecting research participants would be useful to informing decisions about review. Commentators have proposed centralizing or regionalizing review for all studies—not just multicenter—with the goal of increasing review efficiency and quality (Wood, Grady, & Emanuel, 2004). Data on how well each model protects participants, as well as how they differ with regard to efficiency, the major issues identified, and costs, would be valuable in considering these options.

Investigators of the studies reviewed here and numerous commentators yearn for more efficiency and less variation in IRB review (Byrne et al., 2006; Emanuel et al., 2004; Fost & Levine, 2007; Goldman & Katz, 1982; Greene & Geiger, 2006; Helfand et al., 2009; Hirshon et al., 2002; Kimberly et al., 2006; Larson et al., 2004; NIH, 2006; OIG, 1998b; Silverman et al., 2001; Vick et al., 2005; Wagner et al., 2004; Whittle et al., 2004; Wood et al., 2004). There is no apparent reason that IRB review of the same protocol at one institution can be completed in a week, while at another institution it takes thirty or more weeks. Efficiency—achieving a desired outcome with a minimum expenditure of time, effort, and resources—is an important value, as is a responsive process that does not needlessly delay research. Efficiency, however, requires clarity about the desired outcome or result. Simply measuring and shortening the time from IRB submission to approval (although likely to be appreciated for many reasons) may be insufficient for improving efficiency. Measures of IRB quality or metrics to show whether or not an IRB achieves the desired results in an efficient manner are also necessary. As noted in the summary of an NIH workshop on Alternative Methods of Review, “Issues to resolve include a clear understanding of what ‘quality’ really means in a review and how it can be measured” (NIH, 2006).

Furthermore, variation, in and of itself, can be legitimate and not necessarily problematic (Edwards, Ashcroft, & Kirchin, 2004). Unjustified variation, however, can be problematic, and as noted, variation in assessment of risk or application of certain regulations can jeopardize the science or contribute to decreased research productivity and increased expense without enhancing protections. Variation in IRB process and outcomes was repeatedly demonstrated in studies reviewed here. Early studies (Goldman & Katz, 1982; Gray, 1975) conducted with the explicit expectation that IRBs would apply similar criteria and reach similar conclusions about a protocol instead found substantial inconsistencies. Although flaws have been noted in their study (Levine, 1984), in 1982, Goldman and Katz said, “These findings [of inconsistency] cast doubt on the adequacy of IRB decision-making and the effectiveness of regulations” (Goldman & Katz, 1982). Multiple studies over the

ensuing decades have repeated this finding and expressed similar concerns. Some level of dissimilarity in interpretation of the regulations might be expected in a system of local review. What has not been addressed, however, is the justification for variation and how variation contributes to protection of human subjects. Importantly, some variation—for example, where one or more IRBs determine that a particular study is too risky to approve and others expedite the same study as minimal risk—seems irrational and could feasibly impact both safety and scientific rigor (Goldman & Katz, 1982; Mansbach et al., 2007; McWilliams et al., 2003; Rogers et al., 1999; Stark et al., 2009). Since the goal of the IRB is to safeguard the rights and welfare of human subjects in research, it is critical to determine the extent to which differences in review processes or outcomes thwart or uphold this objective.

Alternative models to the single institutional IRB should continue to be utilized and evaluated (Maschke, 2008; NIH, 2006). In addition to concern about wide differences in outcome and the time and resources expended to review and revise a single proposal numerous times for multisite studies, evidence is needed regarding what protections multiple IRB reviews add. Studies are also needed to evaluate how well centralized review achieves quality and efficient review of research. Well-defined metrics will be crucial to making these determinations.

Much can be learned from available data on IRBs in the United States, including that (1) IRB members need and want more guidance or training; (2) IRB membership could be more diverse; (3) the volume of IRB workload varies by institution, and the costs per review are generally lower in programs that review a higher volume of studies; (4) interpretation of federal regulations varies among IRBs, and additional guidance or clarification about how to apply the regulations would be seen as helpful; (5) the review of multisite studies of various kinds by multiple IRBs is inefficient, perceived as burdensome, and results in changes to research proposals that could affect the quality of the science; (6) IRB decisions differ about compensation, risk level, recruitment methodologies, and other important aspects of research without explanation for the differences; and (7) IRBs frequently recommend changes to consent documents, but changes do not always ensure compliance with existing recommendations. In addition to addressing these documented problems and inconsistencies, proposals continue to call for assessment of the impact of IRB review on protection of human subjects (Taylor, 2007; Candilis, Lidz, & Arnold, 2006; Coleman & Bouesseau, 2008). The impact of IRB review on the protection of subjects has proven difficult to measure, yet efforts are needed to identify and test appropriate, well-defined, acceptable, and useful metrics to determine how well IRBs accomplish their objectives, and how they can do so in an efficient and reasonable manner.

Best Practices

IRBs or others that have developed mechanisms or best practices for enhancing efficiency, for documenting the rationale for their decisions, for justifying variation, for reducing costs, or for measuring IRB quality should share these with other IRBs and interested parties.

Research Agenda

As described above, considerable additional research would be valuable in informing the future structure and organization, processes, and outcomes of IRBs in the U.S. An important overarching question is, What do we expect from IRBs? Clarifying expectations is important to being able to measure effectiveness. For example, should we expect IRBs to be more consistent at determining the risk of certain procedures or the risk level of a study? In protecting subjects from risk, could we examine how IRBs minimize risk? Or how changes in study proposals required by the IRB protect participants from risk? Centralized data on

the risks that research participants experience would also be helpful in this regard. We might determine, perhaps through interviewing a subset of investigators in different areas, how often studies are not done (and what kind of studies) or how often investigators or sites chose not to be involved in a study *because* IRB review is a deterrent. It would be helpful to investigate how often IRBs disapprove studies and on what basis; or similarly, how often IRBs approve studies without any stipulations. Ethnographic studies of IRB meetings or studies of IRB minutes could help to identify the primary issues that IRBs consider in their review of research proposals and where there are gaps. Importantly, efforts should be made to identify and test metrics for measuring the quality of IRB review and the effect of IRB review on protection of human subjects. It may be necessary to first develop a consensus view on how IRB quality should be understood. Studies could then investigate how quality differs by IRBs, what accounts for the differences, and how to identify best practices. Studies might compare local review and central review with respect to the quality of IRB discussions, attention to important issues, outcome determinations, costs, efficiency, and other factors. Work should be done to identify areas of variation that might be acceptable or even justifiable between IRBs reviewing the same study. Studies should also be done to investigate what changes are stipulated by IRBs in consent forms, and what impact these changes have on the research participants' understanding. Evaluation of different strategies used at IRB meetings—e.g., the use of primary and secondary reviewers, checklists, or a standard set of specific questions—would be useful. Additional helpful studies could evaluate the effect on investigator satisfaction, IRB efficiency, and the focus of the pre-IRB review by an IRB coordinator or administrator to make sure that the written submission is complete and addresses regulatory issues.

Educational Implications

Data show that IRB members would benefit from, and express a desire for, additional training regarding protection of human subjects. Existing educational programs about the federal regulations and interpretive guidance should be available to all IRB members, and updated periodically. In addition, education about the function and merits of IRB review continues to be valuable for the important stakeholders, including investigators, institutional officials, and research participants.

Acknowledgments

The authors acknowledge Nicole Newman and Karen Smith for their assistance in identifying empirical studies, and Frank Miller for his review and constructive comments. This work was supported by the Clinical Center, Department of Bioethics, National Institutes of Health. The views expressed are those of the authors and do not necessarily reflect those of the Clinical Center, the National Institutes of Health, the Public Health Service, or the U.S. Department of Health and Human Services.

Biographies

Lura Abbott is now retired, but worked on this article when she was on the staff of the Office of Human Subjects Research in the Division of Intramural Research, National Institutes of Health. Her responsibilities at OHSR involved overseeing a number of NIH Intramural IRBs. Previously, she had been an IRB coordinator for many years. She wrote her doctoral dissertation on the evaluation of IRBs.

Christine Grady is the Acting Chief of the Department of Bioethics, Clinical Center NIH. Her primary research interests are the ethics of clinical research, including interest in research oversight and IRB review. She has been a member of an IRB for over 20 years.

References

- Association for the Accreditation of Human Research Protection Programs, Inc. (AAHRPP).
www.aahrpp.org
- Anderson E. A qualitative study of non-affiliated, non-scientist institutional review board members. *Accountability in Research*. 2006; 13(2):135–155. [PubMed: 16827216]
- Bell, J. Evaluation of NIH Implementation of Section 491 of the Public Health Service Act, Mandating a Program of Protection for Research Subjects. National Institutes of Health; Bethesda, MD: 1998.
- Bramstedt K, Kassimatis K. A study of warning letters issued to institutional review boards by the United States Food And Drug Administration. *Clinical Investigational Medicine*. 2004; 27(6):316–323.
- Burman W, Breese P, Weis S, Bock N, Bernardo J, Vernon A, et al. The effects of local review on informed consent documents from a multicenter clinical trials consortium. *Controlled Clinical Trials*. 2003; 24:245–255. [PubMed: 12757991]
- Burman WJ, Reves R, Cohn. D, Schooley R. Breaking the camel's back: Multicenter clinical trials and local institutional review boards. *Annals of Internal Medicine*. 2001; 134(2):152–157. [PubMed: 11177319]
- Byrne M, Speckman J, Getz K, Sugarman J. Variability in the costs of institutional review board oversight. *Academic Medicine*. 2006; 81(8):708–712. [PubMed: 16868423]
- Campbell E, Weissman J, Clarridge B, Yucel R, Causino N, Blumenthal N. Characteristics of medical school faculty members serving on institutional review boards: Results of a national survey. *Academic Medicine*. 2003; 78(8):831–836. [PubMed: 12915379]
- Campbell EG, Weissman JS, Vogeli C, Clarridge BR, Abraham M, Marder JE, et al. Financial relationships between institutional review board members and industry. *New England Journal of Medicine*. 2006; 355(22):2321–2329. [PubMed: 17135585]
- Candilis P, Lidz C, Arnold R. The need to understand IRB deliberations. *IRB: Ethics and Human Research*. 2006; 28(1):1–5. [PubMed: 16680872]
- Clark S, Pelletier AJ, Brenner BE, Lang DM, Strunk RC, Camargo CA. Feasibility of a national fatal asthma registry: More evidence of IRB variation in evaluation of a standard protocol. *Journal of Asthma*. 2006; 43(1):19–23. [PubMed: 16448960]
- Coleman C, Bouesseau M. How do we know if research ethics committees are really working? The neglected role of outcomes assessment in research ethics review. *BMC Medical Ethics*. 2008; 9:6. [PubMed: 18373857]
- Cooke, RA.; Tannenbaum, AS.; Gray, B. A survey of institutional review boards and research involving human subjects. The National Commission for the Protection of Human Subjects of Biomedical and Behavioral Research, DHEW; 1977.
- DeVries RG, Forsberg CP. What do IRBs look like? What kind of support do they receive? *Accountability in Research*. 2002; 9:199–216. [PubMed: 12816128]
- Department of Health and Human Services (DHHS). 45 CFR 46
- Dziak K, Anderson R, Sevick MA, Weisman C, Levine D, Scholle S. Variations among institutional review boards reviews in a multisite health services research study. *HSR: Health Services Research*. 2005; 40(1):279–290.
- Edwards S, Ashcroft R, Kirchin S. Research ethics committees: Differences and moral judgement. *Bioethics*. 2004; 18(5):408–427. [PubMed: 15462024]
- Emanuel EJ, Wood A, Fleischman A, Bowen A, Getz KA, Grady C, et al. Oversight of human participants research: Identifying problems to evaluate reform proposals. *Annals of Internal Medicine*. 2004; 141(4):282–291. [PubMed: 15313744]
- Federal Drug Administration (FDA). 21 CFR 50
- Fost N, Levine R. The dysregulation of human subjects research. *Journal of the American Medical Association*. 2007; 298(18):2196–2198. [PubMed: 18000206]
- Government Accountability Office (GAO). Federal report says protection of human subjects is threatened by numerous factors. *Human Research Report*. 1996; 11(5)
- Goldman J, Katz MD. Inconsistency and institutional review boards. *Journal of the American Medical Association*. 1982; 248(2):197–202. [PubMed: 7045428]

- Gray BH. An assessment of Institutional Review Committees in human experimentation. *Medical Care*. 1975; 13(4):318–328. [PubMed: 1091788]
- Green L, Lowery C, Kowalski C, Wyszewianski L. Impact of institutional review board practice variation on observational health services research. *HSR: Health Services Research*. 2006; 41(1): 214–230.
- Greene S, Geiger A, Harris E. Impact of IRB requirements on a multicenter survey of prophylactic mastectomy outcomes. *Annals of Epidemiology*. 2006; 16(4):275–278. [PubMed: 16005245]
- Greene SM, Geiger AM. A review finds that multicenter studies face substantial challenges but strategies exist to achieve Institutional Review Board approval. *Journal of Clinical Epidemiology*. 2006; 59(8):784–790. [PubMed: 16828670]
- Grodin MA, Zaharoff BE, Kaminow PV. A 12-year audit of IRB decisions. *Quality Review Bulletin*. 1986; 12(3):82–86. [PubMed: 3085041]
- Gunsalus C, Bruner E, Nicholas B, Dash L, Finkin M, Goldberg J, et al. Mission creep in the IRB world. *Science*. 2006; 312:1441. [PubMed: 16763114]
- Hayes GJ, Hayes SC, Dykstra T. A survey of university institutional review boards: Characteristics, policies, and procedures. *IRB: Ethics and Human Research*. 1995; 17(3):1–6. [PubMed: 11654267]
- Helfand BT, Mongiu AK, Roehrborn CG, Donnell RF, Bruskewitz R, Kaplan SA, et al. Variation in institutional review board responses to a standard protocol for a multicenter randomized, controlled surgical trial. *Journal of Urology*. 2009; 181(6):2674–2679. [PubMed: 19375101]
- Hirshon JM, Krugman SD, Witting MD, Furuno JP, Limcangco MR, Perisse AR, et al. Variability in institutional review board assessment of minimal-risk research. *Academic Emergency Medicine*. 2002; 9(12):1417–1420. [PubMed: 12460847]
- Institute of Medicine (IOM). *Responsible research: A systems approach to protecting research participants*. Institute of Medicine; Washington, DC: 2002.
- Jones JS, White LJ, Pool LC, Dougherty JM. Structure and practice of institutional review boards in the United States. *Academic Emergency Medicine*. 1996; 3(8):804–809. [PubMed: 8853678]
- Kimberly MB, Hoehn KS, Feudtner C, Nelson RM, Schreiner M. Variation in standards of research compensation and child assent practices: A comparison of 69 institutional review board–approved informed permission and assent forms for 3 multicenter pediatric clinical trials. *Pediatrics*. 2006; 117(5):1706–1711. [PubMed: 16651328]
- Larson E, Bratts T, Zwanziger J, Stone P. A survey of IRB process in 68 U.S. hospitals. *Journal of Nursing Scholarship*. 2004; 36:260–264. [PubMed: 15495496]
- Lemonick M, Goldstein A, Park A. Human guinea pigs. *Time Magazine*. April 22, 2002
- Levine RJ. Inconsistency and IRBs: Flaws in the Goldman-Katz study. *IRB: A Review of Human Subjects Research*. 1984; 6(1):4–6.
- Loh ED, Meyer RE. Medical schools' attitudes and perceptions regarding the use of central institutional review boards. *Academic Medicine*. 2004; 79(7):644–651. [PubMed: 15234914]
- Mansbach J, Acholonu U, Clark S, Camargo C. Variation in institutional review board responses to a standard, observational, pediatric research protocol. *Academic Emergency Medicine*. 2007; 14(4): 377–380. [PubMed: 17312334]
- Maschke K. Human research protections: Time for regulatory reform. *Hastings Center Report*. 2008; 38(2):19–22. [PubMed: 18457223]
- McCarthy, C. Challenges to IRBs in the coming decades. In: Vanderpool, Y., editor. *The Ethics of Research Involving Human Subjects: Facing the 21st Century*. University Publishing Group, Inc.; Frederick, MA: 1996. p. 127-144.
- McClure K, Delorio N, Schmidt T, Chiodo G, Gorman P. A qualitative study of institutional review board members' experience reviewing research protocols using emergency exception informed consent. *Journal of Medical Ethics*. 2007; 33(5):289–293. [PubMed: 17470507]
- McWilliams R, Hoover-Fong J, Hamosh A, Beck S, Beaty T, Cutting G. Problematic variation in local institutional review of a multicenter genetic epidemiology study. *Journal of the American Medical Association*. 2003; 290(3):360–366. [PubMed: 12865377]
- National Cancer Institute Central IRB Initiative. www.ncicirb.org

- National Institutes of Health (NIH). National conference on alternative IRB models: Optimizing human subjects protection. Paper presented at the National Conference on Alternative IRB Models: Optimizing Human Subjects Protection; 2006.
- Office of the Inspector General (OIG). Institutional Review Boards: Promising approaches. US Department of Health and Human Services; Washington, DC: 1998a.
- Office of the Inspector General (OIG). Institutional Review Boards: A system in jeopardy. Department of Health and Human Services; Washington, DC: 1998b.
- Office of the Inspector General (OIG). Institutional Review Boards: A time for reform. U.S. Department of Health and Human Services; Washington, DC: 1998c.
- President's Commission. Implementing human research regulations. President's Commission for the Study of Ethical Problems in Medicine and Biomedical and Behavioral Research; Washington, DC: 1983.
- Rogers A, Schwartz D, Weissman G, English A. A case study in adolescent participation in clinical research: Eleven clinical sites, one common protocol, and eleven IRBs. *IRB: Ethics and Human Research*. 1999; 21(1):6–10. [PubMed: 11657967]
- Rothstein WG, Phuong LH. Ethical attitudes of nurse, physician, and unaffiliated members of institutional review boards. *Journal of Nursing Scholarship*. 2007; 39(1):75–81. [PubMed: 17393970]
- Sengupta S, Lo B. The roles and experiences of nonaffiliated and nonscientist members of institutional review boards. *Academic Medicine*. 2003; 78(2):212–218. [PubMed: 12584103]
- Shah S, Whittle A, Wilfond B, Gensler G, Wendler D. How do institutional review boards apply the federal risk and benefit standards for pediatric research? *Journal of the American Medical Association*. 2004; 291(4):476–482. [PubMed: 14747505]
- Sherwood ML, Buchinsky FJ, Quigley MR, Donfack J, Choi SS, Conley SF, et al. Unique challenges of obtaining regulatory approval for a multicenter protocol to study the genetics of RRP and suggested remedies. *Otolaryngology–Head and Neck Surgery*. 2006; 135(2):189–196. [PubMed: 16890066]
- Silverman H, Hull SC, Sugarman J. Variability among institutional review boards' decisions within the context of a multicenter trial. *Critical Care Medicine*. 2001; 29(2):235–241. [PubMed: 11246299]
- Sobolski G, Flores L, Emanuel E. Institutional review board review of multicenter studies. *Annals of Internal Medicine*. 2007; 146(10):759. [PubMed: 17502642]
- Stair TO, Reed CR, Radeos MS, Koski G, Camargo CA. Variation in institutional review board responses to a standard protocol for a multicenter clinical trial. *Academic Emergency Medicine*. 2001; 8(6):636–641. [PubMed: 11388939]
- Stark A, Tyson J, Hibberd P. Variation among institutional review boards in evaluating the design of a multi-center randomized trial. *Journal of Perinatology*. 2010; 30:163–169. [PubMed: 19798046]
- Steinbrook R. Improving protection for research subjects. *New England Journal of Medicine*. 2002; 346(18):1425–1430. [PubMed: 11986426]
- Strech D, Synofik M, Marckmann G. Systematic reviews of empirical bioethics. *Journal of Medical Ethics*. 2008; 34(6):472–477. [PubMed: 18511623]
- Taylor H. Moving beyond compliance: Measuring ethical quality to enhance the oversight of human subjects research. *IRB: Ethics and Human Research*. 2007; 29(5):9–14. [PubMed: 17969778]
- Veterans Administration Central Institutional Review Board.
www.research.va.gov/programs/pride/cirb/default.cfm
- Vick C, Finan K, Kiefe C, Neumayer L, Hawn M. Variation in institutional review process for a multi-site observational study. *The American Journal of Surgery*. 2005; 190(5):805–809.
- Wagner T, Bhandari A, Chadwick G, Nelson D. The cost of operating institutional review boards. *Academic Medicine*. 2003; 78(6):638–644. [PubMed: 12805049]
- Wagner T, Cruz AM, Chadwick G. Economies of scale among institutional review boards. *Medical Care*. 2004; 42(8):817–823. [PubMed: 15258484]
- Weil C, Rooney L, McNeilly P, Cooper K, Borrer K, Andreason P. OHRP compliance oversight letters: An update. *IRB: Ethics and Human Research*. 2010; 32(2):1–6. [PubMed: 20373732]

- White MT, Gamm J. Informed consent for research on stored blood and tissue samples: A survey of institutional review board practices. *Accountability in Research*. 2002; 9(9):1–16. [PubMed: 12705242]
- Whitney SN, Alcer K, Schneider CE, McCullough LB, McGuire AL, Volk RJ. Principal investigator views of the IRB system. *International Journal of Medical Sciences*. 2008; 5(2):68–72. [PubMed: 18392146]
- Whittle A, Shah S, Wilfond B, Gensler G, Wendler D. Institutional review board practices regarding assent in pediatric research. *Pediatrics*. 2004; 113(6):1747–1752. [PubMed: 15173501]
- Wood A, Grady C, Emanuel EJ. Regional ethics organizations for protection of human research participants. *Nature Medicine*. 2004; 10(12):1283–1288.

(MeSH terms): (Ethics committees, research (which is used for, and includes, institutional review boards)).
AND
(MeSH terms): Evaluation studies as topic
Program evaluation
OR
Outcome assessment
Observation
Quality assurance
(Subheading): Organization and administration
Standards
OR
Legislation and jurisprudence
(Text word: Empiric* (truncated to pick up variants of the root word) title/abstract)

Fig. 1.
Search Terms Used in the Systematic Review of the Literature on PubMed.

TABLE 1

Studies evaluating IRB Structure.

Author(s)/ Year	objectives	Methods	Results	conclusions
MEMBERSHIP				
Rothstein & Phuong 2007 (27)	Compare attitudes towards IRB review and perceived influence of physician, nurse, unaffiliated, and other IRB members	Survey of 284 IRB members from 27 IRBs (37% physicians, 10% nurses, 21% unaffiliated)	Self-rated influence was lower than influence rated by others for each group (IRB chairs 78%/95%; physicians 62%/94%; nurses 39%/56%; unaffiliated members 24%/52%). Ethical issue most frequently (89%) selected as very important by all groups: "informing subjects of research risks."	Recommend adding more nurses. Adding more unaffiliated members, as recommended by the National Bioethics Advisory Commission and others, is not likely to result in more attention to ethical issues.
Anderson 2006 (25)	Examine the experiences and perceptions of non-affiliated/non-scientist (NA/NS) IRB members	Qualitative interviews of 16 NA/NS members regarding IRB experience, responsibilities, contributions, and training	NA/NS members were mostly white, educated, and professional, and recruited through contacts at the institution. Very few had experience as research participants. NA/NS members asked for more clarity about their role and more respect from other members.	Recommend varied recruitment methods to ensure more diverse NA/NS members, and members with previous research experience; a clearer role description; and tailored training for NA/NS members.
Campbell et al. 2006 (24)	Examine the nature and extent of financial relationships between academic medical center IRB members and industry	Survey of 893 members from 100 academic medical centers in the U.S. (RR = 67.2%)	Overall 36% of IRB members had at least 1 financial relationship with industry. 15% said they had a potential conflict with a protocol presented to their IRB within the past year; of this subgroup, 58% disclosed the relationship to the IRB, and 65% recused themselves from the vote. 85% did not think financial relationships of fellow IRB members adversely affected decisions.	Financial relationships between IRB members and industry are common. Current regulations and policies to reduce possible conflicts of interest should be examined.
De Vries & Forsberg 2002 (22)	Examine the characteristics of multiple IRBs, including level of financial support	Survey of random sample of 89 IRBs selected from the OHRP list. 87/89 (98% response) administrators responded representing 206 IRBs.	IRBs had an average of 13 members who are predominantly white, affiliated, and more often males. Physicians make up the largest discipline. Number of current active studies not always known and varied greatly. 25% of the IRBs, more often at medical schools and universities, had > 500 active studies. At most IRBs, staff support is inadequate for the volume. Chairs reported that few members had ethics training (< 20%). Of 1161 members, only 20 were trained ethicists.	IRBs continue to be comprised predominantly of white scientists affiliated with the institution. Thus, the interests of researchers and institutions may take precedence over the interests of research subjects. Adequate staff support is a glaring need that may compromise oversight and monitoring. Authors recommend including more lay members of IRBs.
Sengupta 2003 (26)	Examine roles and experiences of non-scientist, non-affiliated (NA/NS) IRB members	Telephone interviews of 32 NA/NS members from 11 IRBs regarding IRB education, interactions with scientific members, IRB problems	Participants said their role was to represent the community of human subjects; 94% felt their main role was to simplify consent documents. 88% had occasionally felt intimidated by scientist members. 78% wanted more education and training.	Measures are needed to strengthen the relationship between scientist and lay members. Additional training should be offered to NS/NA members.
Campbell et al. 2003 (21)	Examine the characteristics of medical school faculty members who serve on IRBs	Survey of 2,989 faculty members in 121 four-year U.S. medical schools (response rate = 64%)	Of respondents—73% male, 81% white, 11% had been on the board for 3 years, 94% actively conduct research (71% clinical), 47% consulted with industry. Logistic regression predicted that underrepresented minorities were 3.2 times as likely to serve on an IRB.	Most IRB members are active researchers, almost half serve as consultants to industry, creating possible conflicts of interest. Serving on IRB was not associated with decreased levels of research productivity.

Author(s)/ Year	objectives	Methods	Results	conclusions
Hayes, G. et al.1995 (23)	Examine IRB composition, policies, and procedures	186 questionnaires returned from 129 universities	Mean number of IRB members = 16, 66% male and 90 % white. 62% of IRBs met monthly, wide range of number of proposals reviewed (mean 29); 25% offered no training to members; policies were variable.	Recommend development of policies and strategies for recruitment of IRB members to ensure diversity in demographic characteristics and experience, and minimize groupthink. Note a critical need for training of IRB members, and regular evaluation of IRB effectiveness.
IRB COSTS				
Sobolski et al.2007 (28)	Estimate the potential benefits and cost savings of centralized review for multisite studies	Survey of 73 IRBs regarding number of multicenter studies, and average number of other centers involved	Only 14 of 73 IRBs knew the number of institutions involved in their multicenter studies. 32 IRBs recorded whether a study was multisite.	Only a fraction of IRBs record multicenter status. Centralization of review could yield substantial cost savings, although estimates of savings are not possible due to imperfect tracking of pertinent data.
Byrne et al. 2006 (29)	Explore variability among IRB costs based on protocol type, organizational size, and specific components of oversight	Analyzed resource utilization data from 59 academic medical center IRBs	Overall median cost for reviewing a protocol is \$560. Although protocol review costs decreased with increasing number of protocols, not all components of IRB costs (e.g., board time) are subject to economies of scale. Expedited reviews are no less expensive at low-volume institutions than full review. Staff time accounts for the majority of costs (65%)	IRB costs for oversight are highly variable. In addition to notable economies of scale, there is considerable variation in the operating and per protocol costs of IRBs of similar size. Standards for measuring the quality and efficiency of IRBs are needed.
Wagner et al.2004 (31)	Examine the association of costs with IRB size and number of actions	Cross-sectional survey of 67 VA medical center IRB administrators (response rate = 61%)	The average cost per action for a small IRB (52 actions) is \$2781, \$416 for a medium (431 actions) IRB, and \$187 for a large (2676 actions) IRB. Small IRBs use as many resources as large IRBs.	Large economies of scale exist. The number of IRBs necessary for review should be considered; other arrangements may be economically advantageous.
Wagner et al.2003 (30)	Estimate the costs of operating high-volume and low-volume IRBs	Average cost per IRB action was calculated from 2001 summary data on personnel (both members and staff); space; computers, databases, equipment, and supplies; training and education	Annual cost to operate a high-volume IRB is \$770,674, and \$76,626 for a low-volume IRB; however, the average cost per action is about 3x higher for a low-volume IRB (\$799) than a for high-volume IRB (\$277). Notably more variability in costs for low-volume IRBs than high-volume IRBs.	High-volume IRBs are more economically efficient. Merging smaller IRBs may produce cost savings and economies of scale. Data is needed on the quality of review.
GENERAL				
Bell & associates 1998 (19)	Evaluate volume and adequacy of IRB review	Surveys from 394 IRB chairs; 249 IRB administrators; 400 institutional officials; 435 IRB members; 632 PIs-all from 491 MPA institutions	Continued increase in the volume of protocols to review, but unevenly distributed. IRB chairs and members generally satisfied with their contributions to human subjects protections, whereas investigators were less likely to agree that the IRB focused on its primary mandate in an efficient manner.	The IRBS system continues to provide an adequate level of protection at a reasonable cost, consistent with federal regulations. No major adjustment is needed; improvements possible by honing IRB structures and procedures and providing increased education and training.
DHHS Inspector General 1998 (1)	Summarize challenges facing IRBs and recommend federal action	Review of federal records, interview of federal officials and representatives of 75 IRBs, site visits to 6 IRBs, and FDA IRB site inspections	IRBs face major challenges: they review too much, too quickly, with little expertise; minimal continuing review occurs; conflicts threaten their independence; members and investigators require more training.	Recommended changes in the federal regulations to give IRBs greater flexibility; holding IRBs accountable for results; strengthening some protections; more education and training; protecting IRBs from institutional conflicts; helping to mitigate workload

Author(s)/ Year	objectives	Methods	Results	conclusions
Government Accounting Office 1996 (34)	Determine if federal oversight processes reduced the likelihood of harm to human subjects	Site visit interviews, records review; 40 DHHS compliance cases, 69 FDA deficiency letters, 31 FDA warning letters	High pressures on IRBs: workload increasing scientific complexity; minimal federal oversight.	pressure and evaluating IRB effectiveness. Current oversight mechanisms appear to be adequately preventing major abuses, but measures of IRB effectiveness are needed.
Gray et al. 1978 (32)	Examine the characteristics of IRBs	Probability sample of 61 IRBs from 420 institutions; 2000 PIs, 800 IRB members, 1000 subjects. Site visits, interviews, review procedures, review research projects	IRBs had an average of: 43 protocols per year (major medical university hundreds). 14 members. 10 meetings per year (range 2–51). 14% modified every proposal and 22% modified a third or fewer.	IRB members and PIs believe IRBs add value; consents not improved; highly diverse policies and procedures. A substantial effort goes into the review process. Serious efforts to understand and improve IRB performance are needed.
Gray 1975 (33)	Assess the strengths and weaknesses of a single IRB	Intensive review of a single IRB; interviews with research subjects	IRB was active, conscientious, and mostly effective.	IRB has an impact on research performance, but lack of monitoring and feedback from subjects limits the impact on the actual conduct of the study.

TABLE 2

Studies Involving IRB Process.

Author(s)/Year	objectives	Methods	Results	conclusions
Weil, C. et al. 2010 (38)	Examine the range of determinations of non-compliance or deficiencies as a result of principally for-cause evaluations	Review of 235 determination letters issued to 146 institutions between August 2002 and August 2007	Most institutions were cited for problems with the initial review process (55%) or with IRB approved consent documents (51%)	This analysis provides information useful for OHRP in developing targeted guidance and educational programs and for institutions in taking proactive measures to be compliant.
McClure et al. 2007 (20)	Determine IRB members' experience when reviewing protocols using the informed consent exception in emergency research	Semi-structured telephone interviews of 10 purposively selected IRB members in 4 different regions of the U.S.	<ol style="list-style-type: none"> (1) These studies require lengthy review. (2) Community consultation and notification is vague and difficult to implement. (3) Current regulations protect human subjects. (4) Legal counsel is important. (5) IRB members have little training on regulations. 	Studies requiring an exception to informed consent in emergency research take longer to review. Community consultation is difficult and time consuming but adequately protects human subjects. Specific training for IRB members would be useful.
Bramstedt et al. 2004 (37)	Explore the ethical issues contained in FDA warning letters issued to IRBs	52 FDA warning letters evaluated for the type of violations cited	Hospital/medical center IRBs received the most letters, followed by university IRBs and private IRBs. Regulatory violations were: failure to have and follow adequate written procedures about how research review is conducted (50/52); inadequate documentation of IRB activities (47/52); and failure to provide adequate continuing review of approved studies (36/52). 19 letters were issued for consent form issues.	FDA warning letters consistently indicate weaknesses in review and documentation activities of audited IRBs. Overburdened IRBs who passively monitor studies raise concerns about study oversight and optimal protection of research subjects.
Whittle, A. et al. 2004 (36)	Determine how U.S. IRBs implement the federal assent requirement in pediatric research	Telephone survey of 188 U.S. IRB chairs	One half of IRBs have a method (usually age) for determining when a child can assent; the other half rely on investigators' clinical judgment. IRBs use adult regulations to determine what information should be given to children in assent.	IRBs need additional guidance on how to implement the assent requirement in pediatric research.
U.S. President's Commission 1983 (39)	Test whether IRBs could be evaluated by expert peer review	Site visits and interviews of 12 IRBs, records review	Wide process variability among IRBs, but generally effective.	Site visits of IRBs by expert peer reviewers are feasible and useful.

Table 3

Studies of IRB Variation in Process and outcome for Multicenter Studies.

Author/year	objectives	Methods	Results	conclusions
Helfand, B. et al. 2009 (43)	Examine variation among IRBs in evaluation of the Minimally Invasive Surgical Therapies study for benign prostatic hyperplasia	Analyzed and categorized responses after initial review from 6 of 7 participating IRBs. Resubmitted protocol approved by 1 institution to others for a second review.	Number and type of responses after review of an identical study varied significantly among IRBs. Required changes were most frequently clarification or word choices, followed by technical issues. The number of required changes correlated with time to approval. 95% of the 121 total requested changes were unique to the institution. Many more changes were required in the 2nd round of review.	Recommend central ethical review of surgical protocols to assure appropriate expertise and promote efficiency.
Stark, A. R. et al. 2009 (51)	Examine variation among IRBs in evaluation of a randomized placebo-controlled multicenter study of vitamin A supplementation for low birth weight infants	Analyzed and categorized concerns from the first written response to the site investigator from each of 18 reviewing IRBs	Half of the IRBs (9/18) withheld approval because of major concerns, including: study background and rationale, design, inclusion criteria, interventions, sample size, analysis, data monitoring; while 7/18 had no concerns, and 2/18 only minor concerns. 16/18 requested consent form changes.	Despite extensive review prior to the IRB, IRBs varied substantially in concerns about the study. IRB evaluations seem to vary because of difficulty and inadequate expertise in assessing the appropriateness of the study design for the scientific question.
Mansbach et al. 2007 (44)	Examine variation in IRB review of a multicenter observational study of children presenting with bronchiolitis to the ER	Reviewed submittal records and IRB correspondence from 34 of 37 sites participating in the Emergency Medicine Network; short survey to investigators	91% of IRBs considered the protocol minimal risk. 13 IRBs required no changes, 18 gave conditional approval, and 3 deferred the study. Median approval time was 42 days. Seven sites did not recruit patients, 1/7 said IRB approval was too late, and 4/7 found IRB hurdles insurmountable.	Although some investigators did not recruit patients because of IRB hurdles, the majority of investigators remained enthusiastic about multicenter research. Recommend centralizing IRB review and reducing unnecessary hurdles.
Greene, S. M. et al. 2006 (47)	Examine IRB variability in review of a multicenter mailed survey evaluating psychosocial outcomes of prophylactic mastectomy	Reviewed central log of IRB submission types, dates, approvals, and modification requests for 6 IRBs	IRBs required differences in physician involvement, active vs. passive consent. Modifications were unique to each IRB, did not overlap, and impacted multiple aspects of the study.	Inconsistencies and inefficiencies in review, and site-to-site variability in IRB requirements impacted site participation, study implementation, scientific rigor, and delayed implementation of multicenter studies. Call for centralized IRB review.
Greene, S. M. and Geiger, J. 2006 (52)	Descriptive review of published accounts of IRB review of multisite observational studies	Reviewed 40 peerreviewed papers and 6 reports by governmental commissions	Across all studies the time for approval ranged from 5 to 798 days. Study initiation delays were attributed to long IRB approval times. Costs of delayed approvals were high, e.g., 17% of overall costs went to IRB review. Consent changes often were detrimental and other human subject issues were dealt with differently by different sites.	Strategies for navigating the IRB process included cooperative activities between larger centers and newer entrants, cooperative agreements among centers to limit changes. Recommended structural changes included training and standardization, centralized review models, and comprehensive reform.
Green, L. et al. 2006 (46)	Descriptive review of the process required to obtain IRB approval for a multicenter observational health services research study	Reviewed field notes and document analysis at 43 VA medical centers IRBs	Wide variation in review standards (1 exemption, 10 expedited, 31 convened full board). One IRB disapproved the study based on risk. 76% required at least 1 resubmission, 15% required more than 3 resubmissions. Most revisions were editorial changes to the wording in the consent document. Many processing failures were noted (long approval times - median 286 days; lost paperwork;	Requiring IRB approvals at multiple sites imposes administrative burdens and delays. Paradoxically, human subjects protection may have been compromised due to IRBs' lack of understanding of confidentiality issues inherent in health services research. Central review with local opt-out, cooperative review, or a system of peer review, could

Author/year	objectives	Methods	Results	conclusions
Clark et al. 2006 (48)	Examine IRB variation in evaluation of a proposal for a national fatal asthma registry in 4 states	Compared responses from 4 separate IRBs as well as days to review	IRBs had divergent opinions about major ethical issues (e.g., opt-out vs. opt-in approach to contacting next of kin) and required other minor changes in a common protocol developed to comply with HIPAA and the federal regulations. Review took a mean of 64 days (38 to 89).	reduce costs and improve protection. Recommend centralized review of registry proposals since inconsistent assessment of what is ethically permissible jeopardizes multicenter efforts to examine important regional or national problems.
Sherwood et al. 2006 (53)	Examine variation in time and review of a minimal risk multicenter genetic study	Collected and reviewed process materials from 14 IRBs at children's hospitals over 2.5 yrs	Median time from investigator agreement to approval 14.7 months. Considerable variability in documentation requirements. 2/3 of the IRBs raised issues related to recruitment and 2/3 related to subject privacy. Almost 60% required word changes to the consent document.	Recommend the use of OHRP cooperative agreements and a standardized application package. Recommend that IRBs refrain from mere editorial changes.
Vick et al. 2005 (54)	Examine variation in time and resources for IRB review of a multicenter observational VA study	Compared time to complete application and review and resources expended (staff salary and % time) at 19 sites	Mean time to complete approval was 2.9 +/-1.5 months. Estimated \$53,000 spent on staff salary dedicated to IRB process (24% of first-year study budget).	The IRB process for multisite study is expensive in both time and money. National VA IRB is recommended.
Larson et al. 2004 (42)	Examine IRB variation in evaluation of a study of the effect of ICU working conditions on patient outcomes and healthcare worker safety	Compared processes used by 68 IRBs reviewing the same non-interventional protocol	IRB requirements and efficiency varied widely. 61.8% of the IRBs expedited review of the study. Expedited review took the longest (average 54.8 days, range 1–303) compared to full board review (47.1 days) and exemption (10.8 days). 26.5% of IRBs required human subjects training for investigators and 10.3% required conflict-of-interest statements.	IRB review processes are not standardized and review time varies widely. Absence of efficient streamlined review might unnecessarily impede national studies without improving participant safety.
Dziak et al. 2005 (40)	Examine IRB review process for a multisite study of healthcare quality involving 3000 telephone interviews at 15 primary care sites	Review of IRB application and correspondence records to note type of review, time for review, what the IRB required regarding patient notification, and participation rates	IRBs varied in determination of the type of IRB review required: 9 expedited, 5 full convened board, 2 exempt. Duration of review ranged from 5 to 172 days. 4 sites required patient notification in advance of the study. Refusal rates were lower at opt-out sites (2–11%) than opt-in (37%).	IRB variations in requirements can negatively affect response rates and alter the generalizability of study results.
McWilliams et al. 2003 (41)	Document variability among IRBs in the approval of a multicenter genetic epidemiology study	A 7-question survey sent to 31 cystic fibrosis care centers regarding IRB function and number and type of changes required	Highly variable review process among the IRBs. Evaluation of the risk of the same study ranged from minimal to high - 7 expedited, 24 full board reviews. IRBs required between 1 and 4 consents for this study.	Multisite studies are important to advances in human genome and genetic studies affecting the public health. Recommend centralized review to reduce variability in human subjects protection and promote efficiency.
Hirshon et al. 2002 (49)	Examine IRB variability in response to a minimal-risk survey of emergency room providers and people in the ER waiting area	Research survey submitted to IRBs at 3 different institutions	One institution reviewed the study in 12 business days and waived informed consent, a second took 15 days in expedited review, the third took 77 days and required 3 revisions, including study methodology, sample size, and recruitment strategies.	Review inconsistencies raise questions about the validity and efficiency of the IRB process. There is a need for standardization of the minimal risk review process to decrease variability and therefore improve the validity of the process.
Stair et al. 2001 (45)	Examine IRB variation in responses to a phase 4 multicenter, randomized, double-blind, placebo-controlled trial of	Survey of investigators participating in a single, standard multisite protocol at 44 sites	Median time from protocol delivery to approval was 102 days (8–142); an average of 3.5 changes were requested; 91% of them consent changes; 82% of protocols were returned from 1 to 4 times for	The use of a national IRB could focus on the scientific aspects of the study, leaving the properly local consent issues to local IRBs. Standardized submittal forms

Author/year	objectives	Methods	Results	conclusions
	outpatient therapy for acute asthma		revisions; 9% required only minor consent changes and 9% were unconditionally approved.	would be useful. PIs must plan time for preparing a submission.
Silverman et al. 2001 (50)	Determine extent of variability among IRBs that reviewed a multicenter ARDS Network ventilator protocol	Analysis of survey sent to all 16 approving IRBs and of approved consent forms	Variation noted among IRB research practices; e.g., 1/16 allowed a consent waiver; 5/16 allowed telephone consent; 3/16 allowed inclusion of prisoners. Only 3 consents had all the regulatory requirements, 13 were incomplete.	Differences in IRB practices attributable to differences in interpretation of federal regulations or level of IRB scrutiny rather than differences in local factors. Clarification of the regulations and education for IRB members are called for. Central review of multicenter studies may provide useful standardization of the protocol.
Goldman & Katz 1982 (63)	To assess adequacy of human subjects research review	3 protocols and consents with deliberate ethical flaws submitted for review to 22 IRBs	Responses to identical protocols were dissimilar; unacceptable scientific designs were approved by some; IRBs consistently objected to the consent forms, but objections were inconsistent.	Inconsistencies in reasoning were so large that protocol revision to satisfy the objections would result in approval of flawed studies. Continued examination of IRB decision-making is essential to the rational regulation of human subjects research.

Table 4

Studies of IRB outcomes.

Author(s)/Year	objectives	Sample/Method	Results	conclusions
Kimberly et al. 2006 (55)	Compare compensation and child assent in 3 standardized multicenter research protocols	69 IRB approved informed permission, assent, and consent forms were compared on compensation and assent requirements	Compensation and child assent requirements varied substantially. Not all IRBs approved compensation. Of those that approved, maximum levels of compensation varied 8-fold; the majority did not stipulate the form of payment or who was to receive it. 83% documented child assent in some way.	Extreme nature of the variation among IRBs warrants closer inspection. Guidelines for operationalizing compensation and assent practices would be helpful.
Shah et al. 2004 (56)	Determine how U.S. IRB chairs apply federal risk and benefit categories for pediatric research	Telephone survey of 188 U.S. IRB chairs	Only a single blood draw was categorized as minimal risk by the majority (81%); other procedures were categorized as both minimal and greater than a minor increase over minimal risk by different IRBs; 60% considered psychological counseling and 10% payment a direct benefit.	There is considerable variability in how IRB chairs apply the federal risk and benefit categories to pediatric research.
Burman et al. 2003 (57)	Evaluate the effects of local review on the consent forms of 2 studies from multicenter TB Trials Consortium	Independent review of changes made at 25 sites in the centrally approved consent forms	No changes were made in the protocol as a result of local review. Consent forms became longer and increased the grade level. 82.5% of changes altered wording without affecting meaning. Median time for local review was 104 days.	Local approval of 2 multicenter clinical trials was time consuming, resulted in no protocol changes, and increased both the length and complexity of the consent forms.
White & Gamm 2002 (58)	Examine variation in IRB informed consent requirements for research with stored biological samples and use of IRB Guidebook or NBAC recommendations	Mailed survey of 427 IRBs to describe % of consent forms with information about stored samples recommended by 2 guidances	IRBs, reviewing from 1 to 3000 protocols/year, usually reported review of biomedical protocols; average 17% involved stored samples. Recommended provisions more likely in IRBs with higher review volumes, that use 2 sources of guidance, and had MPAs.	IRB practices are difficult to evaluate, and vary greatly, including variation in practices regarding collection, storage, and future use of biological samples.
Jones et al. 1996 (59)	Describe structure and practices of U.S. IRBs	Questionnaire to 447 institutions regarding IRB size and structure, policies, and review of research proposals	IRBs averaged 14 members, representing 27 medical specialties. Protocols were rejected for problems with consent (54%), study design (44%), unacceptable risk (34%), ethical or legal problems (24%), and lack of scientific merit (14%). 86% of IRBs assisted the investigator in responding to its concerns.	Despite variations in committee structure and representation, IRBs have similar procedures for governing research. Investigators should discuss their proposal with an IRB representative prior to review.
Grodin et al. 1986 (18)	Categorize and describe the decision-making of a large urban hospital IRB	Audit of single IRB decisions over 12 years	8% of studies unconditionally approved, 72% approved with conditions, 20% deferred for major revisions.	The IRB's protocol review standards were constant and unaffected by membership changes and other organizational factors.