Published in final edited form as:

J Empir Res Hum Res Ethics. 2008 December ; 3(4): 57–69. doi:10.1525/jer.2008.3.4.57.

Survey of U.S. Human Research Protection Organizations: Workload and Membership

Joseph A. Catania

Oregon State University

Bernard Lo and Leslie E. Wolf

University of California, San Francisco

M. Margaret Dolcini

Oregon State University

Lance M. Pollack and Judith C. Barker

University of California, San Francisco

Stacey Wertlieb and Jeff Henne

The Henne Group

Abstract

In a national survey of institutions with federally assured human research protection programs, we obtained workload and other relevant data on their Institutional Review Boards (IRBs) and management organizations. The number of IRBs increased substantially from 1995 to 2005/06 (491 to 2,728 IRBs) with a further increase in 2008 to 3,853 IRBs. Nationally, IRBs reviewed over a quarter million research applications in the year prior to our survey, of which 35% were new applications requiring full committee review. Compared to estimates from 1995, current IRBs review more new and full committee review applications, but the relative percentage of new and full committee applications remained about the same. High volume research institutions have IRBs with a substantially larger per person workload, relative to smaller volume IRBs (i.e., members spent nearly seven times more hours reviewing new applications outside formal committee meetings). Virtually all IRBs included community representatives as members (92%); however, a small number may not be compliant with federal regulations. The present findings suggest the need for research to (a) examine workload and its effects on review quality, research costs, and faculty morale, (b) develop methods for determining optimal fit between IRB workload demands and institutional labor and

Address correspondence to: Prof. Joseph A. Catania, College of Health and Human Sciences, Oregon State University, 705 NW Elizabeth Drive, Corvallis, OR 97330. Phone: 541-737-3828; Fax: 541-207-3165; E-MAIL: catania1951@comcast.net..

Author Contributions. Drs. Catania, Lo, Wolf, Dolcini, Barker, and Pollack were involved in the design and execution of the project, as well as in the evaluation of the data and writing of the manuscript. Ms. Wertlieb co-directed field work, helped to analyze the data, and helped to conceptualize and develop the manuscript. Mr. Henne was responsible for the data collection; he also participated in the design and execution of the project, and reviewed the final manuscript.

¹Tier One has 120 Institutional Review Board Organizations (IORGs) among 100 institutions, i.e., 29 IORGs are “extras.” Tier Two has 1,950 IORGs (2,070 total IORGs: 120 Tier One IORGs and 1,950 Tier Two IORGs), of which 99% are in single IORG institutions, or . $99 \times 1,950 = 1,930$ single IORG institutions. Of the remaining 20 IORGs in Tier Two, these are all represented by institutions with only 2 IORGs, so $20/2 = 10$ institutions. Thus, Tier One (100) + Tier Two (1,930 + 10) institutions = 2,040 total institutions.

²The following data were used in computing hours worked in reviewing new applications.

Bell Report: 14.9 hrs/review high-volume IRBs, 7.1 hrs/review low-volume IRBs on “initial” reviews; overall mean hours = 11.0 hrs/review for new applications.

Present Study data: M new applications = 87.8, M committee size = 13.4 members. New applications = $(11.0 \text{ hrs/review} \times 87.8 \text{ new applications per IRB})/13.9 \text{ members} = \text{Number of hours spent on review over all IRBs} = 69.5 \text{ hours per person/year on new applications}$. Tier One new applications = $(14.9 \text{ hours} \times 283.7 \text{ new applications})/15.4 \text{ members} = 274.5 \text{ hours/person}$; Tier Two = $(7.1 \text{ hours} \times 68.9 \text{ new applications})/12.4 \text{ members} = 39.5 \text{ hours per person/year}$.

financing requirements, (c) construct benchmarks for judging reasonable workload for individual IRB members, and (d) examine if the need to recruit IRB scientific expertise from outside the institution, particularly true for smaller research institutions, causes delays in IRB review, and if a more effective way of locating and recruiting experts would improve quality and time to completion.

Keywords

IRB Survey; workload; structure

In the United States, Institutional Review Board Organizations may manage one or more Institutional Review Boards (IRBs).¹ Although IRBs are mandated in the United States by federal regulation (45 C.F.R. 46.101 et seq.), there have been few large-scale representative studies of IRBs or their corresponding management organizations since the 1995 “Bell Report” (Bell et al., 1998; also see Gray, Cooke, & Tannenbaum, 1978; De Vries & Forsberg, 2002).² Most sources agree that the number of IRBs in the United States has increased since 1995, although the exact magnitude of this change is uncertain (for example, see Citro et al., 2003). Although the cause of this growth may be attributable to increases in federal research funding, an added consideration is that IRB growth may reflect institutional pressures (for example, pressures to insure IRB review of all research whether federally or non-federally funded; Gunsalus et al., 2007).

It is unclear if the growth in IRBs has kept pace with workload demand. In this regard, research institutions are not required to reasonably support IRBs, and they have little financial incentive to increase the capacity of IRBs or their managing organizations. The federal government provides no direct funding for IRBs or their management, and overhead support in government research grants does not earmark IRB capacity building. Thus, some institutions may have too few IRBs to cover existing workload. Poor funding may also underlie the fact that research institutions rely heavily on volunteers to fill their IRB memberships. As noted below, this situation may impact review quality (Citro et al., 2003; Gunsalus et al., 2007). It is important, therefore, to determine if IRB workload has increased over time and whether such changes have precipitated increased IRB capacity building.

Current data addressing IRB workload issues are at best indirect. Re-analysis of data from Bell et al. (1998) indicate that the average IRB deliberation time on new applications decreased from 1 hour in 1975 to 8 minutes in 1995 (Citro et al., 2003, pp. 36–39). This suggests that by 1995 workload increased significantly while the IRB labor force failed to increase. Together with studies of investigator complaints about the long time lags associated with IRB application reviews (for example, see Gunsalus et al., 2007; Bledsoe et al., 2007), this latter finding raises concerns about workload and review quality. That is, are IRBs so overloaded that they no longer have sufficient time for deliberation and achieving well-reasoned reviews? Such a situation may produce a rush to judgment, particularly with applications based on methodologies unfamiliar to IRB members (Hamburger, 2005; Citro et al., 2003). Heavy

¹In the United States, the Federal Office for Human Research Protections (OHRP) designates the combined system for human research protection of any given research institution as an Institutional Review Board Organization(s) (IORGs). IORGs may represent one or more Ethical Review Committees. In the U.S., Ethical Review Committees are known as Institutional Review Boards. IORGs are also referred to as Human Research Protection Programs (Rubin & Sieber, 2006).

²De Vries and Forsberg (2002) report IRB survey data for a national random list sample ($n = 89$ IRB administrators) stratified by state. Statistical weights to adjust for the complex sample design (stratification scheme) and unequal probability of selection (larger institutions with more IRBs have higher probability of selection) are not reported being utilized, therefore, their national-level findings may not be directly representative of the population. This produces interpretative difficulties. For example, the present study found 85% of institutions with only one IRB, while De Vries and Forsberg report 59% managing only one IRB. This difference may be due to changes occurring from 2001 to 2005/06 or sampling differences related to unintentionally oversampling larger multi-IRB institutions, and not adjusting for the oversample in the De Vries and Forsberg survey.

workloads are also problematic given the reliance on volunteers to fill IRB membership. For small research institutions, IRB time commitment may be reasonable. However, for larger institutions the level of committee time has, anecdotally, been reported as a significant burden for faculty (Citro et al., 2003). The present study examined IRB workload in the U.S. and estimated the amount of time spent reviewing applications. We examined IRB workload and structure for large and smaller volume research institutions. A consistent observation, albeit based on limited data, is that larger institutions have greater workload strain (Citro et al., 2003; pp. 42–43). If institutions create IRBs in relation to demand then per-IRB member workload at larger and smaller volume institutions should not differ substantially.

Methods

SAMPLE

We obtained the 2004 computerized listing of Institutional Review Board Organization(s) (managing organizations) from the U.S. Department of Health and Human Services' Office for Human Research Protections. We drew a stratified random sample from a universe of 2,070 Institutional Review Board Organizations located in the 50 States and the District of Columbia. Although most U.S. research institutions have only one managing organization, a small minority have multiple managing organizations within the same institution. The sample was divided into two stratum or tiers. Tier One consisted of 120 managing organizations representing the top 100 institutions in terms of NIH funding for fiscal year 2002. Tier Two included a random sample of the remaining managing organizations ($N = 274$; 12 pretest/reviewing organizations were excluded). Sixteen organizations were ineligible because they were deactivated. A detailed sampling report is available from the authors. We obtained 244 (cooperation rates: 65% total; 59% Tier One, 67% Tier Two) completed Institutional Review Board Organization–administrator surveys from eligible organizations.

Although our sample frame is based on the 2004 listings of federally assured management organizations, the IRB estimates are based on the ten-month window during which data were collected (the survey was conducted between October 2005 and June 2006). If significant changes occurred during these time frames in the underlying number of managing organizations, a cross-sectional survey would not be able to capture those changes. However, within the managing organizations sampled, we have reasonable estimates of the number of working IRBs at the time.

PROCEDURES

After we obtained informed consent, administrators of the managing offices were interviewed by telephone with a portion of the interview conducted through several optional methods (telephone, mail questionnaire, secure-Internet questionnaire). Following completion, thank-you letters were sent to administrators. Study protocols were approved by IRBs for the University of California, San Francisco and the Henne Group (San Francisco, CA), the data collection subcontractor.

MEASURES

Administrative interviews (complete interview available from authors) obtained general information on the managing office and on the IRB chairs and committees. The interview asked administrators to (a) describe how many other managing organizations were housed within their institution or associated institutions, (b) describe the workload of their managing organization, (c) enumerate the number of IRBs subsumed under their managing organization (d) enumerate for each IRB under their managing organization the types and volume of protocols received in the past year (total number of protocols [new and prior], number of new [all types] and of new full-committee review protocols), and (e) describe the composition of

each IRB committee administered by their office (total members per committee, number of non-institutional members, number of non-institutional members without a science background). Demographic data were not obtained from the administrators because administrators themselves were not the object of investigation; rather the workings of their management organization and IRBs were the primary focus of study.

Workload estimates represent the one-year period prior to the date the interview was conducted. Thus, our survey represents workload for the time period from October 2004 through June 2006. Our workload estimate should be considered then as representative of workload over a 20-month window.

Counts of numbers of applications were measured using bounded categories (0, 1–25, 26–50, 51–75, 76–100, 101–150, 151–200, 201–250, 251–300, 301+). These assessments, therefore, are estimates. Administrators and/or staff did not typically have the resources to produce an actual count. Definitions of a non-institutional IRB member, and a non-institutional non-science IRB member were supplied to administrators during this interview process:

By non-institutional members we mean people who are not connected to the university (or corresponding parent institution) through employment or unpaid adjunct faculty positions. Non-institutional members might be community advocates or members of community based organizations (CBOs). [The administrators should not be included in this number.]

Non-institutional members might be community advocates or members of community based organizations who do not have science backgrounds, meaning they do not have graduate or professional training in basic or applied sciences of any kind.

STRATIFICATION VARIABLES

We constructed two *a priori* workload stratification variables based on research volume and institutional designation. Research volume was indexed by the number of National Institute of Health grants received in fiscal year 2002 (top 100 funded institutions vs. all others), and institutional designation. Institutional designation was determined from administrator interviews and includes three levels (colleges or universities [42%]; non-university affiliated health institutions [41%]; and other types of research institutions: governmental departments, independent research institutes, independent IRBs [16%]).

STATISTICAL METHODS

Sample weights were developed to adjust for unequal probabilities of selection. Details of the sample weight construction are available from the authors. Our sampling strategy dictates that weights are needed only in specific circumstances wherein Tier One and Tier Two samples are combined and we wish to make generalizations to the universe of Institutional Review Board Organization(s) at the national level. Weights are not used when computations are made only within a given stratum, or when the unit of analysis is the IRB instead of the managing organizations. Data were analyzed using SPSS and SVY procedures in STATA release 9 (to adjust standard errors when using weighted data). Estimating the number of applications processed by the managing organizations was accomplished by utilizing three values from each predetermined category (minimum, midpoint, and maximum). The highest frequency category, which is open-ended (301+ applications), was truncated at 350 for estimating purposes. It should be noted, therefore, that the maximum frequency values might be somewhat larger than presented. The application frequency categories were constructed based on our pilot interviews. Midpoints of the categories provide a moderately conservative estimate and are used in all computations unless indicated otherwise. All other counts were assessed as simple frequency counts.

LIMITATIONS

The present study has several limitations. We can not estimate the effects of nonresponse on our estimates. For instance, though the population is relatively homogenous, it is possible that more heavily burdened administrators did not respond to the survey, and, therefore, our workload estimates might be viewed as lower bounds. Moreover, our work estimates do not include all types of IRB-related work (for example, reviewing violations and reports of harm). This would also lead to underestimating workload. In addition, the workload data are self-reported and may contain sources of error (for example, when relying on memory rather than actual records). Comparing the present study's findings to Bell et al. (1998) are limited because the two studies differ in a number of ways including: (a) the base populations being different between the studies (for example, Bell et al. excluded IRBs reviewing <10 applications per year, and IRBs in 1995 were likely over-represented by universities and top National Institute of Health-funded institutions), and (b) definitional differences. These differences prevent comparison of the total IRB workload between studies.

Results

INSTITUTIONS WITH INSTITUTIONAL REVIEW BOARD ORGANIZATION(S): STRUCTURE AND NUMBER

The number of Institutional Review Board Organizations (managing organizations) in the United States is somewhat in flux. From the 2004 Office for Human Research Protections listings, we found that approximately 4.1% no longer functioned as of 2005–2006 because the institution no longer existed or was closed (for example, due to a natural disaster or no longer maintaining an IRB). Fewer Tier One institutions (top 100 funded institutions) had closed down (0.8%) compared to Tier Two institutions (4.7%). Although most institutions in the United States have only one management organization, approximately 5.4% have multiple managing organizations. Among high-volume (Tier One) institutions, approximately 15% have multiple management organizations (<1% of Tier Two have multiple managing organizations) (see Table 1). These institutions are primarily universities with major medical research centers and/or schools of public health. Among institutions with multiple managing organizations, those with two or three management organizations represent the typical structure, and only a rare few have four or more management organizations.

We computed the number of U.S. institutions with managing organizations using unweighted within-stratum data. The within-stratum data were generalized to all managing organizations in the United States to produce stratum-specific population-level estimates. These estimates were then aggregated across strata to give total population-level results. By working back from the number of managing organizations (2,070) to the number of institutions, we estimate that there were approximately 2,031 institutions with managing organizations (i.e., Institutional Review Board Organizations, see Research Note 1) in the United States for the time period represented.

ESTIMATING THE NUMBER OF IRB COMMITTEES IN THE U.S.

We identified 400 IRBs among the 244 managing organizations in our sample. We estimate that nationally there were 2,728 IRBs among the 2,070 managing organizations identified in the Office for Human Research Protections 2004 listings. Approximately 85% (weighted) of the sampled organizations contained a single IRB. The distribution of IRBs per managing organization is relatively narrow, although one administrator reported managing 10 IRBs. Only 4% of the managing organizations reported four or more IRBs, 4% reported three IRBs, and 7% reported two IRBs. Table 2 provides the distribution of IRBs (single vs. multiple) by institutional strata. Multiple IRB institutions are more frequent among Tier One institutions (73%) and universities (33%).

CHANGES IN IORGS AND IRBS 2004–2008

The Office for Human Research Protections has recently made available online data on Institutional Review Board Organization(s) and IRBs. These data are presented for each state and the U.S. Territories. The number of IRBs and their managing organizations [Institutional Review Board Organization(s)] are provided on-site in addition to information on location and assurance numbers. It is unclear if IRBs that have been deactivated are included in these counts, but this would have only a small impact on the overall count. The 2008 data indicate that for the 50 states and the District of Columbia, the geographic base for the 2004 data, the number of managing organizations has grown to approximately 3,103, an increase of 1,033 organizations, with an associated increase of 1,125 IRBs (41% increase).

INSTITUTIONAL AND NON-INSTITUTIONAL IRB MEMBERS

In the United States, IRBs are composed chiefly of members derived from the parent institution, typically faculty, one of whom is the IRB Chair. In addition, IRBs are required to include membership from outside the institution that in some manner represents the broader community of research participants. Further, IRBs may recruit scientific expertise from outside the institution to supplement key areas. We examined details of IRB composition among the 400 IRBs enumerated by the sample of management organization administrators (see Table 3). In general, the vast majority of IRBs reach outside the institution for additional, albeit few, members. We identified <1% of IRBs ($N = 3$) who had not included non-institutional members (science and community) and only 8% ($N = 32$) who had not included external community membership without scientific training.

We computed the average number of total and non-institutional IRB members by institutional strata (Table 3). Tier Two (and “Other”) institutions have significantly smaller total IRB memberships, reflecting perhaps lighter workloads and/or smaller institutional populations from which to draw members. In this regard, Tier Two (and non-university health/“other”) IRBs have significantly more non-institutional members. Thus, smaller/lower volume institutional strata rely to a greater extent on individuals outside the institution to fill IRB membership.

Given a smaller corpus of institutional scientists to draw from, we might expect that smaller research institutions would draw on external sources to recruit *ad hoc* or regular scientific members. Our data suggest this is the case. First, Tier Two IRBs have more non-institutional members overall than Tier One IRBs. Among the non-institutional IRB membership, Tier Two IRBs and Tier One IRBs are similar in their representation of non-institutional community representatives without scientific training. Therefore, Tier Two IRBs must be drawing on more external (non-institutional) scientific expertise than Tier One IRBs. The “other” institutional IRBs have a pattern similar to Tier Two IRBs, suggesting that they also recruit more external scientific expertise to fill committee seats. Non-university health-related institutions (for example, free-standing healthcare corporations) have IRBs with a slightly different pattern. They have more non-institutional members overall and more non-institutional community members without scientific training than their counterparts on IRBs for universities. That is, they are recruiting more community representatives than scientists from outside the institution.

WORKLOAD VOLUME IN THE U.S. (RESEARCH APPLICATIONS REVIEWED)

We computed the total number of research applications processed by the sampled managing organizations at each institution in the past year with breakdowns by the number of new applications and new applications requiring full committee review (Tables 4–6). As noted previously, we assessed the number of applications received as categorical frequencies. The category boundaries can be understood to represent low- and high-end estimates of the actual

frequencies. Midpoints of the categories are used in all computations unless indicated otherwise.

At the national level, we estimated that a total of 269,740 IRB applications were processed by all managing organizations in the past year. Among these, approximately 181,669 applications were new and 95,702 received a full IRB committee review. This latter figure represents over a third (35%) of the total workload. Tier One institutions (and universities) carry larger workloads (see Tables 4-6). For example, at the national level, Tier One institutions, which represent 4.9% of federally assured managing organizations in the United States, reviewed approximately 20% of all applications (54,394/269,740) and 40% (38,604/95,702) of all new applications.

Inspection of Tables 4-6 indicates that the average number of applications received by Tier One institutions is nearly three times that of Tier Two institutions. Further, Tier One managing organizations, compared to Tier Two organizations, receive nearly four times the average number of new applications and approximately six times the average number of new full-committee review applications. Since Tier One institutions have more IRBs per institution, we adjusted the workload estimates accordingly. Adjusting for the number of IRBs, Tier One averages 101 total applications/IRB/year vs. 103 applications/IRB/year for Tier Two (applications/IRB: Tier One 21,101/209; Tier Two 19,592/191). Thus, overall, Tier One and Tier Two look similar at the level of total average applications per IRB. However, there are substantial differences in the more labor-intensive new applications, particularly those requiring full IRB committee review. Tier One averages 91 new applications/IRB/year vs. 63 new applications/IRB/year for Tier Two (applications/IRB: Tier One 19,003/209; Tier Two 11,981/191), and, respectively, 63 vs. 29 new full-committee review applications/IRB/year (applications/IRB: Tier One 13,209/209; Tier Two 5,529/191). Similar relative differences were found in comparing universities to other types of research institutions.

The disproportionate workload between Tier One and Tier Two institutions (or universities and other institutions) may be a function of Tier One institutions receiving more protocols dealing with sensitive issues or vulnerable populations, particularly medical and mental health-related research protocols (see accompanying article). That is, proportionately more protocols may proceed to full-committee review in Tier One institutions.

As an overall assessment of workload and IRB structure, we computed correlations between the number of IRBs within managing organizations and the number of applications processed (total, new, full committee review). The results indicate moderate associations between workload and the number of IRBs (total applications $r = .41$, new applications $r = .51$, and full committee review applications $r = .69$, all p -values $< .0001$). That is, a strong relationship between workload and committee number would be evidenced by correlations of .90 and larger. Stated differently, as in the case of new applications, only 26% of the variation in total workload is accounted for by the number of IRB committees devoted to the task.

WORKLOAD COMPARISONS TO 1995

We examined change by comparing our findings to those of Bell et al. (1998).³ There are many differences between the two studies (see discussion), but there are also points of comparison. We determined how many applications were “new rather than modifications or renewals,” which encompasses the definition in Bell et al. of “initial” reviews. Both studies also determined the number of applications that required full committee reviews.

³De Vries and Forsberg (2002) report workload data for IRB managing organization(s) in terms of “current active studies” that does not allow comparisons with the present study particularly at the level of protocols needing full committee review. Nevertheless, De Vries and Forsberg (2002) found, as we did, that managing offices of universities and medical schools carried the heaviest workload.

In terms of new and full committee reviews, the universe of federally assured IRBs in 1995 received approximately 105,000 new applications and conducted an estimated 60,900 full-committee reviews (58% of new applications). By 2005, we estimate that, nationally, IRBs received 181,669 new applications and conducted 95,702 full committee reviews (53% of new applications). Thus federally assured IRBs in 2005 reviewed more new and full review applications in total volume than IRBs in 1995, but the relative percentage of new applications needing full review was similar over time.

Changes in total workload may be a function of there being a larger universe of federally assured IRBs in 2005 as compared to 1995. Changes in the size and composition of the base population over time present barriers to trends analyses unless one can adjust for these differences. A typical solution of computing the average workload per IRB is flawed in this case because the 491 federally assured IRBs in 1995 are over-represented by universities and top National Institute of Health-funded institutions (among the first entities developing federally assured administrations). We had insufficient information to adjust differences between studies at this level. Making comparisons without these adjustments would result in artificially divergent findings. For instance, in Bell et al., the 491 IRBs averaged 214 new applications/IRB, while in the present study the average number of new applications is approximately 88/IRB. This difference between studies in average work volume is likely due to differences in research volume between sample frames. Indeed, university IRBs in our study average 230 new applications/year, a value similar to the Bell et al. average for their total sample frame. Adjusting for diverging population and definitional differences between studies and, consequently, the observed differences in workload is necessary, but beyond the scope of the current project.

Despite limitations, Bell et al. provided data that may be combined with current estimates to generate more detailed workload indices. In particular, Bell et al. provided data on the number of hours spent reviewing applications outside formal committee meetings. On average, they found that IRB members reviewing new applications spent approximately 11.0 hours/review outside the IRB meetings. Assuming that time spent per review outside the IRB has remained approximately the same over time (as noted previously it has changed within the IRB), we can estimate this portion of current workload in hours (see Research Note 2). The average number of new applications per IRB in our study is estimated at 87.8/IRB with an average committee size of 13.9 members. Assuming the workload is distributed evenly, we estimate that each member is spending a total of 69.5 hours/year outside IRB meetings on new applications alone. For Tier One IRBs, this figure is substantially higher at 274.5 hours/year spent reviewing new applications outside the committee proper by each member, while for Tier Two IRBs, the figure is lower at 39.5 hours/year.

Discussion

OVERVIEW

The National Research Council (Citro et al., 2003) and Gonsalus et al. (2007) conclude that there are insufficient data on the functioning and effectiveness of the United States' IRB system (also see Institute of Medicine, 2003, pp. 164–165). Moreover, these critiques suggest that the large volume of protocols being reviewed and institutional pressures to regulate federal and non-federal research at review levels higher than those required under the regulations are significant problems that may affect quality of ethical review. In addition, it is unclear if IRB committees have the requisite scientific expertise needed to provide high-quality reviews (Citro et al., 2003; Gonsalus et al., 2007; Hamburger, 2005). The National Research Council report notes that national level data on workload and IRB membership need updating (also see Institute of Medicine, 2003, pp. 164–165). The present study reports findings on IRB workload and membership. This is the first study since 1995 (Bell et al., 1998), to our knowledge, to

obtain a representative national sample of the managing offices that oversee IRBs and correspondingly of their respective IRB committees. Although Campbell et al. (2003) reported national IRB data, their survey is generalizable to medical school IRBs, not the U.S. as a whole. We generated comparative analyses with Bell et al. (1998) that address the changing landscape of IRBs in the United States from 1995–2005.

NATIONAL ESTIMATES

There were approximately 2,040 institutions with Institutional Review Board Organizations (i.e., managing organizations) in the United States in 2004 constituting 2,070 managing organizations. The number of IRBs with federal assurances increased from 491 in 1995 (Bell et al., 1998) to an estimated 2,728 in the 2005/06 survey period.

Data from the Office of Human Research Protection for 2008 suggest there continued to be a substantial increase in federally assured managing organizations and IRBs from 2004 to 2008. This increase may be associated with the Food and Drug Administration's proposed requirements for research under their review to be approved by registered IRBs. Anticipating the proposed directive, many IRBs, particularly those reviewing primarily for the pharmaceutical industry, may have sought OHRP assurances. In addition, since 2004, the Department of Health and Human Services has deliberated on the implementation of a proposed rule, “making the IRB registration system uniform with the proposed IRB registration requirements of the Food and Drug Administration, and creating a single Health and Human Services IRB registration system (CFR Citation: 21 CFR 56.106).” These deliberations may have also spurred unregistered IRBs to make additional effort to register in advance of what they perceived to be an inevitable federal requirement.

Managing organizations in the United States processed approximately 269,740 applications in the year prior to our survey, including approximately 181,669 new applications with 95,702 applications receiving labor-intensive full committee reviews. In general, Tier One research institutions and/or universities account for the majority of institutions with the heaviest workloads and, correspondingly, the most likely to have multiple managing organizations and multiple IRBs within a managing organization. Correlations between workload and number of IRBs available to handle the workload were moderately large. Thus, there appears to be some consistency at the macro-level between organizational structure and potential workload. However, at the level of individual IRBs, we found workload to show substantial differences between smaller and larger research institutions (see below).

Moreover, although the data imply that heavier workload is associated with multiple managing organizations within an institution, the rationale underlying the development of institutions with multiple managing organizations may not be stimulated only by demand. For instance, our screening survey, wherein we queried administrators on the managing organization's structure within their institution, indicated that some larger institutions have managing structures that grew out of a network of institutional associations developed through acquisition or affiliation. These networks may have multiple managing organizations, which may be under the administration of the primary institution, but in some instances are actually a confederation of managing organizations (and IRBs) that grew with the network and may reflect “turf” demands rather than workload requirements. How these different structural arrangements function in terms of efficiency is a relevant question beyond the scope of the present study.

IRB MEMBERSHIP

Including community representation on IRBs is an important goal and a regulatory requirement [45 CFR 46.107(d)]. Our data suggest that virtually all IRBs include community representatives as members. Nevertheless, there may be a small number of IRBs that are not in compliance.

However, one must acknowledge that the IORG administrators reporting membership data may have been in error. Nevertheless, external IRB monitoring may be required. Although IRB compliance monitoring has begun through the Association for the Accreditation of Human Research Protection Programs, Inc. (AAIORG), this effort is relatively new and voluntary, except for Veterans' Affairs IRBs. To date, AAIORG has accredited only 84 managing organizations with another 400 reported to be in the review pipeline.

Non-university health institutions report having greater community representation on their IRBs. It may be relevant to determine if IRBs with greater community representation function differently than those with less representation. Higher levels of community representation may provide for a more supportive environment for community representatives to speak out (Sengupta & Lo, 2003). However, the ability of community representatives to effectively portray the diverse interests of their constituencies is a significant question that past studies have not addressed.

Our findings suggest that IRBs from smaller institutions more often recruit from outside their institutions for scientific expertise than do larger institutions. How this portion of the system operates is an important question. The need to delay reviews in order to recruit necessary expertise may contribute to investigator reported problems with long IRB review processes that increase research costs.

WORKLOAD

Heavy workload may decrease review quality, increase research costs and negatively influence the willingness of faculty to participate on IRBs (see, for example, Citro et al., 2003). Direct links to these outcomes are, however, lacking or anecdotal. Moreover, research on IRB workload is in its infancy. For instance, a formula for determining an optimal fit between work demand and labor at the level of individual committee members has not been constructed in this arena (see Wagner and Barnett, 2000). Moreover, we do not have exact benchmarks for judging what is a reasonable workload for individual IRB members under the varying circumstances of faculty workload distribution across teaching, research, committee, and consulting functions. Although most faculty are expected to contribute some time to internal and external committee work (service), the amount of time considered reasonable is unclear.

The current study examined the relative workloads among different types of institutions and estimated changes in workload over time by comparing our data to Bell et al.'s (1998) 1995 IRB survey. On a rational basis, increased work demand should result in an increase in the number of managing organizations, IRBs, and, to some extent, IRB membership. Thus institutions reviewing a higher volume of applications should have more managing organizations, more IRBs, and larger IRBs. Nearly three-fourth of the highest volume institutions have multiple IRBs and/or managing organizations. The correlations between workload and number of IRBs available to conduct the work are moderately large and in a positive direction suggesting that there is an effort at some institutions to balance workload and labor requirements. However, there is considerable variation from institution to institution in the number of IRBs devoted to the same workload levels. This suggests a lack of standardization across institutions in the work force to workload ratio.

Wagner and Barnett (2000) constructed economic and optimal committee number models for hypothetical medium (370 total applications per year) and high volume (1,380 total applications per year) workloads. They concluded that high-volume workloads would require at least 4 IRBs compared to 2 IRBs for institutions with medium-sized workloads. Based on these estimates, the average high-volume research institution should have at least 2 IRBs. Further, we estimate that approximately 30% of high research volume institutions ought to increase their IRBs by at least one or more additional committees.

On average, high-volume research institutions (Tier One, universities) have significantly larger IRB committee memberships than lower-volume institutions. However, the difference in average per member workload between larger and smaller institutions is substantial. For Tier One institutions, the average time expended outside committee meetings on reviewing new applications is significant, representing approximately seven 40-hour work weeks per year (14% FTE/12 months). For Tier One members, the average time spent reviewing new applications outside IRB meetings is nearly seven times of that for Tier Two institutions. These estimates are likely lower bounds as they do not take into account other duties performed outside regular meetings by committee members. For Tier One institutions, the workload seems consistent with the requirements of a part-time job layered on top of regular institutional duties. The effects of this heavy committee workload on review quality and faculty morale warrant investigation (Wagner and Barnett, 2000).

Summary

We found general consistency at the macro-level between organizational structure and potential workload. Most institutions with the heavy workloads (Tier One and universities) have larger managing organizational and IRB structures. High-volume IRBs, however, confront a substantially larger per person workload relative to smaller-volume IRBs, suggesting that the IRB labor force has not kept pace with work demands at high-volume institutions. Comparisons with prior research suggest that IRBs currently review more new and full-review applications. In terms of IRB membership, virtually all IRBs included community representatives as members, although a small number may not be in compliance with federal requirements for community representation. Lastly, IRBs from smaller institutions more often recruit from outside their institutions for scientific expertise.

Best Practices and Suggestions

Evidence of rapid change and some less-than-optimal committee structures suggest several best practices for alleviating the problems observed. In particular, research administrators should give careful consideration to the impact of IRB workload, particularly at high-volume research institutions, on work stress and review quality. In particular, they should:

- Compensate IRB members adequately for high workloads. Options for doing so include increased compensation in the form of additional salary, adjusting other workloads to compensate for the heavier IRB workload, and greater weighting of IRB service in the promotion process.
- Increase the number of committee members to reduce workload. Approximately a third of high-volume institutions may need to increase the number of IRBs by one or more committees to reach optimal efficiency.
- Assess any other problems of IRBs that appear to be having difficulties, including examination of the problems of recruiting community members.

Research Agenda

- Compare Office of Human Research Protection 2004 data on IRB managing organizations and our survey results with updates from the Office of Human Research Protection for 2008. This comparison suggests a substantial increase in federally assured managing organizations and IRBs. We recommend additional surveys be conducted to evaluate variations in existing management structure and function (for example, staffing, organization, management philosophy, operating procedures, decision making), budget support and needs, work parameters (workload, work strain), and IRB committee structure and function (membership, review quality,

efficiency) across this enlarged universe of IRBs and related management organizations (for example, see IOM, 2003, pp. 164–165).

- Future studies should also examine workload estimates that include other types of work (for example, record keeping, education, informal consultation, dealing with violations of regulations) performed by IRB committees or management office administrators/staff.
- Plans by the Office of Human Research Protections to collect IRB workload data should be reviewed with an eye towards validating those indices in terms of optimal staffing, financing, and workload. Studies examining workload and review quality are necessary, however, to validate such indices. In addition, the ethical review environment is changing rapidly. As recommended elsewhere (IOM, 2003, pp. 164–165), consideration should be given to conducting regular national surveys (for example, at five-year intervals) to assess managing organizational and IRB administrative and committee activity, and review performance.
- We recommend conducting outcome and modeling studies to evaluate the impact on productivity (for example, time to process), review quality (for example, consistency with regulations and use of regulatory flexibility), and research costs (for example, related to delays in obtaining approvals) of management/IRB budgetary support and needs, organizational structure, philosophy, and function, and staffing/membership expertise. Options to reduce the negative impact of workload should be developed and evaluated.
- Models should be developed and evaluated for determining optimal fit between IRB workload, budget support, and labor-force size for a wide array of IRB environments. Further, we recommend development of benchmarks for judging what is a reasonable workload and compensation for individual IRB members. As noted, indices might be developed to monitor quality improvement in these areas.
- Although there is some consistency between organizational structure and workload among high-volume institutions, further research is needed to examine how these types of institutions are functioning, particularly since they may generate more scientifically diverse applications that place higher demand on IRB review committees.
- IRB committees of low-volume institutions deserve careful investigation for the reason that their IRB membership may not have the expertise needed for some types of applications (see accompanying paper in this journal). Research is recommended to examine if the practice of recruiting scientific expertise from outside the institution, a more common practice among smaller research institutions, affects review quality and research costs, and, if so, how this problem might be resolved.

Educational Implications

Investigators, administrators, and IRBs should consider the significant implications of workloads on review quality, research costs, and employee stress, and develop appropriate educational interventions to reduce these problems.

Acknowledgments

This work was supported by NIMH grant RO1-MH064696 to Dr. Catania and Dr. Lo. We wish to thank Caroline Fisher for manuscript preparation and Michael O'Grady for interview work.

Biographies

Joseph A. Catania, Ph.D., is a developmental psychologist and Professor of Public Health at Oregon State University, specializing in research on sensitive topics and vulnerable populations, and research methods including the impact of subject protection requirements on survey quality.

Bernard Lo, M.D., is Professor of Medicine and Director of the Program in Medical Ethics at University of California, San Francisco, and is interested in ethical issues in clinical research.

Leslie Wolf, J.D., is Associate Professor in the School of Law at Georgia State University, where she works primarily on issues in research ethics and the law.

M. Margaret Dolcini, Ph.D., is Associate Professor of Public Health at the University of Oregon and conducts research on health problems of adolescents and ethical issues in research with adolescent populations.

Lance Pollack, Ph.D., is a research specialist and co-director of the cancer center's behavioral methods core at the University of California, San Francisco.

Judith C. Barker, Ph.D., is Professor in Medical Anthropology at the University of California, San Francisco, and examines cultural issues in health.

Stacey Wertlieb, M.P.H., served as a research associate on this project at the University of California, San Francisco.

Jeff Henne heads the Henne Group, a medical and social research company.

References

- American Association of University Professors. Research on human subjects: Academic freedom and the institutional review board. 2006 Retrieved November 1, 2006 from <http://aaup.org/AAUP/comm/rep/A/humansubs.htm>
- Bell, J.; Whiton, J.; Connelly, S. Evaluation of NIH implementation of section 491 of the Public Health Service Act, mandating a program of protection for research subjects. U.S. Department of Health and Human Services; Washington, D.C.: 1998.
- Bledsoe CH, Sherin B, Galinsky AG, Headley NM, Heimer CA, Kjeldgaard E, et al. Regulating creativity: Research and survival in the IRB iron cage. *Northwestern University Law Review* 2007;101(2):593–641.
- Campbell E, Weissman J, Clarridge B, Yucel R, Causino N, Blumenthal D. Characteristics of medical school faculty members serving on institutional review boards: Results of a national survey. *Academy of Medicine* 2003;78:831–836.
- Citro, CF.; Ilgen, DR.; Marrett, CB. Protecting participants and facilitating social and behavioral sciences research. National Academies Press; Washington, D.C.: 2003.
- De Vries R, Forsberg C. What do IRBs look like? What kind of support do they receive? *Accountability in Research* 2002;9:199–206. [PubMed: 12816128]
- Gray B, Cooke R, Tannenbaum A. Research involving human subjects. *Science* 1978;201:1094–1101. [PubMed: 356268]
- Gunsalus C, Bruner E, Burbules N, Dash L, Finkin M, Goldberg J, et al. The Illinois White Paper: Improving the system for protecting human subjects: Counteracting IRB “mission creep.”. *Qualitative Inquiry* 2007;13(5):617–649.
- Hamburger, P. The new censorship: Institutional review boards (University of Chicago, Public Law Working Paper No. 95). University of Chicago; Chicago, IL: 2005.

- Institute of Medicine. Responsible research: A systems approach to protecting research participants. National Academies Press; Washington, D.C.: 2003.
- Rubin P, Sieber JE. Empirical research on IRBs and methodologies usually associated with minimal risk. *Journal of Empirical Research on Human Research Ethics* 2006;1(4):1–4.
- Sengupta S, Lo B. The roles and experiences of nonaffiliated and non-scientist members of institutional review boards. *Academic Medicine* 2003;78(2):212–218. [PubMed: 12584103]
- Wagner, T.; Barnett, P. Human research compliance programs: Optimal operating costs in VA (VA Health Economics Resource Center Technical Report #2). U.S. Department of Veterans Affairs; Washington, D.C.: 2000.

TABLE 1
Number of IRB Management Offices (IMOs) per Institution by Institutional Strata.

	IMOs Among Institutions with Multiple IMOs					
	IMOs		2 IMOs		3 IMOs	
	1 IMO	2+ IMOs	Wt	(Uwt, N)	Wt	(Uwt, N)
<i>Sampling Stratum</i>	Wt	(Uwt, N)	Wt	(Uwt, N)	Wt	(Uwt, N)
Tier One	85%	(81%, 52)	15%	(19%, 12)	10%	(11%, 7)
Tier Two	99.5%	(99%, 174)	0.5%	(0.6%, 1)	0.5%	(0.6%, 1)
<i>Type of Institution</i>						
University	95%	(88%, 86)	5%	(12%, 12)	4%	(8%, 8)
Non-university	100%	(99%, 100)	0%	(1%, 1)	0%	(4%, 4)
Health						
Other	100%	(100%, 40)	0%		0%	

Note: Percentages add across rows. Ns are unweighted values. Tier One = top 100 NIH-funded institutions; Tier Two = all other institutions; Wt = weighted estimate; Uwt = unweighted estimate.

TABLE 2

Number of IRBs by Institutional Strata (N = 244 IMOs).

	Number of IRBs per IMO	
	% with 1 IRB Wt (Uwt, N)	% with 2+ IRBs Wt (Uwt, N)
<i>Sampling Stratum</i>		
Tier One	27% (30%, 21)	73% (70%, 48)
Tier Two	91% (91%, 159)	9% (9%, 16)
<i>Type of Institution</i>		
University	67% (51%, 52)	33% (49%, 51)
Non-university Health	93% (89%, 90)	7% (11%, 11)
Other	96% (95%, 38)	4% (5%, 2)

Note: Percentages add across rows. Sample Ns are unweighted values. Tier One = top 100 NIH-funded institutions; Tier Two = all other institutions; Wt = weighted estimate; Uwt = unweighted estimate.

TABLE 3

Mean Number of IRB Members by Institutional Strata.

	Total Members M (SD)	Non-institutional Members M (SD)	Non-institutional Non-science Members M (SD)
<i>Stratum</i>			
Tier One	15.4 (6.5) ****	2.6 (2.0) **	1.9 (1.3)
Tier Two	12.4 (6.3)	3.1 (2.2)	1.8 (1.4)
<i>Type of Institution</i>			
University	14.1 (6.3) **	2.3 (1.2) ****	1.6 (1.1) ****
Non-university Health	14.5 (6.8)	3.5 (2.0)	2.3 (1.4)
Other	10.9 (6.1)	3.9 (4.1)	1.8 (1.8)

Note: Significance tests based on *t*-tests or *F*-test

*
 $p < .05$

**
.01

.001

.0001.

TABLE 4
Total Applications Submitted to IRB Management Offices in Past Year.

Unweighted Sample Data	Number of Applications			Mean Number of Applications		
	Low*	Mid*	High*	Low	Mid	High
Total Sample	36,193	40,693	44,725	148.3	166.8	183.3
Type of Institution						
University	24,103	26,371	28,725	234.0	256.0	278.9
Non-university Health	9,900	11,544	13,225	98.0	112.1	128.4
Other	2,190	2,778	3,375	54.8	69.5	84.4
Stratum						
Tier One	19,469	21,101	22,800	282.2	305.8	330.4
Tier Two	16,724	19,592	22,525	95.6	111.9	128.7
Population Numbers From Weighted Data						
Total Population						
Type of Institution						
University	234,417	269,740	305,946	113.2	130.3	147.8
Non-university Health	125,287	138,604	152,392	195.9	216.7	238.3
Other	89,476	105,733	122,334	87.1	103.0	119.1
Stratum						
Tier One	19,941	25,695	31,525	49.4	63.7	78.1
Tier Two	50,220	54,394	58,738	282.1	305.5	329.9
	182,186	213,177	244,869	96.3	112.7	129.4

Note: For the 301+ group, low = 301, midpoint = 325, and high = 350.

* Low, Mid, and High refer to estimates based on the range represented by aggregating across that end of the response categories with Mid being the average of the end points.

TABLE 5

New Applications Submitted to IRB Management Offices in Past Year.

Unweighted Sample Data	Number of Applications			Mean Number of Applications		
	Low*	Mid*	High*	Low	Mid	High
Total Sample	26,892	30,984	35,175	1107	127.5	144.8
Type of Institution						
University	21,302	23,462	25,700	208.8	230.0	252.0
Non-university Health	4,150	5,542	6,950	41.1	54.9	68.8
Other	1,440	1,980	2,525	36.0	49.5	63.1
Stratum						
Tier One	17,443	19,003	20,625	256.5	279.5	303.3
Tier Two	9,449	11,981	14,550	54.0	68.5	83.1
Population Numbers From Weighted Data						
Total Population						
Type of Institution						
University	150,222	181,669	213,676	72.6	87.8	103.2
Non-university Health	104,643	117,024	129,797	163.6	183.0	202.9
Other	31,509	45,203	59,017	30.6	43.8	57.2
Stratum						
Tier One	12,831	18,226	23,571	32.1	45.6	59.0
Tier Two	46,428	50,502	54,737	260.8	283.7	307.5
	103,054	130,367	158,075	54.5	68.9	83.5

Note: For the 301+ group, low = 301, midpoint = 325, and high = 350.

* Low, Mid, and High refer to estimates based on the range represented by aggregating across that end of the response categories with Mid being the average of the end points.

TABLE 6

New Applications Submitted to IRB Management Offices Receiving Full-Committee Review in Past Year.

Unweighted Sample Data	Number of Applications			Mean Number of Applications		
	Low [*]	Mid [*]	High [*]	Low	Mid	High
Total Sample	15,162	18,738	22,375	63.2	78.1	93.2
Type of Institution						
University	11,802	13,614	15,475	115.7	133.5	151.7
Non-university Health	2,722	4,006	5,300	27.8	40.9	54.1
Other	638	1,118	1,600	15.9	27.9	40.0
Stratum						
Tier One	11,793	13,209	14,675	173.4	194.2	215.8
Tier Two	3,369	5,529	7,700	19.6	32.1	44.8
Population Numbers From Weighted Data	Number of Applications			Mean Number of Applications		
	Low	Mid	High	Low	Mid	High
Total Population	68,180	95,702	123,476	32.9	46.2	59.7
Type of Institution						
University	44,637	54,621	64,787	68.7	84.0	99.7
Non-university Health	20,902	33,762	46,681	20.4	33.0	45.6
Other	4,077	8,746	13,424	10.3	22.0	33.8
Stratum						
Tier One	34,560	38,604	42,797	181.5	202.7	224.7
Tier Two	36,063	59,700	83,452	19.2	31.8	44.4

Note: For the 301+ group, low = 301, midpoint = 325, and high = 350.

* Low, Mid, and High refer to estimates based on the range represented by aggregating across that end of the response categories with Mid being the average of the end points.