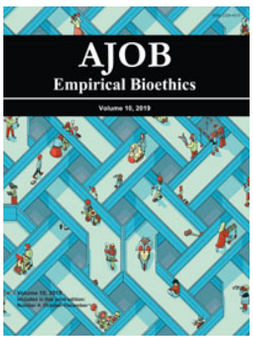

AJOB Empirical Bioethics

ISSN: 2329-4515 (Print) 2329-4523 (Online) Journal homepage: https://www.tandfonline.com/loi/uabr21

Operational Characteristics of Institutional Review Boards (IRBs) in the United States

Genevieve L. Nesom, Iraklis Petrof & Tyler M. Moore

To cite this article: Genevieve L. Nesom, Iraklis Petrof & Tyler M. Moore (2019) Operational Characteristics of Institutional Review Boards (IRBs) in the United States, AJOB Empirical Bioethics, 10:4, 276-286, DOI: 10.1080/23294515.2019.1670276

To link to this article: https://doi.org/10.1080/23294515.2019.1670276

AJOB EMPIRICAL BIOETHICS
2019, VOL. 10, NO. 4, 276–286
https://doi.org/10.1080/23294515.2019.1670276

Taylor & Francis
Taylor & Francis Group

Operational Characteristics of Institutional Review Boards (IRBs) in the United States

Genevieve L. Nesom[a], Iraklis Petrof[b], and Tyler M. Moore[c]

[a]Metabolic Disorders, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA; [b]Gastroenterology, The Children's Hospital of Philadelphia, Philadelphia, Pennsylvania, USA; [c]Psychiatry, Hospital of the University of Pennsylvania, Philadelphia, Pennsylvania, USA

ABSTRACT

Background: Federal Law requires approval from an Institutional Review Board prior to conducting human subjects research to ensure ethical distribution of benefits and harms. Notwithstanding this role and almost no prescriptive requirements about design or operation, there is little systematic research describing the key attributes of IRBs, as reported by IRB personnel themselves. **Methods:** Here, 55 IRB directors completed a survey of 77 questions. The goals of the study were to establish what a typical US IRB "looks like," determine whether IRB characteristics can be summarized by a smaller number of overarching components, determine the best predictors of IRB speed and efficiency, and determine whether IRBs differ by high-level qualitative characteristics such as institution type. The above was explored and tested using the general linear model and principal components analysis, and for the former, dependent variables of interest were, a) the time necessary for an IRB to approve a study, and b) efficiency of the review process for full board and expedited reviews. IVs of interest included multiple IRB characteristics. **Results:** 1) IRB characteristics can be summarized by four key components; 2) IRB speed and efficiency are most strongly determined by tendency to receive biomedical submissions, especially drug-related; and 3) IRBs do vary by institution type on some key variables. **Conclusion:** These results are the first step toward establishing national norms and building a working model of US IRBs to which other IRBs can compare themselves.

KEYWORDS
Institutional Review Boards; bioethics; clinical research; factor analysis; biomedical research

Research is essential to public health and medical practice. Research is also, however, a source of potential harm. Federal law requires prior approval from an Institutional Review Board (IRB) for most human subjects' research to ensure an ethical distribution of potential benefits and harms (cite 45 CFR 46 here). Reviews that are insufficiently rigorous risk harm to participants; reviews that are too restrictive may deter or impede the speed of research. Getting the balance right is difficult, contested, and, as of recently, the subject of considerable reform (Department of Health and Human Services 2011). Notwithstanding the importance of IRBs to health innovation and promotion, there is relatively little systematic research describing their key attributes.

Recent changes to the Common Rule were fueled in part by the burden, delay, and ambiguity felt by research investigators when navigating the IRB submission and approval process (45 CFR 46) (Department of Health and Human Services 2011). Despite the proposed changes, little evidence is available to quantify the actual duration of delays or the context in which they occur. Similarly, revolutionary ideas such as that of a "real-time IRB," where investigators attend IRB meetings and implement proposed changes in real time (Spellecy et al. 2018), have shown to be effective at reducing review times. These methods however are resource intensive and cannot be implemented universally.

Data quantifying IRB delays are notoriously difficult to obtain (Millum and Menikoff 2010; Silberman and Kahn 2011; Abbott and Grady 2011). Available studies are based on study team member retrospective reports, and are often local or type specific (e.g. only pediatric hospital IRBs) (Hirshon et al. 2002; Stark et al. 2010; Abbott et al. 2013; Khan et al. 2014; Varley et al. 2016; Desai et al. 2017). Systematic reviews and meta-analyses evaluating IRBs find high

CONTACT Tyler M. Moore tymoore@upenn.edu Psychiatry, Hospital of the University of Pennsylvania, Philadelphia, PA 19104-4283.
Supplemental data for this article can be accessed on the publisher's website.

variability in review times and attribute the difficulty of explaining the cause of variability to imperfect or unavailable data. A 2011 Milbank Quarterly meta-analysis states that the broadest evidence of overall IRB insufficiencies can be found in studies of IRBs' operating costs and from reports of the time needed to obtain an IRB's approval. However, the most comprehensive data on review times used in their analysis are from an evaluation of IRBs reviewing NIH-funded research conducted in 1995 (Silberman and Kahn 2011). Conclusions drawn from 20-year-old data, particularly data collected prior to rise of the internet, may not be an accurate reflection of current IRB operating timelines.

Health implications

Whether a research study is observational or aims to support new drug or device efficacy, vaccine development, or a behavioral intervention, delays in study approval affect more than the research team. An estimated thirty million Americans (or 10% of the US population) suffer from a disease that is inadequately understood or lacks effective treatment (Global Genes 2018). While researchers are waiting for their studies to be approved, the public is waiting for a new treatment. This is further complicated by data that suggest "difficulty and delays with the local IRB approval process sometimes result in sites or investigators choosing not to participate in research" (Abbott and Grady 2011, p.13, paraphrasing Mansbach et al. 2007).

In this study, we aimed to describe the current state of IRBs in the United States by collecting information regarding the timelines between protocol submission and approval, in addition to other relevant data that we believed might provide useful clues regarding practices that help shorten the review process and increase efficiency. Data were collected (and analyses planned) to address the following specific questions: 1) what do IRBs in the United States "look like," on average, in terms of basic characteristics like number of employees, review processing time, source of funds, etc.?; 2) can IRB characteristics be summarized by a small number (2–5) of overarching characteristics (e.g. size), and if so, what are those characteristics?; 3) which characteristics (or clusters of characteristics) have the greatest impact on IRB review speed and efficiency?; and 4) do IRBs differ in important ways according to large-scale qualitative factors, such as type of institution?

This study was overseen by the Children's Hospital of Philadelphia office of regulatory affairs and compliance.

Methods

Participants

Surveys were distributed to directors at 83 AAHRPP-accredited IRBs in the U.S. in January 2017. These included all IRBs at universities, medical schools, and nonprofit hospitals within the NIH's 2016–2017 top 50 funding recipients; IRBs from all member institutions of the PEDS (Patient-Centered Outcomes Research Institute) Network, a network of 8 high-volume pediatric research institutions; and additional institutions with a record of frequent collaboration with the Children's Hospital of Philadelphia (National Institutes of Health 2017). The final survey responses were received in July 2017.

Measures

Surveys were administered through REDCap (Research Electronic Data Capture), hosted at the Children's Hospital of Philadelphia (Harris et al. 2009). From REDCap, which is a secure, web-based application, data were transferred into SPSS for data management and analysis.

The survey was created specifically for this study, and was designed to assess, a) size of the IRB and institution (e.g. total staff), b) typical timelines for the IRB (e.g. average days from submission to approval), c) types of submissions received (e.g. biomedical), d) procedural details (e.g. how reviews are assigned to boards), and e) other important characteristics (e.g. does the IRB have a website?). The survey contained 77 items including Likert scales, rankings, yes/no questions, multiple-choice questions, and optional open-ended questions (due to low response rate to optional questions, only one was included in the discussion). We estimated the survey would take a minimum of 15-20 minutes to complete. The 32 survey questions included in the analyses are shown in full in Supplementary Table S1. The director of each IRB was identified using their institution's public IRB website, which also offered contact details (email and telephone numbers). Each director was sent an email inviting them to participate in the survey. Participants were given 4 months to complete the survey. Recipients were given a maximum of two email reminders to complete the survey and were informed that they could assign a delegate (e.g. support staff) to

Table 1. Item-Level Descriptive Statistics.

Variable	N	Min	Max	Median	Mean	Std. Dev
Number of active studies	55	102	6800	2200	2581	1873.3
Number of exempt studies received per year	49	10	1500	150	285.2	329
Number of expedited studies received per year	48	30	6000	381	807.2	1173.9
Number of full board studies received per year	50	10	2455	90	246	388
Social/behavioral/educational studies are highest ranked type of study*	51	1	3	2	2.1	0.6
Biomedical studies are highest ranked type of study*	54	1	3	3	2.4	0.8
Percent of biomedical studies that are drug/biological medical submissions	41	0	98	60	51.6	27.3
Funding from industry†	53	1	4	3	2.6	0.9
Funding from institution/internal†	53	1	4	4	3.0	1.1
Funding from US government†	52	1	4	3	2.6	1.0
Number of boards	55	1	9	2	3.1	2.1
Number of chairs	55	1	9	3	3.3	2.1
Number of full-time administrative staff	55	0	14	2	3.1	3.0
Number of full-time analysts/coordinators	55	0	24	6	7.5	5.9
Number of other staff members (not counted elsewhere)	55	0	14	1	2.5	2.9
Number of board meetings per month	55	1	20	3	4.6	4.0
Number of studies reviewed per convened board	53	1	90	8	12.3	13.7
Duration (in hours) of typical convened board meeting	54	1	3.5	2	2.1	0.6
Institution permits use of commercial IRBs‡	55	0	2	2	1.5	0.7
Median number of days from submission of a full board study to review by a convened board	48	3	45	15	18.5	8.0
Median number of days from submission of a full board study to approval	49	15	75	35	40.3	15.8
Median number of days from submission of an expedited study to review by an IRB reviewer	47	3	35	8	10.4	7.9
Median number of days from submission of an expedited study to approval	49	10	50	10	20.1	13.6
Median number of days from submission of an exempt study to determination	48	10	30	0	11.7	5.59

Footnotes.
Std. = standardized.
* = Value is ordinal. Study types are ranked 1 to 3; higher number indicates greater rank (more studies within that category).
† = Value is ordinal. Funding sources are ranked 1 to 4; higher number indicates greater rank (more funding from that source). "Other sources" of funding not included here.
‡ = Value is ordinal. 0 = no, 1 = on a limited basis, 2 = yes.

complete the survey on their behalf. If a response was not received after the 2nd reminder, the institution was not contacted further. Identifiers were optional and removed prior to analysis. The Children's Hospital of Philadelphia IRB Office acknowledged that the study did not constitute human subjects research.

Analysis

In line with the specific goals of the study, the analyses proceeded as follows: 1) descriptive statistics, 2) Principal Components Analysis to determine whether and how the survey responses can be clustered into sub-components, and 3) multiple regression predicting IRB speed and efficiency using both subfactors from the PCA and (separately) all individual variables. The multiple regressions using all variables used forward stepwise variable selection (Henderson and Denison 1989). Analyses described below were performed in SPSS (SPSS Inc. Chicago, IL) or R (R Core Team 2018).

For step #3 above, all dependent variables (DVs) were specifically related to the number of days it takes an IRB to review and approve (or waive approval of) submissions in three review categories[1]: (1) full-board, (2) expedited, and (3) exempt. The number of days to rejection of a submission was not queried.

DVs were either, a) absolute (count of days from submission to approval), or b) relative (*ad hoc* efficiency score defined below). Examples of independent variables of interest include number of boards, number of staff members, sources of study funding, number of studies per category (full board, expedited, exempt), number of studies at each IRB by type (e.g. biomedical, sociobehavioral), and process assigning studies to boards and/or staff members.

Descriptives

IRB directors were asked to estimate the median number of days from study submission to review, and from submission to IRB approval or determination, depending on study category. Table 1 shows the descriptive statistics for the continuous variables and Table 2 shows the descriptive statistics for the dichotomous variables. Bivariate relationships among all continuous variables (Pearson) are also shown in Supplementary Table S2. Note that dichotomous variables were excluded from this matrix but are included in the principal components analyses (PCAs) (see below).

Principal components analyses

Due to the large number of variables collected and the relatively small sample size, our first step was to

[1]Review categories are determined by IRBs per 45 CFR 46 and 21 CFR 56 (2008), which provide direction regarding assignment to review categories based on the research's level of risk.

Table 2. Dichotomous Variables.

Variable	N	Percent endorsed
Administrative changes are addressed with investigators prior to IRB review	55	92.7
Ancillary approvals are required prior to release of IRB approval	55	85.5
Electronic IRB system integrates with ancillary groups	55	56.4
IRB charges a fee for the review of industry funded studies	55	81.8
Protocols are assigned to boards by indication (e.g. oncology, neurology)	55	50.9
Protocols are assigned to boards by study type (e.g. biomedical, social behavioral)	55	27.3
Protocols are assigned to boards by timing of submission (rolling assignment)	55	12.7
Submissions are assigned to staff members by current workload	55	60.0
Submissions are assigned to staff members by indication (e.g. oncology)	55	38.2
Submissions are assigned to staff members by research classification (e.g. biomedical, social behavioral)	55	20.0
Submissions are assigned to staff members by review type (e.g. full board, expedited)	55	38.2
Submissions are assigned to staff members by submission type (e.g. continuing reviews, new studies)	55	43.6

reduce the data using PCA with oblimin rotation (Sass and Schmitt 2010). This is useful for data reduction, but will also reveal patterns of inter-variable relationships that might be of theoretical interest in themselves. For thorough didactic discussion of factor analysis (including PCA), its strengths/weaknesses, and how to interpret results, see Wolfe and Dobria (2008) and Kline (2014).

The first PCA was performed on 5 variables thought to be directly associated with IRB review duration (a variable of interest that is measured in absolute units, days); 1 component was extracted from these variables. The 5 variables (included in Table 1) include median duration of time from: (1) submission of a full board study to review by a convened board; (2) submission of a full board study to approval; (3) submission of an expedited study to review by an IRB reviewer; (4) submission of an expedited study to approval; (5) submission of an exempt study to determination. These 5 items were analyzed alone because they composed a key DV. When *all* items were analyzed together (not shown), these 5 items composed a clear component in addition to the 3 components described below. However, we wished to keep this key DV free of other variables' influence (due to cross-loadings in the PCA), and therefore kept the speed component as its own unidimensional PCA.

A second PCA was run on the remaining 31 variables producing 3 components, which served two purposes. First, from a theoretical perspective, it allowed us to explore the dimensional structure of IRBs—i.e. to the extent that response patterns on the survey were caused by higher-level dimensions underlying IRBs (e.g. size), PCA allowed us to uncover what those dimensions are, and how many. Second, from a practical perspective, PCA allowed us to reduce the number of relevant variables from 36 to 4 (review duration, plus the 3 other components discussed below).

The number of components to extract (3) from the second PCA was determined by visual examination of the scree plot (Cattell 1966) and by the minimum average partial (MAP) method (Velicer 1976). Also note that because the variables in this study were a mix of continuous, dichotomous, and ordinal, the correlation matrix used for the PCA was a mix of Pearson (for continuous-continuous), biserial[2] (for continuous-dichotomous), tetrachoric (for dichotomous-dichotomous), and polychoric (for dichotomous-ordinal and ordinal-ordinal).

Predicting review duration

With the above 4 components extracted, we then examined how the 3 components from the second PCA related to IRB review duration (the component from the first PCA). This was done using linear regression where the speed component was the DV and all 3 of the other components were IVs. No covariates were included, because all potential covariates were already incorporated into the components. Next, to examine more specific relationships of the IRB variables with review duration, we ran a stepwise linear regression predicting review duration with all variables in the data set (i.e. the same variables that went into the second PCA). Note that for both linear models described above, there were some minor violations of general linear model assumptions.

Predicting review efficiency

In addition to examining absolute time from submission to approval/rejection, we examined the efficiency of IRBs' review processes. "Efficiency," developed specifically for this study, was defined by the following equation:

$$E_i = \log_{10}\left(\frac{s_i}{(p_i)(d_i)}\right)$$

Where E_i is the efficiency score for IRB i, s_i is the number of studies received per year (study count), p_i

[2]Note that these are biserial correlations, not point-biserial. See Kemery, Dunlap, and Griffeth (1988) for details.

is the staff size (employee count), and d_i is the duration (days) between submission and approval/rejection. The $\log_{10}$ transformation is necessary due to the extreme skew of the raw score. The units of E are logit_{10} studies per employee-day. Thus, for any given value of p or d, an increase in s is an increase in efficiency; likewise, for any given value of s, an increase in p or d is a decrease in efficiency. For example, if IRB1 has 1 employee and reviews 1 study in 1 day, its E would be log[1/(1*1)] = log(1) = 0. If IRB2 has 1 employee and reviews 2 studies in 1 day, its E would be log[2/(1*1)] = log(2) = 0.30 (more efficient).

There were thus 3 types of scores:

1) an "absolute speed" factor score reflecting review duration for full-board, expedited, and exempt reviews combined. This score was used for the "Review Duration" analyses described in the section above.

2a) relative efficiency (E) of review for full-board submissions

2b) relative efficiency (E) of review for expedited submissions

Comparison by review type

To test whether institutions differed by institution type, we ran an analyses of covariance (ANCOVA). The *p*-value for the associated F-test was Bonferroni-corrected.

Results

Final sample

A total of 71 IRB directors or their representatives started or completed the survey, an 85% response rate. Of those surveys, 55 were included in this analysis; others were excluded from the analysis because less than 50% of the survey were completed. The final sample yielded 32 completed surveys (58% of final sample) and 23 surveys with at least 73% of questions answered. On average, surveys were 90% complete.

Tables 1 and 2 give a rough profile of key characteristics of a "typical" IRB in the United States. It largely addresses our first research question: what do IRBs in the United States "look like," on average, in terms of basic characteristics?

Breakdown of respondent types yielded: 32 universities with medical schools, 6 universities without medical schools, 9 children's hospitals, 7 nonprofit healthcare systems, and 1 anonymous. Item wise descriptive statistics across institutions are shown in the Methods section (Tables 1 and 2). Supplementary Figure S1 shows a map of the participating states, with number of institutions indicated for each.

Principal components analysis

The Principal Component Analysis (PCA) results address our second research question: can IRB characteristics be summarized by a small number (2–5) of overarching characteristics (e.g. size), and if so, what are those characteristics? Table 3 shows these results, extracting 3 components with oblimin rotation. Numbers in Table 3 can be interpreted as correlations of the variables with each factor. Component 1 comprises 14 variables, the majority of which are associated with institution size and, by extension, IRB size. The two items that most strongly define Component 1 are: 1) "How many IRB studies are current active[…]?", and 2) "How many board/committees does your IRB have?". Component 1 is clearly capturing size.

Component 2 comprises 9 variables principally related to methods of assignment of protocols to IRB staff and boards/committees. The two most representative items indicate assignment according to indication (e.g. "oncology") or staff workload. An IRB high on Component 2 would be one that tends to assign work according to indication and workload and tends not to review research protocols funded internally (by the institution itself).

Component 3 comprises 8 items mostly related to the types of studies reviewed. The two items most strongly defining Component 3 are (paraphrased): 1) "Of biomedical research reviewed by your IRB, what percentage is drug/biological?", and 2) "Social/behavioral/educational studies are the most common type of study reviewed," which loads negatively. Component 3 appears to describe orientation toward industry sponsored biomedical research with a focus on drug and medical development, and away from social, behavioral, and educational research.

In addition to the components determined by each item's highest loading (largest absolute value), there are several items that load on multiple factors simultaneously (cross-loading items) in the solution. For example, Component 1 comprises the 14 items at the top of Table 3, but in addition, the following items contribute at least moderately to Component 1: low internal funding, high use of "rolling assignment" (based on submission timing), high number of full-board studies received per year, and policy of charging a fee for industry-sponsored studies. In addition to the 9 items in the middle of Table 3, Component 2 is

Table 3. Exploratory Factor Analysis Results for the 3 Factors Characterizing IRBs in the United States.

Variable	Size	Volume Mgmt	Biomed
Number of active studies	0.89	0.03	0.03
Number of boards (if >1) within the IRB	**0.81**	0.15	0.12
Number of chairs	**0.80**	−0.06	0.17
Number of full-time analysts/coordinators	**0.78**	0.20	0.01
Number of full-time administrative staff	**0.77**	−0.18	0.00
Number of board meetings per month	**0.69**	0.01	0.29
Number of exempt/other studies per year	**0.67**	0.11	−0.34
Submissions are assigned to staff by research classification	**0.66**	0.24	−0.09
Electronic system integrates with ancillary groups	**0.65**	−0.62	0.07
Number of studies reviewed per convened board meeting	**0.57**	0.02	−0.16
Ancillary approvals are required prior to release of IRB approval	**0.52**	0.03	0.09
Duration of board meetings	**0.38**	−0.09	−0.34
Number of other staff members (not counted elsewhere)	**0.37**	0.36	0.36
Funding from US Government	**0.33**	0.12	−0.30
Submissions assigned to staff by indication (e.g. oncology)	0.10	**0.86**	−0.24
Submissions assigned to staff by current staff workload	−0.11	**0.83**	0.35
Submissions are assigned to boards by study type	0.23	**0.75**	−0.04
Funding from institution/internal	−0.37	**−0.56**	0.22
Submissions are assigned to board by timing of submission (rolling assignment)	0.50	**0.52**	0.22
Number of full board studies received per year	0.36	**0.40**	−0.07
Number of expedited studies received per year	0.30	**0.38**	0.01
Submissions are assigned to staff by review type (e.g. full board, expedited, exempt)	−0.09	**0.35**	0.03
Institution permits use of commercial IRBs	0.09	**0.24**	0.17
Percent of biomedical studies that are drug/biological medical submissions	0.12	−0.07	**0.78**
Social/behavioral/educational studies are highest ranted type of studies	0.08	−0.07	**−0.74**
IRB charges a fee for review of industry sponsored studies	0.41	−0.04	**0.72**
Biomedical studies are highest ranked type of study	−0.01	−0.08	**0.55**
Submissions are assigned to boards by indication (e.g. oncology, neurology)	0.20	0.08	**0.50**
Submissions are assigned to staff by submission type (e.g. new studies, CRs)	−0.11	0.42	**0.49**
Administrative clarifications are addressed with investigators prior to review	0.28	0.26	**−0.46**
Funding from industry	0.02	0.11	**0.14**
	Inter-Factor Correlations		
	F1	F2	F3
F1	1.00	0.25	0.19
F2	0.25	1.00	0.03
F3	0.19	0.03	1.00

Note. Rotation = oblimin; Mgmt = management; Biomed = biomedical.

indicated by: absence of an electronic system integrating ancillary groups, high number of staffs classified as "other," assignment of studies according to type, and tendency to clarify administrative issues with investigators prior to review. Finally, in addition to the 8 items at the bottom of Table 3, Component 3 is indicated by: few exempt studies, relatively brief board meetings, high number of staffs classified as "other," and assignment of studies based on workload.

The correlations among the 3 components (size, study assignment pattern, and industry orientation) (not shown) are weak (max = 0.31), suggesting that the components are relatively orthogonal, capturing unique variation in IRBs that would probably not be better captured by a hierarchical solution (e.g. bifactor).

Predicting review duration

The two sections below describe the prediction of efficiency and duration and address our third research question: which characteristics (or clusters of characteristics) have the greatest impact on IRB review speed and efficiency?

Further analyses were performed to determine whether review duration could be predicted by the 3 PCA components. Based on a linear regression analysis, none of the variables was able to predict review duration (all $p > 0.05$); null results shown in Supplementary Table S3.

Next, using stepwise regression, we tested whether any of the original survey variables were associated with review duration (Speed). As shown in Table 4, two variables showed significant associations: 1) Percentage of biomedical protocols that are drug/biological medical submissions ($\beta = 0.624$; $p < 0.0005$), and 2) Biomedical studies are the most common type of study ($\beta = -0.393$; $p < 0.005$). The third variable selected by the algorithm (Number of other staff members (not counted elsewhere)) was not significant in the final model ($p > 0.05$). The first variable (Percent of biomedical […]) is positively related to speed, meaning that institutions with a high percentage of drug/biological studies with a medical orientation tend to have shorter review durations. The second variable (Biomedical studies rated the greatest […]) is negatively associated with

Table 4. Stepwise Regression Results Predicting IRB Review Duration (coded such that higher indicates faster) Using All Available Quantitative Variables.

Predictor	Std. β	t	*p*-value
Intercept	n/a	−0.254	0.800
Percent of biomedical studies that are drug/biological medical submissions	0.624	4.676	0.000
Biomedical studies are the highest ranked type of study	−0.393	−2.948	0.005
Number of other staff members (not counted elsewhere)	0.223	1.694	0.096

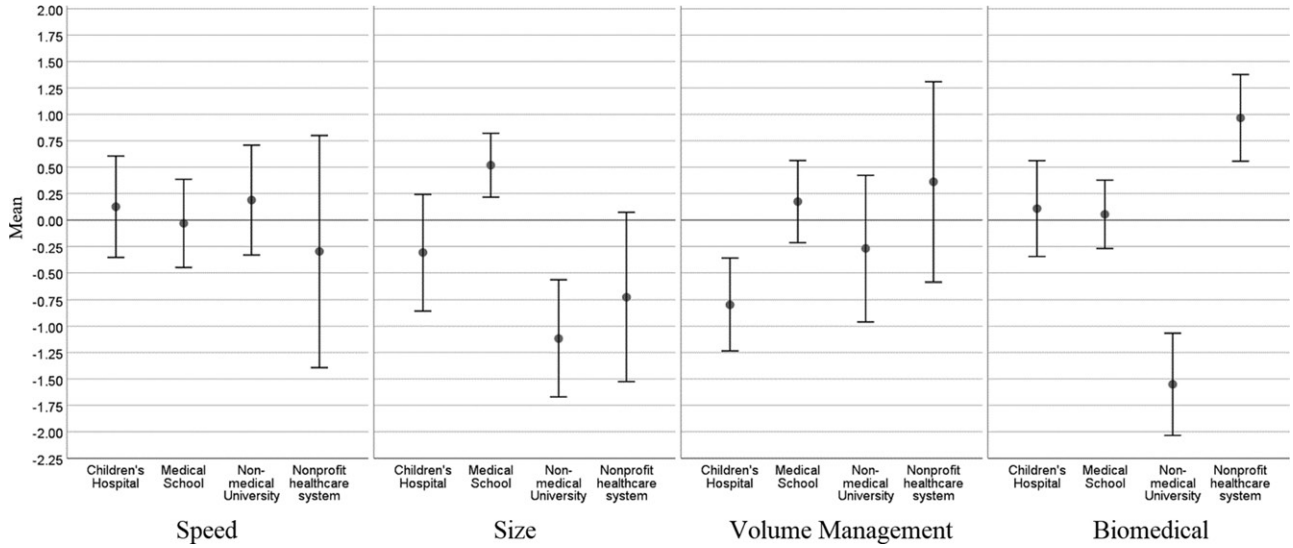

Figure 1. IRB Speed, Size, Volume Management, and Biomedical Orientation, Separated by Institution Type.

speed, indicating that orientation toward biomedical studies, broadly, tends to result in slower review. These two relationships in the opposite direction, with the drug-related orientation strongly associated with faster review, but biomedical orientation in general moderately associated with slower review, suggest that the slowest review boards tend to be those that review many biomedical studies that are *not* drug-related.

Predicting efficiency

Analyses were performed to determine whether efficiency of IRBs could be predicted by the same variables used to predict absolute review duration. Efficiency was unrelated ($p > 0.05$) to institution type, the 3 principal components, and all other variables used to predict it, with one exception. Efficiency of full board review was not significant for any variable. In the stepwise regression using all variables to predict expedited review Efficiency, the requirement of Ancillary Review significantly predicted efficiency, such that boards requiring Ancillary Review were 0.37 SDs less efficient ($t = -2.52$; $P = .016$). Notably, the requirement of ancillary review predicts 13% of the variance in efficiency ($R^2 = 0.134$).

Comparison by demographic factors

Comparing IRBs by demographic factors, specifically by institution type, addresses our fourth research question: do IRBs differ in important ways according to large-scale qualitative factors?

Figure 1 displays the four PCA variables by type of institution. Note that all continuous variables (including components) were z-transformed (to the global mean and SD) and any group mean values thus reflect SDs above (if positive) or below (if negative) the sample mean. Of the four variables, size and biomedical orientation factors showed significant differences among types ($p < 0.05$). Post hoc independent sample t-tests revealed 4 significant size differences between institution types, as follows: between children's hospitals and non-medical universities ($p < 0.05$), children's hospitals and medical schools ($P < 0.01$), medical schools and non-medical universities ($P < 0.005$), and medical schools and nonprofit healthcare systems ($P < 0.01$). Post hoc t-tests (with Bonferroni-corrected p-values) also revealed 5 significant biomedical orientation differences as follows: between children's hospitals and non-medical universities ($p < 0.0005$), children's hospitals and nonprofit healthcare systems ($p < 0.01$), medical schools and non-medical universities ($p < 0.0005$), medical schools and nonprofit healthcare

systems ($p < 0.01$), and non-medical universities and nonprofit healthcare systems ($p < 0.0005$).

There were no significant institution type predictors of review duration (speed). All institutional types (previous analysis, above) were within .5 SDs of the mean ("Speed" section of Figure 1).

Other descriptive analyses

One of the survey's questions that was not included in the PCA asked IRBs the following: "What is the biggest impediment to decreasing review time?" Sixteen IRBs (29%) reported that delays can be attributed to extended time spent with the investigator/study team, 14 IRBs (25%) reported that poor quality of submissions is the main factor for review delays, and an additional 14 IRBs (25%) reported high work volume is the main culprit for long review times.

Discussion

Time is of essence when implementing human subjects research. Researchers, their home institutions, sponsors, and subjects themselves may see the IRB as an impediment during this time, expecting their initial review to add days or months to timelines. Here we report that IRBs across the U.S. are reviewing studies at a consistent pace, such that no specific type of institution has an advantage when predicting IRB review duration. Similarly, IRBs show little variation in terms of efficiency, except for the difference between those requiring or waiving Ancillary Review. This finding, which has been reported previously (Caligury et al. 2017), is surprising and challenging to explain. It is possible that institutions that require ancillary reviews are overall more scrutinizing than those that do not, and that this is also reflected in their IRBs, which may tend to pore over protocols for longer, thus delaying final protocol approvals.

While there is no single major predictor of protocol reviewing time, certain variables are associated with shorter protocol turnaround times. These variables are principally related to the nature of the research proposed and not modifiable by the IRB. More specifically, institutions that reported biomedical studies being the primary type of research reviewed had significantly longer turnaround times. This is not surprising given that biomedical studies carry a greater risk for adverse effects and require review by medical specialists. Other aspects of biomedical studies, such as higher frequency of industry sponsorship, investigator conflicts of interest, and ethical questions regarding the societal benefits of "me too" drug studies and post marketing research may also delay approval (Klitzman 2015). On the other hand, and perhaps somewhat surprisingly, IRBs that review a high percentage of *treatment oriented* (drug or biologic) biomedical research are associated with shorter review times. This observation is harder to explain, and may have to do with the collinearity between biomedical research and treatment-oriented research. Additionally, drug-related interventional studies may be treated with a greater urgency compared to observational studies because they are seen as more likely to have an immediate impact on healthcare.

PCA

The three factors extracted by PCA (size, volume management, and biomedical) point to the primary underlying and measurable qualities of IRBs. They give flesh to a picture of what a typical IRB "looks like," and primary characteristics by which it can be described. Size is not a surprising quality as IRBs included in this survey ranged from 102 to 6,800 active studies (SD 1,873) and institutions with more studies tend to have a greater number of boards, chairs, and other personnel to support larger work volumes. Similarly, the biomedical factor may have strong variable loading as biomedical studies were ranked higher (indicating they were more common) than socio-educational-behavioral studies among all IRBs.

While the factors are orthogonal, volume management, a factor generally describing how studies are assigned to staff for review, may shed light on how institutions are able to handle a large number of submissions (the size factor) beyond hiring additional staff members. These institutions may have better defined or efficient processes for assigning studies to staff members, perhaps due to staff specialization.

Self v. Investigator reported outcomes

The data collected here were self-reported by IRB staff and are therefore subject to reporting bias. Review times may be underreported or overreported depending on whether the source is the reviewing body or the investigator. Previous investigator-reported research found that it takes an average of 36 days (range 2 - 124 days) to receive IRB approval for a minimal risk multicenter study, though high variation was also observed (Khan et al. 2014). IRBs that completed our survey reported expedited and exempt studies take an average of 20 and 12 days to approval,

respectively, which is substantially less than that reported by investigators.

The Association for the Accreditation of Human Research Protection Programs (AAHRPP), the unofficial human research accreditation body that figuratively gives IRBs a "gold seal" of approval, occasionally makes public "metrics" on the activity of its members. Their 2017 metrics included data from over 200 clients (IRBs) and reported the average duration of review for exempt and expedited studies – 10 and 16 days, respectively, for determination or approval (AAHRPP 2017). This is slightly faster than our results, 12 and 20 days, whereas the number of days to receive approval for a full board review study, per AAHRPP, is 38 days, compared to our 40 days.

While study teams are known to complain about the duration of IRB review and its questionable benefit towards the protection of human subjects (Millum and Menikoff 2010), qualitative results from our study may reverse this complaint towards investigators. The majority of IRBs reported that the reason behind long review times lies, at least partially, with investigators, who either need prolonged periods of time to discuss their studies with the IRB, or who submit poor quality protocols which require multiple rounds of clarifications and revisions. At least a quarter of the IRBs we surveyed attributed the delays to their own high workload, but it is unclear whether the high workload is the result of insufficient staff or a high volume of submissions that require further work. In either case, it appears that better education of the investigators in terms of improving the quality of their submissions is likely, at least from the IRBs' point of view, to alleviate reviewing workload.

Limitations and future directions

Despite its strengths, there are several limitations to this study that deserve mention. First, tests for differences among IRBs by, a) geographic region, and b) institution type, were performed separately, meaning the findings cannot be interpreted as controlling for all variables. Differences among institutions by geographic region could be due to regions differing by institution type, and vice versa. Second, results for question 4 (do IRBs differ in important ways according to large-scale qualitative factors, such as type of institution?) were limited. Qualitative questions included in the original survey sent to IRB directors were marked as optional and thus had a low response rate, something for future researchers to keep in mind.

Third, there are multiple things to keep in mind when interpreting the benefits of short reviewing times. Our study did not measure the quality of review by IRBs, for example whether IRB review increased the protection of research subjects, as is their intent, but solely examined the factors characterizing IRBs. We cannot determine whether reviews were equally error-free (in terms of regulatory compliance) when carried out near the slow and fast ends of the spectrum. As noted above, IRBs in our sample reported the low quality of submitted protocols and time spent with the investigator's team as main factors prolonging review times. While this is partially subjective, it is undeniable that the continuous education of investigators and other protocol authors on how to produce thorough and regulatory-compliant submissions will be beneficial in reducing the need for time-consuming corrections and stipulations. Our survey included questions regarding the frequency of one-on-one and group training, office hours, and whether a website with submission instructions was available and updated regularly, none of which were significant or correlated with the 4 PCAs.

Fourth, it remains to be determined how changes to the Common Rule, such as elimination of continuing review for certain qualified protocols, expansion of studies that qualify as exempt, and increased review of multisite studies by central IRBs, will affect IRB reviewing times. While these regulatory changes are expected to help reduce reviewing times, it will be of particular interest to see whether and how they would alter the dynamics we describe here. A replicated survey could show signs of potential improvements, trends, and changes in IRB characteristics.

Finally, an important potential limitation of the summary (PCA) scores used here is that the model contained numerous variables that loaded on multiple components simultaneously (cross-loadings), and relatedly, some components comprised items that do not clearly belong in said component's theoretical conceptualization. The scores are therefore more difficult to intuit and interpret. Further, the relatively small sample used here increases the likelihood that some of the item-component specifications are erroneous—i.e. the PCA-implied configural model could be wrong. Further research using larger samples is necessary to explore this possibility. Authors recommend using the top loading variables within each factor, rather than all variables, when designing future surveys as factor association can be determined with a minority of variables.

Conclusion

Our data offer a glance at the various factors that may or may not affect IRB review times. These results are the first step toward establishing national norms and building a working model of US IRBs to which other IRBs can compare themselves. Researchers and those with investments in research may take heart that, in general, IRBs are functioning at equal efficiency. Whether this efficiency is adequate, or effort should be put forward to improve it further, is a matter of perspective. In either case, a continuous assessment and re-assessment of current methodologies and practices is essential to drive forward a constant effort for improvement. It will be of great interest for a similar investigation to evaluate the impact of the various factors we have examined here, following the full implementation of the Final Rule.

Acknowledgments

The authors appreciate Amy Schwarzhoff, MBA, Director of the Children's Hospital of Philadelphia IRB, for her insight during study design.

Disclosure statement

No potential conflict of interest was reported by the authors.

Funding

This project was supported by NIH/NCRR Colorado CTSI Grant Number UL1 RR025780. Its contents are the authors' sole responsibility and do not necessarily represent official NIH views.

Ethical approval

The Institutional Review Board at the Children's Hospital of Philadelphia acknowledged that this study did not constitute human subjects research.

References

21 CFR 56. 2008. Code of Federal Regulations Title 21, Food and drugs, Chapter I - Food and Drug Administration, Dept. Health and Human Services, Subchapter A, General Part 56 Institutional review boards. https://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfcfr/CFRSearch.cfm?CFRPart=56

AAHRPP. 2017. Metrics on human research protection program performance. Association for the Accreditation of Human Research Protection Programs. Updated May 2, 2018. http://www.aahrpp.org/apply/resources/metrics-on-hrpp-performance

Abbott, D., R. Califf, B. W. Morrison, C. Pierre, J. Bolte, and S. Chakraborty. 2013. Cycle time metrics for multisite clinical trials in the united states. *Therapeutic Innovation & Regulatory Science* 47 (2):152–60. doi: 10.1177/2168479012464371.

Abbott, L., and C. Grady. 2011. A systematic review of the empirical literature evaluating IRBs: What we know and what we still need to learn. *Journal of Empirical Research on Human Research Ethics* 6 (1):3–20. doi: 10.1525/jer.2011.6.1.3.

Caligury, M., K. Allen, N. Buscher, L. Denney, C. Gates, K. Kantelo, A. Magit, R. Sak, G. Firestein, and J. Fontanesi. 2017. A multisite study of performance drivers among institutional review boards. *Journal of Clinical and Translational Science* 1: 192–197. doi: 10.1017/cts.2017.8.

Cattell, R. B. 1966. The SCREE test for the number of factors. *Multivariate Behavioral Research* 1 (2):245–76. doi: 10.1207/s15327906mbr0102_10.

Department of Health and Human Services. 2011. Department of Health and Human Services. Human subjects research protections: Enhancing protections for research subjects and reducing burden, delay, and ambiguity for investigators. *Fed Registrar* 76 (143):44512–31. https://www.federalregister.gov/documents/2011/07/26/2011-18792/human-subjects-research-protections-enhancing-protections-for-research-subjects-and-reducing-burden

Desai, P., P. Nasa, J. Soo, C. Jia, M. L. Berbaum, J. H. Fischer, and T. P. Johnson. 2017. Effects of regulatory support services on institutional review board turnaround times. *Journal of Empirical Research on Human Research Ethics* 12 (3):131–9. doi: 10.1177/1556264617704294.

Global Genes. 2018. Statistics and Figures on prevalence of genetic and rare diseases. Accessed March 7. https://globalgenes.org Web site. https://globalgenes.org/rare-diseases-facts-statistics. Updated 2015. .

Harris, P. A., R. Taylor, R. Thielke, J. Payne, N. Gonzalez, and J. G. Conde. 2009. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *Journal of Biomedical Informatics* 42 (2):377–81. doi: 10.1016/j.jbi.2008.08.010.

Henderson, D. A., and D. R. Denison. 1989. Stepwise regression in social and psychological research. *Psychological Reports* 64 (1):251–7. doi: 10.2466/pr0.1989.64.1.251.

Hirshon, J. M., S. D. Krugman, M. D. Witting, J. P. Furuno, M. R. Limcangco, A. R. Perisse, E. K. Rasch. 2002. Variability in institutional review board assessment of minimal-risk research. *Academic Emergency Medicine* 9 (12):1417–20. doi: 10.1111/j.1553-2712.2002.tb01612.x.

Khan, M. A., M. S. Barratt, S. D. Krugman, J. R. Serwint, and M. Dumont-Driscoll. 2014. Variability of the institutional review board process within a national research network. *Clinical Pediatrics* 53 (6):556–60. doi: 10.1177/0009922814527504.

Kline, P. 2014. *An easy guide to factor analysis*. New York: Routledge.

Klitzman, R. 2015. *The ethics police?: the struggle to make human research safe*. New York: Oxford University Press.

Mansbach, J., U. Acholonu, S. Clark, and C. A. Camargo. Jr. 2007. Variation in institutional review board responses to a standard, observational, pediatric research protocol. *Academic Emergency Medicine* 14 (4):377–80. doi: 10.1197/j.aem.2006.11.031.

Millum, J., and J. Menikoff. 2010. Streamlining ethical review. *Annals of Internal Medicine* 153 (10):655. doi: 10.7326/0003-4819-153-10-201011160-00008.

National Institutes of Health. 2017. NIH research portfolio online reporting tools (RePORT): Awards by location & and organization. Accessed February 1. https://report.nih.gov/award/index.cfm Web site. https://report.nih.gov/award/index.cfm. Updated 2017.

Sass, D. A., and T. A. Schmitt. 2010. A comparative investigation of rotation criteria within exploratory factor analysis. *Multivariate Behavioral Research* 45 (1):73–103. doi: 10.1080/00273170903504810.

Silberman, G., and K. L. Kahn. 2011. Burdens on research imposed by institutional review boards: The state of the evidence and its implications for regulatory reform. *Milbank Quarterly* 89 (4):599–627. doi: 10.1111/j.1468-0009.2011.00644.x.

Spellecy, R., A. M. Eve, E. R. Connors, R. Shaker, and D. C. Clark. 2018. The Real-Time IRB: A collaborative Innovation to Decrease IRB Review Time. *Journal of Empirical Research on Human Research Ethics*: *JERHRE* 13 (4):432–7. doi: 10.1177/1556264618780803.

Stark, A. R., J. E. Tyson, and P. L. Hibberd. 2010. Variation among institutional review boards in evaluating the design of a multicenter randomized trial. *Journal of Perinatology* 30 (3):163–9. doi: 10.1038/jp.2009.157.

Varley, P. R., U. Feske, S. Gao, R. Stone, S. Zhang, R. Monte, R. M. Arnold, and D. E. Hall. 2016. Time required to review research protocols at 10 Veterans Affairs Institutional Review Boards. *Journal of Surgical Research* 204 (2):481–9. doi: 10.1016/j.jss.2016.06.004.

Velicer, W. F. 1976. Determining the number of components from the matrix of partial correlations. *Psychometrika* 41 (3):321–7. doi: 10.1007/BF02293557.

Wolfe, E., and L. Dobria. 2008. Applications of the multifaceted rasch model. In Best practices in quantitative methods, ed. J. Osborne, 71–85. Thousand Oaks, CA: SAGE Publications, Inc.