Laura A. Petersen, MD, MPH; Kate Simpson, MPH; Richard SoRelle, BS; Tracy Urech, MPH; and Supicha Sookanan Chitwood, MPH

Background: The Department of Health and Human Services recently called for public comment on human subjects research protections.

Objective: To assess variability in reviews across institutional review boards (IRBs) for a multisite, minimal-risk trial of financial incentives for evidence-based hypertension care and to quantify the effect of review determinations on site participation, budget, and timeline.

Design: A natural experiment occurring from multiple IRBs reviewing the same protocol for a multicenter trial (May 2005 to October 2007).

Participants: 25 Veterans Affairs (VA) medical centers.

Measurements: Number of submissions, time to approval, and costs were evaluated; patient complexity, academic affiliation, size, and location (urban or rural) between participating and nonparticipating VA medical centers were compared.

Results: Of 25 eligible VA medical centers, 6 did not meet requirements for IRB review and 2 declined to participate. Of 17 applications, 14 were approved. The process required 115 submissions, lasted 27 months, and cost close to $170,000 in staff salaries. One IRB’s concern about incentivizing a particular medication recommended by national guidelines prompted a change in our design to broaden our inclusion criteria beyond uncomplicated hypertension. The change required amending the protocol at 14 sites to preserve internal validity. The IRBs that approved the protocol classified it as minimal risk. The 12 sites that ultimately participated in the trial were more likely to be urban and academically affiliated and to care for more complex patients, which limits the external validity of the trial’s findings.

Limitation: Because data came from a single multisite trial in the VA system that uses a 2-stage review process, generalizability is limited.

Conclusion: Complying with IRB requirements for a minimal-risk study required substantial resources and threatened the study’s internal and external validity. The current review of regulatory requirements may address some of these problems.

Primary Funding Source: Veterans Affairs Health Services Research and Development and the National Heart, Lung, and Blood Institute.

Many authors have documented variability in the process of obtaining human subjects approval from local institutional review boards (IRBs) for multisite studies. This variability includes standards of review (1–4), consent documents and requirements (1–7), and the time from submission to approval (2, 3, 8). For example, 1 observational health services research protocol at 43 sites noted that the time from submission to approval ranged from 52 to 798 days (3). In a well-known case, a quality improvement study led by Pronovost and colleagues (9) illustrated how regulations meant to protect human subjects were interpreted by the Office for Human Research Protections in a way that seemed contrary to their intent (10). A recent systematic review of evidence from 52 studies concluded that some decisions made by IRBs are not consistent with federal policy (11).

The gold standard for generalizable research is a multisite, randomized, controlled trial. However, such trials are relatively rare in health services research, and IRBs may lack experience in reviewing them. In this article, we focus on the variability in review determinations across IRBs for a multisite trial that sought to improve delivery of evidence-based hypertension care. We also seek to quantify the effect on the type of site ultimately participating, budget, timeline, and project staff. To our knowledge, ours is the first study to evaluate the IRB approval process for this type of research and to highlight the effect on both the internal and external validity of the study’s findings. We hope that our findings will help to inform current efforts to solicit public comment about the need to revise the Common Rule (12).

Methods
This study is a natural experiment occurring from multiple IRBs reviewing the same protocol for a multisite trial. We reviewed records detailing the IRB approval process from May 2005 through mid-October 2007.

Description of the Trial
The trial was designed to test whether explicit financial incentives (also termed “pay for performance”) (13) improved hypertension guideline adherence (14). The study
Effect of Variability in the IRB Review Process on Health Services Research

methods are described elsewhere (15). Briefly, 12 Veterans Affairs (VA) medical centers were randomly assigned to 1 of 4 study groups according to the type of incentive rewarded: physician level, health care provider group level, physician and group level, and none (control). Participants in all 4 groups received audit and feedback on their performance. Primary care physicians who worked at least 0.6 full-time equivalents (approximately 3 days per week related to clinical activities) or had a panel size of at least 500 patients were eligible to participate. At the 6 study sites randomly assigned to the group-level incentive, the physicians invited up to 15 nonphysician colleagues, including other clinicians (for example, nurses and pharmacists) and administrative support staff (for example, clerks) to participate.

Procedures for Obtaining IRB Approval

Multisite research studies conducted within the VA system are required to designate a local principal investigator (PI) and to obtain approval from both the local IRB and the local VA Research and Development (R&D) committee. In this trial, the process of identifying a local PI at each site included contacting local leadership to identify potential site investigators, obtaining site PI assent to participate, educating the site PI about the project and his or her role, and ensuring that he or she had an academic appointment and current research training. After identifying a site PI, research staff from the coordinating center in Houston, Texas, prepared the IRB and R&D applications for each site. A certified IRB professional was hired 8 months after beginning the submission process to help submit the regulatory paperwork; the need for this was not initially anticipated.

Site Selection

Leaders of 5 VA networks (consisting of 45 facilities) agreed to participate (Figure 1). The study power calculations required at least 5 physicians per site. We pursued the 25 facilities with at least 8 full-time physicians to account for potential attrition, and we intended to randomly select 12 facilities before study group randomization to achieve a representative sample. Of the 25 facilities with at least 8 full-time physicians, 3 did not have an on-site or affiliated IRB. At 2 of the 22 facilities with IRBs, the local hospital directors declined to participate. At another facility, we were unable to recruit a site PI. We prepared IRB and R&D applications for the 19 remaining facilities. However, at 2 of these, the site PI was unable to complete the academic credentialing or research certification process. Applications were therefore submitted to and reviewed at only 17 of the 25 eligible facilities.

Record Review and Measures

At least 2 authors independently reviewed the regulatory submission materials and correspondences at each site to gather data on submission requirements, board structure, study review category, and submission and approval dates. When their observations were in disagreement, they consulted a third author.

We calculated the time from initial submission to approval as the number of calendar days from the date the application was submitted to a site’s PI until the date of the initial approval letter from that site. One site’s IRB had a concern that could only be addressed by changing the study methods, necessitating a modification of the protocol at all sites to preserve the internal validity of the study (details of the change are described in the next section). Because this delayed commencement of the project, we also calculated the time from the submission of this modification request to its approval. For the 13 sites that approved the study before the modification was submitted, we evaluated the relationship between the date when the application was submitted and the number of days to initial approval. Site identification numbers reflect the order in which applications were submitted (applications were submitted to site 1 first and to site 17 last).

We enumerated the submissions the team made to each site’s regulatory board from the submission date of the
initial application until the date of the initial approval letter and from the submission date of the protocol modification until the date of the modification approval letter. We considered a submission to be any of the following: initial application, protocol modification, renewal, a response to an IRB or R&D committee decision requiring application modifications, or a response to any IRB or R&D request that involved a substantial amount of team effort.

We estimated the amount of staff time that was involved in the IRB approval process. In addition to the certified IRB professional, we employed a team of 3 other project coordinators (master’s degree level) and 2 other research assistants (bachelor’s degree level) who spent a portion of their overall work effort on IRB- and R&D-related tasks. To calculate the cost of these human resources, we multiplied staff time by staff salary, including benefits.

We compared characteristics between participating and nonparticipating facilities by using Mann–Whitney U tests. Using methods published elsewhere (16), we summarized the complexity of the patients cared for at each study site (where a higher complexity index corresponds to a more complex patient), number of resident slots per 10 000 patients to assess each facility’s academic mission, number of hospital operating beds to quantify facility size, and number of hospital beds in the community to distinguish between urban and rural areas. Analyses of facility characteristics were performed by using Stata, version 11.2 (StataCorp, College Station, Texas).

Role of the Funding Source

Veterans Affairs Health Services Research and Development and the National Heart, Lung, and Blood Institute provided funding for this study. The study sponsors played no role in the design, conduct, or analysis of the parent trial or this record review, nor did they have any role in the preparation, review, or approval of the manuscript.

Results

The original premise of the study was to determine whether financial incentives to physicians could improve the translation of the findings from the ALLHAT (Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial) trial into outpatient practice. Specifically, the study planned to incentivize the ALLHAT findings about the effectiveness of thiazide diuretics in lowering blood pressure in most hypertensive patients and the recommendations that blood pressures below 140/90 mm Hg (<130/80 mm Hg in patients with diabetes) be considered controlled. In the summer of 2006, one IRB stated that their Office of General Counsel was concerned that the study involved antihypertensive medications. For the performance measures to remain consistent across sites and thus to preserve the internal validity of the study, we submitted this modification as a protocol amendment to all 13 sites that had approved the study, and we revised and resubmitted the initial application to 1 site where the study was undergoing a review decision. We chose to make the changes because we were concerned that these issues might arise at other sites after the intervention began, such as during a protocol renewal, when changing the methods would have threatened the entire project.

Of the 17 IRBs to which we submitted the application, 15 required full board review and 2 granted expedited review. Institutional review board and R&D committee reactions to the study varied markedly by site. Some sites appreciated the novelty and timeliness of the proposal. The IRB at site 8 noted, “It is well known that significant (and expensive) research such as the ALLHAT trial often fails to translate into changes in provider behavior. Financial incentives are a proposed mechanism for facilitating this translation, and it is important to evaluate them prior to wide-spread adoption.” The R&D committee at site 4 found the study to be “an interesting and exciting project.” However, the intervention’s novelty caused concern at some sites. The site 9 IRB questioned, “Is this legal for a research study? If legal, it seems to lead to unethical behavior similar to paying finder fees.” The IRB at site 2 said, “Offering money to people to do what is expected of them is not ethical.” The study team responded to these concerns by noting that several public and private health plans already are implementing “pay-for-performance” models, yet there are few empirical studies of their effectiveness (13). Sites 2 and 9 ultimately approved the proposal.

The IRBs at sites 11 and 13 granted approval by means of expedited review. At site 13, the R&D committee then tabled the protocol, stating, “Address why patients are not being told (through written consent) that their physicians were being paid to follow a specific protocol for their care.” The study team explained that the physician was being incentivized for providing high-quality care in accordance with national guidelines and that, because each assessment of the physician’s care delivery was based on a random subset of his or her hypertensive patients, it was not feasible to obtain patient consent beforehand. Because the IRBs at several other sites expressed similar concerns about patient awareness of the study, the study team agreed to notify patients that their physician may be participating in this study by placing flyers in the clinic area of these sites. The IRB at site 3 insisted that patients be informed
individually if their physician was participating, despite our concerns about breaching physician confidentiality and introducing bias through patient activation (because a patient questioning his or her physician about treatment may affect the care provided). This IRB ultimately disapproved the study, stating, “The potential risk to hypertensive patients is too great to justify their involvement.” The site 13 R&D committee also ultimately disapproved the study; ironically, despite an expedited IRB approval at this site, the application had to be formally closed. Despite the variability in initial reactions to the study protocol, among the 15 sites at which the IRB approved the study, 14 categorized it as minimal risk; the remaining site’s IRB did not determine its risk category.

Fourteen (82%) of the 17 submitted applications received the IRB and R&D approvals required for implementation. Additional submissions were required at all 17 sites, for a total of 80 submissions before the study application either was approved at or, if not approved, withdrawn from all 17 sites (median number of submissions per site, 4; mean, 5; range, 2 to 10). Among the 14 sites where the application received full approval, 35 additional submissions were required to approve the protocol modification, resulting in a total of 115 submissions before the study could be implemented (median number of submission per site, 6; mean, 7; range, 4 to 14). Among the 14 sites that received full approval, the number of days required for initial approval plus the number of days required for approval of the protocol modification ranged from 57 to 400 days per site (median, 168; mean, 181). There were no significant differences between VA- and university-affiliated IRBs in the average number of submissions per site, the average time from submission to approval, the percentage of sites receiving IRB approvals, the percentage of sites receiving R&D committee approvals, or the percentage of sites where the protocol received expedited review (Table 1). Most IRBs required paper submissions; all 3 IRBs with electronic submissions were university-affiliated.

Figure 2 shows the number of days from initial submission to approval, arranged by date of initial submission, for the 13 sites that approved the study before the modification was submitted. In July 2005, we submitted the application to the first site (site 1), which approved the study in 36 days. We submitted the application to the final site (site 17) in August 2006; it was approved in 76 days. The shortest time to approval (36 days) occurred at sites 1 and 7. Sites 9 and 10 took the longest to approve the study (198 and 230 days, respectively), suggesting that the protracted approval process was not due to a “learning curve” on the part of the study team. The total time spent in the IRB approval process before the study could be implemented, from initial submission to the first site to approval of the protocol modification at the final site, was 827 days, or more than 27 months. This is 21 months longer than we had proposed and 23 months longer than the time for which we had received a budget (Figure 3). Staff spent an estimated 6729 hours working on IRB- and R&D-related tasks, costing approximately $168 229 in salaries. This estimate does not include the salary for the PI or site PIs.

We began recruiting participants at a site as soon as its IRB fully approved the protocol modification. By the time the study was approved at all sites, 7 physician participants had withdrawn because of a position change, transfer, maternity leave, or retirement. In December 2007, during random assignment of sites to study groups, the IRB for 1 of the sites where we had exceeded our physician recruitment goal shut down, preventing the study from continuing at that site. We had to replace that site with another where we had not met our recruitment goal; after 4 more months of recruiting, we met the goal at the replacement site.

Of the 25 sites initially eligible for inclusion in the study, only 12 ultimately were included. The average patient complexity index at included facilities was significantly greater than that at the 13 excluded facilities ($P = 0.017$) (Table 2). Included sites also had a significantly greater mean ratio of medical resident slots to 10 000 unique patients ($P = 0.004$) and a significantly greater average number of hospital beds in their community ($P = 0.005$) than sites that could not be included, suggesting that included facilities were more urban.

**Discussion**

The Department of Health and Human Services recently called for public comment on human subjects research protections (12). Our experience suggests that this

---

**Table 1. Submission Process and Results, by IRB Structure**

<table>
<thead>
<tr>
<th>Site Characteristic</th>
<th>IRB Structure</th>
<th>P Value*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>VA</td>
<td>University-Affiliated</td>
</tr>
<tr>
<td>Sites where applications submitted</td>
<td>Total, n</td>
<td>8</td>
</tr>
<tr>
<td>Received expedited IRB review, n (%)</td>
<td>0</td>
<td>2 (22)</td>
</tr>
<tr>
<td>Approved by IRB, n (%)</td>
<td>8 (100)</td>
<td>7 (78)</td>
</tr>
<tr>
<td>Approved by IRB and R&amp;D, n (%)</td>
<td>8 (100)</td>
<td>6 (67)</td>
</tr>
</tbody>
</table>

IRB = institutional review board; R&D = Veterans Affairs Research and Development; VA = Veterans Affairs.

*From 2-tailed Fisher exact test (binomial data) or 2-tailed t test (continuous data).
†Number of days from the initial submission to the approval of the application at each site plus the number of days from the submission of the protocol modification to its approval.

---
The human subjects review process at 17 sites involved 115 submissions, consumed more than 6700 staff hours, and lasted almost 2 years longer than planned. The time to initial approval was shortest for the first submission, and the longest times occurred at the midpoint of the submission process, suggesting that the protracted approval process was not due to a “learning curve” on the part of the study team. This process greatly impacted the trial. First, changes required at 1 site necessitated a protocol modification to all sites to preserve the study’s internal validity. Second, the approval process had a profound financial effect on the project, costing close to $170 000 in staff salaries. Third, delays in approval affected participant recruitment and retention; 7 physician participants had left their primary care setting before all approvals were received. Finally, requirements for local site PIs and for IRB and R&D committee approvals effectively resulted in inclusion of more highly affiliated, urban sites that were treating more complex patients, potentially affecting the external validity (generalizability) of the study findings.

All 14 IRBs that approved the study and provided risk determinations classified the study as minimal risk, making the time and costs involved in the review process seem incongruous, especially when compared with studies done...
in other research disciplines. For example, genome-wide association studies routinely use more than 100,000 single-nucleotide polymorphisms to genotype individuals, yet an individual can be uniquely identified with access to fewer than 100 single-nucleotide polymorphisms (17). Surprisingly, according to the Office for Human Research Protections, these data are not considered identifiable and no IRB oversight or informed consent is mandated, nor does the Health Insurance Portability and Accountability Act necessarily provide protection for participants (17).

To our knowledge, this is the first health services research study to examine the IRB process for a randomized, controlled trial of an intervention to improve the provision of evidence-based care and the first to quantitatively evaluate the effect of human subjects requirements on the external validity of study findings. We conducted a PubMed search of empirical studies of the IRB process in the implementation of multisite studies in the United States. Other studies have documented marked site-to-site variation in the time from submission to approval (2, 3, 6, 8, 18–20), in the number of resubmissions required (3, 6, 19), and in IRB review decisions (2, 3, 6, 8, 11, 18–21). Two studies also estimated the costs involved; one cited $17,000 spent on coordinating center personnel, space, and supplies for an 8-site study (21), and a 14-site study estimated that staff salary spent on the IRB process cost more than $53,000 within the first year (8). We found that salaries for the staff involved in securing IRB approvals for this study (a process that took more than 27 months) amounted to almost $170,000. Several prior studies also have cited concerns about the generalizability of their research due to changes mandated by the IRB. In 1 survey of patients in a study on improving health care quality, opt-in and opt-out procedures imposed by several IRBs resulted in a loss of up to 37% of potential patients (2). The authors noted that such hurdles to participation may have disproportionately affected racial and ethnic minorities and low-income patients (2). In another study, changes imposed by the IRB resulted in a protocol that was not translatable into clinical practice, prompting the authors to ask, “Is it ethical to involve humans in research if the research is not likely to yield valid answers to the proposed research questions?” (22). They noted that the implied social contract between researchers and society is to maximize the impact of research by making the study as generalizable to clinical practice as possible (22). In our study, the 12 sites that ultimately participated in the study were more likely to be urban and academically affiliated and to care for more complex patients than excluded sites. External validity is a concern because of the expectation that health services and comparative effectiveness research will yield findings that are directly implementable and translatable into improvements in patient care (23). Although some sites were excluded because of their inability to fulfill requirements for IRB review, others were excluded because the IRB and R&D committee either disapproved the study or provided conflicting rulings. When multisite studies receive very different IRB determinations as we experienced in this study, regulations do not provide clear guidance on how to resolve conflicts (10).

Many of the variations in IRB processes are due to the system of local review, whereby a multisite study has to be evaluated by local IRBs to ensure that the protocol addresses any problems that might arise from local contexts (24). Some variation in review may be appropriate because of local values in assessing human subjects’ risks and benefits. However, many of the revisions requested by local IRBs, when compared with what was approved by the IRB of a multisite study’s coordinating center, have been shown to add little in terms of local context or essential protections and usually make few, if any, substantive changes to the study protocol (21, 24). Our experience confirms this finding. One underlying issue responsible for the type of local variation we had is that IRBs do not seem to agree on the limits of their sphere of human research protections and do not confine themselves to reviewing the ethical issues related to them (25). For example, 1 IRB required that we provide documentation of union approval and then asked whether we were providing any incentives to the institution itself.

Several authors have made recommendations for easing the burden of IRB review in multisite studies. These suggestions include increasing standardization of the review process across IRBs (1, 26); centralized IRB review, in which the coordinating center’s local IRB or an independent IRB reviews the protocol and takes responsibility for human subject protections for all sites (24, 27); and the use of a single, central IRB (3, 4, 7, 8, 21, 24). Although the last option may seem to be a logical answer to standardizing reviews, 1 study estimated that the cost of running the National Cancer Institute’s central IRB exceeded the amount of money it saved (28). One study has suggested several methods for streamlining the IRB process, includ-

**Table 2. Characteristics of Facilities That Were and Were Not Included in the Trial**

<table>
<thead>
<tr>
<th>Facility Characteristic</th>
<th>Included (n = 12)</th>
<th>Excluded (n = 13)</th>
<th>P Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mean patient complexity index (SD)*</td>
<td>1584 (1191)</td>
<td>541 (821)</td>
<td>0.017</td>
</tr>
<tr>
<td>Mean ratio of medical resident slots to 10,000 unique patients (SD)</td>
<td>3.0 (10.2)</td>
<td>8.0 (10.1)</td>
<td>0.004</td>
</tr>
<tr>
<td>Mean number of hospital operating beds (SD)</td>
<td>174 (118)</td>
<td>103 (70)</td>
<td>0.115</td>
</tr>
<tr>
<td>Mean number of hospital beds in community (SD)</td>
<td>6814 (5594)</td>
<td>2626 (2909)</td>
<td>0.005</td>
</tr>
</tbody>
</table>

* The patient complexity index is a measure of patient complexity based on the relative weight and frequency of each diagnosis-related group.
ing creating model IRB applications, starting the process and communication with the IRB before one’s application is selected for funding, maintaining that communication throughout the study, and being prepared for changes during the IRB application process (29). Although the authors met their timeline, they acknowledged that their procedures could have placed undue burden on the sites that ultimately were not selected to participate in the study and that the prospect of receiving funding could have caused sites to unduly influence their local IRBs (29).

Several limitations of our study must be addressed. First, the VA health care system uses a 2-stage review process. Research must be approved by both the local IRB and the local VA R&D committee before implementation. However, multisite studies in non-VA settings (24), including minimal-risk and health services research studies (1, 2), have also reported encountering substantial costs and delays due to the IRB review process. Second, the trial tested a novel intervention, providing financial incentives for high-quality care. Although noninvasive, the lack of a precedent may have prompted regulatory boards to err on the side of caution in granting approval. Finally, the data from this study are derived from a single multicenter trial involving regulatory submissions to only 17 sites.

Our study shows that obstacles presented by the IRB review process exist even for a minimal-risk health services research study using a randomized, controlled design. Institutional review board rulings that affect study design can threaten the internal validity of a study, and the barriers to obtaining IRB approval may favor studies taking place at highly selected sites that do not necessarily reflect health care delivery in most of the United States. Furthermore, the Office for Human Research Protections regulations seem inappropriate for minimal-risk studies (10). An overall review of the standards for research as planned by the Department of Health and Human Services is welcome.

From the Houston Veterans Affairs Health Services Research and Development Center of Excellence, Michael E. DeBakey Veterans Affairs Medical Center, and Baylor College of Medicine, Houston, Texas.

Disclaimer: Laura A. Petersen, MD, MPH, had full access to all of the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Acknowledgment: The authors thank Eric J. Thomas, MD, MPH, for his comments on an earlier version of the article.

Grant Support: The Veterans Affairs Health Services Research and Development Investigator-Initiated Research 04-349, National Institutes of Health R01 HL79173-01, and Houston Veterans Affairs Health Services Research and Development Center of Excellence HFP90-020 funded the design and conduct of the study; the collection, management, analysis, and interpretation of the data; and the preparation of this manuscript.

Potential Conflicts of Interest: Disclosures can be viewed at www.acponline.org/authors/icmje/ConflictOfInterestForms.do?msNum=M11-2377.

Requests for Single Reprints: Laura A. Petersen, MD, MPH, Health Services Research and Development (152), Michael E. DeBakey Veterans Affairs Medical Center, 2002 Holcombe Boulevard, Houston, TX 77030; e-mail, laurap@bcm.edu.

Current author addresses and author contributions are available at www.annals.org.

References

Current Author Addresses: Dr. Petersen, Ms. Simpson, Mr. SoRelle, Ms. Urech, and Ms. Chitwood: Health Services Research and Development (152), Michael E. DeBakey Veterans Affairs Medical Center, 2002 Holcombe Boulevard, Houston, TX 77030.