

Cluster Randomized Trials

Opportunities and Barriers Identified by Leaders of Eight Health Plans

Kathleen M. Mazor, MS, EdD, James E. Sabin, MD,†‡§ Denise Boudreau, PhD,¶
Michael J. Goodman, PhD,|| Jerry H. Gurwitz, MD,* Lisa J. Herrinton, PhD,**
Marsha A. Raebel, PharmD,††‡‡ Douglas Roblin, PhD,§§ David H. Smith, RPh, PhD,¶¶
Vanessa Meterko, BA,* and Richard Platt, MD, MS‡§*

Background: Cluster randomized trials (CRTs) offer unique advantages over standard randomized controlled clinical trials (RCTs) and observational methodologies, and may provide a cost-efficient alternative for answering questions about the best treatments for common conditions.

Objectives: To describe health plan leaders' views on CRTs, identify barriers to conducting CRTs, and solicit recommendations for increasing the acceptability of CRTs.

Research Design: Qualitative in-depth telephone interviews with leaders from 8 health plans.

Subjects: Thirty-four health plan leaders (medical directors, pharmacy directors, Institutional Review Board leaders, ethics leaders, compliance leaders, and others).

Measures: Qualitative analysis of interview transcripts to identify barriers, factors influencing leaders' views, ethical issues, aspects of CRTs that appeal to leaders, and recommendations for increasing acceptability of CRTs.

Results: Multiple barriers were identified, including financial costs, concerns about stakeholders' perceptions of CRTs, impact on physicians' prescribing habits, and formulary changes. Most leaders recognized the potential value of studying the comparative effectiveness of therapeutics, and many stressed the need for head-to-head trials. Leaders' views would be influenced by variations in study design and implementation. Recommendations for increasing

acceptability of CRTs included ensuring that the fiscal impact of a CRT be budget neutral, and that researchers educate stakeholders and decision-makers about CRTs.

Conclusions: Overall, health plan leaders recognized the need for studies of the comparative effectiveness of therapeutics under real world conditions, and many expressed support for CRTs. However, researchers seeking to conduct CRTs in health plans are likely to face numerous barriers, and preparatory work will be essential.

Key Words: cluster randomized trials, comparative effectiveness

(*Med Care* 2007;45: S29–S37)

Cluster randomized trials (CRTs) are characterized by randomization at the level of the cluster or group.^{1–3} Physicians, practices, health plans, or even geographic regions (eg, states) can be defined as clusters. In a CRT, all individuals within a given cluster are assigned to the same study arm. CRTs can offer unique advantages over standard randomized controlled clinical trials (RCTs) and observational methodologies for certain research questions under certain circumstances. In the context of studying therapeutics, an important advantage of CRTs is that they typically focus on effectiveness, evaluating outcomes under conditions of actual use. CRTs can also offer considerable cost and time efficiencies when implemented in health plans that have extensive existing information about members and their treatments and outcomes, along with an existing research infrastructure. Although clustering imposes costs in terms of statistical power (eg, limited degrees of freedom and variance inflation), the benefits associated with CRTs can outweigh these limitations in some circumstances. Although CRTs generally require larger sample sizes than RCTs to achieve statistical power, the cost per subject in CRTs is often less than in RCTs, thereby allowing larger CRTs, and sufficient power. For instance, if the intervention is a change to a formulary, guideline, or policy, it may not be feasible to recruit and consent each patient individually. Further, in many health plans the necessary data collection is ongoing and routine; when that is the case, the cost and time efficiencies noted above can be realized, and the cost of the study will

From the *Meyers Primary Care Institute, University of Massachusetts Medical School, Fallon Clinic Foundation, and Fallon Community Health Plan, Worcester, Massachusetts; Departments of †Psychiatry and ‡Ambulatory Care and Prevention, Harvard Medical School, Boston, Massachusetts; §Harvard Pilgrim Health Care, Department of Ambulatory Care and Prevention, Boston, Massachusetts; ¶Group Health Center for Health Studies, Seattle, Washington, ||[no affiliation], Woodbury, Minnesota; **Division of Research, Kaiser Permanente, Oakland, California; ††Kaiser Permanente Colorado, Clinical Research Unit, Denver, Colorado; ‡‡School of Pharmacy University of Colorado at Denver, Denver, Colorado; §§The Center for Health Research/Southeast, Kaiser Permanente, Atlanta, Georgia; and ¶¶The Center for Health Research, Kaiser Permanente, Portland, Oregon.

Supported by a supplement from the Agency for Healthcare Research and Quality to the Health Maintenance Organization Research Network Center for Education and Research in Therapeutics.

Reprints: Kathleen Mazor, MS, EdD, Meyers Primary Care Institute, 630 Plantation St., Worcester, MA 01605. E-mail: Kathleen.Mazor@umassmed.edu.

Copyright © 2007 by Lippincott Williams & Wilkins
ISSN: 0025-7079/07/4500-0029

be relatively low. If an entire health plan participates in a study, and the intervention is a formulary change or copayment change, all patients experience the intervention. Because of this, CRTs can yield results that are both generalizable and relevant for making “real world” decisions about changes in practice guidelines, formularies, and copayment tiers. Although CRTs have often been used to evaluate management strategies and public health interventions,^{4–8} they have been used less frequently to study the comparative effectiveness of therapeutics.^{9–11}

Acquiring information about comparative effectiveness under conditions of actual use is important to guide recommendations about treatment regimens and inform coverage decisions for these regimens. Currently, many questions remain about the best treatment regimens for many common chronic conditions. New medications are tested against placebo rather than existing therapies, and clinical trials providing direct comparisons of 2 or more medications are rare. Yet, direct comparisons are important. For example, the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT) demonstrated that direct medication comparisons can provide valuable and unexpected information. Before this study, angiotensin-converting enzyme (ACE) inhibitors, or calcium-channel blockers, were widely believed to be a more effective initial therapy than thiazides for hypertension. ALLHAT findings demonstrated that chlorthalidone, a thiazide, was as efficacious as amlodipine, a calcium channel blocker, or lisinopril, an ACE inhibitor, for initial therapy for hypertension.¹² This study had a large impact on the management of this common chronic disease.

The ALLHAT example also illustrates that although RCTs are the gold standard in clinical research, the costs in both time and resources can be prohibitive. The ALLHAT study extended over 8 years and the cost was more than \$80 million. Cost-effective alternatives for studying the comparative effectiveness of widely used therapeutic regimens are worth considering. CRTs offer such an alternative.

Although CRTs can offer significant advantages over both RCTs and observational methodologies to address some questions, they pose unique ethical, practical, and logistical challenges. Understanding these challenges is a prerequisite to designing CRTs that are acceptable to patients, prescribers, purchasers, and health plan leaders. We conducted a multisite study to assess views regarding CRTs among patients, providers, purchasers, and health plan leaders. Findings from the patient, provider, and purchaser interviews will be reported in separate articles. In this article, we report results of the in-depth interviews conducted with health plan leaders to ascertain their views on CRTs, identify potential barriers to conducting CRTs in health plans, understand how variations in study design and implementation would influence leaders’ views, and determine how CRTs can optimally be designed and presented to health plan leaders.

METHODS

Study Setting and Sample

The study was conducted in the context of the Health Maintenance Organization Research Network (HMO Re-

search Network) Center for Education and Research in Therapeutics (CERT). Eight sites participated. The HMOs associated with these sites serve geographically and ethnically diverse populations with a broad age range. All of the participating plans were located in metropolitan areas; the following geographic regions (as defined by the U.S. Census Bureau) were represented: Northeast, Midwest, South, and West. Enrollment of participating plans ranged from approximately 200,000 to more than 3 million; the combined population was more than 6 million members. Plan structures varied, and included group/staff models, and mixed models.

The study design called for interviews with 4 leaders at each site: a medical director, a pharmacy director, the head of the Institutional Review Board (IRB), and an ethics or compliance leader. At some sites, more than 1 person had a given title (eg, some organizations had multiple medical directors); in these instances, the CERT site investigator identified the person regarded as most appropriate to interview, given the purpose of the study. If an invitee was unable to participate, we asked for suggestions for an appropriate substitute. If more than 4 leaders were suggested or volunteered to participate from a given site, we accepted the additional volunteer as long as they were in a leadership role, or were designated as a substitute by a leader.

The study was reviewed and approved by the IRB of each participating site.

Conduct of the In-Depth, Semistructured Interviews

After interviewees agreed to participate, a short letter was sent confirming the date and time of the interview. A 1-page summary of key differences between CRTs and standard RCTs, and a fact sheet describing the study were included. Written informed consent was obtained where required by the site IRB.

All interviews were conducted via telephone by one of the authors (V.M.). The interviewer introduced herself, reiterated the purpose of the interview, and confirmed consent to participate. She then played a brief (3-minute) audio vignette. In the vignette, actors portrayed a health plan leader and her colleague discussing the health plan’s possible participation in a CRT. The key issues raised by speakers in the vignette dialogues are presented in Table 1. Vignettes were used to provide all interviewees with a common point of reference^{13,14} and a “listenable” description of CRTs.

We developed 2 versions of the vignette and assigned a version randomly before the interview. One version referred to a CRT comparing 2 selective serotonin-reuptake inhibitors; the other referred to a CRT comparing 2 antihypertensives (Table 1). After playing the vignette, the interviewer asked several open-ended questions, soliciting leaders’ reactions, concerns, questions, likely barriers, favorable views, and recommendations (Table 1). The interviewer also asked several “what if” questions, exploring responses to possible variations in design. The interviewer used the interview script as a guide, but because these were qualitative interviews, rather than standardized interviews, she was allowed to vary the question order, and occasionally question wording. For instance, if an interviewee raised an issue that would be

TABLE 1. Interview Content**Vignette Main Points**

- Both study drugs are approved, widely used, and on the market for “quite a while”
- No direct comparative data exists
- Plan would be randomly assigned to drug
- New users only to be included
- Informed consent mentioned but not resolved
- Doctors free to prescribe out of study
- Guidelines and formulary would depend on the arm assigned to
- Possible outcome measures mentioned (ie, adherence, hospitalizations)

Interview Main Points

- Reactions to the possibility of a cluster randomized trial in general
 - What is your initial reaction?
 - What concerns would you have?
 - What questions would you have?
 - What barriers would you foresee?
 - Anything you would be pleased about?
 - What recommendations would you have?
- Reactions to study variations
 - Would your opinion change if the study were changed so that ...*
 - it involved switching medications instead of only involving new users?
 - 1 drug is considerably less expensive than the other?
 - purpose of the study were to determine a cost effective alternative?
 - it involved 2 short-term drugs instead of 2 long-term drugs?
 - 1 study drug was a newer drug (“on the market for 1 yr”)?
 - comparison was between a drug and nondrug alternative?
 - practices or sites were randomized instead of health plans?
- Views on ethical issues
 - How would you feel about ...*
 - changing formulary to change physician behavior, but not informing physician (or patient)?
 - not informing patients of study?
 - need for informed consent?
 - reducing copayment to influence prescribing; is this coercive?

addressed in a later portion of the interview, the interviewer might further explore that issue, including posing the subsequent questions, at that time. Follow-up questions were asked as needed to clarify responses. The interviews were audio-taped and transcribed verbatim. The interviewer also collected demographic data (eg, highest degree, official title, years in position). Interviews lasted approximately 30 minutes.

Data Analysis

Qualitative analysis of the transcripts proceeded iteratively. The goals of the study and the interview questions provided an initial organizing framework. To begin the analysis, 1 investigator (K.M.) read a set of 5 transcripts and generated an initial list of themes and coding categories. A second investigator (J.S.) read 3 other transcripts and suggested modifications. Applying grounded theory,¹⁵ we repeated this process, with successive readings of the transcripts. We coded interviewee responses through application of the coding categories, with modifications suggested and discussed. This process continued until the interviewer and

both investigators agreed that the listing captured all major issues raised by interviewees.

We trained a research assistant to code the transcripts by question and code category and to flag all substantive comments that were not captured with this listing. After coding, we sorted all interviewee responses by question and response code and reviewed them for additional themes or subthemes. Two investigators (K.M., V.M.) reread the entire set of transcripts to ensure that no relevant themes had been missed. The major thematic categories identified were: potential barriers, impact of study variations, ethical issues, favorable views, and recommendations. The study design and analysis were not intended to provide quantitative data; rather, they were intended to provide a thorough and accurate description of the range of opinions and reactions expressed during the interviews.

RESULTS

A total of 34 interviews with health plan leaders were completed. The final number of completed interviews varied by site: at 1 site, we were able to recruit only 3 leaders; at 4 sites, we met our goal of recruiting 4 leaders; and at 3 sites, we exceeded our goal, and recruited 5 leaders. The completion rate for initial invitees was 74%. For leaders who declined, a substitute within the same plan was identified; 83% of the substitutes completed interviews. Characteristics of interviewees are presented in Table 2. No systematic differences in themes raised was found in comparing responses to the 2 vignettes, except that some subjects noted the existence of effective, relatively low cost treatments for hypertension, and implied that treatments for other conditions would be of greater interest.

Potential Barriers

Subjects identified numerous possible barriers to a CRT (Table 3). Some questions that were raised would not be answerable before instituting a study. For instance, some subjects worried that CRT participation could have a negative impact on patient satisfaction. This could not be known at the outset and might depend on the results of the study. Subjects were also concerned about financial costs, possible negative publicity, and physician resistance. Several noted that asking prescribers to change their practices for a study, and possibly change back afterward, could encourage “bad” prescribing habits, confuse prescribers, and undo ongoing efforts to encourage preferred prescribing. Several expressed concern about making formulary changes as part of the study. Subjects also raised questions that could be answered or negotiated during the planning phase. For instance, many would want to know the trial’s funding source, and suggested that pharmaceutical funding would be a potential barrier for some plans. A requirement for written informed consent would also be a potential barrier for some interviewees.

Impact of Study Variations

We asked subjects whether certain variations in study design or implementation would affect the acceptability of a CRT (Table 4). All respondents said that switching medica-

TABLE 2. Interviewee Characteristics

Characteristic	No. (%) N = 34
Role	
Medical director	9 (26)
Pharmacy director	8 (24)
Institutional Review Board leader	7 (21)
Ethics/compliance leader*	6 (18)
Other	4 (12)
Degree	
MD	15 (44)
PharmD	6 (18)
JD	4 (12)
PhD	2 (6)
Other	7 (21)
Ever worked as a clinician?	
Yes	24 (71)
No	10 (29)
Ever been involved in conducting research?†	
Yes	23 (68)
No	11 (32)
Gender	
Male	20 (59)
Female	14 (41)
Age	
<40	2 (6)
40–49	12 (35)
50–59	16 (47)
60–69	3 (9)
Declined	1 (3)
Race/ethnicity	
White	31 (91)
Asian/Asian-American	2 (6)
Declined	1 (3)
Total interviewees	34

*Ethics and compliance leaders were distinct from IRB leaders. Specific position titles varied but most included the word *ethics* or *compliance* (eg, “Compliance Officer”).

†Time spent involved in research ranged from 3 months to more than 25 years.

tions for patients already on therapy would be unacceptable or extremely difficult; however, for all other variations presented, views were mixed. Subjects tended to discount the potential impact of large cost differentials between the study drugs, and instead stressed the importance of whether the plan was already using a cost-effective therapy for the condition. They pointed out that if existing evidence suggested that 2 or more drugs had comparable efficacy and safety profiles, and that 1 drug was significantly less expensive, then the plan would already be using the less expensive alternative. As 1 subject said, “At least part of the motivation would be that one of the arms of the study had a less expensive alternative than the one we have right now.” It was also noted that if the more expensive drug was more effective, evidence of that would be valuable. A number of subjects would be less receptive to participating in a study that involved a newer drug (ie, a drug that had been on the market for only 1 year),

because of concerns that previously unknown problems with the newer drug might occur over the course of the study.

Ethical Issues

Informing Physicians and Patients of the Study

During the interview, the interviewer suggested the possibility that a CRT might occur without the physicians and patients being informed. Subjects expressed varying opinions on this possibility. Most respondents thought it was unethical and dishonest not to notify physicians. Subjects predicted that if physicians were not informed of the study at the onset, they would learn about it eventually and would respond with mistrust, anger, and confusion. The health plan would lose credibility, and the relationship between the physicians and the plan would be damaged. One subject said, “I could see where you might want to say that this is going to be our preferred drug and one of the reasons that it is preferred is because we’re trying to determine whether it is more highly effective. That sounds more ethical to me.”

Some subjects who opposed the informing of physicians and patients noted that doing so could undermine the study or introduce bias. A number of subjects (including IRB heads) thought that informing patients and providers might be unnecessary under some conditions. One subject noted, “We would say, ‘We’re moving this drug to the second tier.’ We don’t tend to give people a reason why we do that . . . When we put out our formulary updates, we don’t say, ‘We are moving a drug from 3 to 2 because the company ponied up another 5%.’ We say, ‘This drug is now going to be on the preferred tier with the others.’”

Written Informed Consent

Subjects expressed mixed views on the need to obtain written informed consent. Five of the 7 IRB leaders thought that written informed consent would not be needed under some conditions (eg, as long as both treatments constituted standard care); 1 IRB leader felt that written informed consent would be required; and 1 was noncommittal. One ethics leader said, “If there is no evidence-based reason for starting someone on 1 agent versus another, then it comes down to preference or custom. I do not see a call for each person to sign an informed consent because they are not taking experimental medication and the care process isn’t being distorted.” Another subject thought that informed consent would not be required for a CRT any more than it would be when a clinical guideline was implemented. Some subjects noted that a CRT might be viewed as a quality assurance effort rather than research. One subject commented that requiring signed informed consent would be “totally self-defeating” and another said, “The clear value of the cluster-randomized trial is that it’s an organizational decision—it takes it out of [the physicians’] hands so they don’t have to be pushing it to the individual patient.” These findings, which might seem surprising in light of the well-known emphasis on informed consent for experimental research, reflects uncertainty among the leaders as to whether to regard CRTs as “experiments” (which would require informed consent) or as akin to clinical guidelines or a tiered formulary (which would not). The

TABLE 3. Potential Barriers

Major Themes	Description
Impact on satisfaction	Members, physicians, formulary committee members, and purchasers might become dissatisfied or displeased because of study.
Physician resistance	Physicians may object to constraints on prescribing, especially constraints imposed because of a study, rather than new evidence; may also object if study imposes additional burden on physicians in terms of time or effort.
Impact on public relations	Plan participation might be viewed negatively by the public; impact could depend on study results (eg, negative perceptions more likely if plan were assigned to the drug that was ultimately found to be less effective); perceptions of pharmaceutical company involvement could also have negative impact; potential for positive publicity also noted.
Formulary changes	Formulary changes viewed as undesirable, difficult, impossible to implement quickly, and sources of dissatisfaction. Formulary decisions are made only after a careful review of evidence and reasoned deliberations; changes are not arbitrary and are not made lightly. Contracts with pharmaceutical companies may limit ability to implement formulary changes, and may affect cost of participation. Impact of formulary changes on prescribing viewed as questionable.
Financial costs	Costs include costs of administering the study, costs of the drugs, and costs of offering a reduced copayment (if that was part of the study design). Participation could also affect pharmacy rates and price negotiations, including rebates or contracts on other agents. Costs would not necessarily cease at the close of the study—additional costs could accrue with further changes to the formulary, or with continuing patients on a study drug after the study ended. Loss of membership due to dissatisfaction related to study could also cost plan. Patients and employer groups might also incur additional costs.
Impact on prescribing practices	Asking physicians to change practices for a study, and possibly change back at the end of the study, could encourage “bad” prescribing habits, and undo substantial work on the part of the plan to encourage and sustain appropriate prescribing. Changes could also confuse physicians; confusion could generalize beyond the study drugs.
Motivation for study	Is the purpose of study scientific (not veiled marketing)? Is there a real absence of data?
Treatment comparability	No evidence that the study drugs differ in terms of efficacy and side effect profiles.
Funding source	How will the study be funded? Is funding from a nonindustry source?
Study design	Doubts that cluster randomized trial would overcome recruitment and self-selection challenges, and take into account or control for differences between plans or practices. Favor observational study or standard randomized controlled trial.
Outcome measures	Concern that the outcome measures be meaningful, relevant, appropriate, and sufficient to confidently answer the study question and be worth involvement in the study.
Study details	How many patients would be involved, for how long? Would physicians have flexibility to prescribe an alternative medication? What data would be collected and how? Would existing administrative data systems be sufficient? How would results be reported?
Burden	Would participation involve additional paperwork, administrative effort and or physician effort?
Need for informed consent	Would informed consent be required? How would consent be determined? Requiring informed consent would be burdensome; not requiring it might be unethical. (Informed consent is discussed in detail under Ethical Issues.)
Potential conflicts	How will study procedures fit with existing regulations, guidelines, standards, and contracts?

variation found in the leaders' responses mirrors the range of perspectives expressed in the ethical analyses of CRTs that have been done to date.^{16,17}

A number of subjects thought it would be appropriate to notify the patient of the study, but that written informed consent would not be necessary (assuming the IRB concurred). The physician could notify the patient at the time of prescribing, or general notifications to health plan members could be issued or posted. Some subjects thought that physicians should have discretion about informing the patient, and others were uncertain about the need for informed consent and would defer to the IRB or ethicists.

Other factors identified as affecting the need for informed consent included the risk to the patient, the equivalence of the drugs (based on current evidence), the level at which randomization occurred (some subjects

thought informed consent would be required if randomization occurred at the practice level, but not required if randomization occurred at the plan level), the question of whether medications would need to be switched, the issue of whether the drugs had been approved for the indication, and the risk to the plan.

Adjusting Copayments

Subjects were asked about reducing copayments to encourage use of the preferred agent. Some subjects judged this to be ethical and appropriate, but they were split as to whether it would be an effective way to influence prescribing. One subject commented, “I think in general physicians are still pretty confused about what patients actually pay . . . Even within a health plan there are many pharmacy benefits that people have.” Other respondents worried that copayment

TABLE 4. Impact of Variations in Study Design and Implementation

Variation	Leaders' Responses
Switching drugs (vs. new starts only)	More problematic; objectionable. More difficult to implement; informed consent critical; potential harm to patients; greater financial costs; time-consuming explanation to the patient, patients and providers would object.
Cost differential (1 drug is much less expensive; goal is to determine whether less expensive is as effective)	<p>Would not be a significant factor; medication effectiveness and safety would be the overriding concerns. Drug prices can change dramatically over time, limiting the impact of large cost differentials at the time of the study.</p> <p>Attractive, but costs are only 1 of several issues that would need to be considered. More expensive drug could be found to be more effective, with financial implications after the close of the study (ie, if the plan then switched to favor the more expensive drug).</p> <p>Status quo important: does the plan currently have a cost-effective alternative?</p>
Compare drug to nondrug alternative (vs. drug-drug comparison)	<p>Possibly acceptable if the treatments were reasonably expected to be equivalent in effectiveness; would depend on the specifics of the alternatives and whether the alternative was within accepted clinical practice guidelines.</p> <p>Cluster randomized trial may not be an appropriate design for drug/nondrug comparison.</p> <p>A drug/nondrug comparison would require different sorts and sources of data.</p> <p>Physician factors would be especially relevant (eg, physicians might resist drug/nondrug comparison, or have strong personal preference type of therapy); physician education would be particularly important.</p> <p>Important to have an "opt out" option.</p> <p>Nondrug alternative might impose a greater burden for the patient, and patients might be less likely to engage in the treatment requiring greater effort.</p>
Newer drug (1 drug on the market for only a year vs. both drugs on the market for "quite a while")	<p>More problematic. Would not have similar evidence-base for both drugs; 1 yr not sufficient to evaluate safety or side effect profile of newer drug.</p> <p>New (negative) information could come out on the newer drug once the study was underway. Some leaders noted that newer drugs are not typically included on their formulary.</p> <p>Studying the newer drug could be particularly important because less information would exist already.</p> <p>One year would be sufficient trial; including the newer drug would not be problematic.</p>
Short-term medications (eg, comparing 2 antibiotics vs. 2 long-term medications such as antidepressants or antihypertensives)	<p>Issues would be the same whether short-term or long-term drugs were involved.</p> <p>Comparison of short-term drugs would raise fewer issues; patients would be less likely to be concerned or to question the drug choice; easier process and an "easier sell."</p> <p>Including short term drugs would be more challenging. Harder to impact behavior through cost adjustments if there was only 1 medication purchase involved; assessing outcomes and determining whether there were in fact differences in effectiveness would be more difficult; physician education about the study could be even more important.</p>
Randomization at practice level (ie, units within a plan would be assigned to different study arms) versus randomization at plan level	<p>Randomization at the practice or site level would be more problematic than randomizing at the plan level. Logistical and practical issues more challenging (eg, patients and providers may receive/give care at multiple practices); study would be more difficult to implement; errors in implementation more likely; effective communication more critical; may violate contractual obligations and regulations (eg, varying copayments may be prohibited within a plan); may increase physician and patient confusion or dissatisfaction; practices may not be comparable (eg, may attract different types of patients).</p> <p>Randomization at the practice or site level would raise similar issues as randomizing at the plan level; would not be more or less acceptable.</p> <p>Randomization at the practice or site level would more acceptable than randomizing at the plan level. Some leaders had participated in studies involving randomization of practices previously. Could reduce the potential for negative publicity, since the plan would not be in the position of having <i>only</i> used the less effective drug. Differences in guidelines possibly feasible in some settings.</p>

adjustments could limit the generalizability of study findings (eg, adherence rates might be high at the lower copayment level implemented for the study, but could drop if the copayment were increased). It was also noted that automated systems used to manage copayments do not adjust well to special circumstances, and that this could be a barrier.

Some subjects simply did not like the idea of adjusting the copayments as part of the study, and characterized this approach as manipulative. Some expressed general negative opinions (ie, "Incentives are bad things"), whereas others explicitly referred to reducing copayments as coercive. Some

subjects would be influenced by the magnitude of the reduction and the length of time over which it was in effect. Others thought that the reduction of copayments should not be necessary. Subjects also questioned what would happen to patients who were started on a study drug that would become more expensive after the study was completed. As 1 subject said, "They are now on this drug that turns out to be not the cost-effective alternative, and now you have enticed them . . . What do you do at the end of the study with that patient and that copay?" Another subject questioned how reducing the copayment for a study

drug would affect copayments for nonstudy drugs. “You are making one drug your preferred drug; does that also mean you are going to take the other drug and make it your nonpreferred drug? If you do that, then you are penalizing the people who are on your previously preferred drug, which is now nonpreferred?”

A number of subjects, however, noted that copayment manipulation is standard practice. As 1 subject said, “That’s how we do it anyway. And that’s how the formulary works. I mean, a formulary is intended to demonstrate the clinical and economic value of a drug and provide incentives to the member to use the one with the greatest clinical and economic value.” Another said, “If it is [coercive], then what I do all day is coercion . . . I mean, one of my major responsibilities is to decide which drugs belong in which tiers, so if it is coercion then I spend all day being a ‘coercioner.’”

Favorable Views

Although subjects voiced many concerns about CRTs, most also saw their potential value in comparative effectiveness studies. One respondent said, “We all want to see this badly, badly.” Several stressed the need for head-to-head studies and noted how rarely these sorts of comparisons are currently done. One medical director said, “Most of the time [drugs] come on the market with a trial against a placebo or an entirely different class of drug . . . but head-to-head comparisons of different drugs in the same class would be a benefit.” Another subject viewed CRTs as “an appealing way to solve an issue that we have encountered a lot . . . typically what happens is the drug is released and it’s compared against nothing.” Subjects also noted CRTs’ potential for providing more generalizable findings than standard RCTs or single-site studies. Several felt a CRT could overcome recruitment and self-selection issues. One IRB leader liked “the idea of being able to answer a question in a meaningful manner without resorting to cumbersome randomization of patients.” Subjects also mentioned other advantages, including the potential to include large numbers of patients, the ability to collect data and make it available more quickly than in an RCT, and the potential for physicians to gain satisfaction through contributing to academic research.

Recommendations

Many subjects made implicit recommendations for increasing the acceptability of CRTs, and the interviewer also explicitly solicited recommendations (Table 5). Some of the most strongly and frequently expressed recommendations were that study drugs be equivalent as far as is currently known (ie, that there be no evidence of differential effectiveness), that outcome measures be sound and appropriate, that the study be scientific and objective, that there be a real absence of data, and that an important clinical question be addressed by the study. Subjects also stressed the importance of involving plan leaders, particularly formulary committee members, early in the process. Financial considerations were also important to subjects, as noted above. Many stated or implied that plans are businesses, and that the fiscal impact of a CRT would influence their

TABLE 5. Recommendations

Choose study drugs (treatment arms) for which there is no evidence of differential effectiveness or safety (ie, presumption of equivalence or comparability is reasonable).
Choose appropriate, sound, relevant outcome measures.
Clarify the need for the study (absence of data) and that study findings will answer an important clinical question.
Ensure that study is scientific, objective, and unbiased.
Differentiate from pharmaceutical industry study; no pharmaceutical industry involvement.
Conduct due diligence; get ethics and research expertise from around the country for this approach (ie, using cluster randomized trials to study comparative effectiveness of therapeutics).
Provide examples of prior, successful cluster randomized trials.
Cluster randomized trial proposal should come from a reputable source.
Garner support from external bodies, including government agencies (eg, Agency for Healthcare Research and Quality).
Gain buy-in of key stakeholders (formulary committee; employer groups; physicians).
Troubleshoot problems with plan leaders.
Provide full information; transparency about the study and the plan’s participation.
Provide feedback to the organization after the study.
Make plan “financially whole”; cover the costs of the study.
Minimize the burden to the plan and physicians (eg, paperwork, time, effort).
Provide incentives to physicians.
Make clear that physicians have flexibility to prescribe nonstudy drugs.
Provide education and training about cluster randomized trials to stakeholders (eg, leaders, Institutional Review Board members, physicians, reviewers).

support for participation. Subjects suggested that the plan be made “whole” financially and that participation costs should be covered. Other respondents strongly favored observational studies over CRTs and recommended avoiding CRTs altogether.

DISCUSSION

The 34 health plan leaders interviewed for this study expressed diverse opinions on the use of CRTs for studying the comparative effectiveness of therapeutics. None expressed unqualified opposition to CRTs, but a small number seemed unconvinced that this approach would provide significant benefits over observational designs and preferred the latter. More than 1 subject suggested that researchers use observational techniques instead of CRTs by seeking out health plans with preexisting differences between their formularies and prescribing practices, and comparing outcomes for these plans. At the same time, many of the subjects participating in this study were keenly aware of the potential value of CRTs and of the need for head-to-head studies.

All subjects identified potential barriers to using CRTs to study the comparative effectiveness of therapeutics in health plans. Concerns about financial costs were prominent. Subjects noted that health plans are businesses and that it would be important for the plan to be made whole financially with respect to its participation. Subjects were also concerned about how stakeholders, including patients, physicians, and purchasers, would respond to the plan’s participation. Sub-

jects stressed that study participation might affect their ability to meet existing guidelines, commitments, and contracts. Concerns about noncompliance involved ethical, as well as practical and legal, considerations.

Subjects would have many questions about any proposed CRTs, ranging from general questions about the need for the study to specific questions about implementation. Researchers seeking to conduct CRTs to study the comparative effectiveness of therapeutics will need to be prepared to answer leaders' questions and address their concerns.

Many leaders were unfamiliar with the concept of CRTs before our study. Education of all stakeholders, including IRB committees, health plan leaders, physicians, patients, and the public, will be important. One subject said, "We make decisions now in rather odd ways . . . This approach [could be] the counterbalance for that and so if people can begin to sort of look behind the curtain and see that this is really needed . . . It's an educational challenge for the community to understand why this approach is warranted or is important."

The ethical concerns raised in this study are important, and a comprehensive ethical analysis would help to address them. Some subjects perceived the ethical requirements for CRTs as distinct from those for standard RCTs. For instance, most IRB leaders thought written informed consent would not be required in a CRT. Views on changing the formulary as an intervention to influence prescribing were also diverse. Many subjects recognized that health plans routinely implement formulary decisions and influence prescribing, and some thought that a study-related formulary change might be considered a possible quality improvement effort. For other subjects, connecting a formulary change to a research study would be problematic, and all requirements typically associated with the conduct of research would need to be met.

This study focused on practical and ethical issues surrounding CRTs, but researchers using CRTs must be cognizant of other issues as well.^{2,18} The decisions of whether to randomize at the cluster level and how to define the cluster involve consideration of scientific, practical, and ethical issues.^{2,16} In addition, recent reviews suggest that studies using CRTs often do not attend to critical analytic issues.^{19,20} Statistical guides need to be developed to facilitate sample size calculation for regimens with various degrees of difference in outcome, taking into account the fact that actual practice will vary considerably as a result of the intention-to-treat design.

An important strength of this study is the participation of leaders from 8 health plans that varied in geographic location, ethnic and racial diversity of membership, organizational structure, and management practices. The study does have some limitations, however. Beginning the interviews with a vignette ensured that the subjects had a common point of reference, but may also have influenced subjects' responses by mentioning certain issues. As in any qualitative interview, the interviewer may inadvertently have influenced subjects' responses. In addition, different investigators might have identified different themes.

Overall, health plan leaders seem to appreciate the need for real world, head-to-head comparisons of therapeutics, but significant barriers exist. Logistical barriers such as costs, although relatively straightforward, are significant. Ethical issues are more complex, but responses of IRB heads and ethics leaders suggest that these are surmountable. A comprehensive ethical analysis may increase the acceptability of a CRT to health plan leaders. Health plan leaders' unfamiliarity with CRTs suggests that education and examples of successful CRTs may also increase acceptability. In addition, studies are underway to understand how patients, prescribers, and purchasers would respond to CRTs.

Researchers seeking to conduct CRTs in health plans are likely to face numerous barriers, and preparatory work with plan leaders will be essential. However, most leaders are acutely aware of the need for direct comparisons of therapeutics, and recognize the potential value of CRTs. Many would support participation in a CRT, if the study were appropriately framed, designed, sponsored and administered. We hope that the findings reported here will provide a foundation for subsequent studies on the use of CRTs in comparative effectiveness studies of therapeutics for the treatment of common chronic conditions.

REFERENCES

1. Cornfield J. Randomization by group: a formal analysis. *Am J Epidemiol.* 1978;108:100–102.
2. Donner A, Klar N. *Design and Analysis of Cluster Randomization Trials in Health Research.* London, England: Arnold; 2000.
3. Murray DM, Varnell SP, Blitstein JL. Design and analysis of group-randomized trials: a review of recent methodological developments. *Am J Public Health.* 2004;94:423–432.
4. Martiniuk AL, O'Connor KS, King WD. A cluster randomized trial of a sex education programme in Belize, Central America. *Int J Epidemiol.* 2003;32:131–136.
5. Ornstein S, Jenkins RG, Nietert PJ, et al. A multimethod quality improvement intervention to improve preventive cardiovascular care: a cluster randomized trial. *Ann Intern Med.* 2004;141:523–532.
6. Thomas RE, Grimshaw JM, Mollison J, et al. Cluster randomized trial of a guideline-based open access urological investigation service. *Fam Pract.* 2003;20:646–654.
7. Whelan T, Levine M, Willan A, et al. Effect of a decision aid on knowledge and treatment decision making for breast cancer surgery: a randomized trial. *JAMA.* 2004;292:435–441.
8. Yealy DM, Auble TE, Stone RA, et al. The emergency department community-acquired pneumonia trial: methodology of a quality improvement intervention. *Ann Emerg Med.* 2004;43:770–782.
9. Jick H, Jick SS, Walker AM, et al. A comparison of wax matrix and microencapsulated potassium chloride in relation to upper gastrointestinal illness requiring hospitalization. *Pharmacotherapy.* 1989;9:204–206.
10. Schneeweiss S, Maclure M, Carleton B, et al. Clinical and economic consequences of a reimbursement restriction of nebulised respiratory therapy in adults: direct comparison of randomised and observational evaluations. *BMJ.* 2004;328:560.
11. Maclure M, Carleton B, Schneeweiss S. Designed delays versus rigorous pragmatic trials: lower carat gold standards can produce relevant drug evaluations. *Med Care.* 2007;45(Suppl 2):S44–S49.
12. ALLHAT Officers and Coordinators for the ALLHAT Collaborative Research Group. Major outcomes in high-risk hypertensive patients randomized to angiotensin-converting enzyme inhibitor or calcium channel blocker vs diuretic: the Antihypertensive and Lipid-Lowering Treatment to Prevent Heart Attack Trial (ALLHAT). *JAMA.* 2002;288:2981–2997.
13. Finch J. The vignette technique in survey research. *Sociology.* 1987;2: 105–114.
14. Hughes R. Considering the vignette technique and its application to a study of drug injecting and HIV risk and safer behaviour. *Sociol Health Illn.* 1998;20:381–400.

15. Glaser BG, Strauss AL. *The Discovery of Grounded Theory: Strategies for Qualitative Research*. Chicago, IL: Aldine; 1967.
16. Hutton JL. Are distinctive ethical principles required for cluster randomized controlled trials? *Stat Med*. 2001;20:473–488.
17. Medical Research Council. *Cluster Randomized Trials: Methodological and Ethical Considerations*. 2002. Available at: <http://www.mrc.ac.uk/Utilities/Documentrecord/index.htm?d=MRC002406>. Accessed February 20, 2007.
18. Donner A, Klar N. Pitfalls and controversies in cluster randomization trials. *Am J Public Health*. 2004;94:416–422.
19. Donner A, Brown KS, Brasher P. A methodological review of non-therapeutic intervention trials employing cluster randomization, 1979–1989. *Int J Epidemiol*. 1990;19:795–800.
20. Varnell SP, Murray DM, Janega JG, et al. Design and analysis of group-randomized trials: a review of recent practices. *Am J Public Health*. 2004;94:393–399.