Issues in Improving Effectiveness Research

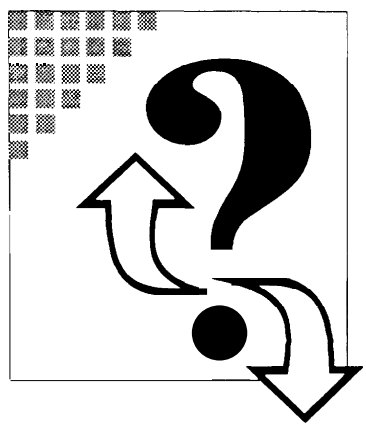
4

The federal effectiveness research effort has gone far in raising questions of the comparative effectiveness of existing technologies and strategies to manage health problems. It has been less successful at answering the questions it has raised.

This chapter addresses the issue of how the federal government can improve effectiveness research. To do so, it first discusses some of the major gaps in effectiveness research as it is currently carried out, and the barriers and possibilities in filling these needs. It then reviews the part the various federal agencies and departments play in this research effort, and how the roles of the different agencies affect the implementation of strategies to address problems in the current effort.

GAPS IN THE EXISTING FEDERAL RESEARCH EFFORT

As described in chapters 2 and 3, **the effectiveness research activities sponsored by the federal government have yielded valuable insights about the relationships between the outcomes and processes of care, but they have been less successful at making clear statements about the relative effectiveness of alternative medical technologies and services.** Among the clear gaps in the existing federal effectiveness research are:



1. **The lack of a systematic assessment of what has already been studied.** Despite the enormous and ever-increasing size of the medical literature, exhaustive reviews of past studies in areas such as treating back pain have sometimes found almost nothing useful. In some cases, however, systematic reviews have demonstrated the existence of unrecognized but relevant studies. A coordinated means of assessing the results of past

studies could help ensure that useless duplication of extensive reviews are reduced, while making better use of knowledge from past studies.

2. **The absence of valid comparative studies of existing technologies.** Effectiveness research has proven adept at raising appropriate qucs-(ions to study and fostering a climate conducive to comparative clinical research on existing medical technologies, but research to address these questions has until now received little real support or commitment from federal agencies.
3. **The inability to prevent the problem of poor evidence from accumulating.** The fast pace of biomedical research, and the relatively small proportion of new technologies ever exposed to rigorous testing before introduction, mean that our collective ignorance about the most effective technologies and strategies may be growing rather than declining.

Each of these areas raises its own issues and possibilities, discussed below.

Systematic Reviews: Making Use of Existing Knowledge

Making the most efficient use of health research resources requires first knowing what has already been studied. Sometimes our lack of knowledge regarding the safety and effectiveness of technologies is due not to a lack of studies but to our lack of awareness about them. The true tragedy of diethylstilbestrol (DES), described in chapter 2 (box 2-1), is not only that it ultimately proved very harmful but that it was never effective, and that its ineffectiveness could have been known from the beginning if clinicians had heeded the results of the more rigorous studies of the drug. Even some of DES harmful effects could have been detected, had contemporary analysts examined the results of their own studies more critically (106).

Systematic reviews of the literature, including meta-analysis, have proved to be a powerful tool of effectiveness research.] The contributions of systematic reviews are threefold. First, they have encouraged a more rigorous approach to defining and conducting a literature search and review than was the norm in the past, making reviews more reliable and providing a needed tool for managing the enormous size of the medical literature. Second, they have added strength to existing evidence, and sometimes added new findings to the existing evidence, through the quantitative reanalysis of previous research results. Third, they can demonstrate areas in which the existing literature is especially weak, an important criterion in targeting resources toward the research questions most in need of investigation.

A powerful demonstration of both the need for systematic reviews and the contributions they can make was a set of meta-analyses by Lau, Antman, and their colleagues, who examined the results of published trials of treatments for acute myocardial infarction (27,442). Their findings implied that thousands of lives have been lost because physicians did not know of, or did not believe the results of, studies that had already been done. Streptokinase, for example, was little used until the late 1980s, when the introduction of a higher priced, genetically engineered alternative kindled new interest in this older drug. Early trials of streptokinase were small and had contradictory results. These researchers showed that had the results of these small studies been combined in a meta-analysis, clinicians could have known by the end of the 1970s that streptokinase, administered soon after a heart attack, saved lives. Yet as late as 1984, most major textbooks and reviews of the field made no mention of the therapy, or argued against its use. Conversely, lidocaine is still being advocated as routine therapy in textbooks, even though 20 years ago a meta-analysis of published trials

[1] "Systematic review" here encompasses both meta-analysis and other comprehensive, highly structured literature reviews that are not able to combine quantitatively the results of individual studies (e.g., because the outcomes measured are too different).

would have raised major questions about its effectiveness (27).

Despite the rising popularity of meta-analyses, the conduct of systematic reviews is also often a frustrating, inefficient, and disjointed exercise. Agency for Health Care Policy and Research (AHCPR) has funded meta-analyses primarily through its Patient Outcomes Research Teams (PORTS), with a few additional methodological studies also receiving grant funding. The PORT reviewers were frequently frustrated with the considerable resources spent on extensive literature collection that nonetheless resulted in few useful studies (807). This experience suggests that better ways of identifying the relevant literature would be a great efficiency. Furthermore, it suggests that recording and updating such reviews where they have been done could prevent others from duplicating the effort.

It might be presumed that areas in which considerable randomonized controlled trial (RCT) activity is being undertaken would be promising areas for systematic reviews of previous trials. NIH conducts and sponsors many clinical trials, but it sponsors few formal research overviews or meta-analyses and almost no methodological activities on this topic. The National Institute for Child Health and Human Development (NICHD) does have one project to study the use of meta-analytic techniques for combining the results of nonrandomized studies, and it and a few other Institutes have one or two meta-analyses that are ongoing or recently completed, but this activity receives little emphasis overall.

Nor is it likely that many researchers conduct a formal meta-analysis with their own resources before proposing a clinical study. Those who have conducted meta-analyses report that the commitment required in terms of researcher expertise and researcher and computer time can be substantial (473). Encouraging the production of systematic reviews thus is likely to require at least some external support, as well as collaboration among a number of investigators.

In summary, meta-analyses and other systematic reviews are ways of making better use of existing knowledge, gaining new knowledge, identifying important questions for future research, and preventing the squandering of resources on previously researched questions. Such systematic reviews can be costly but at present have few sources of federal funding, and there is little to encourage researchers to conduct them before undertaking new research projects. Some of the costs and potential duplication in systematic reviews could be substantially reduced if review efforts were more coordinated, and if there were better mechanisms to help reviewers identify relevant studies more systematically.

The Cochrane Collaboration

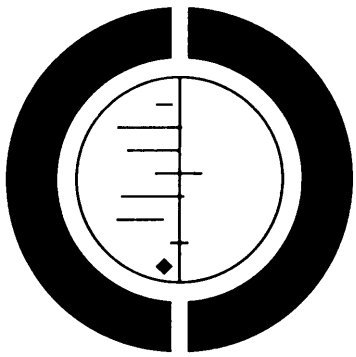
One response to the need for better understanding of what existing studies can already tell us has been the establishment of the Cochrane Collaboration, a remarkable international effort whose goal is to "prepare, maintain and disseminate systematic. up-to-date reviews of RCTs of health care, and, when RCTs are not available, reviews of the most reliable evidence from other sources" (13 1,666) (box 4-1).

The model for these collaborative reviews is a comprehensive review of interventions in pregnancy and childbirth, which includes systematic reviews of about 600 separate topics in the field (131). Reviewers participating in the group addressing the subject—about 30 individuals from 8 countries—prepare systematic reviews and update them as more trials on those topics are conducted.

A unique feature of the Collaboration is that the results of reviews, disseminated electronically, are not copyrighted (108). ". . .[A]lthough those contributing to the Collaboration are named in its electronically published output. the Cochrane Collaboration itself belongs to all of the contributors, collectively" (131).

BOX 4-1: The Cochrane Collaboration Logo

The Cochrane Collaboration, a cooperative international network of researchers, is dedicated to the preparation, maintenance, and dissemination of systematic reviews of the effects of health care.



The Cochrane Collaboration logo illustrates a systematic review of data from seven randomized controlled trials (RCTs) Each horizontal line represents the results of one clinical trial (the shorter the line, the more certain the result), and the diamond represents their combined results, The vertical line indicates the position around which the horizontal lines had similar effects, if a horizontal line touches the vertical line, it means that particular trial found no clear difference between the treatments. The position of the diamond to the left of the vertical line indicates that the treatment studied is beneficial.

This diagram shows the results of a systematic review of RCTs of a short, inexpensive course of a corticosteroid given to women expected to give birth prematurely. The first of these RCTs was reported in 1972. The diagram summarizes the evidence that would have been revealed had the available RCTs been reviewed systematically a decade later: it Indicates strongly that corticosteroids reduce the risk of babies dying from the complications of immaturity By 1991, seven more trials had been reported, and the picture in the logo had become still stronger This treatment reduces the odds of babies of these women dying from the complications of immaturity by 30 to 50 percent.

Because no systematic review of these trials had been published until 1989, most obstetricians had not realized that the treatment was so effective. As a result, tens of thousands of premature babies have probably suffered and died unnecessarily (and cost the health services more than was necessary). This is just one of many examples of the human costs resulting from failure to perform systematic, up-to-date reviews of RCTs of health care

SOURCE The Cochrane Centre, "The Cochrane Collaboration pamphlet, Oxford, United Kingdom, 1993

Six centers around the world, including one in the United States. have been established to support the reviewers who participate in the Cochrane Collaboration (108).[2] In addition to coordinating, compiling. and disseminating the reviews in general topic areas, these centers maintain registries of systematic review's undertaken by others (131). There are no central sources of funding for either the centers or the reviewers; all of the contributors to the Collaboration are responsible for finding their own sources of support. The only U.S. Cochrane Center thus far is located in Baltimore, Maryland, It is presently subsisting on a small one-year grant from NIH's Office of Medical Applications of Research (169).

Improving the Efficiency of Systematic Reviews

One of the most time-consuming tasks of performing a meta-analysis, or any systematic literature review. is the identification of all relevant studies to be reviewed (11 O). This task is also an inefficient one; different reviewers may each spend time trying to separately identify essentially the same studies.

The task of identifying published studies is made somewhat easier by the existence of MEDLINE®, an electronic database of the medical literature maintained by the National Library of Medicine (NLM) at NIH. Unfortunately, however. this database has several limitations that make it unreliable as a source to identify all, or even the great majority. of publishcd RCTs. Among its most prominent constraints are:

- It includes only citations to articles in the medical literature published after 1965.
- The 3,700 journals it covers represent less than 20 percent of all medical journals (and it includes few publications from related fields. such as health services research).
- The search heading used to identify RCTs (the main types of publications used in meta-analyses) was very restrictive before 1990 and did not identify the full range of trials of interest to reviewers.
- Even since 1990, RCTs are often not labeled as such on MEDLINE, because the persons entering the information on the database cannot tell easily from the published articles that they in fact are this type of study (680).

Authors can inadvertently compound the difficulties of conducting literature searches via MEDLINE. Articles in which the authors have made poor choices of key words, or have abstracts that do not clearly identify the article as an RCT, can be difficult for reviewers to identify in a MEDLINE search (865).

The extent of MEDLINE's limitations is demonstrated by a search of RCTs relating to vision treatments published in 66 journals in 1988. In this case, it was already known that all 66 journals were among those indexed on MEDLINE, so the retrieval rate using that database should have been very high. Of over 1,500 trials identified and examined, 201 were clearly randomized controlled trials. Another 18 turned out to be RCTs, but this was not obvious from the published articles and had to be confirmed in other ways. Of this total of 219 trials, 30 could not be identified using MEDLINE (168).

Literature searches of clinical trials can be even less successful if they are not restricted to recent trials published in journals known to be indexed on MEDLINE. On average, even searches conducted by an experienced medical librarian yield only about one-half of all relevant RCTs (173).

A major stride towards improving the efficiency of systematic reviews occurred in late December 1993. In a commendable example of a voluntary response to a clearly defined problem,

[2] The first Cochrane Centre was established in the United Kingdom, in O\ ford. England. The U.S. center is in Baltimore, Maryland. Other centers are 1 ocated in Canada, Denmark, 1 taly, and Australia.

NLM has committed its resources towards creating an augmented clinical trials database (173). This database will be parallel to MEDLINE, and searching MEDLINE for clinical trials will alert users to its existence. It will include:

- tags to all clinical trials already indexed on MEDLINE;
- abstracts of clinical trials published in journals held by NLM in its collection, but not currently indexed on MEDLINE; and
- clinical trials published before 1966.

The Baltimore, Maryland Cochrane Center is helping to coordinate the effort, which will result in an expanded database available to users beginning in 1995 (173).

Filling in the Knowledge Gaps

The State *of Comparative Effectiveness Trials*

Most of the effectiveness research sponsored as part of the federal government effectiveness research initiative has been descriptive, using administrative databases and other observational data to describe patient outcomes. A great disappointment of this research is that although it has identified important questions for comparative research, neither the funding nor the research communities have proved able to capitalize on this. Having created an environment potentially amenable to good comparative research studies, effectiveness research has been largely unable to carry out those studies. The deficiencies include comparative research on existing practices where uncertainty has been shown to exist; the broader incorporation of outcomes that measure patients' quality of life into RCTs; and more research in settings and on patients that are "ordinary," on questions that matter to patients and could further help them and their care providers make better decisions.

AHCPR has very few comparative effectiveness studies underway. The agency is contributing to a few RCTs sponsored primarily under the aegis of the Department of Veterans Affairs (VA) or NIH, and it has plans to fund a few more on its own through the PORT and other grant programs (82 1). The agency has also funded a followup case control comparative study to see whether the findings of the cataract PORT regarding retinal detachment can be confirmed (724,821). This study, and a few other small randomized and nonrandomized studies, comprise its investment in comparative effectiveness research. The agency does not consider that its current funding level permits much more than this (821).

NIH, in contrast, sponsors hundreds of clinical trials, but most of these are believed to be basic safety and efficacy trials of predominately new technologies. NIH does sponsor at least some comparative studies (both RCTs and nonrandomized studies) of existing technologies. Examples include:

- a comparative trial of alternative treatments for acute ear infections in children,
- a comparative trial of behavioral treatments for urinary incontinence in elderly persons,
- a study assessing the outcomes of temporal mandibular joint (TMJ) surgery,
- a large, simple effectiveness trial on the effects of digitalis on survival in patients with congestive heart failure, and
- a multicenter trial of treatment for early glaucoma (846,853).

The VA is another sponsor for a number of comparative effectiveness studies. The VA Medical Research Service's Cooperative Studies program, for example, has five large, multicenter randomized controlled trials that are ongoing or recently completed. All could be considered "effectiveness trials" in some sense, and all but one of them are cosponsored by other federal agencies. They are:

- the Prostate Cancer Intervention Versus Observation Trial (PIVOT) trial of early intervention for prostate cancer (cosponsored by AHCPR) (935),
- a trial to evaluate a new drug to reduce drug cravings in persons who are opiate dependent (co-sponsored by the National Institute on Drug Abuse) (978),
- a large trial of digitalis for heart disease (cosponsored by the National Heart, Lung and

Blood Institute and Burroughs Wellcome) (877),
- a continuing study of the role of zidovudine (AZT) in preventing progression of AIDS (co-sponsored by the U.S. Army Medical R&D Command) (876), and
- an evaluation the comparative effects of a num - ber of antihypertensive agents (497).

Neither NIH's nor VA's trials, however, are linked in any way to the priority areas for research on existing technologies that emerge from the descriptive "effectiveness research" work of AHCPR.

The inability of the existing research structure to carry out the full range of studies implied by "effectiveness research" is eloquently captured in the saga of clinical trials on treatments for benign prostatic hyperplasia (BPH), the noncancerous enlargement of the prostate (box 4-2). In this instance, descriptive effectiveness research sponsored by AHCPR raised specific questions about the relative effectiveness of common treatments for BPH. The clinical community came to accept the need for a comparative trial of the treatments and actually proposed the trial. Yet the trial went unfunded by AHCPR due to lack of money, and unfunded by NIH due to lack of interest. **One of the prime justifications for descriptive effectiveness research is to identify important research questions and illuminate medical uncertainty that would enable an RCT to take place, yet in the case where this has most clearly happened, the needed study has never materialized.**

Improving the conduct of comparative effectiveness research requires improving the way trials on existing technologies are conducted, so that the results of the trials are as broadly applicable and as relevant to patient and clinician decisionmaking as possible. Many improvements are possible; three that are particularly closely tied with the goals of effectiveness research are discussed in this chapter. They are incorporating broader measures of health outcomes in clinical trials, m'here relevant and possible: improving the public's knowledge of clinical trials, to broaden participation: and improving the research infrastructure so that large-scale, practice-based research becomes not only feasible but efficient.

Equally important to improving the conduct of effectiveness trials is improving the federal government's sponsorship of such research. Establishing high-priority questions to study, and improving the research infrastructure to study them, is useless if no federal agencies consider it one of their major responsibilities to support this infrastructure and fund research within it. This issue is discussed later in this chapter.

Incorporating Broader Outcome Measures

The topic of incorporating quality-of-life assessments in clinical trials has been the subject of three separate NIH workshops (847,848,852). Despite the variety of trials in which patient functioning or quality-of-life measures are used, however, these trials probably represent a minority of NIH-sponsored trials. and the proportion apparently varies among Institutes. The National Institute for Allergies and Infectious Disease, for example, estimates that "at least 10 percent" of its trials incorporate such measures. Several other Institutes report using such measures but list only a few examples, suggesting that these measures are not major endpoints in most trials (846).

A few of these trials incorporate generic quality-of-life instruments, such as the SF-36 and the Sickness Impact Profile. that incorporate the patient's self-assessment. These instruments have proved useful in enabling more consistent comparisons of disease and treatment impact across conditions, and in enabling the treatment-specific impacts of care on a patient life to be detectable even when the patient has multiple health conditions (see chapter 3). The National Eye Institute, for example, uses one or both instruments in at least three of its clinical trials, and several trials of AIDS treatments use a variation of the SF-36 adapted for that particular condition (846).

Most NIH trials that incorporate patient functioning or quality-of-life as an outcome measure. however. apparently use disease-spccif:c instru-

BOX 4-2: The Elusive Prostate Treatment Trial

The outstanding example of a comparative effectiveness trial that did not happen was the result of efforts to investigate alternative therapies for benign prostatic hyperplasia (BPH), the subject of one of the first four Patient Outcomes Research Teams (PORTS) funded by the Agency for Health Care Policy and Research (AHCPR).

Research results from the BPH PORT documented enormous variation in the rates at which physicians chose to treat this condition with early surgery (as opposed to "watchful waiting, " then surgery only if symptoms worsened) (918) Results from database analyses also raised questions about the relative effectiveness of new transurethral surgical procedures compared with traditional open surgery (920). Although early suggestions that the transurethral procedure was actually less safe were probably unwarranted (1 40), the research nonetheless raised significant questions about the effectiveness of alternative management strategies that prompted the urological community to consider a randomized study of alternative treatments for the first time.

In fact, the American Urological Association (AUA) proposed such a clinical trial and applied to AHCPR for trial support (41). The AUA also conducted a pilot study of 400 patients to demonstrate the feasibility of the idea (913).

AHCPR concluded that the study initially proposed was too expensive to be feasibly funded out of the agency's small budget. The AUA then submitted a second scaled-down proposal, but it was deemed by reviewers unlikely to be large enough to answer the questions being investigated (41), The National Institutes of Health (NIH), on its part, was apparently uninterested in funding a study that involved only existing treatments and offered little opportunity for new insights into the underlying biological mechanisms of the disease

Paradoxically, other studies of treatments for BPH are taking place that in their way highlight the current inadequacies of effectiveness research Both NIH and the Department of Veterans Affairs (VA) are conducting randomized clinical trials testing finasteride, a newly approved drug that is

ments. In fact, NIH staff note the contributions of NIH-sponsored research in developing instruments for such disease areas as cancer and rheumatology. They also note the current efforts to develop a standard vision function questionnaire suitable for patient self-assessment of the broad spectrum of vision function deficits.

Although the diversity of measures being studied and applied has advantages, some agreement on common measures is desirable for the sake of making cross-study comparisons, at least within the same disease. The efforts of the National Eye Institute in supporting development of an instrument useful for studies of eye disease, for example, are aimed at this goal (852).

The fact that a majority of comparative clinical studies still apparently do not include data collection on patient functioning and quality of life, whether through generic or disease-specific instruments, suggests that opportunities are being lost to provide important information for future patient decisionmaking. Equally troubling is that efforts to fill this need do not seem to be proceeding easily. AHCPR and NIH both clearly have much to contribute on the topic, yet cross-fertilization across the agencies in this area is more

BOX 4-2 continued: The Elusive Prostate Treatment Trial

believed to reduce symptoms in patients with this condition. The NIH trial, currently in the pilot phase, Involves 150 men at six medical centers The proposed full trial will be larger and more extensive, lasting six years (859a,879) In contrast, the VA trial will involve 1,200 men at 30 VA medical centers for only a year (564a) Trial designs are somewhat different as well The NIH trial is comparing finasteride against an alternative drug and a placebo and will revolve a number of tests and measurements aimed at better understanding the underlying disease The VA trial is likewlse testing the drug against both an alternate drug and a placebo but with a much simpler protocol and fewer clinical measurements (132)

There are two main differences between these funded studies and the unfunded one that was proposed to answer some of the questions raised by the PORT The first is that the funded studies involve a new technology, the drug flnasteride NIH's interest in funding a trial is piqued much more by new than by existing technologies, particularly when the trial offers possibilities for additional biochemical research as well Second, the funded studies involved a drug rather than a procedure Drugs unlike procedures, must be approved by the Food and Drug Administration (FDA), and manufacturers are accustomed to the routine of clinical trials Drugs also have identifiable "owners" who can profit from the results and thus sometimes may be willing to help support a study, in fact, the VA study is receiving support from the manufacturers of both drugs being tested (564a)

Both trials have worthwhile goals, and some duplication in research can add to the validity of the overall findings Still, in an area of research in which the gaps are so great, and the resources being made available to fund clinical trials on existing therapies so limited, it is ironic that the federal government is funding two simultaneous studies of a single therapy for benign prostatic hyperplasia, when another study of the same disease that was clearly needed has been unable to find funding

SOURCE Off Ice of Technology Assessment 1994 based on sources as shown Full citations are at the end of the report

notable for the exceptions than the norm. (In one exception. for example, AHCPR and NHLBI are collaborating to incorporate a quality-of-life measure into the ALLHAT[3] clinical trial of hypertension therapies (821).) Many institutes seem to have relatively little interest in the methodological work done at AHCPR; and where there is interest, AHCPR seems to perceive it as interest in that agency's resources rathcr than real interest in intellectual collaboration. [4]

Enhancing Know/edge of Ongoing Trials

Making the most use of ongoing clinical trials requires knowing that those trials are happening. and something about their characteristics. It has been suggested that one way to do this is to create a register of all ongoing clinical trials (166). Potential purposes of such a register are:

1. to foster more efficient research spending. by promoting collaboration among investigators

[3] AI, I. HAT is the acrony m for the Antihypertensive and Lipid Lowering Treatment to Prevent Heart Attack Trial.

[4] A recently published Italian study of the comparative effectiveness of different followup regimens for breast cancer patients underscores the feasibility and potential benefits of incorporating quality-of-life measures in effectiveness trials (280).

considering similar trials and preventing unnecessary duplication of research,
2. to enable better methodological research about the way that trials are undertaken and used (e.g., studies into publication bias of research results);
3. to recruit patients and providers into clinical trials more effectively and efficiently; and
4. to enhance scientific reviews of the literature, including meta-analyses (166).

The third of these reasons addresses the needs of effectiveness research in an especially direct manner. If broad groups of patients from across geographic areas and care settings are to be included in trials so that trial results areas generally applicable as possible, patients and their clinicians must know about trials. And, if large, multisite trials are to be completed in time for their results to be useful, patients must be enrolled as quickly as possible.

A number of registries of ongoing clinical trials in particular topic areas do exist. The AIDS and cancer communities have been particularly active in supporting registries so that patients and clinicians can learn about ongoing trials for which they might qualify. The PDQ database of the National Cancer Institute, for example, contains information on ongoing and completed clinical trials of cancer therapies in the United States. Information on AIDS treatment trials is available through MEDLINE, and numerous regional AIDS information services include additional detail on trials in their areas (167,418,757).

The AIDS database of ongoing clinical trials is unique because it relies on a special statutory except ion for information to be released by the Food and Drug Administration (FDA). Until 1988, FDA was prohibited from releasing information on ongoing clinical trials funded by private industry, information that industry generally considers to be confidential (418). This prohibition was lifted for information on AIDS-related trials only, due to the urgency of research on this disease. Thus, both NIH and FDA contribute information regarding ongoing trials, so that both publicly and privately sponsored clinical trials are included in the database. Information on private trials conducted under FDA auspices is summarized at FDA to protect as much confidential information as possible (191). In contrast, the PDQ database includes all NCI-sponsored trials, but it includes information on other trials only if that information is volunteered by the sponsoring organization (757).

The only cross-topic registries of ongoing controlled clinical trials in the United States are the clinical trials databases held by VA and NIH, respectively, to keep a comprehensive list of the clinical trials they sponsor. The NIH database is of special interest, because NIH is such a prominent sponsor of clinical trials, and because these trials are less 1 inked to a particular demographic population (i.e., veterans).

NIH maintained an inventory of its clinical trials from 1974 until 1979, when it discontinued the inventory for budgetary reasons (864). The inventory was re-established in 1985, through the Office of Medical Applications of Research (OMAR). Its road, however, has been rocky. Data collection was onerous; rules were changed in 1988 to require Institutes to report only data on controlled clinical trials (data on uncontrolled trials became optional). Data on mechanisms and sources of financial support has been particularly difficult to collect consistently (235). Data on cancer trials does not correspond precisely to data on trials from other Institutes, due to NCI's own internal trial database (864).[5]

Recent legislation requiring NIH to compile a cross-disease registry of clinical trials that involve women has helped stimulate interest in assembling a comprehensive database of ongoing clini-

[5] For all of these frustrations and limitations, summary data from the database, available for 1989, are interesting. In that year, NIH supported 440 controlled c1 inical trials, at an average annual cost per trial of just under $800,000 (864). NCI trials were excluded from the calculation of average annual cost per trial due to data inconsistencies.

cal trials (235). One barrier to any comprehensive registry is the lack of incentive for privately funded trials to be included: manufacturers consider this information confidential. FDA regulations protect this confidentiality, in order to protect manufacturers' financial incentives to develop new products. Short of withdrawing this protection, any comprehensive database must rely on the voluntary participation of private sponsors.

The other major barrier to a comprehensive clinical trials database is cost. One suggested solution is to use electronic technology to link existing trial databases. rather than initiating new ones, although this suggestion suffers the constraint of being limited in scope to the topics of existing (or new) registries (865).

Although there is clearly some interest in comprehensive, or linked, registries of ongoing trials, there is as yet no consensus about what form such an effort should take if it happens, or even about the extent of information such a database should contain. As NIH staff point out based on their experiences trying to maintain an NIH-wide trial database, collecting more detail on each trial would make the database more useful to researchers, trial participants, and policy makers alike, but greater detail comes at the cost of greater difficulty obtaining complete, accurate. and consistent data from the Institutes themselves (235).

Improving the Clinical Research Infrastructure

An important component of effectiveness research is the effort to make study results relevant to ordinary practice and the population at large. To do so, studies must address issues that arise in everyday care. and they must include an array of patients and providers representative of the overall population. For many questions, such as those in the area of primary care. undertaking comparative effectiveness trials can require large numbers of patients and physicians who are not currently affiliated with research institutions. The financial and administrative barriers to setting up such trials are substantial, and a major reason why these trials are not more common (box 4-3).

Furthermore, the barriers to large-scale, community-based trials must be overcome anew for each new trial proposed. NHLBI, for instance. is investing considerable resources in establishing a research network with as many as 300 practice sites for its ALLHAT trial of antihypertensive and cholesterol-lowering therapies (846). Once the trial is over, however, the network may well dissolve.

Conducting broad community-based trials would be substantially more streamlined if a network of providers already existed who had previously agreed to participate in research of interest to them and their patients. Establishing, maintaining, and supporting such networks is one way that the federal government could enhance the efficiency of comparative effectiveness trials, the generalizability of their results, and researchers' ability to carry them out.

To increase both provider and patient participation in clinical trials, trial enrollment and data collection requirements may need to be simpler than they are in many current trials. Thus, those designing and funding clinical trials may need to pay more attention to the techniques of large, simple trials described in the previous chapter. In addition, however, researchers and sponsors must find ways to recruit, train. and support a much broader set of very busy practicing clinicians.

Some of the best known examples of standing research networks are the infrastructures created for the various large, simple trials of therapies for heart disease. The GISSI and ISIS trials, described in chapter 3, both created an infrastructure of participating hospitals in their first respective studies that could be used on future trials as well; the fifth trial using the ISIS network is now underway. The important point about these networks is that they include many centers that are not teaching institutions and otherwise might not participate in detailed clinical trials. The GISSI network is an interesting model because it is so comprehensive: most of the coronary care units in Italy have participated in the GISSI trials.

Several U.S. examples of community-based medical research networks exist as well. A num-

BOX 4-3: The Need for Community-Based Research Networks: A Hypothetical Example

The administrative barriers to conducting large-scale, community based research are substantial. Establishing and carrying out such a study usually requires a major investment in recruiting providers and patients to participate. The investment is especially a great barrier for comparative testing of technologies already in use, since there are few eager sponsors for experiments involving existing interventions

As an example of the administrative difficulties a potential trial might face, imagine a researcher wishing to conduct a clinical trial of the comparative effectiveness of two common medications (inhaled cromolyn vs Inhaled steroid) in enabling the maintenance of normal activities in children with mild asthma Most of these children would be managed by primary care physicians, and many would never have even been hospitalized for their condition. Furthermore, effectiveness could well vary according to characteristics of children (e. g., cromolyn might be presumed to require more doses per day to be equally effective, and compliance with this stiffer regimen might differ according to a child's age)

Thus, the trial would have to recruit a large number of children covering a wide range of ages and other characteristics The researcher would need to identify these children, recruit their physicians, train these physicians, and have funding sufficient to give them the support they need to follow the study protocol and collect data for the trial. Simply getting the trial underway and convincing physicians to participate in the study would require a major investment of time and resources

Even when administrative barriers to such a trial are overcome, financial support may not be forthcoming. The American Urological Association, for example, tentatively established a network of physicians willing to participate in an ongoing series of trials of therapies for prostate disease (1 32) The network has never become fully operational, however, because the initial trial was never funded (see also box 4-2)

SOURCE Off Ice of Technology Assessment, 1994

ber of small practice networks exist that are loosely organized but enable clinicians and researchers to connect as needed to address particular research questions (122). The VA Cooperative Studies program is a standing multisite program that routinely involves practicing clinicians in clinical trials (523). Three additional examples illustrate in more detail the potential and experience so far with practice networks.

The *Community Clinical Oncology Program (CCOP),* sponsored by the National Cancer Institute (NCI), supports patients and physicians in community hospitals who wish to participate in cancer trials. NCI provides funding to cover administrative and data collection costs, without which community hospitals might not be able to participate in trials. The trials themselves are coordinated by NCI-supported teaching and research hospitals (260). About 50 CCOPs, representing about 300 community hospitals, receive funding from NCI to support their participation in cancer trials through this network. CCOP patients represent roughly one-third of all patients enrolled in NCI trials (260).

The *Ambulatory Sentinel Practice Network (ASPN),* a private effort supported in part by the American Academy of Family Practice, is another longstanding U.S. community research network. Established in 1982, its purpose is "to increase and refine the primary care knowledge base by

studying the problems that occur in primary care" (11). It includes 72 participating medical practices (including over 300 practitioners) in the United States and Canada. An overwhelming majority of the participants are family practice physicians (12).

Because its members are largely community-based primary care physicians, ASPN's data collection has been very simple: basic demographic data on patients seen in the practice, with data collection on the study question through a weekly mailed card. Many study questions originate with the practitioners themselves, and most are descriptive studies. Funding for individual studies is sought from whatever sources are available: sponsors have included such federal agencies as the Centers for Disease Control and Prevention and NHLBI. Examples of recent and ongoing studies include depression in primary care; management of carpal tunnel syndrome; acute low back pain; and the effect of digitalis on mortality (12). Research is administered through a central headquarters in Denver, Colorado.

A possible concern of this and other research networks is that because the participants are self-selected, their patient populations may not be representative of patients overall. ASPN researchers addressed this concern by comparing detailed characteristics of patients and visits to ASPN practices with the characteristics reported on a national survey of ambulatory care (the National Ambulatory Medical Care Survey) (297). They found considerable similarity in visit characteristics (e.g., patient diagnoses) but some differences in patient demographics.

The *Vermont Trials Network is* a newer private effort, an innovative network of hospital neonatal intensive care units established to perform collaborative clinical research in neonatology and integrate research into daily practice (383). As of February 1994, there were 111 neonatal centers participating in the network, many of which had no affiliations with universities (729). The great majority of these are centers in U.S. hospitals, but recently hospitals in Australia, Germany, Japan, and other countries have also expressed an interest in participating (729).

The centers collect basic data on the medical and demographic characteristics of infants. They also collect information on the prevalence of some conditions and on the use of particular technologies and services (e.g.. the use of ventilators and surfactant). These data are intended to provide information for planning clinical trials and to pm-nit centers to compare their outcomes with each other as an aid in quality management (354). The database and trials facilitation service are administered through a central office in Vermont, which operates with temporary grant funding from a private foundation.

The first clinical trial to be implemented in the network centers, which began in January 1992, was a randomized comparison of two commercially available surfactants (preventive treatment for lung disease in premature infants). Both surfactants have been proven effective in previous trials, but direct comparisons of the two drugs are not available (354). The participating researchers hope to find an answer of practical importance to community neonatologists and to be able to compare the costs and results of this trial to those of a smaller, NIH-funded trial on the same topic being carried out only at university centers (354).

These three examples differ considerably in their sophistication. sources of funding. and size. They range from research by office-based family practitioners to large clinical trials in neonatal care units. What all have in common is that they involve an underlying structure through which non-academic as well as academic health care providers can participate in clinical research of interest to them and their patients. Indeed, in the case of the ASPN network, the providers themselves suggest some of the research questions.

None of these examples are of "firms" research infrastructures, which may require more intensive effort and investment on the part of the health care institution. The emergence and growth of managed care providers and the interest in methods for continuous quality improvement. however, might

make firms structures an attractive form of research for many institutions, particularly if they received some startup financial support. The VA is exploring the establishment of a research structure of this type (930).

An aspect of comparative effectiveness trials largely uncommented on, in either the literature or the health policy debate, is the relationship between effectiveness trials conducted within a committed infrastructure and the goals of continuous quality improvement, a topic that is very much the subject of current discussion. This relationship is particularly marked in the GISSI large, simple trials, which included most of the coronary care units in Italy. As the trials were completed, units could incorporate the findings, and new trials begun to achieve the next level of quality improvement. Questions of generalizability of findings were almost irrelevant, since most units and patients participated. Firms trials have accomplished this objective on an institute-specific basis; as an intervention proved effective, it was adopted by the other firms in the institutions and became the new level against which future improvements would be measured.

The Comparative Evaluation of New Technologies

A major contributor to the current state of ignorance about what works best, and under what circumstances, in health care is the fact that many—probably most-new medical technologies need not undergo rigorous review of their effectiveness before being adopted by practitioners and patients. Furthermore, of those that are reviewed for their effectiveness, most need not prove that they are actually more effective than other alternative technologies already on the market.

There are three avenues through which new technologies can be identified and enrolled in comparative evaluations:

- **Manufacturers.** Those producing new technologies could be encouraged to identify them and conduct comparative assessments directly. This avenue could take the form of increased regulatory oversight, such as a broad extension of current FDA requirements for new drugs; or it could take the form of inducements (e.g., favored regulatory treatment for manufacturers willing to sponsor comparative postmarketing studies).
- **Payers.** Health insurers, including government payers, could offer insurance coverage for new technologies only if they had met explicit standards of evaluation and effectiveness.
- **Government.** Some researchers have suggested that the most efficient way to increase the number of direct comparative studies on new as well as existing technologies is for the federal government to conductor sponsor such studies directly (624,625).

These avenues are not mutually exclusive; all three could be pursued simultaneously.

An underlying question implicit in choosing among these options is who should be paying for the evaluation of new technologies. Manufacturer-sponsored evaluations could come about either through regulatory incentives or pressure by payers. Alternatively, payers could withhold coverage from unevaluated new technologies but could also help fund their evaluations, by paying for some of the costs of the studies. Government-sponsored evaluation would clearly increase the proportion of studies of new technologies funded by taxpayers generally.

Three issues are especially prominent in considering how to enhance the number and quality of comparative evaluations of new with existing technologies. The first, especially important in strategies that depend on manufacturers to conduct evaluations, is the role the FDA plays in the evaluation of new technologies. The second issue, associated with payer-dependent strategies, is the role of health insurers in paying for new and experimental technologies. Both of those issues are discussed in this section. The third issue is the potential role of different federal agencies in conducting or supporting evaluations. This issue extends to current effectiveness research efforts

comparing existing technologies as well, and it is discussed in the final section of this chapter.

Role of the Food and Drug Administration

The charge of the FDA is to ensure that new drugs and medical devices are safe and efficacious--i.e., that the medical benefits outweigh the medical risks—before they are marketed to the public. Its regulatory authority extends not only to whether a product can be put on the market, but what claims the manufacturer can make about that product. In reviewing evidence about the efficacy of a product. FDA gives strong weight to evidence from randomized clinical trials as the most valid basis for making efficacy claims.

FDA's authority over medical devices is slightly different from its authority over drugs. All drugs that involve new chemical formulations must show proof of efficacy, with a stringent level of evidence to provide that proof. Most often they are compared in randomized trials with placebos, although new drugs in certain categories, such as new antibiotic and anticancer drugs, are commonly tested against accepted existing drugs instead (748). In contrast, new medical devices are categorized by FDA staff into one of three classes, according to the types and controllability of risk associated with the device in its intended use, with each class subject to a different standard. Class I and II devices considered to involve only low or moderate risk-e. g., new wheelchairs—must be registered with FDA, and their producers must conform to good manufacturing standards. Class III devices such as x-ray machines also must meet performance standards. In addition, however, Class III devices—generally those posing a potentially higher risk to patient health—must meet standards similar to those for new drugs. [h] Class 111 devices account for roughly 10 percent of medical devices (922).

Over time, FDA policies have changed somewhat in the kinds of outcomes considered the most relevant for regulatory decisions. For medical devices, the agency has historically placed a strong emphasis on what FDA terms "functional utility:" i.e., whether the device does what the manufacturer claims it does (e.g., remove plaque in blood vessels). In 1990, an internal FDA policy guideline established 'clinical utility''—the ability of the device to produce a desirable treatment outcome—as a preferable standard. Under this standard, for example, home uterine monitoring devices would have to prove not only that they could detect uterine contractions, but that clinical outcomes (e.g., the number of premature births) were improved (922).

Trends in the standards for evaluating new drugs have some differences from those for devices. In some areas, for example, the trend has been to emphasize clinical endpoints that can be measured quickly. In particular. the urgency of the need to identify drugs that might be efficacious in treating AIDS has led to greater use of "surrogate endpoints" in the approval of anti-AIDS drugs for marketing (e.g., endpoints such as showing a difference in the rate of certain biochemical markers that indicate the progression of disease). The use of surrogate endpoints has its own well-known hazards; in a recent example, a drug approved for marketing by FDA on the basis of improvements in surrogate endpoints could not be shown, in a longer European trial, to have any effect on total mortality from AIDS (142). FDA staff cite this example as a reason to conduct post-marketing studies of such drugs, so that effects on ultimate endpoints can be measured as well (839).

In other areas, however, there are examples of a greater attention to ultimate outcomes (e.g., mortality) as a factor in FDA decisionmaking. In the clearest example, quinidine--a drug originally approved for marketing on the grounds that it was shown to be efficacious in reducing atria] fibrillation (irregular heartbeats)—was later required to be relabeled or withdrawn from the

[6] Class III devices considered by FDA to be "substantially equivalent" to a device already on the market in 1976. when the regulatory authority over medical devices was added, are not immediately required to meet these standards but can be required to do so in the future (784a).

market after a clinical trial showed that this drug actually increased, rather than decreased, mortality rates in some groups of patients (839).

The randomized clinical trial continues to be the gold standard for the assessment of a product efficacy, but some of the more recent innovations in conducting and analyzing clinical trials are occasionally finding their way into FDA decisionmaking. Large, simple trials, for example, have not been used as a basis for approving a drug, but they have been used to support approved changes in a drug's label or advertising (839). Similarly, the results of a meta-analysis have been used as the basis for insisting that a drug be relabeled, after the meta-analysis showed treatment groups to have a higher overall mortality (839).

An interesting example of "effectiveness" trials required by FDA involves the transition of a drug from prescription-only to over-the-counter availability. Manufacturers interested in marketing drugs for nonprescription uses must conduct "usage trials." In a typical usage trial, several thousand patients are given the medication, with its proposed labeling and instructions for use, and are monitored to determine whether the drug is safe and effective as actually used by these patients (327).

Thus, FDA plays three strong roles in the comparative evaluation of new medical products.

- First, it requires that the underlying efficacy of all new drugs, and some new devices, is established—i.e., that the product works under at least some conditions. Efficacy sometimes involves direct comparisons with existing drugs, as in the case of antibiotics, but even direct comparisons often do not provide broad information on comparative effects in ordinary practice.
- Second, although much of FDA's role in drug and device approval focuses on approving new products for marketing, the agency also plays a role in the post-marketing monitoring of the effects of products in general use, and it plays a strong role in the effectiveness claims that manufacturers can make when advertising their products.
- Third, FDA establishes acceptable levels of evidence for showing that a product works. Hence, FDA's greater emphasis on ultimate health outcomes, and on results from randomized trials, have trickle-down effects on health research. Standards have generally required that a trial show a very strong ability to reject a hypothesis that the new treatment made no difference, an issue that has potential repercussions for FDA oversight of later comparative effectiveness and cost-effectiveness claims.

Issues in Insurance Coverage of Newly Introduced Technologies

The role of payers in covering (or withholding coverage from) new or experimental technologies is an issue that has been growing in prominence. Its importance to government policy makers contemplating changes to the health care system is demonstrated by proposals to include experimental services as a health insurance benefit under certain conditions. The State of Maryland, for example, is devising a basic benefit package, which insurers who market health insurance to small employers must offer, that includes coverage for technologies offered as part of authorized clinical trials (409). The health reform proposal of the Clinton Administration included a provision that would have required coverage of "routine care" associated with experimental therapies and permit coverage of the therapies themselves if they met certain conditions (S 1757).

Historically, insurers have relied on the term "medically necessary" to broad] y describe the services covered by their health policies and "experimental" to define at least some of the services beyond the boundary of health care coverage. Since the 1980s, the definition of these terms has been an increasingly contentious issue. Experimental services are particularly controversial because they often involve potentially life-saving treatments for desperate y ill patients who are personally willing to take the risk that the service may prove to be unsafe or ineffective. Today, the interpretations of "medically necessary" and "experimental" are hotly contested among insurers,

researchers, physicians, manufacturers of drugs and devices. and patients and are often mediated (albeit inconsistently) in the courts.

A typical insurance contract defines a service or supply to be "medically necessary," and therefore covered, if: (a) it is ordered by a doctor; (b) it is commonly and customarily recognized throughout the doctor's profession as appropriate in the treatment of the sickness or injury; and (c) it is "neither educational nor experimental in nature nor provided primarily for research purposes" *(322).7* The point at which a new treatment moves from the investigational or experimental stage and into the realm of '-state of the art" medically necessary treatment is not at all clear (25).

There are no data that systematically document those services commonly excluded by insurers because of their experimental nature. Autologous bone mm-row transplant with high-dosage chemotherapy (ABMT/HDC) for breast cancer is perhaps the most widely contested and well-known experimental treatment (box 4-4). Other examples of technologies typically excluded from coverage on the grounds that they are currently experimental, or covered only case-by-case, include growth hormone for children with short stature, pancreas transplants, and home uterine monitoring for the prevention of premature births (178).

Despite explicit contract language to the contrary, it appears that some insurers sometimes allow coverage of certain experimental treatments on a case-by-case basis. For example, five major carriers reported in a recent telephone survey that, given certain criteria and conditions. they would pay for a number of "experimental" treatments including ABMT/HDC. pancreatic transplant, growth hormone for short-stature children, home uterine monitoring, and radial keratotomy (178). Researchers conducting a clinical trial comparing ABMT/HDC with conventional treatment point out that coverage decisions across and even within insurance companies for this therapy are inconsistent (601).

Recently, a few insurers have taken an unprecedented step into the controversy surrounding: coverage of newly introduced technologies. They have agreed to pay for ABMT/HDC for insured patients who are enrolled in an NCI-approved randomized controlled clinical trial to compare ABMT/HDC to standard treatment for breast cancer (404) (box 4-4). An important component of these trials is that they are randomized: thus, many of the patients enrolled in the trial will not receive the experimental therapy. For this and other reasons, patient accrual to the trials has been disappointing to researchers (135).

The issue of coverage for experimental technologies has begun to receive attention from insurers at a national level as well. The Health Insurance Association of America (H IAA), for example. has endorsed a policy to encourage their membership to pay for the patient care costs related to NIH-sponsored and certain other officially endorsed randomized clinical trials. The National Association of Insurance Commissioners has established a working group on the topic. whose goals include drafting a model regulation or statute to address off-label use of prescription drugs, and researching the impact of experimental treatment exclusions. Several States (e.g., Washington, Florida, New Hampshire) have already passed insurance regulations related to experimental treatment (552).

Despite the increasing interest and movement to change the link between insurance coverage and the experimental status of a technology', proposals to address the connection between coverage and the degree to which a technology has been proven effective face a number of competing interests and concerns:

7 An interesting facet of the definitions that exclude experimental technologies from insurance coverage is demonstrated by contract language for insurance contract exclusions developed by Towers, Perrin, Inc. for its c 1 ients use in their health benefit plans (178). I n this contract language, the fact that a technology is the subject of a controlled clinical trial merits the label of experimental. This categorization might present a problem for randomized c1 inical trials comparing two technologies already in common use.

BOX 4-4: Autologous Bone Marrow Transplants in the Treatment of Breast Cancer

Conventional therapy for women with advanced breast cancer consists of mastectomy followed by radiation, chemotherapy, or both, But conventional therapy frequently fails, and approximately 46,000 women die of breast cancer each year (777),

A limitation of conventional chemotherapy is that it cannot be administered at high dosages without killing the patient's own bone marrow as well as the cancer cells. Autologous bone marrow transplant with high-dose chemotherapy (ABMT/HDC) is a technique aimed at enabling higher doses of chemotherapy to be given. In this procedure, the patient's bone marrow is removed before chemotherapy is administered, and then reinfused after the chemotherapy regimen is complete,

ABMT/HDC has engendered great enthusiasm in the medical world (670) and is now being tried for other solid tumor cancers as well (e.g., testicular and colon cancers). Still, the efficacy of ABMT/HDC over standard treatment for advanced breast cancer has not yet been definitively demonstrated (850), and one assessment of the technique based on past studies expresses skepticism that it wiII prove effective in this population (195).

Most payers view ABMT/HDC for metastatic breast cancer as experimental, Recently, however, some major insurers and HMOs (including Metropolitan Life, Prudential, CIGNA, Travelers, U.S. Healthcare, Kaiser, and some Blue Cross Blue Shield plans) have begun to cover ABMT/HDC for breast cancer under certain conditions (66,1 78,404). It is not clear that these private health insurers have taken this step because they now accept ABMT/HDC as nonexperimental or state-of-the-art therapy. Rather, there is much anecdotal evidence to suggest that they are motivated by the desire to avoid legal action, expense, and negative publicity. Efforts by insurers to refuse reimbursement for ABMT/HDC for breast cancer have been widely contested in the courts; in one recent and well-publicized case, a jury awarded $89 million to the patient of a California HMO that had refused to cover the procedure (135).

Still, insurers' response to the pressure to pay for ABMT/HDC have varied widely, not only among carriers but within companies as well. For example, in an effort to develop the clinical data necessary to assess ABMT/HDC's efficacy, U.S. Healthcare[1] and 17 Blue Cross Blue Shield plans[2] are currently supporting several National Cancer Institute (NCI) randomized controlled clinical trials to compare ABMT/HDC to standard treatment for breast cancer (66,404). The trials are continuing to accrue patients (with varying rates of success) and final results are not expected for at least three years (135). Other insurers, including Metropolitan, CIGNA, Prudential, Travelers, and Aetna, are now reimbursing for ABMT/HDC on at least a case-by-case basis, but apparently do not require that patients participate in the NC I trials.

---

[1] U S Healthcare is a 16 million member HMO based In Philadelphia, PA

[2] These 17 plans include approximately half of the nation's Blue Cross and Blue Shield's membership (282)

SOURCE Office of Technology Assessment, 1994, based on sources as shown Full citations at the end of the report

- Manufacturers generally support coverage for technologies that are still at the experimental stage as a way to minimize the time lag between a product's development and its availability to patients (461). At the same time, a proposal that linked payment to rigorous proof of efficacy would probably meet producer resistance, since many existing technologies cannot meet this standard (and many new ones need not at present). Whether a standard of proof as rigorous as the RCT is even necessary for all technologies is very much a matter of debate.
- Patients and providers likewise generally support coverage for experimental technologies, since it would increase the treatment options financially available to them.
- Insurers, and those who pay the insurance premiums, tend not to support coverage of investigational interventions, on the grounds that coverage would increase costs without any assurance that the interventions would be either safe or effective for those who would receive them.
- Some observers also express concern that opening insurance coverage to investigational therapies could lead to a worsening of the problem of poorly conducted studies, unless strict controls and monitoring of the investigational protocols were also in place (581).

ISSUES IN FEDERAL FUNDING AND SUPPORT

The Roles of the Federal Agencies

The self-perceived roles and goals of agencies have a strong influence over the part each plays in the current debates over how to improve the effectiveness, quality, and costs of health care. They also explain a great deal about where and why duplication or gaps in effectiveness research appear among agency activities.

The federal organizations that currently sponsor effectiveness research (and other evaluative activities) do so for three reasons. The first is to provide information to the public and to private insurers and providers, in order to improve the private sector's ability to deliver effective care. The second is to support the government own health care financing and delivery programs. such as Medicare, the veterans' health system, and the myriad preventive and other public health programs. The third reason is to provide information that can enhance public policy decisionmaking generally, for purposes ranging from distributing research resources to helping Congress decide whether to establish new Medicare benefits.

Most of the federal organizations involved are sprinkled throughout the U.S. Department of Health and Human Services (DHHS). They include the Agency for Health Care Policy and Research (AHCPR), the National Institutes of Health (NIH), the Centers for Disease Control and Prevention (CDC), the Health Care Financing Administration (HCFA), and a small office under the Assistant Secretary for Health, the Office of Disease Prevention and Health Promotion (ODPHP) (figure 4-1, see p, 105). In addition, however, the Department of Veterans Affairs (VA), which operates a health care system for veterans of the U.S. armed services, has its own entirely autonomous research arm to investigate health care services and technologies.[8]

These organizations all are concerned in some way with health care research and delivery, but they have greatly differing purposes and orientations. These differing purposes affect their approaches to identifying effective and cost-effective medical services and technologies, their methods for assessing technologies for clinical and public policy purposes, and the degree of their activity in these areas.

[8]To identify relevant activities currently conducted by these agencies, OTA asked administrators in each organization to provide information on those studies and activities they considered relevant. Their responses form the basis for discussions of activities presented in this report, The Department of Defense also conducts medical research, but its activities were not investigated in detail in this report.

BOX 4-5: The Components of the Agency for Health Care Policy and Research

1. The **Center for General Health Services *Extramural* Research** carries on much of the legacy of general research into the interactions of the health delivery system inherited from the National Center for Health Services Research (NCHSR) The center supports research on such topics as health care costs and financing, and Improving the delivery of health care services to special populations.
2. The **Center for General Health Services *Intramural* Research** is another NCHSR legacy. It performs in-house general health services research, drawing heavily on data from federal health care databases.
3. The **Center for Medical Effectiveness Research** is the focal point for the federal government's investment in effectiveness research. This center supports the Patient Outcomes Research Teams (PORTS), as well as many other extramural research projects on medical practice and outcomes variation, the effectiveness of particular medical interventions, and the refinement of some of the tools of effectiveness research.
4. The **Office of Science and Data Development** oversees activities related to the enhancement of databases and the Implications of advances in medical information systems. Its activities support the effectiveness research infrastructure—e.g., by supporting efforts to link large databases together, and sponsoring research to stimulate the development computer-based patient records
5. The **Office of the Forum for Quality and Effectiveness in Health Care** is responsible for the development of clinical practice guidelines. It organizes the guidelines panels and provides them with staff support and supplemental contracted expertise.
6. The **Office of Health Technology Assessment** is another direct holdover from the old NCHSR, although its responsibilities have expanded somewhat. It conducts In-house assessments of individual medical technologies for the Medicare and CHAMPUS programs.
7. The **Center for Research Dissemination and Liaison** has the primary responsibility for disseminating clinical practice guidelines developed by the Forum to clinicians, patients, and other Interested parties It also operates the Users Liaison program, which runs informational conferences and provides technical assistance to State personnel and other consumers of AHCPR's work

SOURCE Off Ice of Technology Assessment 1994, based on documents provided by the U S Department of Health and Human Services, Public Health Service Agency fc. Health Care Policy and Research, Rockville. MD 1993

Agency for Health Care Policy and Research

Implicit in Congress' creation of AHCPR in 1989 was a statement that the federal government should actively support and promote effectiveness research and health technology assessment. AHCPR is primarily a research-sponsoring organization. with an agenda that tends to focus on health services and current therapies, the legacy of its inheritance of the National Center for Health Services Research (NCHSR), its predecessor agency. It contains seven centers and offices (BOX 4-5). Three of these are direct holdovers from NCHSR. The remaining four were newly created specifically to carry out AHCPR's new mission.

In conformance with its small budget, its health services research orientation, and its legislative mandate, AHCPR's investment in effectiveness research has leaned heavily towards the development of effectiveness research tools (e.g., developing databases and health status measurement instruments) and descriptive research on the outcomes associated with particular technologies

and patterns of care. Many of these activities take place through the Patient Outcomes Research Teams (PORTS), the central research program of the federal government's effectiveness research initiative As described in chapter 3, AHCPR supports very few controlled trials of clinical interventions.

AHCPR's budget grew from $97 million in fiscal year 1990 to $128 million in fiscal year 1993, All of that increase went to effectiveness research and guideline development efforts, whose funding grew from $37 million $73 million during the same period (J. Clinton, at AHSR, June 1993). The budget is broken into three activities:

- Program support. This component receives the smallest portion of the budget—$2.5 million in 1993, or about 2 percent of the total
- Research on health costs, quality, and access. This component is the continuation of the health services research efforts previously carried out through NCHSR and amounted to $53.1 million in 1993, or 41 percent of the total. It supports both intramural and extramural general health services research and supports the National Medical Expenditures Survey, a major source of medical cost data.
- Medical **Treatment Effectiveness Program** (MEDTEP). This funding line accounted for $73.0 million of AHCPR's 1993 budget, or 57 percent of the total. It supports not only the acti - vities sponsored by the Center for Medical Effectiveness Research, which funds extramural effectivencss research, but also the guidelines activities of the Forum, the resource development activities of the Office of Science and Data Development, and the activities of the Center for Research Dissemination and Liaison.[9]

Like most agencies, the great bulk of AHCPR's funding (85 pcrccnt,or$109 million in fiscal year 1993) clinics from federal gcncral revenues funding. AHCPR's authorizing legislation also permits substantial transfers from the Medicare Trust Fund for medical effectiveness research and guidelines development. In fiscal year 1993. $103.6 million was authorized from this source, but only $5.8 million was appropriated. making up 4 percent of the *agency's* budget (814). In addition, AHCPR is authorized to draw funds from the Public Health Service Evaluation Set Aside ("One Percent Funds"). These funds account for a significant proportion of the agency's total budget ($1 3.2 million in fiscal year 1993. or 11 percent of the budget), but they are earmarked to fund the National Medical Expenditures Survey and cannot be used for other purposes under current law.

The National Institutes of Health

The National Institutes of Health (NIH), with a budget of approximately $10 billion in fiscal year 1993, is the primary sponsor of biomedical research in the United States (844). From its origins in a public health service research laboratory (box 4-6). NIH has come to comprise 24 relatively independent research institutes. NIH coordinates an extensive intramural research agenda as well as funding extramural research conducted at 1,700 institutions nationwide.

Most NIH institutes conduct a great amount of basic "'bench" research as well as some applied clinical research, primarily on developing new therapies. NIH spent $864 million, or just under 10 percent of its budget ($8.4 billion), on clinical studies in 1992 (844). Although in recent years a significant proportion of NIH funding has been "earmarked" (e.g., for AIDS or women health research), its agenda is still largely investigator-driven and heavily influenced by the makeup of its "study sect ions," the groups of outside researchers who review grant applications.

Three things are notable about NIH clinical studies. First, it is a widely held opinion that most

[9] The PORTs take up only about one-fourth of the MEDTEP program. The program also includes investigator-initiated grants to examine specific issues relating to variation, outcomes, and methods development; and a program to support research centers on minority populations. Eleven such research centers are currently funded (821).

BOX 4-6: The History and Components of the National Institutes of Health

The National Institutes of Health traces its origins to a single small laboratory, the Laboratory of Hygiene, that was established in 1887 within what was then the Marine Hospital Service (eventually to become the Public Health Service), for the purpose of investigating infectious diseases such as cholera The laboratory officially became the National Institute of Health by congressional fiat in 1930 and proceeded to undertake basic research into such widespread health problems of the day as tooth decay, undulant fever, and pellagra The National Cancer Institute was established separately by legislation in 1937 and for many years was functionally separate from the National Institute of Health

NIH formally became the National *Institutes* of Health in 1948, when four new institutes were created to work on heart problems, dental research, microbiological studies, and experimental biology and medicine Construction on the NIH clinical center, to further efforts to test the clinical applications of research, was begun at this time as well Additional institutes and centers were added over the following decades In 1994 NIH comprised the Office of the Director (which includes the Office of Medical Applications of Research) and *24* institutes, centers, and divisions

- National Cancer Institute
- National Eye Institute
- National Heart, Lung, and Blood Institute
- National Institute on Aging
- National Institute on Alcohol Abuse and Alcoholism
- National Institute of Allergy and Infectious Diseases
- National Institute of Arthritis and Musculoskeletal and Skin Diseases
- National Institute of Child Health and Human Development
- National Institute on Deafness and Other Communication Disorders
- National Institute of Diabetes and Digestive and Kidney Diseases
- National Institute of Dental Research
- National Institute on Drug Abuse
- National Institute of Environmental Health Sciences
- National Institute of General Medical Sciences
- National Institute of Mental Health
- National Institute of Neurological Disorders and Stroke
- National Institute for Nursing Research
- National Center for Human Genome Research
- National Center for Research Resources
- Clinical Center
- Fogarty International Center
- National Library of Medicine
- Division of Computer Research and Technology
- Division of Research Grants

SOURCES D Pugh, *The National Institutes of Health A Bethesda Landmark Celebrates Its Centennial* (Rockville, MD U S Department of Health and Human Services, Public Health Service, National Institutes of Health 1987) U S Department of Health and Human Services, Public Health Service, National Institutes of Health "Origins of the National Institutes of Health " pamphlet Bethesda, MD, undated U S Department of Health and Human Services Public Health Service, National Institutes of Health, Investment for Humanity, " Bethesda MD 1993

NIH RCTs focus on new technologies rather than existing therapies. Second, NIH does sponsor at least some comparative clinical studies of existing therapies. Third, however, there is at present no way to know the extent to which this is actually true. because NIH's data on its own RCTs are not complete enough nor detailed enough for anyone to examine the question.

As one example of clinical trials ongoing at one Institute. the National Eye Institute documents 21 ongoing studies (853). At least eight of these studies compare two or more technologies already widespread before the study began. NEI trials may be unusual in a number of ways (e. g., a substantial proportion address surgical interventions). Still. the} demonstrate that NIH does sponsor a presumably small but possibly significant number of comparative clinical trials of existing technologies. Even if such trials comprise only one-tenth of NIH clinical trials budget of not quite $900 million, as a federal financial commitment they would surpass the entire MEDTEP budget ($78 million) of AHCPR.

Other than NIH sponsorship of some comparative clinical trials that focus on existing technologies and broad populations. the NIH activities most directly tied to effectiveness research arc its development of health status and quality-of-life measures for certain diseases, and database resource activities. During the past six years, for example. the National Cancer Institute and the National Eye Institute (NEI) have held workshops on quality-of-life assessment in their respective areas, and the National Center for Nursing Research held a conference on methods to measure the effectiveness of nursing practice. A number of Institutes also maintain disease and procedure registries, a resource for researchers interested in augmenting databases (e.g., the Huntington Disease Research Roster, and the Vascular Surgery Registry.)

A notablc effectiveness research resource activity is a collaborative effort to link Medicare administrative data on patient scr}'ices with cancer epidcmiological data from NC I SEER (Surveillance, Epidemiology and End Results) registry. a database that collects detailed, verified clinical data on persons with cancer in 11 areas across the country. The resultant merged database includes both data on tumor size and cancer severity and data on clinical services received by Medicare beneficiaries with cancer, as well as information on the costs of those seri'ices (610). The linked HCFA-SEER database will be used, for example, for a study of the patterns and outcomes of cancer care in the Medicare population (796). SEER data have also been used in AHCPR cancer outcomes studies (479).

The Centers for Disease Control and Prevention

In 1946, the Office of Malaria Control in War Areas was replaced by the Communicable Disease Center, whose primary goal was to reduce the transmission of venereal diseases from homecoming soldiers (828). Now the centers for Disease Control and Prevention, CDC has grown to encompass 11 individual centers and offices whose common goal is disease and injury prevention (box 4-7).

In accordance with its mission, CDC stresses the epidemiology of disease and the identification of new disorders. Legionnaire disease, toxic shock syndrome, AIDS, and most recently a new deadly outbreak of a previously unknown virus in the American southwest have all been traced, described, and studied by CDC scientists. The agency's role includes some public health and prevention-related research (e.g.. into infectious diseases), but it is at least as much a service sponsor as a research agency; much of its role is in funding prevention programs.

The agency began to emphasize "prevention effectiveness" in the early 1990s (750,754). The focus for this effort was the establishment in 1992 of a Prevention Effectiveness Activity. with its own chief, within CDC Epidemiology Program Office. CDC staff specifically intended this activity to parallel the medical effectiveness initiative, with CDC assessing the effectiveness of population-bawl prevention efforts while others (primarily AHCPR) assessed the effectiveness of

BOX 4-7: The Components of the Centers for Disease Control and Prevention

The 11 operating units that collectively make up the Centers for Disease Control and Prevention demonstrate the agency's focus on population-based preventive and environmental health They are.

- National Center for Chronic Disease Prevention and Health Promotion
- National Center for Environmental Health
- National Center for Health Statistics
- National Center for Infectious Diseases
- National Center for Injury Prevention and Control
- National Center for Prevention Services
- National Institute for Occupational Safety and Health
- National Immunization Program
- Epidemiology Program Office
- International Health Program Off Ice
- Public Health Practice Program Off Ice

SOURCE U S Department of Health and Human Services, Public Health Service, Centers for Disease Control and Prevention,organizational chart unpublished document, Atlanta, GA February 1994

"medical procedures" (749). In one example of the influence of this activity, a recent CDC statement on suicide prevention included a comment on the lack of information on the relative effectiveness of different strategies to prevent suicides (837).

CDC's focus on "population-based" interventions leads it to emphasize environmental and behavioral programs (e.g., lead abatement and public safety campaigns), although it also supports population-based programs involving clinical preventive services (e.g., programs to increase the rate of screening for particular diseases). However, the distinction between "population-" and "individual-based" clinical preventive services is not always clear-cut. One major current study taking place in clinical settings is a public-private collaborative effort, in which six health maintenance organizations (HMOs) are sharing existing HMO data on "prevention strategies for assessing the effectiveness of prevention activities in an HMO and community setting" (836a). Possible services to be examined include mammography utilization, antibiotic treatment of otitis media, diabetes management and prevention, and the treatment and prevention of domestic violence (751,753).

CDC is also a major compiler of health care registries and databases, although most of its registries have not so far played a major part in effectiveness research. Its compilation of mortality statistics and population-based health indicators through the National Center for Health Statistics, however, are fundamental to many studies.

The Office of Disease Prevention and Health Promotion

ODPHP, a small office located within the Office of the Assistant Secretary for Health itself, was created by statute in 1977. Allocated a budget of somewhat less than $5 million for fiscal year 1994, its purpose is to establish national public health goals and strategies to achieve those goals, to act as a clearinghouse for information on disease prevention and health promotion, and to coordinate departmental activities in these areas (Public Law 98-55 1). To do this, ODPHP undertakes such activities as monitoring progress to-

wards the goals of its *Healthy People* **2000** report, operating the National Health Information Center, and coordinating health promotion and prevention activities among federal agencies and between the federal government and nongovernmental organizations.

ODPHP also occasional] y undertakes activities to fill perceived gaps in prevention activities undertaken by other federal agencies. This office, for example, helped develop the dietary guidelines (the "food pyramid") subsequently promoted through the U.S. Department of Agriculture. It also produces the Surgeon General Report on Nutrition and Health and is currently in the process of a structured literature review to support recommendations for dietary fat intake (325).

For the most part, ODPHP's activities draw on the results of effectiveness research, rather than sponsoring or conducting research itself. Two of these activities are discussed in more detail later in this report: sponsoring the U.S. Preventive Services Task Force and the Cost-Effectiveness Panel on Clinical Preventive Services. and convening an interdepartmental discussion group on cost-effectiveness of clinical preventive services.

The Health Care Financing Administration

HCFA's mission is to administer the Medicare and Medicaid programs, the two massive programs that provide health insurance to elderly, disabled, and poor persons. As part of that responsibility, the agency includes within it an Office of Research and Demonstrations. This office sponsors such activities as pilot projects of novel approaches to delivering care to its constituent populations, evaluations of demonstration projects, and research into new methods of paying for services.

HCFA's largest contribution to effectiveness research is its enormous Medicare databases, which include detailed data on hospital care, outpatient care, health care institutions, and other factors. The potential of these databases to be rich sources of information on care patterns and outcomes was a major motivation for the federal government's "effectiveness initiative" (65 1). Their main disadvantages for descriptive purposes are that they do not include much of the information researchers want to discriminate among patients with different levels of health need. they often cover only a small slice of an individual health care experience (e.g.. inpatient care), and the clinical progression of disease can be inferred only indirectly, as a consequence of the procedures recorded in the data.

To address some of these issues, HCFA is currently involved in two separate efforts to provide greatly augmented databases. One of these is the linkage of the SEER-Medicare databases. The other, the Medicare Beneficiary Health Status Registry, will create a new database based on a survey of a large sample of Medicare enrollees (766). The survey. a mailed questionnaire, asks beneficiaries about their current health status. health risk factors, and socioeconomic characteristics. The survey is presently being pilot-tested.

HCFA also sponsors some descriptive studies to document outcomes associated with particular conditions or particular care practices in the Medicare and Medicaid populations. One major set of studies, for example. is examining the appropriateness and outcomes of care provided to Medicaid patients for conditions such as pediatric asthma, complicated delivery, and hysterectomy (638). Outcomes of care in Medicare patients who have end-stage renal disease, and outcomes of care in patients who have had hip surgery. also fall into this category. A few other studies deal with patterns of care and the examination of particular outcome measures. Examples are studies of post-hospital outcomes and studies analyzing the application of mortality and hospital readmission (796).

The Veterans Health Administration

The Veterans Health Administration, located within the U.S. Department of Veterans Affairs (VA), provides for much of the health care of veterans of the U.S armed forces. In addition to its hospital system. VA has a long-standing set of supporting research programs that encompass prosthetics, medical care, and health services re-

search. Its budget for these activities in 1992 was $858 million. with a small but significant part of that budget derived from NIH through interagency transfers.

As with HCFA, the primary purpose of the health organizations within VA is to assure that the population for which it is responsible (i.e., veterans) are covered for their health care needs. Unlike HCFA, however, VA delivers these services directly. Consequently, it has developed in-house the clinical research and health services research capabilities that outside of VA are carried out by NIH and AHCPR.

Most research protocols are conducted by VA staff at one of its 172 medical centers around the country (many of which are affiliated with local medical schools). The three VA programs that directly sponsor research into the effectiveness of medical and mental health treatments, as well as the design and development of rehabilitative devices and systems of care are:

- **The Health Services Research and Development (R&D) Service,** the VA's in-house analog to AHCPR, performs most of the organization effectiveness research activities. Health Services R&D sets research priority areas and encourages research into these areas (523). "Outcomes and effectiveness research" was one of these priority areas in 1993 and included a number of projects analogous to the kinds of studies being supported in AHCPR's MEDTEP (e.g., the development of measures of health status and a feasibility assessment for collection of outcomes data on VA patients). In addition, building on the Medical Research Service's Cooperative Studies Program, Health Services R&D is funding multisite projects on such diverse topics as cardiac surgery outcomes, the clinical and cost impact of clozapine treatment on refractory schizophrenia, and a multisite randomized trial of team-managed hospital-based home care (875a).
- **The Medical Research Service,** VA's internal analog to NIH, sponsors its biomedical research and clinical trials projects. Many of its clinical trials are limited to VA patients (mainly elderly and disabled men), but some are potentially of broader applicability. Clinical trials particularly likely to fall into this latter category are large trials that are cosponsored by other DHHS agencies. The VA has considerable history and experience in multisite clinical studies.
- **The Rehabilitation Research and Development Service** has no real analog in the Public Health Service. It primarily conducts basic rehabilitative research, specialized product development, and tests of treatment or device efficacy. However, it also conducts a few descriptive and comparative effectiveness studies of interventions in the area of rehabilitation (e.g., alternative rehabilitation therapies for patients with multiple sclerosis).

Coordinating Research Activities

Congress has designated AHCPR as its lead agency for research on improving the effectiveness of medical care through the evaluation of existing technologies and practices. The agency has had some successes at doing so but has encountered some substantial barriers as well.

Successful examples of intra-agency research coordination include the establishment of six "work groups" that enable research personnel from its various PORTS to meet periodically and discuss methodological issues. such as the use of health status survey instruments, common to all of the teams. As hoped, there has also been some natural coordination between a PORT and a guideline panel on the same topic; for instance, the researcher who was the consulting methodologist to the cataract guideline panel was the principle investigator of the PORT (724a). Guideline panels have several times been influenced by previous or con-

current work done by PORT teams (e.g., the incorporation of the prostate PORT's work on patient preferences into the recommendations of the prostate guideline panel.[10]

Interagency coordination of activities is more demanding and less successful. Underlying mechanisms that exist for coordinating among agencies include:

- **Representation on other agencies' advisory** bodies. AHCPR's advisory committee, for example, includes the administrators of seven other health-related agencies or departments as ex officio members.[11]
- Formal observer status for planning groups or task forces. NIH's OMAR, for example, regularly convenes a group of representatives from the different NIH Institutes to discuss issues for consensus conferences and other concerns. A roster of designated observers from other agencies. including AHCPR. are also routinely invited to these meetings, although the actual attendees from any particular agency may vary from meeting to meeting (78).
- Formal interagency coordinating groups. There appear to be no formal interagency coordinating groups on effectiveness research itself. There is a group of representatives from DHHS agencies that meets regularly to discuss issues in cost-effectiveness methods, convened by ODPHP to complement the work of its advisory task force (see chapter 7).
- **Conferences and workshops with invited participants and observers from other agencies.** Both AHCPR and other agencies frequently sponsor conferences and workshops on specific topics, to which staff and researchers associated with other agencies are often invited as either participants or observers. One classic example was a 1990 conference on primary care research, which was not only attended by but cosponsored by AHCPR. two NIH institutes, two other DHHS agencies. and a private nonprofit research foundation (798). Such efforts require that staff have foreknowledge of other agencies, and departments, interest, and that they act on it, which is not always the case. At a recent workshop on how to include cost-effectiveness considerations in clinical guidelines, for example, it did not occur to AHCPR staff to invite staff from ODPHP. who had been involved for some time in an effort to improve cost-effectiveness methodology. ODPHP staff knew to attend only because they found out about the workshop second-hand (869).
- **Interagency solicitation of cofunding for a study.** This has probably been one of the most successful mechanisms. It led, for example. to a collaborative study of the management of acute ear infections in children. a randomized trial being cofunded by the National Institute for Child Health and Human Development and AHCPR. According to AHCPR staff, this collaboration came about after NICHD sought cofunding for the study from other agencies (824). AHCPR agreed to help fund it after requiring some additions to the study protocol. Among several other examples of cooperative funding include the National Institute of Mental Health cosponsorship of the schizophrenia PORT; AHCPR's input into a prostate treatment study at VA; and AHCPR's funding of a quality-of-life component to be added on to an NHLBI-sponsored trial of treatments for hypertension and high cholesterol
- **Informal contact between staff.** The director of AHCPR'S Office of Medical Effectiveness

[10] AHCPR has several times deliberately assigned the same medical condition to both a PORT and a guidelines panel. This dual attention is in part the result of the priority AHCPR staff have placed for both activities on high-frequency procedures and conditions that have correspondingly high costs to the Medicare program.

[11] The seven agencies and departments are the Food and Drug Administration, the Health Care Financing Administration, the Substance Abuse and Mental Services Administration, the National Institutes of Health, and the Center\ for Disease Control. in the U. S. Department of Health and Human Services; the Department of Veterans Affair\; and the Department of Defense (52).

Research, for example, was previously affiliated with the VA. Informal contact can be a particularly important mechanism for coordination. although it is unreliable over time if more formal mechanisms do not exist because it is dependent on individual staff.

At present, instances of significant, productive cooperation among agencies and their activities seem more the exceptions than the rule, with cosponsorship of studies one of the more successful mechanisms. These both enhance AHCPR's resources to do broader effectiveness research and, presumably, help that research address questions of interest to other constituencies as well. The examples, however, are few.

Despite the problems with coordination among agencies in effectiveness research, formal mechanisms to increase coordination can present their own problems. Numerous people with whom OTA spoke during this study expressed doubts about the usefulness of formal interagency coordinating groups that meet periodically, because the activity tends to be viewed as relatively unimportant by participants, and the individuals participating tend to vary over time. Mechanisms to formally notify agencies about each other's activities also are viewed with skepticism because they tend to be considered a bureaucratic burden that would simply increase paperwork and discourage actual activity. Thus, relying on these mechanisms to increase cooperation may not be effective unless they are very limited and very targeted to specific purposes.

CONCLUSIONS

The crucial question for the next stage in effectiveness research is how to address the gaps that currently exist in this research. Some of these gaps include:

1. Improving the efficient production of systematic reviews of existing studies, to make the best use of past efforts at clinical evaluation and to help identify important areas for research. Possible mechanisms for improvement include increasing funding for meta-anal - ysis (e.g., through specific grants, PORTS, or the U.S. Cochrane Center); requiring investiga[ors proposing new studies to demonstrate. through references to meta-analyses, that the research is not unnecessarily redundant: and maintaining a commitment to establishing a comprehensive database of controlled clinical trials.
2. Conducting more, and more efficient, clinical trials that yield valid comparative information on existing technologies, with results directly useful to patient and clinician decisionmaking. Possible mechanisms for achieving this objective are encouraging the use of patient-oriented outcome measures in more NIH-sponsored clinical trials; establishing and maintaining a comprehensive database of ongoing clinical trials sponsored by the federal government (and, where possible, private industry); and investing in a community-based research infrastructure that could be used for conducting large, community-based clinical trials on topics of broad interest to practitioners and patients.
3. Encouraging more comparative evaluations of newly introduced technologies. Possible mechanisms include offering incentives to manufacturers to conduct comparative studies; encouraging or requiring payers, including government insurers, to link health insurance coverage for new technologies with evaluation of those technologies; and expanding the government role in sponsoring comparative evaluations of new technologies.

At present, attempts to address any or all of these gaps face two major barriers. First, expanding the funding of comparative effectiveness research requires either new resources, which are extremely hard to come by, or a shift of existing resources, which faces the substantial opposition of those currently benefiting from those resources. And second, at present, there is no federal agency within the U.S. Department of Health and Human Services that considers the funding of comparative clinical trials of existing technologies to be one of its major responsibilities. Thus. although the feder-

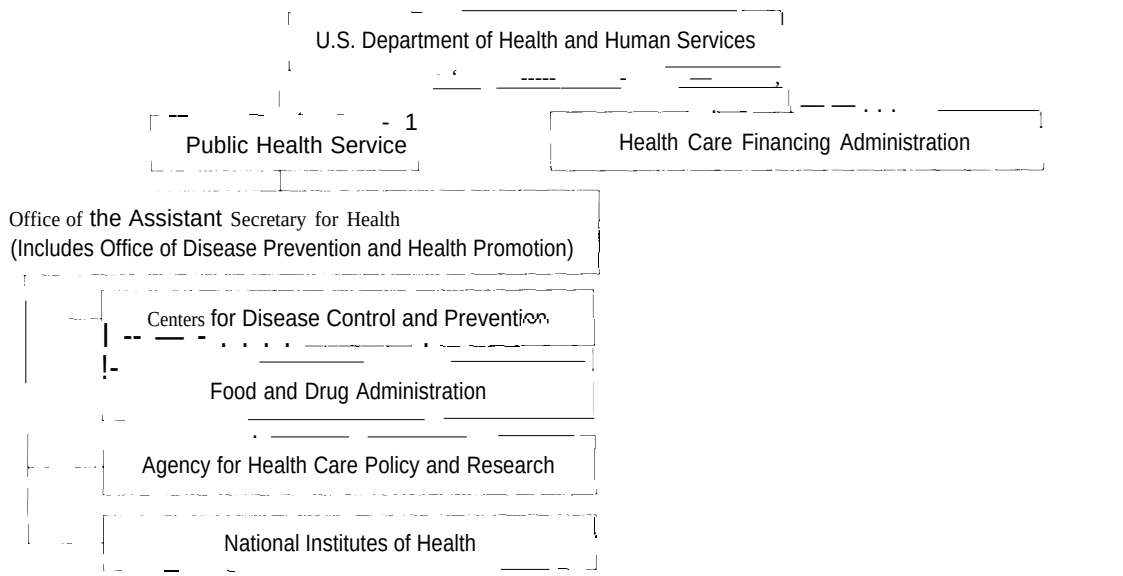
al government is investing resources in identifying high-priority questions about current medical practice, there is no real link between those high-priority questions and the actual comparative research that is being conducted.

The gaps in the federal government's current effectiveness research effort cannot successfully be addressed without assigning responsibility to fill them to a lead agency. Although in some respects the natural lead agency is AHCPR, its current level of funding is insufficient to fill these gaps unless it has the committed cooperation of larger agencies. At present, AHCPR has a moderatc level of interest in comparative clinic al trials. but it has neithcr the funding nor the leverage to ensure cooperation. NIH greater resources for conducting clinical trials are a natural target, but NIH does not view its role as primarily one of supporting evaluations of current technologies. Changing eithcr AH CPR's funding and leverage, or NIH's priori tics. will probably require congressional interest and intervention.

Although its trials arc not linked to AHCPR-generated research priorities, NIH does probably conduct a significant number of relevant clinical trials of existing technologies. The inability at present to identify relevant NIH trials, how its clinical trials funds are allocated, and other questions related to the characteristics of NIH studies deserves attention.

The VA is an under-recognized resource for federally sponsored effectiveness research. Although the population served by the VA has unique characteristics, many of the questions it faces are the same as those faced in the broader non-VA health care system. Consequently. the VA might well be a practical test in: ground for the potential to conduct effectiveness research and translate the results of that research into practice guidelines that can be implemented and evaluated. The VA is also well-organized for large, practice-integrated clinical trials. and for combining health services and clinical research aspects in single studies. Some examples of collaboration between the VA, AHCPR, and NIH exist, and greatcr collaboration might well prove worthwhile. Mechanisms for greater collaboration among these agencies, HCFA, and CDC regarding effectiveness research activities deserve emphasis and exploration.

FIGURE 4-1: Relevant Agencies in the U.S. Department of Health and Human Services



SOURCE Office of Technology 1994