Chapter 17

A Brief History of Evidence-Based Practice and a Vision for the Future

H. Stephen Leff, Ph.D.
Director, The Evaluation Center@HSRI

Introduction

"Evidence-based practices is a new metaphor to me. I thought evidence was the source of mental health practices. Are you saying it was not? May still not be? On what are practices then based? At no time have I ever heard that people in the mental health profession based their acts upon anything other than evidence. Was that a lie? Is this 'new' metaphor also a lie? And how would I tell the difference?" (H. A. Maio, personal communication, May 17, 2002.)

The Evaluation Center@HSRI is a Federally funded technical assistance center for the evaluation of adult mental health system change. The Center encourages dialog with the public about its activities. Recently, the knowledge assessment page of our Web site received the thought-provoking message above, eloquently stating the questions and concerns many people have about evidence-based practices in mental health. This chapter addresses the various questions and concerns raised in that message.

First, the chapter briefly describes the concept of and reviews the history of evidence-based practice. It then describes the types of individuals and organizations currently focusing on evidence-based practices in mental health and the nature of the information they provide. Next, the concerns raised about evidence-based practices in mental health are considered. The chapter concludes with a vision of how evidence-based practices should be pursued in the future, taking into account the concerns that have been raised.

This chapter focuses on psychosocial interventions for adults with severe mental illness treated in the public sector. Nevertheless, this discussion of the concerns about evidence-based practices and its vision for the future is applicable to evidence-based practices in general.

It is important to review the concept of evidence-based practices critically because the shift to such practices has become a movement. "Movement" refers to an organized effort of leaders and followers to identify, disseminate, and cause the adoption of certain practices believed to be different and better than current ones. In summary, this chapter speaks to the following questions:

- What are evidence-based practices?

- Where did evidence-based practices come from?

- How might the concept of evidence-based practices be applied to psychosocial practices in mental health?

- Why are some people so concerned about evidence-based practices?

What Are Evidence-Based Practices?

Evidence-based practices are practices that have been tested employing specified scientific methods and shown to be safe (or relatively safe, since they may have side effects judged acceptable, given their positive impacts), efficacious, and effective[1] for most persons with a particular disorder or problem. These methods are similar, but not identical, to ones required by the Food and Drug Administration (FDA) and will be discussed further. However, there is no FDA for mental health treatments other than drugs, and there are some disputes about how close we can come to scientific standards in testing mental health interventions.

---

[1] The distinction between efficacy and effectiveness is an important one for discussions of evidence-based practices. This chapter is concerned with practices demonstrated to be effective as well as efficacious. This concept is discussed further below in the section on Professional/Trade Organizations.

For those who believe we cannot truly approximate such standards, evidence-based practices are only a metaphor, the products of processes that have some qualities of science, but are not literally science-based.

That a practice is evidence-based, in the sense intended here, is clear if a body of scientific research shows that the practice has specific effects that are replicable independent of who does the research. Establishing an evidence base involves either consulting secondary reviews of studies or synthesizing the results of single studies. Ideally, there should be guidelines for identifying evidence-based practices, involving meta-analyses of research findings that quantitatively synthesize the available information.

As H. A. Maio's message suggests, many mental health professionals follow practices for which they believe there is evidence. However, the evidence may be in the form of their own experiences, the experiences of their teachers, or the experiences of their clients (often referred to as anecdotal evidence). This type of evidence is limited in its usefulness. First, it does not distinguish between changes that happen as a result of treatment and those that happen because of factors such as maturation; the assistance of friends, family, and community caregivers; and the passage of time. Second, anecdotal evidence may describe the experiences of only a self-selected group of persons, not the experience most people will have. Third, anecdotal evidence is subject to bias. Caregivers and service recipients wish treatments to work for many reasons—some humanitarian, some financial, and some ideological (Chambless and Ollendick, 2001). The perception that a treatment has worked can therefore be the result of wishful thinking. We are surer that a treatment has worked when independent observers agree that it has. In relying on anecdotal evidence, mental health caregivers are not that different from providers of physical health care. Millenson (1997) estimates that 85 percent of everyday medical treatments have never been scientifically validated.

Today the mental health field is coming to a different understanding of what is acceptable evidence. This understanding is based on the evolution of the medical, social, and behavioral sciences.

The History of Evidence-Based Practices

We think of medicine as being based on scientific knowledge. However, if we define scientific knowledge as knowledge derived from true experiments (referred to as "randomized clinical trials" in medicine) or quasi-experiments that address the threats to validity in other than randomized clinical trials (Campbell and Stanley, 1966), then this has not always been the case. In medicine, analysts point to three landmarks on the road to evidence-based practices.

One is the Flexner Report (Millenson, 1997) which created a blueprint for medical education based on a rigorously scientific curriculum. The second is medicine's first randomized clinical trial, a study of the efficacy of streptomycin in treating tuberculosis, which appeared in a 1948 issue of the *British Medical Journal,* and placed clinical judgment within a new scientific framework (Millenson, 1997). The third is the establishment of the FDA and related governmental organizations with the mission of testing the safety and effectiveness of medical interventions.

The FDA, as we know it today, evolved over time. First, partly in response to the publication of Upton Sinclair's novel, *The Jungle*, Congress passed the 1906 Food and Drug Act, specifying that the main ingredients of foods and drugs had to be identified on package labels and that the labels could not be misleading (Healy, 1997). The next major event in the evolution of the FDA was the Food, Drug and Cosmetics Act passed in 1939. This Act prohibited the marketing of any preparation of a compound until it had been accepted as safe by the newly created agency, the Food and Drug Administration (Healy, 1997). The final defining event was the passage of the 1962 Kefauver-Harris drug amendments. This legislation, prompted by birth defects caused by the drug thalidomide, put in place "more rigorous requirements on manufacturers to satisfy the FDA that a new compound was both safe and efficacious for the ailment for which it was designed" (Healy, 1997, p. 26). A major consequence of these amendments was to institutionalize the view that randomized, placebo-controlled, double blind trials are the gold standard to establish the efficacy of an intervention (Healy, 1997).[2] Another major consequence was to underline the importance of testing interventions for their safety. The FDA does

---

[2] The 1962 amendments addressed not only prescription but also over-the-counter drugs. Healy (1997) reports that it was estimated that there "might be up to half a million OTC products on the market. … A preliminary investigation of five hundred suggested that anywhere between half and three-quarters were ineffective. As a result of FDA scrutiny, a large number of 'antidementia' drugs and 'antidepressants' vanished" (p.27).

not simply evaluate the efficacy and effectiveness of interventions; it also asks what risks are associated with this intervention and how those risks compare with its benefits. This idea that both the safety and effectiveness of interventions need to be evaluated is another underlying theme in the move to evidence-based practices.

Prior to this recognition, medical practitioners were perceived to have special knowledge, but that knowledge was not necessarily based on science. As Freidson (1970) notes,

> The professional is an expert because he is thought to possess some special knowledge unavailable to laymen who have not gone through his special course of professional training. His special professional knowledge may not be demonstrably and consistently efficacious, but it is the best available to the times, and it is taught to all members of the profession in order to prepare them for the proper performance of their work. (p. 338)

More specifically, as Healy (1997) notes, "…impressions of both efficacy and safety hinged on the testimonials of a few clinicians rather than on demonstrable effects from multicenter studies and a systematic cataloguing of adverse events" (p. 26).

Two other advances are important in the evolution of evidence-based practices. The first is the highly influential work by Donald Campbell on approaches to deriving causal inferences from quasi-experiments. Campbell's best known work, *Experimental and Quasi-Experimental Designs for Research*, appeared in 1966.

The second advance was the emergence of meta-analysis. Hunt (1997) describes meta-analysis as

> a means of combining the numerical results of studies with disparate, even conflicting, research methods and findings; it enables researchers to discover the consistencies in a set of seemingly inconsistent findings and to arrive at conclusions more accurate and credible than those presented in any one of the primary studies. More than that, meta-analysis makes it possible to pinpoint how and why studies come up with different results, and so determine which treatments—circumstances or interventions—are most effective and why they succeed. (p. 1ff)

Meta-analysis was anticipated in early work by Karl Pearson, a British mathematician studying the effectiveness of inoculation against typhoid fever (Hunt, 1997). In 1937, William Cochrane, a British biostatistician, developed a key technique of meta-analysis, a method for combining the effect sizes reported in different studies (Hunt, 1997). In an influential book published in 1972, Archibald Cochrane, a British epidemiologist, drew attention to the fact that people who wanted to make more informed decisions about health care did not have ready access to reliable reviews of the available evidence. Cochrane was highly influential in Britain, and the world's preeminent nongovernmental organization for conducting research syntheses and meta-analyses, The Cochrane Collaboration (http://www.cochrane.org/; see the appendix for a list of Web sites for all organizations, centers, and groups referenced in this chapter), was named after him. The Cochrane Collaboration is discussed later in this chapter. However, what Hunt (1997) refers to as the "meta-analysis movement" (p. 12) is generally acknowledged to have begun in 1976 with a speech and then a publication by Gene V. Glass (1976) of a method for combining studies of psychotherapy.

Campbell's contribution was important because he gave coherence and credibility to the idea that, given certain analytical methods, causal inferences could be drawn from studies that were not strictly experimental. This idea is important to studying the safety and effectiveness of certain nonpharmacological interventions in health and mental health for which conducting true experiments has proven problematic. These problems arise from difficulties in areas such as defining placebo controls, gaining acceptance for random assignment, and controlling the fidelity of practices.

The development of meta-analysis was important to the emergence of evidence-based practices for two reasons. First, research and evaluation activities in science are "anarchic" (Hunt, 1997, p. 4); they are not organized prospectively into logical steps, except when these are required to meet the requirements of the FDA. Meta-analysis offers a methodology for retrospectively synthesizing uncoordinated studies in a systematic manner that approximates a logical process.

The second reason meta-analysis was important to the emergence of evidence-based practices is that it offers a route to a relatively unbiased synthesis of the evidence for interventions. Prior to the development of meta-analysis, research was synthesized in narrative review articles (Hunt, 1997) that left substantial room for bias to intrude. As Chalmers and Lau (as cited in Hunt, 1997) write,

> Too often, authors of traditional review articles decide what they would like to establish as the truth either before starting the review process or after reading a few

persuasive articles. Then they proceed to defend their conclusions by citing all the evidence they can find. The opportunity for a biased presentation is enormous, and its readers are vulnerable because they have no opportunity to examine the possibilities of biases in the review (p. 7).

These advances in education, science, and the role of government, reinforced by the need for efficient treatments given rising health care costs, have led to a belief in the paradigm of evidence-based practices. Millenson (1997) provides a concise statement of this paradigm:

> A health care delivery system characterized by idiosyncratic and often ill-informed judgments must be restructured according to evidence-based medical practice, regular assessment of the quality of care and accountability. The alternative is a system that makes life and death treatment decisions based on conflicting anecdotes and calculated appeals to emotion (p. 6).

This paradigm has also entered into mental health, as exemplified in *Mental Health: A Report of the Surgeon General—Executive Summary* (U.S. Department of Health and Human Services, 1999), which states the following:

> A wide variety of effective, community-based services, carefully refined through years of research, exist for even the most severe mental illnesses yet are not being translated into community settings. Numerous explanations for the gap between what is known from research and what is practiced beg for innovative strategies to bridge it (pp. xix–xx).

The balance of this chapter focuses on the questions posed earlier of how the concept of evidence-based practices might be applied to psychosocial services in mental health and why some people are so concerned about evidence-based practices.

The individuals or organizations that should be responsible for supporting this process are considered first. For medications, the FDA is responsible for defining the tests a medication must pass to be accepted as safe and efficacious. Drug companies engage in the arduous task of attempting to meet these tests. Treatments that pass are allowed on the market; ones that fail are prohibited from use. If we wish to know about an available drug treatment, we can find out at least some information about its safety and efficacy from the packaging. Currently, however, no single locus of responsibility exists for

identifying evidence-based psychosocial interventions for persons with severe mental illness, maintaining a registry or database of such practices, disseminating information about the practices, or updating the registry. Instead, multiple actors sporadically and independently support such activities. This chapter considers who these actors are and whether the current situation meets stakeholder needs.

Second, I describe the concerns in identifying evidence-based practices in mental health and indicate how they might be addressed. Will evidence-based practices bring safer and more effective interventions to stakeholders or just provide a new label or "metaphor" for services that have no more foundation in evidence than previous ones?

Third, I discuss the process that should be used to determine the degree to which psychosocial interventions for persons with severe mental illness are evidence-based. Put differently, how can H. A. Maio tell the difference between mental health practices that are more evidence-based and ones that are less so? Note, though, that this question cannot be answered simply. Different amounts of evidence support different psychosocial interventions. Ultimately, we need a system for grading the quality, strength, and consistency of evidence on a continuum. The history of the evidence-based practices movement suggests that the process should involve research synthesis and meta-analysis. However, other factors must also be considered.

Sources of Evidence-Based Practices

In the absence of an FDA for psychosocial mental health practices, various individuals and organizations have assumed the responsibility of identifying evidence-based practices. They include academic researchers, trade organizations, organizations of scientists committed to synthesizing research results, some government agencies with scientific missions, and some advocacy organizations.

This field of evidence-based practices can be compared to a baseball game, although the individuals and groups described in this section may not cover all the players in this serious game. Their activities are well intentioned and have advanced the field, but their activities are also "anarchic" (Hunt, 1997) and therefore confusing to consumers, providers, and other nonscientists. A number of individuals and organizations (players) are behaving accord-

ing to a loosely specified, and to some extent diverse, set of rules. Moreover, there is no "league" in the sense of an organization responsible for defining the rules and assisting people to play by them. And there are no umpires, or individuals who have been given the authority to apply the rules. The problem, therefore, is knowing which interventions are winning and which are losing. Another important point is that it's not clear whether the game is fair, that is, whether all promising practices have had equal opportunity to be tested. It can be argued that this game is market-driven; however, markets cannot work if participants do not have reliable information, and reliable information is difficult to identify. These qualifications result in the important concerns about evidence-based practices described below.

Researchers and Evaluators in Academic or Other Settings Pursuing Their Own Research Interests

Individual researchers, often in academic settings, who prepare narrative reviews or meta-analyses for publication in journals and books have probably taken the lead in identifying evidence-based practices. Pikoff (1996), for example, has compiled summaries and analyses of 242 clinical research reviews published in mental health and substance abuse journals.

These syntheses have no authority beyond the reputations of their authors and the journals in which they appear (peer-reviewed journals having the highest credibility) and the scientific quality of the works themselves. Different syntheses adhere to different rules; some are narrative syntheses, others use meta-analysis. Some include only investigations that were randomized trials, whereas others include quasi-experimental and pre-experimental studies. Over time, syntheses will include different studies. The list of ways in which syntheses can differ is quite long. Not surprisingly, different syntheses can and do reach different conclusions. There is no accepted process for updating syntheses and resolving differences among them, nor is any organization charged with offering the public updated information from the latest syntheses.

Voluntary Organizations of Scientists Committed to Evidence-Based Practices

The Cochrane Collaboration

As noted above, Archibald Cochrane was a pioneer in the development of methods for synthesizing evidence. In 1992, a group of British scientists at Oxford University established a collaboration to identify evidence-based practices in medicine named in honor of Cochrane. The Collaboration was formally established in 1993 by 77 individuals from 11 countries. Today the Collaboration consists of 50 collaborative review groups composed of researchers, health care professionals, consumers, and others.

Cochrane reviewers employ methods of synthesizing evidence from work developed by the Cochrane Methods Groups, created to improve the validity and precision of systematic reviews. Currently, methods groups are formed in the following areas: Applicability and Recommendations, Health Economics, Health-Related Quality of Life, Individual Patient Data Meta-Analyses, Methodology Review Group, Nonrandomized Studies, Prospective Meta-Analysis, Reporting Bias Methods Group, Screening and Diagnostic Tests, and Statistical Methods.

The purpose of this international body is to help people make informed decisions about health care by "preparing, maintaining and ensuring the accessibility of systematic reviews of the effects of health care interventions" (The Cochrane Collaboration, 2002). The collaboration is based on 10 principles:

(1) Collaboration

(2) Building on the enthusiasm of individuals

(3) Avoiding duplication

(4) Minimizing bias

(5) Keeping up to date

(6) Striving for relevance

(7) Promoting access

(8) Ensuring quality

(9) Continuity

(10) Enabling wide participation

Following these principles, the Cochrane Collaboration created and maintains the Cochrane Library, which consists of almost 1,500 syntheses of medical and behavioral interventions. The key words or phrases "schizophrenia," "affective disorder," "mental health," and "psychosocial interventions," bring up 20 entries for psychosocial intervention for adults with serious mental illness. This is by far the largest number of syntheses available from any source. Many more interventions are contained in the library for psychoactive medications and for interventions for children.

The Cochrane Collaboration's principles are ones to which any organization charged with informing the public should adhere. It is striking that the Cochrane Collaboration, a voluntary organization, has accomplished as much as it has. Nevertheless, the Cochrane Collaboration's reviews are highly technical and usually involve little input by nonscientist stakeholders, such as consumers and advocates. The Collaboration also makes no effort to reconcile its reviews with those of others that reach different conclusions.

The Campbell Collaboration

In 1999, a group of American scientists founded and named a collaboration, modeled after the Cochrane Collaboration, in honor of Donald Campbell. The Campbell Collaboration (http://www.campbellcollaboration.org), designed to identify evidence-based practices for interventions, includes researchers from the United States, Great Britain, Canada, and Sweden. It is pledged to prepare and maintain systematic reviews of studies of the effects of policies and practices in education and the social and behavioral sciences. Using standards for quality of evidence considered transparent and criticizable, the Collaboration has solicited contributions by researchers in such fields as criminal justice and substance abuse that meet the needs of those with a strong interest in high quality evidence of "what works." The Campbell Collaboration sees itself as paving the connection between knowledge assessment and policymaking. The extent to which the Campbell Collaboration produces or supports syntheses on psychosocial interventions for persons with serious mental illness, and if it does, how it will coordinate with the Cochrane Collaboration, not to mention other groups, remains to be seen.

Professional/Trade Organizations

One way professional organizations contribute to the identification of evidence-based practices is by supporting the publication of systematic reviews and meta-analyses in journals. For example, *Psychiatric Services*, a journal of the American Psychiatric Association, has included a special section, "Focusing on Evidence-Based Practices," in its issues. The production of these articles usually depends on the initiative of individual scientists.

Another way professional organizations participate in identifying evidence-based practices by issuing practice guidelines. Many professional organizations, such as the American Psychiatric Association, the International Society of Psychiatric-Mental Health Nurses, and the National Association of Social Workers, issue practice guidelines. Some may assess the evidence for practices as a part of practice guideline development. There are too many such groups to describe all their efforts here. However, various sites with guidelines can be found on the Web, and the National Guideline Clearinghouse (NGC) maintains a searchable database of evidence-based clinical practice guidelines (http://www.guideline.gov/body_home.asp). NGC is sponsored by the Agency for Healthcare Research and Quality, the American Medical Association, and the American Association of Health Plans. Its site can be searched by disease, intervention, and organization. Relevant guidelines can be identified and compared. Guideline comparisons include information on methods used to collect evidence, to analyze evidence, and to assess the quality and strength of evidence.

These efforts by professional organizations tend not to measure the strength of evidence for practices in depth (West et al., 2002). They rarely conduct systematic reviews or meta-analyses, relying heavily on expert opinion or narrative reviews. Moreover, their conclusions tend to be influenced by the perspective of the particular professional group making the guideline recommendations.

What happens when a professional organization presents evidence that some of its members view as harmful to the interests of the profession is illustrated by the case of the American Psychological Association? Chambless and Ollendick (2001) present a detailed analysis of this example.

In 1995, the Task Force for Promotion and Dissemination of Psychological Procedures, a task force of the Clinical Psychology Division (Division 12) of the American Psychological Association, issued the first of three reports identifying a number of psychological interventions as empirically supported

treatments (ESTs). The Division 12 Task Force had little to say about the treatment of serious mental illness; it listed only three therapies that "some evidence suggests" are useful in treating schizophrenia and other severe mental illness: (1) family interventions, (2) social skills training, and (3) supported employment. Nevertheless, the history and evolution of ESTs in psychology say much about the problems of expecting provider groups to identify and disseminate evidence-based practices. The reports of the Division 12 Task Force "reaped both praise and opprobrium" (Chambless and Ollendick, 2001, p. 2). Eventually, the American Psychological Association decided it would not pick up the work of creating and maintaining the list of evidence supported treatments. These activities are being continued by a standing committee of Division 12. The Division disseminates information on evidence-based practices through its quarterly publication of *The Clinical Psychologist* along with a guide to empirically supported treatments in the areas of mental health, including the areas of "Anxiety Disorders and Stress," "Depression," and "Schizophrenia and Other Severe Mental Illnesses." Under the "Depression" and "Schizophrenia and Other Severe Mental Illnesses" treatment categories, for example, brief descriptions of the disorder are given along with narrative summaries of psychological interventions with proven results (http://www.apa.org/divisions/div12/rev_est/index.shtml).

The Division 12 list of ESTs ran afoul of a number of issues. Among them were "guild or economic" concerns about how managed care might use such a list, fears that practitioners of psychotherapies not on the list would be "disenfranchised," and the worry that such lists would make practitioners more vulnerable to malpractice suits (Chambless and Ollendick, 2001). Any movement to evidence-based practices will have to confront this type of "guild" resistance when it might be contrary to the interests of providers.

The list of empirically supported treatments ran into technical criticisms as well. One criticism was that the criteria for deciding what is evidence-supported and the methods for reaching decisions were unclear. A related concern was that the criteria were too lenient. Additional concerns were that the available research focused on outcomes that were too narrow and reflective of only certain perspectives. Finally, there was concern that evidence is lacking about whether evidence-supported treatments work with subgroups that did not participate in the testing of the treatments and, consequently, about how these treatments might have to be modi-

fied to be relevant to these subgroups. These criticisms have been raised not only by psychologists about ESTs but also by different stakeholder groups about evidence-based practices generally.

Division 12 also commissioned *A Guide to Treatments That Work,* edited by Nathan, Gorman, and Salkind (1999). This book was produced by a task force of experts separate from the ESTs. A similar publication, *What Works for Whom*, was prepared by Roth and Fonagy (1996) pursuant to a commission from the National Health Service Executive of the English Department of Health. Chambless and Ollendick (2001) note that these different workgroups did not use the same categories for indicating the degree to which treatments were evidence supported, nor did they define evidence in exactly the same ways.

A comparison of the psychosocial interventions for severe mental illness reviewed by the Cochrane Collaboration, the Division 12 Task Force, Nathan and colleagues (1999), Roth and Fonagy (1996), and Pikoff (1996) shows the unevenness with which these interventions are covered by the different reviewers and how findings differ for several interventions reviewed by most sources. However, before making that comparison, one more set of players, Federal agencies, should be considered.

Federal Agencies

Although there is a history of evidence-based government regulation of pharmacotherapies, Federal government agencies generally have not taken a similar approach in the past when it came to psychosocial interventions for persons with severe mental illness. The theory in this area seems to be that the identification and dissemination of evidence-based practices in mental health can be left to the market place in which consumers, armed with information supplied by providers (and occasionally the Government) about their choices, decide what mental health care they desire. The Government does take some responsibility for influencing practice by licensing practitioners, providing limited funding for research to develop and test new interventions, and funding promising practices. However, with few exceptions, these efforts are influenced by providers, scientists, and advocates acting independently and with different information. Consequently, although they often make real contributions to care, Government efforts tend to be reactive, loosely connected to the evidence, and un-

systematic, rather than evidence-based, orderly, and coordinated.

A completely market-based approach to psychosocial services seems problematic given the evidence that providers can supply biased information and that persons with severe mental illness are a particularly vulnerable group. Several governmental agencies contribute to and support the identification and dissemination of evidence-based practices, and the roles of these agencies might be enlarged.

The National Institute of Mental Health

The mission of the National Institute of Mental Health (NIMH) is to diminish the burden of mental illness through basic scientific research. It carries out its mission by funding and disseminating research on the nature of mental illness and its treatment, focusing on the areas of basic neuroscience, behavioral science, and genetics. The agency consists of five divisions, one of which, the Division of Services and Intervention Research, houses the Services Research and Clinical Epidemiology Branch (http://www.nimh.nih.gov/dsir/index.cfm). This branch supports and conducts research programs to improve the quality and outcomes of treatment and rehabilitation services. Most of the psychosocial research NIMH funds is field initiated. NIMH takes no particular role in reviewing or synthesizing evidence, and its mission is not to create or maintain an "official" list of evidence-based practices.

The Substance Abuse and Mental Health Services Administration

The Substance Abuse and Mental Health Services Administration (SAMHSA), located in the Department of Health and Human Services, is the Federal Government's lead agency for improving the quality and availability of substance abuse prevention, addiction treatment, and mental health services in the United States. It exercises leadership by providing strategic funding to increase effectiveness and availability of services. It also develops and promotes quality standards for service delivery, models for training, and effective data collection and evaluation. Mental health services are addressed by Center for Mental Health Services (CMHS) SAMHSA's. Although periodically CMHS does support the synthesis and dissemination of knowledge to improve mental health services, it has not routinely supported the production, maintenance,

and dissemination of this information. That may change as CMHS implements its new "sciences to services" mission (Curie, 2002). Two concrete manifestations of this agenda are the CMHS-supported Evaluation Technical Assistance Center@HSRI (http://www.tecathsri.org), which maintains a searchable database of narrative, systematic, and meta-analytic reviews of interventions for persons with severe mental illness, the EbPMetabase, and the National Association of State Mental Health Directors Research Institute Center of Evidence-based Practices, Performance Measurement, and Quality Improvement (http://nri.rdmc.org/RationaleEBP-CenterReview.pdf).

Other programs within SAMHSA attempt to further the concept of using evidence to guide practice, such as the Addiction Technology Transfer Center (http://www.nattc.org), supported by the Center for Substance Abuse Treatment (CSAT). A particularly noteworthy program from the perspective of identifying evidence-based practices is the National Registry of Effective Programs (NREP; http://preventionpathways.samhsa.gov/nrep/default.htm), supported by SAMHSA's Center for Substance Abuse Prevention (CSAP). NREP is a national registry of effective drug and alcohol prevention programs to guide stakeholders in identifying effective prevention programs. NREP funds an independent contractor to have outside experts in the evaluation of prevention programs review and rate the evidence for the effectiveness of prevention programs. Ratings are made according to 15 criteria: (1) basis in theory, (2) intervention fidelity, (3) process evaluation, (4) sampling strategy and implementation, (5) attrition, (6) outcome measures, (7) missing data, (8) data collection, (9) analysis, (10) other plausible threats to validity, (11) replication potential, (12) dissemination capability, (13) cultural and age-appropriateness, (14) integrity, and (15) utility of the intervention. Prevention programs are asked to supply the necessary information on a voluntary basis. The NREP program comes closer than any other Government activity to providing, maintaining, and disseminating information on the evidence-base for psychosocial interventions; however, it does not include psychosocial interventions for severe mental illness, and its criteria for rating the evidence base for programs do not require rigorous syntheses of all available evidence. Efforts to expand the NREP program are under way, and it may emerge as a major governmental effort to promote a science-to-services agenda.

Agency for Health Care Research and Quality Evidence-Based Practice Centers

The mission of the Agency for Health Care Research and Quality (AHRQ), formerly the Agency for Health Care Policy and Research (AHCPR), is to support research designed to improve the outcomes and quality of health care, reduce its costs, address patient safety and medical errors, and broaden access to effective services. In the area of identifying and disseminating evidence-based practices for severe mental illness, AHRQ (then AHCPR) joined with NIMH to sponsor the development of the Schizophrenia Patient Outcomes Research Team (PORT) Treatment Recommendations (Lehman and Steinwachs, 2001) and the Depression Patient Outcomes Research Team Recommendations (Agency for Health Care Research and Quality, 2003, para. 1). These recommendations were based on narrative reviews of the treatment literature. For example, the Schizophrenia PORT provides brief, narrative reviews of the psychosocial interventions: family interventions, vocational rehabilitation, case management, and assertive community treatment. The PORT recommendations have been widely disseminated and are currently being updated.

AHRQ has stepped back from guideline development because of controversy stimulated by various provider groups (Hermann, unpublished 2002); however, it has continued its leadership in identifying the evidence-base for practices. A current AHRQ program that relates directly to the identification and dissemination of evidence-based practices is its Evidence-based Practice Program. Under this program, the agency has funded 13 Evidence-based Practice Centers (EPCs; http://www.ahrq.gov/clinic/epc) to develop evidence reports and technology assessments on clinical topics using rigorous, comprehensive syntheses and analyses of relevant scientific literature. The EPCs employ meta-analyses and cost analyses to report on clinical topics that are considered common, expensive, or significant for subscribers to federally funded Medicare or Medicaid programs.

In 2002, AHRQ released a report on systems to rate the strength of scientific evidence (West et al., 2002). This report included sections describing systems for grading the strength of bodies of evidence. These systems incorporate judgments of both study quality and whether the same findings have been detected by others using different studies or different people. The report proposes that any system for rating the overall strength of a body of evidence should address three general areas:

- The **quality of findings,** measured as the quality of all relevant studies for a given intervention, where "quality" is defined as the extent to which study design, conduct, and analysis has minimized selection, measurement, and confounding biases.

- The **quantity of findings,** measured as the magnitude of treatment effect, the number of studies that have evaluated the intervention, and the overall sample sizes of the studies considered.

- The **consistency of findings,** measured as the extent to which similar findings are reported from work using similar and different study designs.

More specifically, the report proposes that systems for measuring the strength of evidence be rated in terms of the domains and elements shown in table 1.

AHRQ would seem to be well suited to identifying, disseminating, and maintaining a registry of evidence-based practices. However, its responsibility is all health care. Psychosocial treatments for severe mental illness are a very small part of all health care; therefore, it is not surprising that no EPC has conducted any syntheses related to psychosocial treatments for severe mental illness. Nor does the recent report on methods for rating the strength of evidence contain any discussion of the special problems that might be associated with measuring the strength of evidence for psychosocial interventions in mental health.

A Comparison of Different Reviews

Table 2 compares interventions reviewed by six different sources: The Cochrane Collaboration; Pikoff (1996); Chambless and colleagues (1998); Nathan, Gorman, and Salkind (1999); Roth and Fonagy (1996); and Lehman and Steinwachs (2001). Of 23 interventions reviewed, only 3 were reviewed by more than half the sources.

Table 3 summarizes the findings by different sources for the three interventions reviewed by more than half the sources. The language used to summarize reviews in this table is intended to be similar to the language in the reviews. This table illustrates that except for individual and group psychodynamic psychotherapy, which all reviewers eschew for persons with severe mental illness, reviews

Table 1. Domains and desirable elements for systematic reviews based on AHRQ report by RTI/UNC EPC

Domain	Desired Elements[1]
Study Question	Clearly specified and appropriate
Search Strategy	• ***Sufficiently comprehensive and rigorous with attention to possible publication biases*** • *Search restrictions justified (e.g., language or country of origin)* • Documentation of search terms and databases used provided • Sufficiently detailed to reproduce study
Inclusion and Exclusion Criteria	• **Selection methods specified and appropriate, with a priori criteria specified if possible**
Interventions	• **Intervention(s) clearly detailed for all study groups**
Outcomes	• **All potentially important harms and benefits considered: symptoms, functioning, service use, satisfaction**
Data Extraction[2]	• Process is rigorous and consistent • Number and types of reviewers is described • Blinding of reviewers • Measure of agreement or reproducibility are provided • Extraction of clearly defined interventions/exposures and outcomes for all relevant subjects and subgroups
Study Quality and Validity	• ***Assessment method specified and appropriate*** • Method of incorporation specified and appropriate
Data Synthesis and Analysis	• ***Appropriate use of qualitative or quantitative synthesis, with consideration of the robustness of results and heterogeneity issues as a function of subject and study characteristics*** • Presentation of key primary study elements sufficient for critical appraisal and replication
Results	• **Quantitative summary statistics and measures of precision, as appropriate**
Discussion	Conclusions supported by results with possible biases and limitations taken into consideration
Funding or Sponsorship	• ***Type, sources of support for included studies and review detailed***

[1] Elements in italics are those with an empirical basis. Elements in bold are those considered essential to give a system a "yes" rating for the domain.
[2] Domain for which a "yes" rating requires that a majority of elements be considered.

present their results using different terms and in some cases reach different conclusions. The reviews also used different methods for determining the evidence for interventions. These facts point to the need for a standard approach to assessing the evidence for psychosocial interventions for persons with severe mental illness, one that adheres to some set of standard criteria (rules, back to the game metaphor), such as the ones suggested in the AHRQ report, and that presents conclusions using standardized terms. They also imply the need for an organization (a league) to support the identification and maintenance of a registry of evidence-based practices that presents the latest reviews and decides how to reconcile conflicting conclusions (umpires). The absence of these and other elements in how we now identify and disseminate information about evidence-based practices has raised a number of concerns about this movement.

Table 2. Psychosocial interventions reviewed by six different sources

Intervention	Sources					
	Cochrane Collaboration (2002)[1]	Pikoff (1996)[2]	Chambless et al. (1998)	Nathan et al. (1999)	Roth and Fonagy (1996)	Lehman and Steinwachs (2001)
Schizophrenia						
1 Assertive community treatment	X[1a]			X		X
2 Case management for people with severe mental disorders	X[1b]					X
3 Cognitive behavior therapy	X[1c]	X[2a,2b]			X	
4 Cognitive rehabilitation	X[1d]					
5 Community mental health teams (CMHTs)	X[1e]					
6 Consumer operated and self-help					X	
7 Crisis intervention for people with severe mental illnesses	X[1f]					
8 Day centers for severe mental illness	X[1g]					
9 Day hospital versus outpatient care	X[1h]					
10 Early intervention		X[2c]				
11 Family psychosocial (but not educational) intervention	X[1I]		X		X	
12 Family psychoeducation	X[1j]	X[2d,2e]		X	X	X
13 Individual and group psychodynamic psychotherapy and psychoanalysis	X[1k]	X[2f]			X	X
14 Patient-held clinical information for people with psychotic illnesses	X[1l]					
15 Prompts to encourage appointment attendance	X[1m]					
16 Psychoeducation	X[1n]					
17 Seclusion and restraint	X[1o]					
18 Short stay hospitalization	X[1p]					
19 Social skills (life) programs, social learning	X[1q]	X[2g,2h]		X	X	
20 Supported housing for people with severe mental disorders	X[1r]		X			
21 Token economy and behavior therapy	X[1s]			X		
22 Treatment programs for people with both severe mental illness and substance misuse	X[1t]					
23 Vocational rehabilitation (Supported employment [SE], prevocational training [PVT], standard care [SC])	X[1u]		X			X

[1] The interventions reviewed by The Cochrane Collaboration (2002) are found in The Cochrane Library (http://www.update-software.com/Cochrane/default.HTM), a database of multiple sources. These sources will appear as separate footnotes. [1a]Marshall & Lockwood (2002). [1b]Marshall, Gray, Lockwood, & Green (2002). [1c]Cormac, Jones, & Campbell (2002). [1d]Hayes & McGrath (2002). [1e]Tyrer, Coid, Simmonds, Joseph, & Marriott (2002). [1f]Joy, Adams, & Rice (2003). [1g]Catty, Burns, & Comas (2002). [1h]Marshall, Crowther, Almaraz-Serrano, & Tyrer (2002). [1i]Pharoah, Mari, & Streiner (2002). [1j]Pekkala & Merinder (2002). [1k]Malmberg & Fenton (2002). [1l]Henderson & Laugharne (2002). [1m]Reda & Makhoul (2002). [1n]Pekkala & Merinder (2002). [1o]Sailas & Fenton (2002). [1p]Johnstone & Zolese (2002). [1q]Nicol, Robertson, & Connaughton (2002). [1r]Chilvers, Macdonald, & Hayes (2002). [1s]McMonagle & Sultana (2002). [1t]Ley, Jeffery, McLaren, & Siegfried (2002). [1u]Crowther, Marshall, Bond, & Huxley (2002).

[2] The interventions reviewed by Pikoff are found in Pikoff (1996). [2a]Alford, & Correia (1994). [2b]Bentall, Haddock, & Slade (1994). [2c]Birchwood (1992). [2d]DeJesus Mari & Streiner (1994). [2e]Strachen (1986). [2f]Mueser & Berenbaum (1990). [2g]Benton & Schroeder (1990). [2h]Liberman, Kopelowicz, & Young (1994).

Table 3. A Comparison of reviews of interventions reviewed by more than half of sources

Intervention	Sources				
	The Cochrane Collaboration (2002) (Date of study depends on review)	Pikoff (1996) (Date of study depends on review)	Nathan et al. (1999)	Roth and Fonagy (1996)	Lehman and Steinwachs (2001)
Family psycho-education	Possibility of usefulness (2000)	"Moderate success" (1994)	"Several well-controlled studies support use" (1999)	Does not reduce relapse	Reduce one year relapse rate
Individual and group psycho-dynamic psycho-therapy and psychoanalysis	No evidence of any positive effect (2002)	Overall results discouraging (1990)	Psychodynamically oriented therapy not shown to be especially effective	Evidence appears to be largely negative	Lack of evidence makes a strong case against use
Social skills training	No clear effect demonstrated (2002)	Gains in specific skills, marginal effects on general functioning and symptoms (1990)	Evidence of helpfulness excellent	Generalized benefit remains to be determined	

Concerns About the Identification and Dissemination of Evidence-Based Practices

This section discusses six of the most important concerns about evidence-based practices. These concerns will have to be addressed as we move to evidence-based practices.

The Democratic Concern

The democratic concern is the idea that consumers of mental health services should participate, but have not, in research and evaluation about these services. This concern stems from the democratic principle that persons should participate in the decisions that affect their lives. It is a concern commonly voiced by consumer advocates and advocates for ethnic minorities (Bernal and Scharron-Del-Rio, 2001; Frese et al., 2001; Marzilli, 2002).

The democratic principle includes a corollary: that the values of consumer groups are likely to differ from those of scientists, and that these values should influence what interventions are investigated, how risks and desired outcomes are defined, and other aspects of research methods. Bernal and Scharron-Del-Rio (2001), for example, are invoking the democratic principle and its corollary when they argue that mainstream therapeutic approaches promote individualistic rather than collectivist or interdependent values, which minorities often endorse, and that these differences require the development of "methodological pluralism" in testing interventions. Gomory (1999) is invoking the principle and its corollary when he states, "professionally defined expectations of client change can be coercive and patronizing, and ultimately harmful" (p. 7).

This principle has been challenged by those who argue that traditional science is the best method we have for determining what services work and that the views of nonscientists have nothing to do with this. This view has been fostered by the development of evidence-based practice. Healy (1997)

states that another consequence of the emergence of the FDA was the establishment of the view that medical experts should decide what constitutes a disease and what constitutes outcomes for the treatment of that disease. He also notes that since the founding of the FDA, critics have been troubled by the cooperation between Government agencies and manufacturers in setting the rules of the evidence-development game. Kitcher (2001) refers to the situation in which all research decisions are made by "scientific subcommunities" as "internal elitism." He refers to the situation in which all research decisions are made by scientists and "a privileged group of outsiders, those with funds to support the investigations" as "external elitism" (p. 133). In a later section, a scientific process that blends scientist and stakeholder participation in a manner that Kitcher labels "well-ordered science" is described.

The Concern that Traditional Science Is Limited

This concern is that the traditional, or conventional, scientific model is unable to capture the evidence for many practices that providers, ethnic subgroups, and consumer groups believe are efficacious and effective interventions. This concern has been around for some time. For example, in an 1835 book, *Research on the Effects of Bloodletting,* Louis wrote that patients who were bled remained sicker for longer and had higher death rates (Millenson, 1997). As Millenson notes,

> Not surprisingly, outraged leech users questioned Louis's methods. [One], for example, warned that mathematical calculations threatened to substitute "a uniform, blind and mechanical routine for the action of the spirit and individual genius of the [physician] artist." (98ff)

More recently, some psychologists (Chambless and Ollendick, 2001) and "social constructionist" therapists have echoed similar themes. The most extreme members of this group, "radical-critical" therapists, are "dismissive of empirical research" and argue that the performance of their methods requires no justification beyond itself (Neimeyer and Raskin, 2001, p. 420).

Bernal and Scharron-Del-Rio (2001) provide an example of this concern from representatives of ethnic minority groups. They state that evidence-based practices "developed within the conventional model of science" are of "questionable utility for ethnic mi-

norities, and many of their limitations are a result of or have been rooted in questionable assumptions of the conventional scientific model" (p. 335). This concern must be addressed by studies of interventions that include adequate numbers of persons from minority groups, include culturally competent interventions, and address outcomes that members of minority groups desire. However, this does not require abandoning the conventional scientific model and its specific methods.

Concerns About Technical Problems in Identifying Evidence-Based Practices

Another set of concerns has to do with technical problems in identifying evidence-based practices. A recent article by Leff and Mulkern (2002) reviews a representative group of these problems. One technical problem is the need to develop ways of identifying and judging evidence that match the developmental stages of interventions. The argument is made below that interventions should be developed following a logical progression from clarifying the nature of the interventions to testing them. Different types of evidence will be required for the different developmental stages; however, different guidelines for different interventions at different stages have not been formalized. Another technical problem is the lack of guidelines for identifying and combining appropriate program contrasts. It is difficult to operationally define placebo-like controls for psychosocial interventions (for example, defining what a placebo control for supported housing would be), as it is difficult to fully understand the nature of controls described with such terms as "services as usual." When different studies use different control groups, we will need guidelines for how the data they provide can be synthesized. A third technical problem is that of designing strategies for sampling persons and treatment settings that provide us with evidence about treatment effectiveness, not just efficacy (Wolff, 2000). We cannot synthesize data about how interventions work in a variety of routine clinical settings with diverse subgroups if such data are not collected. Collecting data on how well interventions work in different settings and with diverse groups, however, is logistically difficult and resource intensive. Psychosocial interventions for persons with severe mental illness, such as Assertive Community Treatment, include many ingredients. A fourth technical problem is we need methods for synthesizing data to identify the active ingredients of interventions. Otherwise, we run the risk of pro-

moting programs that devote resources to unnecessary or even harmful ingredients. A fifth technical problem stems from the fact that much of the evidence about psychosocial interventions for severe mental illness comes from quasi-experimental studies that are analyzed using multilevel models. Guidelines are needed on how to synthesize data from such models and how to weight such syntheses, as opposed to syntheses from randomized clinical trials, in making policy and clinical decisions. A final technical problem is that articles in mental health journals do not contain standardized information to facilitate syntheses or to aid replications. Instead, the information is shaped by the preferences of particular authors and reviewers and the pressures on journal space. If there were an FDA-like process that required interventions to pass a standardized series of tests, this would not be as much of an issue. However, if syntheses must rest on independent and uncoordinated studies, then the information in journal articles will be crucial. Efforts have been made to improve the reporting of randomized clinical trials in medicine, but such efforts have not been made for psychosocial interventions for severe mental illness (Begg et al., 1996).

The Overstatement Concern

The overstatement concern is that advocates of evidence-based practices, in their zeal to disseminate and inform the public about these practices, do not sufficiently stress that there is very little evidence about the efficacy or effectiveness of even our most well-researched interventions for specific subgroups (e.g., ethnic groups) and for specific outcomes (e.g., recovery) (Anthony, 2001; Bernal and Scharron-Del-Rio, 2001; Chambless and Ollendick, 2001). This concern is shared by representatives of ethnic minority and consumer groups and some scientists. For example, Bernal and Scharron-Del-Rio (2001) state that, "from the perspective of the conventional scientific model, we know very little if anything about the efficacy of treatments for ethnic minorities" (p. 333). "Thus, the mission to disseminate and inform the public should not over-state the…applicability of [evidence-based practices]. [A simple] list [of evidence-based practices] may actually misinform a significant sector of society" (p. 332).

The Concern that Untested Will Be Interpreted as Equivalent to Ineffective

This is primarily a concern of providers who, for a variety of reasons, prefer to continue to implement practices that are not evidence-based because they have not been subject to scientific testing. Their concern is that payers and potential clients will confuse the fact that their interventions have not been fully tested scientifically with the idea that their interventions are known to be ineffective (Chambless and Ollendick, 2001).

The Concern that Knowing Is Not Equivalent to Practicing

This concern is that even if we identify evidence-based practices, this does not mean that providers will adopt them or implement them with fidelity to the practices that were scientifically tested. As Freidson (1970) and others have noted, there is a lack of "equivalence between knowing and doing" (p. 339). The result of this gap is a need to develop technologies for motivating and training individuals to implement evidence-based practices with fidelity.

A Vision of the Future for Evidence-Based Practice: Well-Ordered Science

The concerns raised above can be addressed by a process referred to as "well-ordered science," a term borrowed from Kitcher (2001, p. 133). Kitcher contrasts well-ordered science with "internal and external elitism" (p. 133).

In well-ordered science applied to determining evidence-based practices, individuals representing diverse stakeholder groups with different initial values would be brought together by some supporting organization(s) to discuss the available courses of inquiry. The idea that a supporting organization should guide the process of bringing psychosocial interventions to the public is not a recent one (Rotter, 1971; Godfried, 1999). "The product of the consideration [sponsored by a supporting organization would be] a collection of lists of outcomes the deliberators would like scientific inquiry to promote (and adverse events to be avoided) coupled with some index measuring how intensely they desire those outcomes" (p. 118). Through iterative discussions, these preferences would be modified to "absorb the needs"

of others with collective preferences reached by consensus or vote. Possible interventional and scientific strategies would then be "assessed" to determine how interventions might be developed and tested. These assessments would be used to select a course for developing an intervention for the priority outcomes. The course selected would follow the rules of traditional science, but it could begin by using observational data to identify candidate practices for testing. These practices might not be evidence-based, but would be "best practices." Best practices are ones that appear best on the basis of all available information, including consumer and provider anecdotes and expert judgment.

The supporting organization or organizations would then fund phased research and evaluation to develop and test the candidate practices (Leff and Hollen, 2002; Leff and Mulkern, 2002). The development phase would consist of writing manuals and workbooks, crafting fidelity measures, selecting or devising outcome measures, and designing training programs (Torrey et al., 2001). The testing phase would address the efficacy and effectiveness of the candidate practices. Testing would involve comparison with no-treatment groups or groups receiving placebo, or with competing interventions that had been tested. Intervention developers would be required to provide detailed and complete reports of their tests, either in technical reports or publications. Next, the supporting organization or organizations would convene consensus groups to use systematic methods for assessing the quality, strength, and consistency of the evidence for the practices studied (West et al., 2002). Evidence would be considered from all scientific tests, ideally synthesized by meta-analysis. The group would then give the practice a score or rating. This rating would inform providers and consumers about the probable effectiveness of the intervention and could be changed by additional research and evaluation. Thus, the supporting organization would have to support ongoing monitoring and be able to convene additional consensus groups. For those interventions that achieved a certain evidence grade, the supporting organization would then fund a dissemination phase in which materials and methods for training providers would be developed to spread the intervention. These materials would include revised and refined manuals, fidelity measures, and outcome measures. Finally, the supporting organization would implement a system for postdissemination monitoring to identify additional uses or risks associated with the interventions discovered in widespread, routine use. This monitoring could also require reconsideration of the evidence for interventions.

Well-ordered science responds to the democratic concern by making a place for consumer and citizen participation in getting to evidence-based practices. The response asserts that science should primarily be about means, whereas consumer participation should be about ends. "Ends," in the case of evidence-based practices, consist of desired outcomes to be sought and undesirable or adverse events to be avoided. The participatory concept in well-ordered science is consistent with consumer, advocacy, research, and evaluation activities in public mental health services (Leff and Mulkern, 2002).

Well-ordered science, however, does leave us with additional issues about participation. One is the problem of representation (Kitcher, 2001). To implement well-ordered science, we will need to operationally define what constitutes adequate representation for stakeholder groups (e.g., how many consumers have to be involved in a consensus group, how should they be selected, and how should their ongoing involvement be structured?). Another issue has to do with the time that well-ordered science takes. The more groups are involved in a process, the longer that process will usually take (Leff and Mulkern, 2002). Yet all stakeholder groups are impatient to identify and implement evidence-based practices.

To respond to the concern that the traditional or conventional scientific model is unable to capture the evidence for many practices that ethnic subgroups and consumer groups believe are efficacious and effective, well-ordered science makes a place for the observational evidence that nonscientists find so compelling in the initial phase of the scientific process. To respond to the overstatement concern, well-ordered science envisions supporting organizations, "tutored" by diverse stakeholder groups convened to play the role of umpire in the testing game, guiding the development and testing of interventions. These organizations, with stakeholder (including scientific) input, decide what should and can be claimed for interventions. To respond to the concern that untested interventions are ineffective, the well-ordered science model provides a process accessible to all to have their interventions tested. This still may leave a period of time when providers cannot claim that their interventions have been tested; however, no alternative exists to this short of "grandfathering" certain interventions into the list of evidence-based practices. Given that untested interventions may be ineffective at best and unsafe at worst, this is not a desirable course for most inter-

ventions. To respond to the concern that knowing is not practicing, the well-ordered science model includes a dissemination phase in which tools are developed to raise the probability that interventions will be implemented with fidelity.

A last and major problem with the vision of well-ordered science is that it does not specify what supporting organizations will have responsibility for convening groups, funding research and evaluation, or deciding when the testing is complete and what its results are. In the development of medications, the pharmaceutical companies fund the research (with some governmental participation) and the FDA attends to the other tasks. As we have seen, there are no funders with comparably deep pockets to finance the development of psychosocial interventions and there is no individual organization or collaborative that consistently fulfills all the other roles of the supporting organizations. The resources and supportive infrastructure for such a process, not to mention concerns of efficiency, require a coordinated governmental response. Recent steps by the Federal government to promote a mental health science-to-services agenda that encourages interagency collaboration provide hope in this regard.

Acknowledgments

REFERENCES

Agency for Health Care Research and Quality. (2003). *Depression PORT publishes latest findings*. Retrieved February 7, 2003, from http://www.ahrq.gov/research/jul99/799ra8.htm.

Alford, B. A., & Correia, C. J. (1994). Cognitive therapy of schizophrenia: Theory and empirical status. *Behavior Therapy, 25*, 17–33.

Anthony, W. A. (2001). The need for recovery-compatible evidence-based practices. *Mental Health Weekly, 5*(11), 5.

Begg, C., Cho, M., Eastwood, S., Horton, R., Moher, D., Olkin, I., et al. (1996). Improving the quality of reporting of randomized controlled trials. *Journal of the American Medical Association, 276*(8), 637–639.

Bentall, R. P., Haddock, G. & Slade, P. D. (1994). Cognitive behavior therapy for persistent auditory hallucinations: From theory to therapy. *Behavior Therapy, 25.*

Benton, M. K., & Schroeder, H. E. (1990). Social skills training with schizophrenics: A meta-analytic evaluation. *Journal of Consulting and Clinical Psychology, 58*, 741–747.

Bernal, G., & Scharron-Del-Rio, M. R. (2001). Are empirically supported treatments valid for ethnic minorities? Toward an alternative approach for treatment research. *Cultural Diversity & Ethnic Minority Psychology, 7*(4), 328–342.

Birchwood, M. (1992). Early intervention in schizophrenia: Theoretical background and clinical strategies. *British Journal of Clinical Psychology, 31*, 257–278.

Campbell, D. T., & Stanley, J. C. (1966). *Experimental and quasi-experimental designs for research.* Boston: Houghton Mifflin Company.

Catty, J., Burns, T., & Comas, A. (2002). Day centres for severe mental illness (Cochrane Review). *The Cochrane Library, Issue 3.* Retrieved July 29, 2002, from http://www.update-software.com/abstracts/ab0001710.htm.

Chambless, D. L., Baker, M. J., Baucom, D. H., Beutler, L. E., Calhoun, K. S., Crits-Christoph, P., et al. (1998). Update on empirically validated therapies, II. *The Clinical Psychologist, 51*(1), 3–16.

Chambless, D. L., & Ollendick, T. H. (2001). Empirically supported psychological interventions: Controversies and evidence. *Annual Review of Psychology, 52*, 685–716.

Chilvers, R., Macdonald, G. M., & Hayes, A. A. (2002). Supported housing for people with severe mental disorders (Cochrane Review). *The Cochrane Library, Issue 3.* Retrieved July 29, 2002, from http://www.update-software.com/abstracts/ab000453.htm.

The Cochrane Collaboration. (2002). *The ten principles of the Cochrane Collaboration.* Retrieved July 3, 2002, from http://www.cochrane.org/cochrane/cc-broch.htm#PRINCIPLES.

Cormac, I., Jones, C., & Campbell, C. (2002). Cognitive behavior therapy for schizophrenia (Cochrane Review). *The Cochrane Library, Issue 3.* Retrieved July 29, 2002, from http://www.update-software.com/abstracts/ab000524.htm.

Crowther, R., Marshall, M., Bond, G., & Huxley, P. (2002). Vocational rehabilitation for people with severe mental illness (Cochrane Review). *The Cochrane Library, Issue 3.* Retrieved July 29, 2002, from http://www.update-software.com/abstracts/ab003080.htm.

Curie, C. (2002). Testimony to the House Appropriations Subcommittee on Labor/HHS/Education Hearing. Washington, DC.

DeJesus Mari, J., & Streiner, D. L. (1994). An overview of family interventions and relapse on schizophrenia: Meta-analysis of research findings. *Psychological medicine, 24*, 565–578.

Freidson, E. (1970). *Profession of medicine: A study of the sociology of applied knowledge.* New York: Dodd, Mead & Company, Inc.

Frese, F. J., III, Stanley, J., Kress, K., & Vogel-Scibilia, S. (2001). Integrating evidence-based practices and the recovery model. *Psychiatric Services, 52*(11), 1462–1468.

Glass, G. V. (1976). Primary, secondary, and meta-analysis of research. *The Educational Researcher, 10*, 3–8.

Godfried, M. (1999). The Pursuit of Consensus in Psychotherapy Research and Practice. *Clinical Psychology: Science and Practice, 6*(4), 462–466.

Gomory, T. (1999). Programs of Assertive Community Treatment (PACT): A critical review. *Ethical Human Sciences and Services, 1*(2), 1–17.

Hayes, R. L., & McGrath, J. J. (2002). Cognitive rehabilitation for people with schizophrenia and related conditions (Cochrane Review). *The Cochrane Library, Issue 3*. Retrieved July 29, 2002, from http://www.update-software.com/abstracts/ab000968.htm.

Healy, D. (1997). *The anti-depressant era.* Cambridge, MA: Harvard University Press.

Henderson, C., & Laugharne, R. (2002). Patient held clinical information for people with psychotic illnesses (Cochrane Review). *The Cochrane Library, Issue 3*. Retrieved July 29, 2002, from http://www.update-software.com/abstracts/ab001711.htm.

Hermann, R. C. (under review 2002). Advancing quality measurement in health care: A need for leadership amid a new federalism. Manuscript submitted for publication.

Hunt, M. (1997). *How science takes stock: The story of meta-analysis.* New York: Russell Sage Foundation.

Johnstone, P., & Zolese, G. (2002). Length of hospitalization for people with severe mental illness (Cochrane Review). *The Cochrane Library, Issue 3*. Retrieved July 29, 2002, from http://www.update-software.com/abstracts/ab000384.htm.

Joy, C. B., Adams, C. E., & Rice, K. (2003). Crisis intervention for people with severe mental illnesses (Cochrane Review). *The Cochrane Library, Issue 1*. Retrieved February 7, 2003, from http://www.update-software.com/abstracts/ab001087.htm.

Kitcher, P. (2001). *Science, truth, and democracy.* New York: Oxford University Press.

Leff, H. S., & Hollen, V. (2002, Spring). Talking about translating research into practice. *Outlook. Outlook is the quarterly newsletter of the Evaluation Center at HSRI.* Cambridge, MA: The Evaluation Center@HSRI and National Association of State Mental Health Program Directors Research Institute, Inc.

Leff, H. S., & Mulkern, V. (2002). Lessons learned about science and participation from multisite evaluations. *New Directions for Evaluation, 94*, 89–100.

Lehman, A., & Steinwachs, D. (2001). *Measuring conformance to treatment guidelines: The example of the schizophrenia PORT.* Cambridge, MA: The Evaluation Center@HSRI.

Ley, A., Jeffery, D. P., McLaren, S., & Siegfried, N. (2002). Treatment programmes for people with both severe mental illness and substance misuse (Cochrane Review). *The Cochrane Library, Issue 3*. Retrieved July 29, 2002, from http://www.update-software.com/abstracts/ab001088.htm.

Liberman, R. P., Kopelowicz, A., & Young, A. S. (1994). Biobehavioral treatment and rehabilitation of schizophrenia. *Behavior Therapy, 25*, 89–107.

Malmberg, L., & Fenton, M. (2002). Individual psychodynamic psychotherapy and psychoanalysis for schizophrenia and severe mental illness (Cochrane Review). *The Cochrane Library, Issue 3*. Retrieved July 29, 2002, from http://www.update-software.com/abstracts/ab001360.htm.

Marshall, M., Crowther, R., Almaraz-Serrano, A. M., & Tyrer, P. (2002). Day hospital versus out-patient care for psychiatric disorders (Cochrane Review). *The Cochrane Library, Issue 3*. Abstract retrieved July 29, 2002, from http://www.update-software.com/abstracts/ab003240.htm.

Marshall, M., Gray, A., Lockwood, A., & Green, R. (2002). Case management for people with severe mental disorders (Cochrane Review). *The Cochrane Library, Issue 3*. Retrieved July 29, 2002, from http://www.update-software.com/abstracts/ab000050.htm.

Marshall, M., & Lockwood, A. (2002). Assertive community treatment for people with severe mental disorders (Cochrane Review). *The Cochrane Library, Issue 3*. Retrieved July 29,2002, from http://www.update-software.com/abstracts/ab001089.htm.

Marzilli, A. (2002, Winter). Controversy surrounds evidence-based practices. *National Mental Health Consumers' Self-Help Clearinghouse, 7*(1), 5.

McMonagle, T., & Sultana, A. (2002). Token economy for schizophrenia (Cochrane Review). *The Cochrane Library, Issue 3*. Retrieved July 29, 2002, from http://www.update-software.com/abstracts/ab001473.htm.

Mueser, K. T., & Berenbaum, H. (1990). Psychodynamic treatment of schizophrenia: Is there a future? *Psychological Medicine, 20*, 253–262.

Millenson, M. L. (1997). *Demanding medical excellence.* Chicago: The University of Chicago Press.

Nathan, P. E., Gorman, J. M., & Salkind, N. J. (Eds.). (1999). *Treating mental disorders: A guide to what works.* New York: Oxford University Press.

Neimeyer, R. A., & Raskin, J. D. (2001). Varieties of constructivism in psychotherapy. In K. S. Dobson (Ed.), *Handbook of cognitive-behavioral therapies* (pp 393–340). New York: The Guilford Press.

Nicol, M. M., Robertson, L., & Connaughton, J. A. (2002). Life skills programmes for chronic mental illness (Cochrane Review). *The Cochrane Library, Issue 3*. Retrieved July 29, 2002, from http://www.update-software.com/abstracts/ab000381.htm.

Pekkala, E., & Merinder L. (2002). Psychoeducation for schizophrenia (Cochrane Review). *The Cochrane Library, Issue 3*. Retrieved July 29, 2002, from http://www.update-software.com/abstracts/ab002831.htm.

Pharoah, F. M., Mari, J. J., & Streiner, D. (2002). Family intervention for schizophrenia (Cochrane Review). *The Cochrane Library, Issue 3*. Retrieved July 29, 2002, from http://www.update-software.com/abstracts/ab000088.htm.

Pikoff, H. B. (1996). *Treatment effectiveness handbook: A reference guide to the key research reviews in mental health and substance abuse.* Buffalo, NY: Data For Decisions.

Reda, S., & Makhoul, S. (2002). Prompts to encourage appointment attendance for people with serious mental illness (Cochrane Review). *The Cochrane Library, Issue 3*. Retrieved July 29, 2002, from http://www.update-software.com/abstracts/ab002085.htm.

Roth, A., & Fonagy, P. (1996). *What works for whom? A critical review of psychotherapy research.* New York: The Guilford Press.

Rotter, J. (1971). On the evaluation of methods of intervening in other people's lives. *Clinical Psychologist,* Spring, 1–2.

Sailas, E., & Fenton, M. (2002). Seclusion and restraint for people with serious mental illnesses (Cochrane Review). *The Cochrane Library, Issue 3*. Retrieved July 29, 2002, from http://www.update-software.com/abstracts/ab001163.htm.

Strachen, A. M. (1986). Family intervention for the rehabilitation of schizophrenia: Toward protection and coping. *Schizophrenia Bulletin, 12*, 678–698.

Torrey, W. C., Drake, R. E., Dixon, L., Burns, B. J., Flynn, L., Rush, A. J., et al. (2001). Implementing evidence-based practices for persons with severe mental illnesses. *Psychiatric Services, 52*(1), 45–50.

Tyrer, P., Coid, J., Simmonds, S., Joseph, P., & Marriott, S. (2002). Community mental health teams (CMHTs) for people with severe mental illnesses and disordered personality (Cochrane Review). *The Cochrane Library, Issue 3*. Retrieved July 29, 2002, from http://www.update-software.com/abstracts/ab000270.htm.

U.S. Department of Health and Human Services. (1999). *Mental Health: A Report of the Surgeon General-Executive Summary*. Rockville, MD: U.S. Department of Health and Human Services, Substance Abuse and Mental Health Services Administration, Center for Mental Health Services, National Institutes of Health, National Institute of Mental Health.

West, S., King, V., Carey, T. S., Lohr, K. N., McKoy, N., Sutton, S. F., et al. (2002). Systems to rate the strength of scientific evidence. (Evidence Report/Technology Assessment No. 47; AHRQ Publication No. 02-E016). Rockville, MD: Agency for Healthcare Research and Quality.

Wolff, N. (2000). Using randomized controlled trials to evaluate socially complex services: Problems, challenges and recommendations. *The Journal of Mental Health Policy and Economics, 3*, 97–109.

Appendix
Links for Organizations, Centers, and Groups Mentioned
(in order of appearance)

The Cochrane Collaboration
http://www.cochrane.org/

The Campbell Collaboration
http://www.campbellcollaboration.org/

The National Guideline Clearinghouse (NGC)
http://www.guideline.gov/body_home.asp

American Psychological Association (APA),
Society of Clinical Psychology
"A Guide to Beneficial Psychotherapy"
http://www.apa.org/divisions/div12/rev_est/index.shtml

The National Institute of Mental Health (NIMH),
Division of Services and Intervention Research,
Services Research and Clinical Epidemiology
Branch
http://www.nimh.nih.gov/dsir/index.cfm

The Evaluation Center@HSRI
http://www.tecathsri.org/

National Association of State Mental Health
Program Directors (NASMHPD) Research Institute,
Inc. (NRI), Center for Evidence-Based Practices,
Performance Measurement, and Quality
Improvement
http://nri.rdmc.org/RationaleEBPCenterReview.pdf

The Addiction Technology Transfer Center
http://www.nattc.org/

National Registry of Effective Programs (NREP)
http://preventionpathways.samhsa.gov/nrepp/default.htm

The Agency for Healthcare Research and Quality
(AHRQ), Evidence-based Practice Centers (EPCs)
http://www.ahrq.gov/clinic/epc/