Persons, J. B. (2001). Conducting effectiveness studies in the context of evidence-based clinical practice. *Clinical Psychology: Science and Practice, 8*, 168-172.

Conducting Effectiveness Studies in the Context of Evidence-based Clinical Practice

Jacqueline B. Persons

San Francisco Bay Area Center for Cognitive Therapy

and

University of California, Berkeley

Abstract

It is proposed that effectiveness studies might be conducted in an infrastructure consisting of groups of evidence-based practitioners. Clinicians utilizing evidence-based approaches to clinical work have well-developed systems for collecting outcome data that are fully integrated into routine clinical practice, a mode of work that is ideally suited to effectiveness studies. A formal, NIH-funded structure might be developed for establishing clinical practice infrastructures that might host or conduct effectiveness studies. These clinical practice infrastructures might also speed the dissemination of evidence-based methods of clinical practice. Finally, clinical practice infrastructures would enhance clinicians' ability to make contributions to research. In particular, evidence-based practitioners utilize empirically-based methods of hypothesis-testing and session by session outcome monitoring that are not well-represented in many of the protocols that are evaluated in randomized controlled trials.

Borkovec, Echemendia, Ragusea, and Ruiz (this issue) describe an innovative and important effectiveness study that is the culmination of years of careful work. Their efforts move the field closer to the goal of conducting effectiveness studies that have both internal and external validity. Borkovec and his colleagues describe the infrastructure they established as Phase I of a research program leading to controlled effectiveness studies in Phase II. Their infrastructure was designed to maximize external validity of effectiveness studies, by collecting data from clinicians working in routine clinical practice (where they make decisions in their usual ways, not following structured protocols) and by studying a heterogeneous unselected sample of patients (not the highly selected patients usually treated in randomized controlled trials). Borkovec and colleagues propose a thoughtful and efficient approach to Phase II effectiveness research that compares, in randomized controlled trials, treatment conditions that are carefully chosen so that data yield information that is both clinically useful and theoretically informative.

The Borkovec et al. study raises many interesting and important issues. I focus here on the infrastructure needed to support effectiveness studies. After a brief account of the infrastructure Borkovec and his colleagues developed, I describe another type of clinical setting (the evidence-based clinical practice) that also has the potential to serve as a infrastructure for the collection of effectiveness data, and I briefly discuss its strengths and weaknesses.

The Pennsylvania Practice Research Network

Borkovec and his colleagues recruited clinicians in the state of Pennsylvania who agreed to participate in the effectiveness study. Two hundred five clinicians expressed interest and, remarkably, eighty attended a day-long meeting at Pennsylvania State University to plan the study. Participants agreed to ask their patients to complete assessment packages at pre-treatment, mid-treatment (7 weeks), post-treatment, and 6-month followup intervals. After collecting data from clients, clinicians mailed the data to a central data management site, which analyzed them and returned them, generally within seven days, to the clinicians (and researchers). Measures were chosen by a committee which searched the assessment literature and selected the COMPASS system of outcome monitoring developed by Kenneth Howard and his collaborators. The principle guiding the design of the Borkovec et al. infrastructure was to create a data collection mechanism that would intrude minimally on the clinician's workload and routine mode of practice and would make minimal demands on clients, in order to maximize the external validity of studies conducted in this setting.

An Alternative Infrastructure for Effectiveness Studies: Evidence-based Clinical Practice

I propose an alternative infrastructure for the conduct of effectiveness studies: evidence-based practice (EBP). Evidence-based practice is clinical work in which therapists rely on the methods and findings of science. In evidence-based practice, the clinician utilizes a hypothesis-testing mode of clinical work in which each treatment is viewed as an experiment with an N = 1. The evidence-based practitioner (the scientist-practitioner) begins by conducting an initial assessment to obtain a comprehensive problem list, diagnoses, and a case formulation. The clinician works with the client to make a mutually-agreed upon set of treatment goals and proposes a treatment plan based on the clinician's assessment findings and review of the outcome literature, particularly the randomized controlled trials (RCTs). Treatment outcome is monitored at every therapy session using objective measures. If outcome is poor, the clinician returns to the assessment phase, with a view to attempting a reformulation of the case that suggests some new treatment interventions that might produce a better outcome. The origins of evidence-based clinical practice lie in behavioral assessment (cf. Hayes, Barlow, & Nelson, 1999 Haynes & O'Brien, 2000), evidence-based medicine (cf. Sackett, Haynes, Guyatt, & Tugwell, 1991) and practice evaluation (cf. Bloom, Fischer, & Orme, 1995).

Two components of evidence-based practice are particularly relevant to the present discussion: outcome monitoring, and reliance on the randomized controlled trials. I discuss each in some detail by describing the approach to these components of EBP that my colleagues and I use at the San Francisco Bay Area Center for Cognitive Therapy.

Outcome monitoring. At our Center, all clients who schedule a consultation session are asked to complete, prior to the session, a packet of assessment materials consisting of the SCL-90 (Derogatis, Lipman, & Covi, 1973), the Beck Depression Inventory (BDI; Beck, Steer, & Garbin, 1988), the Burns Anxiety Inventory (BAI; Burns & Eidelson, 1998), a modification of the

CAGE substance abuse screening measure (Mayfield, McLeod, & Hall, 1974), and a brief demographic questionnaire. We use the SCL-90 because it is an extensively-used and well-validated measure that assesses a wide variety of symptoms and problems. We assess depressive and anxiety symptoms because most of our patients seek treatment for those problems. We use the BDI because it is extensively used and normed and because it is widely used in the RCTs. We use the BAI because we find it to be a clinically useful measure that is sensitive to changes during treatment; David Burns, its originator, recently published some normative and validation data (Burns & Eidelson, 1998). The CAGE is a brief, clinically useful and well-validated measure of substance abuse; we have modified it by adding items asking patients to report the types and amounts of alcohol and illicit drugs they use.

When we first began to systematize our outcome assessment strategies, we established a system similar to the one established by Borkovec et al. Modelling our data collection system on those utilized in the randomized controlled trials (RCTs), we asked clinicians to administer a standard package of assessment materials to their patients at intake, 12 weeks later, and at the end of treatment. After experimenting with this method for several months, we abandoned it because attrition rates were high and—most important—it did not yield clinically useful data. Attrition was high because clinicians frequently forgot to administer the measures at the correct point in time. In addition, it proved to be difficult to obtain post-treatment assessments. We found that once a patient left our office without an appointment to return at a later date, the probability that he or she would return a packet of measures was low. Moreover, a sizeable minority of patients ended treatment by simply drifting away, without stating that they planned to end treatment, and this made collecting post-treatment assessments problematic.

Most important, the system was not clinically useful, for several reasons. Clinicians did not find it useful to readminister, twelve weeks after treatment had begun, a set of measures that frequently included scales assessing symptoms that were not problematic for a particular patient. Moreover, even if the patient had had clinically significant symptoms on one or more measures at the beginning of treatment, an assessment at the 12 week point was of limited use; instead, clinicians needed a system that would allow them to monitor progress each session (Kazdin, 1993).

In response to these observations, we developed an outcome monitoring system that is both clinically more useful and simpler to implement. We continue to utilize the initial packet of assessment measures. When the initial assessment is complete and treatment goals are set, the therapist and patient select which (if any) of the measures in the initial packet would be useful in assessing progress toward the patient's treatment goals at each therapy session. If none of the measures in the original packet address the patient's treatment goals (e.g., the patient seeks treatment for trichotillomania), the clinician obtains another measure from a constantly-expanding collection of measures we maintain, by consulting Corcoran and Fischer (1994), or by developing an idiographic measure.

Large numbers of our patients set treatment goals of reducing depressive, anxiety, or obsessive-compulsive symptoms. To monitor progress toward these goals, clinicians collect weekly scores on the Beck Depression Inventory, the Burns Anxiety Inventory, and/or the Yale-Brown Obsessive-Compulsive Scale (Y-BOCS; Goodman et al., 1989). We keep clipboards holding these measures in the waiting room, and patients are asked to come five minutes prior to their session to complete the measures relevant to their treatment in the waiting room. At the beginning of the session, the therapist scores the measures and enters the scores on a plot that charts the patient's progress. Patient and therapist review the plot and if there is a prominent deterioration or improvement in the score, this is noted and may become an agenda item for the therapy session (cf. Davidson, Persons, & Tompkins, 2000). Thus, progress monitoring occurs every session and the results of the assessment feed immediately and directly into the treatment.

Reliance on the randomized controlled trials (RCTs). We rely in several ways on the results of the randomized controlled trials (RCTs). Our initial case formulations are based on the nomothetic formulations that underpin therapies that have been shown effective in RCTs (Persons, Davidson, & Tompkins, 2000). We base treatment plans on RCT-validated interventions. We frequently rely on outcome measures used in the RCTs, so that we can use the results of the RCTs as benchmarks for our clinical work (cf. Wade, Treat, & Stuart, 1998).

However, as noted above, we deviate from the outcome monitoring systems used in the RCTs, and we also rely more on individualized case formulations than do most RCT-validated protocols.

The Process of Conducting Effectiveness Studies in an EBP Infrastructure

A formal structure might be established for conducting effectiveness studies in the context of an EBP. An NIH grant mechanism might be established to encourage practitioners to establish group practices utilizing evidence-based methods that could provide an infrastructure supporting Phase II effectiveness studies. Evidence-based practitioners, with the help of NIH funding, might hire a part-time data manager who would work on site to maintain high compliance rates and quality control of data collected routinely in the practice. Ethical issues would require careful attention and the EBP might need to establish a review board to evaluate its procedures. Development of a system to code charts by number rather than name might facilitate the research. Incentives (for patients, clinicians, or both) may be needed for significant data collection needs of the researchers that are not part of routine clinical work, particularly when patients are asked to provide data after treatment has ended.

Investigators wishing to conduct effectiveness studies could be invited to submit protocols to evidence-based practices, which could elect to participate in studies that address questions of interest to the practitioners. Practitioners' commitment to the research is particularly important if they will be asked to modify their work in some way (e.g., add an additional assessment measure or two or modify their intervention strategies, as in the example proposed by Borkovec et al., this issue). Some financial incentives would be needed to induce practitioners to revise their routine practice.

Strengths and Weaknesses of Evidence-based Practice as an Infrastructure for Effectiveness Research

The EBP infrastructure for effectiveness research has several strengths and (at least) one weakness. An important strength is that because, in the EBP infrastructure, the collection of outcome data is an integral part of the treatment process, little in the way of incentives is needed to maintain high rates of data collection. Therapists will not forget to administer outcome measures at designated time points, as the clinically relevant measures will be routinely administered at every therapy session. Because patients will provide data at every session, even if they leave treatment abruptly or simply drift away, data from the final treatment session will always be available.

Another advantage is that evidence-based practitioners are familiar with and utilize many RCT-validated protocols. This is likely to be important when investigators wish to conduct Phase II effectiveness studies, as they are likely to wish to evaluate the effectiveness in clinical settings of RCT-validated protocols or modifications of those protocols.

In addition to providing an infrastructure that supports effectiveness research, the EBP may promote the dissemination of evidence-based modes of work into routine clinical practice. Clinicians who participate in evidence-based practices will change their way of working every day, not just with research subjects. And clinicians who adopt an evidence-based mode of practice are likely to communicate that way of thinking and practicing to junior clinicians they train or supervise in other settings. Thus, the EBP infrastructure has a generative quality. In addition, because they are collecting data as part of their routine clinical practice, the evidence-based practitioners can begin to make their own contributions to the professional and empirical literature. As a result of our own routine outcome monitoring procedures, we have contributed two open trials examining the effectiveness of cognitive therapy for depression in private practice samples (Persons, Burns, & Perloff, 1988; Persons, Bostrom, & Bertagnolli, 1999) as well as other studies testing theories of psychopathology and process of change. The EBP infrastructure also does not have the disadvantage of reinforcing clinicians' maladaptive beliefs that outcome monitoring is extraneous to routine clinical practice.

A really exciting feature of the EBP infrastructure for effectiveness research is that it increases the influence that clinicians can have on researchers. It is ironic that the outcome monitoring systems utilized in the RCTs, as described above, are not optimally clinically useful and do not facilitate an empirical approach to clinical work. In fact, in important ways the outcome monitoring systems typically utilized in RCTs are less evidence-guided than those utilized by evidence-based practitioners. In many RCTs, outcome data are collected too infrequently to have a significant effect on the treatment process. Also, the RCT-protocol-guided

clinician is guided by a protocol, not by a hypothesis (the case formulation). Assessment and treatment are too frequently independent of one another (Persons, 1991). In contrast, in the EBP mode of treatment, outcome data inform the treatment process session by session, and treatment is guided by a hypothesis (the case formulation). Increased communication of clinicians and researchers might lead to the dissemination of some modes of evidence-based practice into research protocols.

For example, some of the measures we have imported from randomized trials into our clinical practice have proved to be cumbersome. Scales with reverse-scored items are an example. Reverse-scoring makes it difficult for the clinician to score the measure in the therapy session, and unless this is done, it is all too likely that the measure will be placed in the clinical record and forgotten. Another obstacle we frequently encounter, and that Borkovec and his colleagues encountered, is difficulty obtaining the most desirable outcome measures. My colleagues and I are are eager to utilize the measures used in the RCTs so that we can use the RCT findings as benchmarks for our clinical work. However, many measures utilized in RCTs are copyright-protected, which makes them expensive for routine use. Solutions to this dilemma that would meet the needs of the investigators who developed the scales while also addressing the need for wide dissemination of evidence-based modes of assessment and treatment would make a major contribution to our field.

A weakness of the EBP infrastructure is its lack of generalizability. Because most clinicians are not operating in an evidence-based way, results of effectiveness studies collected in the EBP infrastructure cannot be assumed to generalize to other types of clinical practice settings. However, an increasing number of clinical settings (cf. the Beck Institute for Cognitive Therapy and Research, the Group Health Cooperative of Puget Sound, the Center for Cognitive Therapy at the University of Pennsylvania, some departments of the Duke University Medical Center, and no doubt many others not mentioned here) are utilizing evidence-based modes of practice. And, as I indicated above, the utilization of evidence-based clinical practice as an infrastructure for effectiveness studies may itself contribute to solving this problem.

Author Note

Several of my colleagues at the San Francisco Bay Area Center for Cognitive Therapy made important contributions to the design and implementation of our outcome monitoring system. Simone K. Madan was the most important of these, and others include Alexandra Matthews and Michelle Hatzis. My colleagues Joan Davidson, Michael A. Tompkins, Chad W. LeJeune and Kristen L. Valus share in the daily work of keeping our system up to date and functioning smoothly.

References

Beck, A. T., Steer, R. A. & Garbin, M. G. (1988). Psychometric properties of the Beck Depression Inventory: Twenty-five years of evaluation. Clinical Psychology Review, 8, 77-100.

Bloom, M., Fischer, J., & Orme, J. G. (1995). Evaluating practice: Guidelines for the accountable professional. Boston: Allyn and Bacon.

Burns, D.D. & Eidelson, R.J. (1998). Why are depression and anxiety correlated? A test of the tripartite model. Journal of Consulting and Clinical Psychology, 66, 461-473.

Corcoran, K., & Fischer, J. (1994). Measures for Clinical Practice: A Sourcebook. Vols. 1 (Couples, Families and Children) and 2 (Adults). New York: Free Press.

Davidson, J., Persons, J. B., & Tompkins, M.A. (2000). Cognitive-behavior therapy for depression: Structure of the Therapy Session [Video]. Washington, DC: American Psychological Association.

Derogatis, L. R., Lipman, R. S., & Covi, L. (1973). SCL-90: An outpatient psychiatric rating scale--preliminary report. Psychopharmacological Bulletin, 9, 13-28.

Goodman, W. K., Price, L. H., Rasmussen, S. A., Mazure, C., Fleischmann, R. L., Hill, C. L., Heninger, G. R., & Charney, D. S. (1989). The Yale-Brown Obsessive Compulsive Scale I: Development, use, and reliability. Archives of General Psychiatry, 46, 1006-1011.

Hayes, S. C., Barlow, D. H., & Nelson-Gray, R. O. (1999). The scientist practitioner: Research and accountability in the age of managed care. Boston: Allyn and Bacon.

Haynes, S. N., & O'Brien, W. H. (2000). Principles and practice of behavioral assessment. New York: Kluwer Academic/Plenum Publishers.

Kazdin, A. E. (1993). Evaluation in clinical practice: Clinically sensitive and systematic methods of treatment delivery. Behavior Therapy, 24, 11-45.

Mayfield, D., McLeod, G., & Hall, P. (1974). The CAGE questionnaire: Validation of a new alcoholism screening instrument. American Journal of Psychiatry, 131, 1121-1123.

Persons, J. B. (1991). Psychotherapy outcome studies do not accurately represent current models of psychotherapy: a proposed remedy. American Psychologist, 46, 99-106.

Persons, J. B., Bostrom, A., & Bertagnolli, A. (1999). Results of randomized controlled trials of cognitive therapy for depression generalize to private practice. Cognitive Therapy and Research, 23, 535-548.

Persons, J. B., Davidson, J., & Tompkins, M. A. (2000). Essential components of cognitive-behavior therapy for depression. Washington, DC: American Psychological Association.

Sackett, D. L., Haynes, R. B., Guyatt, G. H., & Tugwell, P. (1991). Clinical epidemiology: A basic science for clinical medicine. Boston: Little, Brown and Company.

Wade, W. A., Treat, T. A., & Stuart, G. L. (1998). Transporting an empirically supported treatment for panic disorder to a service clinic setting: A benchmarking strategy. Journal of Consulting and Clinical Psychology, 1998, 66, 231-239.