Several issues concerning stepped care are discussed: the constraints of using Diagnostic and Statistical Manual of Mental Disorders diagnoses in randomized clinical trials (RCTs), the importance of basic and process research, the unintended negative effects of exaggerated claims of effectiveness and efficiency, the limits of RCTs in evaluating improvement and deterioration, the self-correcting nature of stepped care, the link between stepped care and empirically supported treatments, clinical judgment in clinical work, the concept of the least restrictive alternative, the costs of using low-intensity but ineffective psychosocial treatments, and the costs of both ineffective and effective psychotropic drug therapy. An analysis of stepped care can lead to an appreciation that the dialectic operating between science and practice affords an opportunity to synthesize the seemingly irreconcilable standards and needs of researchers and clinicians.

Doing more with less was the mantra of the 1990s. Business cut back on the number of (lower and middle-level) employees. The U.S. military was downsized. Universities reduced the hiring of tenure-track faculty and began to rely increasingly on outside contract instructors (whose compensation is often less than that of a graduate assistant for the same work). And germane to the topic of stepped care, managed care has markedly reduced the number of sessions covered for psychotherapy services as well as the compensation that practitioners can expect for a given amount of professional effort (Kent & Hersen, 2000).

One argument is that, like physicians and other medical practitioners, we mental health professionals had it coming. For too long, some contend, we have been able to continue raising our fees and extending with little restraint the numbers of sessions billed to our patients or their insurance companies, and we have sometimes done so with interventions and assessments lacking in empirical support. On the other hand, as suggested by several of the articles to be discussed here, cutting back in this way and trying to conform to a managed care formula may result not only in vitiating our ability to meet our responsibilities to our patients but may also end up costing more in terms of actual monies spent by people—and their insurance carriers—and lead to limited positive impact on the personal, social, and occupational functioning of our patients.

Sometimes one can do more with less—"more" at least in the sense that we can serve more people because of less time devoted to each one. Sometimes we do less with more. Sometimes we end up doing much less with less. And sometimes having less time to work with patients ends up costing more in the long run. These are the conundrums and challenges posed by stepped care, which refers to the practice of beginning one's therapeutic efforts with the least expensive and least intrusive intervention possible and moving on to more expensive and/or more intrusive interventions only if deemed necessary in order to achieve a desired therapeutic goal.

I have extracted and constructed a number of themes from the four articles in this special series on stepped care in this journal (Newman, 2000; Otto, Pollack, & Maki, 2000; Sobell & Sobell, 2000; Wilson, Vitousek, & Loeb, 2000) in an effort to comment in ways that may lead to an elaboration of the worthwhile questions raised by these authors.

**The Constraints of Randomized Clinical Trials (RCTs) That Define Treatment Groups in Terms of Diagnostic and Statistical Manual of Mental Disorders (DSM) Diagnoses**

As argued elsewhere in a critique of RCTs (see Goldfried & Davison, 1994, p. 292), assigning RCT participants to treatment groups on the basis of a DSM diagnosis discourages, if not makes impossible, the kind of idiographic functional analysis that is the lifeblood of responsible clinical practice as well as of sound experimental research. Newman (2000) discusses this well and reviews some research that she is involved in that, for example, applies functional analysis and tailored intervention to treatment failures from RCTs—especially failures in experimental conditions that, on average, effect statistically and even clinically significant improvements. Another way to appreciate the importance of this research strategy is to look at it as studying the variance that is unexplained in RCTs. I agree fully with Newman that it is the DSM focus that has created the problem, and she proposes constructive strategies for complementing the now-familiar RCT research paradigm with studies that permit functional analysis of individual cases and flexible utilization of therapy techniques rather than adherence to a treatment manual applied to heterogeneous research participants who are regarded as homogeneous per a DSM diagnosis.

In this context, however, it is important to note what Linehan, Moras, Pilkonis, and Rehm (1998) referred to as a second generation of treatment manuals, which emphasize strategies more than the strict application of specific techniques. Linehan's (1993) widely used dialectical behavior therapy for
people with borderline personality disorder, "who present invariably with multiple, pressing problems. . . . She [Linehan] provide[s] hierarchies of problems, sequential strategies for treating them, and a mix of specific tactics within strategies, and it is unlikely that the treatment of any two patients will be identical, although both treatments will be recognizable as dialectical behavior therapy [italics added]" (Calhoun et al., 1998, p. 154). More explicit use of therapy manuals as strategic tools rather than formulaic prescriptions can reduce the danger of, as Goldfried and Wolfe (1996) cautioned, manuals being used "as more of a straightjacket than a set of guidelines" (p. 1014).

One of the questions raised by practitioners is whether RCTs are relevant to the challenges they face (see, e.g., the debate in Persons & Silberschatz, 1998; see also Seligman’s, 1995, report of the Consumer Reports survey and the critique by Jacobson & Christensen, 1996). In my view, those who dismiss the relevance of RCTs to the everyday work of clinicians are no more correct than are researchers who focus all their evidentiary efforts on group studies without taking as seriously the "error variance" in the outcomes of all RCTs (more on variances within groups below). Going back to the data from an RCT and looking more idio graphically at the nonresponders, as Newman and her associates are doing, is a welcome enhancement to our efforts to understand therapeutic change. It is also consistent with idiosyncratic applications of nomothetic principles as articulated by a number of authors (e.g., Eifert, Evans, & McKendrick, 1990; Lazarus & Davison, 1971; Persons, 1989).

Sobell and Sobell (2000) also deal with this issue in their discussion of what has come to be called attitude treatment interaction (ATI), the matching of particular kinds of patients with putatively appropriate treatments. In discussing Project MATCH (Allen et al., 1999), they remind us of the failure to show much in the way of ATI—in the MATCH case, the matching variables included motivation to abstain and psychiatric comorbidity—and suggest that, in general, the problems finding ATIs might lie at least in part with the use of the DSM to categorize participants and then randomly assign them to treatment conditions. The question is simply this: What is being matched? Reliance on the DSM to categorize research participants for randomized assignment to experimental conditions may underlie the difficulties of RCTs in uncovering ATIs not to mention more straightforward differences in efficacy of various treatment modalities.

Basic Research in Applied Psychology and the Importance of Process Research

As Wilson et al. (2000) note, "distilling the essential elements of [an expensive, expert] treatment package requires that we understand the mechanisms of change of that treatment" (p. 566). Process research should shape our thinking about stepped care as it does our views of psychotherapy generally. Wilson et al. highlight the importance of process research when they report that cognitive-behavioral therapy (CBT) for bulimia achieves most of its gains in the first few sessions (cf. Tang & DeRubeis, 1999, for similar findings in cognitive therapy for depression). If most improvement takes place in the earliest sessions of, say, a 20-session treatment program, then (a) the whole treatment package has elements that are unnecessary and (b) what happens in those first few sessions needs to be examined very closely with an eye to extracting those procedures and understanding those processes that are operating. It has been an article of faith in psychotherapy research that there is great need for a focus on mechanisms of change, for greater knowledge of these not only enhances our science but also can and should lead to more effective and more efficient, cost-effective interventions.

But the issue goes further. Several years ago I commented on the role of basic research in clinical psychology (Davison, 1994) and had occasion to develop the argument further during a conference sponsored by the National Institute on Drug Abuse (NIDA) concerned with untapped opportunities to use basic research in developing clinical procedures de novo (Davison, 1997). Simply put, searching for change mechanisms in existing effective techniques is to work after the fact, and although such process research is very important (as just stated), working in the other direction may be even better. The take-home message from the NIDA conference was that the rhetoric of the scientist-professional model and especially of behavior therapy—moving from experimentally established principles of change to novel and effective clinical applications—is an inadequately explored strategy for developing new therapeutic procedures that, from the outset, will have known mechanisms of change (because such research begins with principles of change). In her article, Newman (2000) articulates the promise of following this strategy.

Exaggerated Claims of Effectiveness and Efficiency

When I was in graduate school in the early 1960s, claims of rapid improvement from the emergent field of behavior therapy were widespread. The then-manageable behavior therapy literature, most of it composed of case studies and theoretical treatises, reported that unrealistic fears could be completely eliminated in under 6 sessions; people who were inhibited all their lives from expressing their needs and wants were enabled in just a handful of sessions to become more assertive; sexually dysfunctional couples were, in fewer than 10 sessions, rendered bunnylike in their zest for procreative activities; and so forth. It was a brave new world. I'm glad I entered that world, but as happened with others of my generation, clinical realities later required a more sober and modest appraisal of what we could do and how quickly we could do it.

Managed care companies love claims of quick and thoroughgoing cures. And why not? The only problem is that, once the initial hypomania from a new kind of therapy diminishes—and it always does—the profession and the people it serves are done no service, for these high expectations cannot be met and limits imposed by third-party payers on the basis of the earlier, unrealistic claims of effectiveness and efficacy can compromise the access that most people have to effective mental health services.

Assessing Improvement in Research Participants

Newman (2000) reminds us of the importance (probably more salient to practicing clinicians than to researchers) of assessing improvement along parameters that go beyond amelioration in DSM-delineated areas. For example, a participant diagnosed as having a major depressive disorder may benefit from an intervention in undetected ways if the researcher is focused on symptomatic improvement. Consider the following scenario. A depressed person's mood is improved significantly by a particular treatment
under study. All well and good. But what about his or her occupational functioning? Is the formerly depressed college professor getting better course evaluations from her students? Are more students coming to her posted office hours to discuss the course? Is she spending more time on campus as contrasted with retreating to her home as soon as classes are over? What is the quality of her committee work? Although improvement in mood is of obvious importance to depressed patients and those close to them, there are other aspects of a person’s life that should be evaluated.1

On Treatment Failures in the Experimental Group and Treatment Successes in the Control Group

The close-up, idiographic approach that all the authors advocate as a complement to RCTs is consistent with an important and seldom-discussed limitation of RCTs that Lazarus and I have noted (Davison & Lazarus, 1995). Consider one of the simplest type of therapy study, involving an experimental group and some sort of attention-placebo control group. We have become accustomed over the years to expect some degree of improvement in placebo groups, sometimes even to the degree that within-condition changes are significant (recall, e.g., Paul’s, 1966, classic study on desensitization). The researcher, of course, hopes that any such improvement will, on average, exceed by positive changes in the experimental condition.

But consider the following situation, which is probably not infrequent. Participant A in the experimental group improves significantly, and Participant B in the placebo control group improves to the same degree. Can we with confidence attribute the improvement of Participant A to a particular feature of the experimental condition? Another way to put the question is as follows: Given that Participant A improved in the experimental condition, can we say he would not have improved to the same degree if he had been randomly assigned to the control condition (for Participant B showed the same improvement, and it is common to find some degree of improvement even in placebo conditions)? Furthermore, because placebo elements are admittedly a part of the experimental condition—hence the inclusion of a placebo control group—can we say with confidence that Participant A’s improvement was not due to the placebo elements inherent in the experimental condition? I suggest that the answers to these questions is “no.”

Reports of comparative outcome research imply, if not assert outright, that improvements in experimental participants are due to something particular about that condition vis-à-vis a control group, even though there is always variance in change scores in both groups. But recall the caution issued many years ago by Bergin (1966, 1970) that there is usually some deterioration among some participants in experimental conditions, even when the group, on average, improves significantly with respect to pretreatment status and more than control conditions. How frequently do authors attribute this worsening to something special about the experimental condition? (See also Haaga & Stiles, 2000, for a discussion of the need for post hoc analyses of such variability in response to treatment conditions.)

The Self-Correcting Nature of Stepped Care

The Sobell and Sobell (2000) article nicely demonstrates that an inherent feature and advantage of stepped care is that it self-corrects; that is, it forces one to monitor constantly the effects of one’s interventions and to adjust subsequent strategies based on what has just happened. This sounds reasonable enough. It is, however, not the way that many practitioners have operated, at least until recently.

To give an example, when I was on my internship at a Veterans Affairs hospital 35 years ago, I served for half a year as ward psychologist. When a new patient came onto the ward, I, of course, would read his file, especially the just-conducted intake report. I was struck by the modal answer of the intake worker to the question regarding treatment recommendations. It was almost always something like “antipsychotics” and/or “psychotherapy.” And given the separation of the assessment–intake operation from what happened to the patient on the ward, it was unclear to me what guidance the intake report really provided to the treatment team. And most certainly the predictive validity or utility of the intake-based treatment recommendation was seldom if ever checked even informally. Not to put too fine an edge on it, but the people doing the intakes seldom followed up on what happened to the patients once a treatment disposition was made. And the contents of the intake report were of limited use (to put it charitably) in deciding how to treat the patient.

As has been noted for many years (e.g., Mischel, 1968), this is not a good model for mental health care—it is neither effective nor efficient. I happened to learn a different way of doing things from Walter Mischel, Albert Bandura, and Arnold Lazarus during my graduate studies. In their own ways they taught that assessment should be inextricably tied to intervention, that assessment is not synonymous with diagnosis, that assessment never stops, and that assessment should be an intrinsic part of intervention in that it provides continuing feedback to therapists about the progress of treatment and about the need to make adjustments as one proceeds.

Sobell and Sobell (2000, p. 574) provide a figure depicting the nature of stepped care. I believe that the figure should be a bit more complex by including within the treatment boxes an explicit acknowledgment that treatments themselves are highly complex affairs and that feedback-based adjustments are made on a session-by-session—indeed, on a moment-to-moment—basis. To take another example from my own clinical and research experience, when one conducts systematic desensitization (Wolpe, 1958), the patient’s progress up the anxiety hierarchy is tied to his or her mastery of anxiety at lower levels of aversiveness. In other words, the treatment itself is made up of many small procedural steps, each of which depends on what has happened at the previous step. I know that getting this molecular runs the risk of getting outside the boundary conditions of what is meant by stepped care, but I believe the example is apt.

The Relationship Between Stepped Care and Empirically Supported Treatments

In their discussion of Project MATCH and of the treatment of alcohol problems generally, Sobell and Sobell (2000) empha-
size that, whenever possible, the treatment used should enjoy some empirical support. But they indicate also that "stepped care... adds the guideline that decisions to change treatments should be performance based" (Sobell & Sobell, 2000, p. 578).

The implications of this comment are considerable and need to be drawn out. A given intervention may enjoy little if any empirical support (from RCTs), and yet, in a particular idiographic context, it may show a positive effect. Such an (interim) outcome can and should inform the clinician in his or her future decision making on behalf of the patient. For a particular patient at a particular time, a nonvalidated intervention may have a positive effect. This is a datum. It makes no sense to disregard it because that intervention lacks empirical support in RCTs. When one is operating with a stepped care cognitive schema, these individual, unique effects provide information on how far one has come and perhaps also how far one needs to go with the patient and how one might try to get there.

The Role of "Clinical Judgment" and "Clinical Experience"

Sobell and Sobell (2000) open up an important can of worms in the following statement:

Recommendations based on efficacy studies may need to be qualified by... factors not usually considered as part of controlled trials... clinical judgment and the provider's experience can be invaluable. The importance of clinical judgment and sensitivity to case management and idiosyncratic issues suggests that the application of a stepped care approach requires treatment decisions to be made by experienced clinicians and not by relatively untrained personnel or on the basis of questionnaires. (p. 573)

To which I say "Bravo" but also "Be careful out there." The caution lies at the heart of the nature of our field as a science and an art, and it goes way back to the very beginnings of our field (e.g., Meehl, 1954). Lazarus and I have recently referred to this as "the dialectics of science and practice" (Davison & Lazarus, 1995). That part of our argument of relevance here is that clinical psychology and the other helping professions (and sciences!) face no greater challenge and enjoy no greater opportunity than to deal with the necessary and potentially constructive dialectical tension between controlled findings and empirically supported principles, on the one hand, and the art of clinical practice, on the other hand. Under the best of circumstances, the behavior of clinicians is informed by empirically based knowledge, but specific interventions with individual clinical cases always require going beyond what the science tells us. Empirically based decision making is always informed by the particularities of an individual case, and clinical judgment—the art of clinical work—is required to implement a general principle.

Least Restrictive Alternative

It is useful to make explicit the connection between stepped care and the legal principle of least restrictive alternative. The latter concept is germane to issues of institutionalization. In general terms, mental health professionals are, under certain circumstances, legally required to provide that treatment which restricts the patient's liberty to the least possible degree while remaining workable and effective (In Re: Torpey, 1991). It is, for example, unconstitutional to confine a nondangerous mentally ill patient who is capable of surviving on his or her own or with the help of willing and responsible family or friends (Project Release v. Pre- vost, 1983). More generally, as pointed out by Sobell and Sobell (2000), treatment should entail as little intrusion as possible. The link to stepped care is that the most intrusive/intensive/expensive interventions should be implemented only when less intrusive ones have failed or, in light of the evidence, are unlikely to serve the patient's best interests (e.g., by being ineffective and thereby prolonging distress). What stepped care does is make explicit the professional and perhaps also the moral applicability of the least restrictive alternative principle to noninstitutional treatment settings.

The Costs of Ineffective Low-Intensity Interventions

Wilson et al. (2000) rightly caution that "failure to respond to an initial low-intensity level of care [could] discourage patients from seeking subsequent treatment or undermine their response to such treatment" (p. 564). Thus, there are risks associated with beginning with low-intensity treatments across the board. It is better—as these authors understand—to construct the steps according to a judgment of the minimally intensive/intrusive treatment that a given patient is likely to respond favorably to. This is not to say that the experimental literature always provides sure guidelines (for the most part, it does not), but the caveat is well made.

And it is related to an issue that is not acknowledged sufficiently, namely, the consequences of patients getting worse from ineffective treatments. I have discussed this in the context of the ethics of offering sexual conversion treatment to gay men and lesbians:

There may be one sense in which efficacy relates to the ethical issues. If an ineffective reorientation therapy is undertaken, the patient is going to be disappointed and likely therefore to feel even worse about his predicament—he has not only failed to achieve the reorientation goal but comes away continuing to believe that his homosexuality is bad and sick and perhaps feeling even worse about being gay. (Davi- son, 1991, p. 148)

Thus, not only do ineffective treatments not bring about desired change (by definition), they may also have a negative impact on the patient's self-esteem and reduce his or her motivation to continue trying to change.²

The Challenge of Briefer and Less Expensive Treatments

Wilson et al. (2000) comment that very few patients with bulimia avail themselves of the best empirically supported treatments, CBT and Weissman and Klerman's interpersonal therapy (Weissman, 1995). Given that these treatments require specialized

² The statement quoted here is from a paper that argues that therapists should not engage in sexual conversion programs with gay and lesbian patients, whether such efforts are effective or not. What can be done is separate from what ought to be done. But one place where an ethical issue does relate to an empirical one is designated in the aforementioned comment; that is, ineffective treatments raise ethical issues in and of themselves because they end up making patients feel worse about themselves.
training and given the shortage of such training, they conclude that "briefer and simpler" (Wilson et al., 2000, p. 564) treatments are called for. Their review of self-help strategies gives reason to be optimistic that, as research finds its way into such widely disseminable sources as self-help books, less expensive and more widely available treatments for bulimia may provide benefits to greater numbers of people than is now the case.

The problem, however, is that all manner of self-help books and tapes are sold aggressively in mass market outlets, and the value of such nonprofessional modes of intervention is widely (or wildly) variable (Rosen, 1987). Again, the consequences of a failed change effort are likely to be far more negative than merely no change; the person is likely to come away from such experiences with an impaired sense of self-efficacy vis-à-vis the problem he or she is seeking remediation of. Moreover, the time spent not improving from an ineffective treatment—whether it be inexpensive, unobtrusive, or both—is time that could have been devoted to a more appropriate intervention, albeit a more "stepped-up" one. Like any other intervention, a self-help manual will be no more effective than the utility and appropriateness of both the underlying principles and the power of the procedures to effect desired change.

The Costs of Ineffective and Effective Psychotropic Drug Therapy

Otto et al. (2000) point out that the dropout rates for drug treatment of panic disorder are considerably higher than those for CBT. Moreover, as is the case with most psychotropics, relapse is the rule when drugs are withdrawn—patients either get worse or have to remain on the drugs indefinitely for there to be any chance at all of maintaining whatever improvement the drugs have resulted in. If they do the latter, there are costs associated with long-term physical damage as well as personal, financial, and societal costs. If they stop taking the drug and relapse, their lives and those of others continue to be negatively affected by limitations in their effectiveness as social beings and by the cascading negative effects of their anxieties in the form of personal unhappiness and lower quality relationships with others. Thus, from the perspective of cost-effectiveness, even though pharmacotherapy may at first require fewer therapist contacts and appears (clearly to managed care companies) to be the least expensive treatment consistent with at least a modicum of effectiveness, people dropping out of it continue to have their lives interfered with by their anxieties, and as Newman (2000) points out, there are major costs associated with such problematic outcomes.

Furthermore, the study reported by Otto et al. (2000) at their specialty anxiety disorders clinic provides evidence dispelling the widely held assumption that drug treatment for psychological disorders is less expensive in dollar terms than CBT, whether CBT is administered individually or in groups. Quite the contrary, the per-patient costs are considerably lower for empirically supported psychosocial treatment, and the short- and long-term outcomes are at least as good as what people derive from various psychotropics. And other research reviewed by Otto et al., including the recently completed multicenter trial (Barlow, Gorman, Shea, & Woods, 1997, as cited in Otto et al., 2000), gives no support, regardless of cost, for combining CBT with drugs. Such a combination, of course, adds considerably to cost and intrusiveness and is therefore difficult to justify within the stepped care framework.

Concluding Comment

One common theme in these articles reflects a core controversy in contemporary mental health care, namely, how clinicians justify the use of particular assessment and intervention procedures. The reports from the Division 12 Task Force on Promotion and Dissemination of Psychological Procedures (e.g., Chambliss et al., 1996) are as bothersome and meaningless to some as they are invigorating and, well, validating to others. The matter is not an easy one to resolve, but inherent to any rational discussion of stepped care is the assumption that the clinician has a basis other than personal preference for how particular interventions are rank ordered in terms of cost and intrusiveness. In other words, stepped care has to be regarded within the context of some kind of consensus as to what treatments are effective. It is no accident that all the articles I have discussed here entail careful reviews of the efficacy and/or effectiveness of treatments for particular disorders.

However this controversy is resolved (or debated, because I doubt it will be resolved), the fact is that the concept of stepped care has meaning only to the degree that the levels of intervention available for selection be judged as likely to benefit the particular patient one is working with at a given place and time. Idiographic considerations are inevitable; indeed, they would seem to be professionally required (cf. Fishman, 1999; Persons, 1989). As noted above, the continuing monitoring of interventions beginning with the least expensive and least restrictive promises to provide an essential guide to both clinician and patient as to whether the treatment being implemented is worth staying with—worth the money and the effort relative to the impact it is having on the problem for which the patient has contracted with the clinician for assistance.

Managed care has doubtless been a factor in the growing interest in stepped care and all that it implies. We do not, however, have to embrace current models and practices of managed care to appreciate the ethical imperatives inherent to stepped care: Clinicians should intrude as little as possible into the lives of their patients while at the same time providing the maximum clinical benefits. These goals are to be achieved by clinicians learning and practicing the best that the science can offer and by carefully and continually monitoring during treatment the changes shown by the patient in those psychological domains that there is joint agreement on as foci for intervention.

References


Received January 11, 2000

Revision received February 29, 2000

Accepted February 29, 2000