A Critical Review of Marital Therapy Outcome Research

Stephen Wesley, MD\textsuperscript{1}, Edward M Waring, MD, FRCPC\textsuperscript{2}

**Objective:** This review critically examines conceptual and methodological issues of outcome research designed to evaluate the efficacy of marital therapy (MT). Behavioural marital therapy (BMT), cognitively orientated marital therapy (COMT), emotionally focused marital therapy (EFMT), and insight-oriented marital therapy (IOMT) have provided evidence for efficacy.

**Method:** The initial literature search involved the use of the MEDLINE and Psychlit CD-ROM data bases. A secondary literature search based on citations in articles discovered in the initial search was also conducted. Three broad categories involving the delineation of the client sample, study design considerations, and determination of treatment effects were addressed to investigate the quality of evidence concerning the efficacy of MT.

**Results:** None of the 4 approaches has proven superiority in efficacy, and the research does not provide evidence for the superiority of any one approach in distinct types of marital discord. There is some evidence to indicate that future outcome research should avoid the use of “waiting list control groups,” since their use appears to be neither ethical nor humane in marital discord research.

**Conclusions:** The authors suggest that the concept of treatment efficiency may be more clinically relevant to outcome research. A gold standard for “efficiency” would suggest that a marital therapy approach should produce subjective and objective improvement in 50% of eligible couples, and this improvement should be maintained in half of these couples at the one-year follow-up assessment. This standard would be comparable to the data on the 4 effective approaches described in this review.

(Can J Psychiatry 1996;41:421–428)

**Key Words:** marital therapy, outcome research

The past decade has seen a major increase in both the quantity and the quality of marital therapy outcome (MTOC) research. Although MTs have been used in the treatment of individual adult psychiatric disorders such as depression, phobias, and alcohol abuse, the present review focuses solely on the studies that evaluate the effectiveness of the treatment of distressed couples. Also, since the purpose of this review is to evaluate MTs that may be representative of clinical practice, treatments of 4 sessions or fewer are not considered. Finally, the review is limited to treatments designed to have applicability to a broad range of maritally distressed couples and does not include treatments designed for more focused concerns, such as particular sexual dysfunctions.

This review is organized into sections. The first section, on methodology, defines the characteristics of MTOC research as the term is used in this review; the domain of published reports is delimited in this section by the specification of certain methodological prerequisites. The next section, on research, summarizes the substantial body of literature regarding MT efficacy. Finally, the discussion section critically evaluates and explores the issues and themes that develop from our review of research on MTOC.
Methodological Considerations

Progress in addressing the essential question of MT efficacy depends very much on how psychotherapy investigations are conducted in general. In this section, methodological issues that relate to the quality of evidence provided about efficacy are addressed. These issues can be divided into 3 broad categories: 1) the delineation of the client sample, 2) study design considerations, and 3) determination of treatment effects.

1. Delineation of the Client Sample

The delineation of the client sample involves describing characteristics of the study couples, such as the methods by which they were recruited and selected, information regarding demographics, and the character and severity of their marital discord. These characteristics 1) can be factors that moderate or determine the effects of treatment and are required, therefore, for the accurate interpretation of study findings, 2) are important in noting for which couples a particular treatment works and for what level of severity of distress, 3) allow for comparisons to be made with client samples across clinical or research settings, and 4) allow for replicability of research findings.

2. Study Design Considerations

Delineation and Specification of Treatment. A critical requirement of outcome research is the delineation of the treatment. This includes a description of the treatment, its rationale, and the process expected to alter the problem. Beyond a description of the conceptual rationale and process are the concrete procedures that are thought necessary to bring about therapeutic change. The concrete specification of treatment procedures allows for the replication of treatment in subsequent investigations and applications. In addition, procedures need to be specified so that their correct execution can be ascertained and so that the therapists using the treatment can be properly trained.

Treatment Integrity. The extent to which treatment has been carried out as intended is referred to as treatment integrity. Treatment integrity is crucial since the interpretation of treatment effects depends on the evidence that the procedures were carried out as intended. Ensuring treatment integrity involves proper training of therapists, procedures for coding therapist and/or client behaviour, and the identification of ways in which adherence to treatment manuals is evident.

Control Groups. In order to evaluate treatment outcome, it is necessary that the effects produced by treatment be compared with a baseline provided either by couples who have received no treatment or by a gold standard. Such a baseline has traditionally been obtained through the use of control groups, which pair couples on waiting lists. Waiting list (WL) control groups have been necessary not only for outcome studies investigating one type of MT but also for outcome studies that compare 2 different types of MT in the absence of a gold standard.

In addition to demonstrating treatment effects relative to WL controls, there exists the necessity of proving that these effects are attributable to the specific procedures proposed by the treatment rationale. As noted by Kazdin:

These factors refer to features of therapy that are common to most techniques such as having the patient come to sessions at a treatment setting, providing a statement about the problem and how treatment will affect its amelioration, engaging in some prescribed tasks or activities that are promoted as therapeutic, and others. These factors have been referred to as nonspecific because they are common to virtually all therapies.

It is, therefore, necessary that investigators control for nonspecific factors in order to assess whether the particular form of treatment being investigated produces effects over and above these factors. This can be accomplished by comparing the active treatment being investigated with a treatment not emanating from any specific theoretical orientation. The purpose of the latter treatment condition is to serve as a control for receiving attention and focusing on problem areas. The use of such “attention placebo” control groups allows for the efficacy of the treatment to be studied.

Randomization. The random assignment of subjects to treatment or control conditions is another essential requirement of MTOC investigations. The comparison of 2 or more groups of subjects that are significantly different on some dimension precludes the definitive interpretation of results in any scientific investigation. Randomization minimizes the confounding of results by maximizing the likelihood that the groups being compared will be equivalent, on the average, on all dimensions.

3. Determination of Treatment Effects

Outcome Measures. Two broad categories of therapeutic change measures exist in MTOC research: self-report and behavioural observation. Although self-report measures are more frequently used, some investigators assume that behavioural observation is more scientific and objective. When a couple’s actual behaviour is of interest, it is clear that direct observation is a more appropriate measure. For dependent variables like marital satisfaction, however, self-report measures may be more relevant. As noted by Johnson and Greenberg (3): “this kind of measurement seems to be particularly appropriate in the sense that marital satisfaction or well-being is a qualitative, subjective factor rather than an externally quantifiable phenomena, and thus it is the perception of behavior that is salient to marital satisfaction.”

Many measures exist to assess various dimensions of a couple’s relationship. Certain measures are specified by the treatment being investigated. For example, studies involving a BMT that emphasizes communication skills will necessarily assess changes in communication. Specific measures, however, are only part of the assessment package that is now required of quality outcome investigations. It is generally recognized that the complex and multifaceted nature of MTOC requires multiple, distinct measures to assess more
completely and objectively the effects of treatment. Assessment of the clinical status of a couple and their relationship at both pretreatment and posttreatment for the purpose of measuring therapeutic change is ideally accomplished by the use of a multimodal–multimethod approach.

**Clinical Significance.** Treatment effects in MTOC studies are typically based on statistical comparisons of mean changes resulting from treatments under study. This use of statistics to evaluate efficacy is limited in a least 2 respects: 1) the tests provide no information on the variability of response to treatment within the sample, and 2) whether a treatment effect exists in a statistical sense has little to do with the size or clinical significance of the effect (4). Clinical effectiveness continues to be a function of client satisfaction and therapist opinion. Evidence suggests that subjective and objective outcome do not always correspond.

Supplementary information regarding the degree of change in couples has been provided by the effect-size method of metaanalysis. For a given measure, an effect size is equal to the mean difference between the treated and control couples after therapy divided by the standard deviation of the control couples. Thus the effect-size statistic is an improvement over standard statistical analysis, in that, unlike standard significance tests, this analysis does reflect the size of the treatment effect. As Jacobson and Truax (4) have pointed out, however, although large effect sizes are more likely to be clinically significant, they are not necessarily clinically significant.

Jacobson and Truax have developed a method that is now routinely used to asses clinical significance directly in the reporting of study results (4). They suggest that when the magnitude of change for a given couple exceeds 1.96 standard errors on a key dependent measure, then the change is statistically unlikely to be due to measurement error and that a statistically reliable change has occurred. The percentage of couples who manifest reliable change, called the Reliable Change Index, is a measure of variability in response to treatment. A second measure, termed the Clinical Change Index, is based on the assumption that the dependent measure has 2 distributions, one for the normal population and one for the clinical population of interest. The point at which the 2 distributions intersect can be calculated, and this criterion score is more likely to have been drawn from the normal than from the clinical population. Subjects exceeding criterion score are regarded as no longer clinically distressed.

**Follow-Up.** Related to the issue of clinical significance is the question of whether improvements resulting from treatment are maintained. Measurement of treatment effects at a point in time after treatment has been completed is termed a follow-up assessment. These evaluations are important not only for determining the durability of treatment effects but also for investigating possible differential treatment effects such as those not evident at termination.

**Summary of MTOC Research**

According to Baucom and Hoffman (5), the most popular frameworks for conceptualizing and treating marital distress include behavioural, experiential, cognitively oriented, communication, systems, and psychodynamic or insight-oriented approaches. The current review evaluates the empirical status of each of these orientations. Unfortunately, there is little research on the experiential and systems approaches; consequently, few conclusions regarding their relative effectiveness can be drawn.

**BMT**

The BMT approach is based on a theory of marital discord that emphasizes faulty behavioural exchange operations as important determinants of marital distress. These deficits are conceptualized in terms of distressed couples using aversive control strategies rather than positive reinforcement strategies in attempting to bring about changes in their relationships.

Behavioural intervention strategies are designed to teach couples how to bring about changes in their relationship more effectively. The skills taught are communication and problem solving (CPS) and behaviour exchange (BE). CPS skills are emphasized so couples learn to talk more effectively around issues that produce conflict. BE or contingency-contracting consists of interventions designed to have immediate effects on the relationship. These interventions feature therapist directives regarding homework assignments and the implementation of contracts.

Eleven studies of the treatment of clinically distressed couples have compared BMT to WL control couples, and all but 2 cases showed statistically significant decreases in negative communication. Fourteen studies have assessed the efficacy of BMT relative to WL controls in regard to improvement on measures of presenting problems, and all but 2 showed BMT to yield significant changes. Thirteen of seventeen BMT studies showed significantly more improvement of marital adjustment (MA) and satisfaction. Four studies comparing BMT to various attention control groups have provided inconsistent results on measures of communication skill, target problem resolution, and marital satisfaction and adjustment. In a 2-year follow-up comparing a complete BMT (CBMT) package with 2 of its major components, BE and CPS, 25% of couples receiving CBMT had deteriorated in comparison with 55% and 66% for CPS and BE, respectively (6). Thus, while the components were relatively interchangeable in their immediate effects, both were required to prevent couples relapsing to the same level of distress as experienced before treatment.

The investigations indicate that BMT is effective in statistically significantly reducing requests for behaviour change, altering some aspects of communication, and increasing MA. These results are limited in describing the magnitude of change that has occurred and the percentage of couples re-
Cognitive restructuring (CR) is one such strategy. CR involves identifying and modifying maladaptive relationship beliefs. Three studies have demonstrated the effectiveness of this treatment in reducing destructive relationship cognitions and in improving MA. The results of 3 different studies, however, indicated no significant improvement of BMT efficacy with the addition of CR.

Two new approaches have recently been described: emotional expressiveness training (EET) and affect exploration. EET is a communication skills approach that teaches self-disclosure of emotions. Affect exploration, as described by Jacobson and Holtzworth-Monroe (9), involves exploration of the individual partners’ experiences of patterns of emotionally charged marital interactions. Therapist interventions are used to encourage cognitive and affective change in this context.

Five studies have investigated the EET approach. The findings suggest that couples can be taught these specific communication skills, but increases in MA do not occur when skills are taught alone. A study by Baucom and others (10) determined that the addition of EET, and both EET and CR, to BMT did not result in a significant enhancement of BMT on any measures. Halford and others (11) compared BMT with a treatment that included affect exploration, CR, BMT, and a component designed to enhance the generalizability of behaviour changes across settings and time. Again, no enhancement of BMT was demonstrated.

**COMT**

Two forms of COMT have been empirically investigated. One form, CR, has been discussed in the previous section on BMT. Cognitive marital therapy (CMT) is a brief, structured, couples therapy based on Kelly’s theory of personal constructs. Kelly suggested that one’s ideas about people and events, hence the term “cognitive,” may define one’s personality (12). His theory of personal constructs emphasized the human tendency to think about relationships to extend one’s personal understanding of experiences (12). CMT is designed to enhance marital intimacy through facilitating self-disclosure of personal constructs. Self-disclosure is the process of revealing one’s thoughts, beliefs, or past experiences to another person, and it is considered the major process through which intimacy between people develops.

In 2 studies, Waring and others (13,14) compared CMT with both a WL and an attention placebo control group in couples referred to psychiatric outpatient settings for severe marital discord. No significant differences were found between CMT and the WL control. For women, however, improvement on a composite measure of symptoms, intimacy, and adjustment was demonstrated. Perhaps of more importance was the failure to observe any evidence for the spontaneous remission of marital discord in the WL control group. Many couples dropped out of the project or separated while on the waiting list.

Johnson and Dandeneau (15) compared CMT with EFCT in the treatment of couples with moderate marital distress. Both treatments showed significant gains over WL controls on measures of MA, intimacy, and target complaints. Differential effects in favour of EFCT were found in MA at 10 weeks after the termination of therapy. The differential success of CMT in the treatment of severely versus moderately distressed couples illustrates the importance of carefully describing characteristics of the client sample.

**EFCT**

EFCT is a form of marital therapy, developed by Johnson, that arises from an integrated intrapsychic and interpersonal perspective (16). It is based on the experiential tradition of psychotherapy, which emphasizes the role of emotion in change, and the systemic tradition, which emphasizes the roles of communication and interactional cycles in the maintenance of discord (16).

The therapist identifies the negative interactional cycles and guides the couple in accessing the unacknowledged feelings underlying each person’s position. The therapist uses interventions to access and heighten emotional responses and then reframes the interactional problems in these emotional responses. Thus attention to out-of-awareness experience and feelings is paired with active therapist efforts to reframe and modify overt behaviour.

Five studies have investigated the efficacy of EFCT in the treatment of marital discord. Johnson and Greenberg (3) compared EFCT and BMT, with both groups showing significant gains over ML controls. EFCT was significantly higher on MA, intimacy, and reduction of target complaint scores at posttest; at the 8-week follow-up session, MA scores for the EFCT couples remained significantly higher than for the BMT couples. A partial replication of this study, in which WL couples from the original study were treated with EFCT, found significant changes on most of the dependent measures. In a study comparing EFCT to an integrated systemic couples therapy (ISCT), both treatment groups showed more improvement than WL controls. Differential effects were found in favour of ISCT at the 4-month follow-up on measures of MA, goal attainment, and target complaint level. Considerable posttest regression is typical in the EFCT condition. In another study, EFCT was compared with EFCT plus communication training (CT). Both treatment groups
achieved superior gains at posttest compared with WL controls. The only differential effects were EFCT plus CT's superiority on a measure of communication and EFCT's superiority on target complaint levels at follow-up. Finally, the study by Johnson and Dandeneau (15) comparing EFCT and Waring's CMT was discussed in the previous section (p 424).

**IOMT**

Although psychodynamic, or IOMT, approaches to MT represent a major theoretical and clinical tradition, their effectiveness in helping couples improve on global marital accord in well-designed experimental studies has been largely unexamined. There have been five published studies of "insight" MTs that meet the minimum requirements of randomization and inclusion of control groups. Interpretation of these studies is complicated by the diversity of their treatment methods and conceptualizations of insight. As noted in Gurman and others' review (17), Epstein and Jackson compared an insight interaction group to a CT group. Through feedback on the impact of couples' verbal and nonverbal behaviour, the insight condition sought to increase spouses' awareness of their behavioural interaction patterns. Relative to a WL control set, the insight group was significantly improved on only 1 of 11 categories of verbal behaviour and showed no differences at posttest on self-reported MA. Crowe (17) compared BMT, group analytic couples therapy, and a support control group. He found no significant differences on individual or target complaints between the 2 treatments at posttest and no differences in MA between the group analytic couples therapy set and the support group at follow-up. O'Farrell and others (17) compared BMT, an insight interaction therapy (ICT), and a no marital therapy (NMT) control group in the treatment of outpatient male alcoholics and their wives. ICT emphasizes mutual support, the sharing of feelings, problem solving through discussion, and verbal insight. ICT was superior to NMT on measures of degree of relationship change and positive communication but showed no significant changes on measures of MA. Boelens and others (76) compared group-based behavioural contingency contracting and a "strategic" therapy designed to "provide partners insight into their overt and covert power struggles." Both groups improved significantly on self-report and assessor ratings of MA, but not on measures of problem-solving skill; there were no significant differences between the 2 measurements.

Gurman and others (17) have criticized these studies for not adequately representing insight therapies by limiting directive therapist interventions. They conclude that the only information these studies provide is that "insight"—whether based on the interpretation of conflicts and defences (Crowe's study), underlying power struggles (Boelen and others' study), or the relational impact of observable behaviour (Epstein and Jackson's and O'Farrell and others' studies), in the absence of concrete therapist directives for behaviour change—is not a particularly effective approach to the treatment of marital problems (17).

In contrast to these 4 studies, Snyder and Wills (18) compared BMT to an IOMT that emphasized not only resolving hidden sources of conflict but also included "instruction in listening and empathy as well as modification of grossly destructive communication patterns." This study represents a well-designed empirical investigation of 2 major forms of marital therapy. Among its notable methodological features are the use of treatment manuals, documentation of treatment integrity, extensive training and use of professional therapists without strong theoretical biases, a comprehensive multidimensional assessment battery, and an unprecedented 4-year follow-up with 96% of the original couples. In addition, this investigation included a relatively large sample population for each treatment group.

The MTs examined in this study represent 2 major clinical orientations with very different theoretical approaches to marital discord. This fact makes it one of the most important studies for illustrating the efficacy and comparability of these 2 MTs. In our report of their results, we will include traditional statistical analyses as well as indices of reliable and clinically significant change.

Both treatments produced significant changes in global marital accord relative to a WL control group, with no significant differences between treatment conditions observed. These changes were largely maintained at a 6-month follow-up assessment. Using an index of reliable change, the investigators classified couples as improved, unimproved, or deteriorated based on a self-report measure of global marital accord. Findings indicated that IOMT and BMT resulted in statistically reliable change for 73% and 62% of couples, respectively, in contrast to 10% improvement in WL couples. When a second criterion of clinical significance was applied requiring both statistically reliable improvement and a change from the dysfunctional to the functional range, 55% of BMT compared with 40% of IOMT and 5% of WL couples joined the ranks of the nondistressed. Eighty-eight percent of BMT and 79% of IOMT couples maintained their posttest status at 6-month follow-up. The proportions of couples at follow-up showing statistically reliable improvement and a change from dysfunctional status relative to intake in IOMT and BMT groups were 43% and 50%, respectively.

Thus, according to results derived from indices of reliable and clinically significant change, BMT would appear to be a somewhat more efficacious treatment. By 4-year follow-up, however, a significantly higher proportion of BMT couples had experienced divorce (38% of BMT and 3% of IOMT couples). In addition, couples in BMT showed significantly higher rates of deterioration relative to their status at intake or termination. The deterioration rates for BMT of 35% to 46% are consistent with the Jacobson and others' (6) 2-year follow-up of a component analysis study, which yielded
deterioration rates of 25% to 66% and divorce or separation rates of 9% to 55%. The significant differences in outcome at long-term follow-up suggest that, in the long term, IOMT is a more effective treatment of marital discord than BMT (19–21).

Discussion

Although most MTOC research has focused on BMT, recent studies have begun to examine the efficacy of other major models of MT. Apart from communication training (EET), treatments tested against a WL control group have shown statistically significant change on self-report measures of MA and satisfaction. The only exception to this generalization is a study that examined the effectiveness of Waring’s CMT in the treatment of severe marital discord. The differential success of CMT helps to qualify the documented success of MTs that have been empirically investigated. The severity of marital discord in couples treated in the CMT study would have excluded them from virtually every study reported in this review. In fact, research has consistently found that severely distressed couples are less likely to benefit from therapy.

The most compelling observation in reviewing these studies may be that WL control groups show no tendency for remission of marital discord. The improvement rate in the absence of treatment is so low that even small changes in an experimental treatment are likely to be statistically significant. Quite often, couples’ relationships deteriorate without treatment. Typically, they remain unchanged during the waiting period (20). In addition to these methodological considerations, there is the question of whether the use of WL controls is ethical or humane. Evidence from this review suggests that, in the future, researchers might consider attention placebo psychotherapy control groups rather than WL controls.

Treatment efficacy refers to the clinical significance of these treatment effects: How often is this approach effective? Are the effects clinically meaningful? The results are generally less impressive than statistical comparisons with WL controls would seem to indicate. When studies use indices of reliable and clinically significant change to determine results, most treatments report no better than 50% success in producing happily married couples. This success rate is quite uniform across both studies and treatment modalities in mild to moderate marital discord. Thus it appears that treatments help some couples with mild to moderate marital discord but are leaving substantial numbers of couples unchanged or still distressed. The 50% success rule, however, may have potential as a gold standard for comparison of new MT approaches.

Although the success rates of MT may not seem impressive, there are several factors that suggest that these rates may not accurately reflect those obtained in clinical practice. For instance, clinicians typically formulate a treatment for each couple, taking into account the unique characteristics of that couple. In MTOC studies, distressed couples are randomly assigned to treatments regardless of the types of complaints they present. Thus an inadequate matching of couples’ specific needs with appropriate treatments may result in a lower rate of success. Another point involves the defining of treatment success. For some couples, separation is a desirable and appropriate treatment; its use as a measure of the lack of success of therapy may not be appropriate. For other couples, separation is a desirable and appropriate outcome, and increasing marital satisfaction is an unrealistic goal. Inadvertent inclusion of such couples in MTOC studies may also be a factor that decreases the MT success rate.

Thus far, only the immediate effects of MT have been considered. Follow-up assessments beyond 6 months have been uncommon. Results from these assessments have been inconsistent, but they suggest that deterioration of treatment gains does occur. One 2-year follow-up of a BMT component study found that about 30% of those couples who recovered during the course of therapy had relapsed. In a 4-year follow-up study conducted by Snyder and Wills (18) comparing IOMT with BMT, 10% and 46% of couples, respectively, had deteriorated since the termination of MT. Also, 3% of IOMT and 38% of BMT couples had divorced at 4 years. Thus at least 2 studies have found substantial relapse in BMT at the time of a long-term follow-up, and one study suggests that IOMT is particularly effective at fostering long-term couple stability (21). Perhaps a gold standard for relapse and/or deterioration may also be determined.

The possibility that differences in outcome can be identified only at long-term follow-up begs the question: Why can immediate differences not be found? In addition, why are no differences found on measures that assess relationship dimensions specifically targeted by one treatment but not another? One explanation may be that labels describing MT are misleading in that they imply that specific interventions have specific effects; as Halford and others (11) have pointed out, the reciprocal influences among behaviour, cognition, and affect make it unlikely that any one treatment affects just one aspect of marital distress. It would seem that all marital therapies affect the entire marital relationship, rather than specific components of marital interaction, and can thus produce similar specific and overall results.

In addition to the generalization effects of specific interventions, nonspecific factors involved in all MTs may act to mask the differences among treatments. This explanation is supported by the MTOC studies that compare active treatments to attention placebo controls not consistently yielding significant differences on any outcome measures. Kazdin (2) has noted that when control conditions do generate high levels of credibility that equal the treatment conditions, treatments and nonspecific control conditions rarely differ in outcome. Presumably, the effects of nonspecific factors deteriorate faster than those induced by the more specific interventions,
thereby allowing differences in MT treatments to become evident only at longer-term follow-up.

A different yet not mutually exclusive explanation for why no short-term differences can be found among MTs pertains to the concept of statistical power. Statistical power refers to the likelihood of finding differences between treatments when, in fact, the treatments are truly different in their outcomes. For a given criterion level of statistical significance (alpha), statistical power is a function of sample size and the differences that exist between groups (effect size). A power of 0.80 has been recommended as an acceptable margin of protection against the null hypothesis, with higher levels (0.90 to 0.95) often encouraged as more appropriate values (22). Given the power of 0.80, the number of couples required in a treatment or control group to determine a certain minimum effect size can be calculated. According to such calculations, 63 couples per treatment group would be the minimum required to detect, 4 out of 5 times, an effect size (ES) classified as medium (ES = 0.50).

No controlled MTOC study to date has included enough couples (range 15 to 86) to detect the medium ES of 0.50 given this recommended power of 0.80. To give an indication of the magnitude of ES demonstrated in MTOC studies, a recent metaanalysis of 17 BMT studies resulted in a mean ES of 1.02 for BMT treatment groups compared with WL controls (8). Thus inadequate sample sizes (low levels of statistical power) could account for the absence of immediate or short-term treatment differences in comparative outcome studies. As the treatment differences (effect sizes) increase with time, the sample size required to detect the difference decreases, and so these differences can eventually be detected. As Kazdin and Bass (22) have noted for psychotherapy in general, it seems that the absence of differences in comparative outcome studies of MTs cannot be interpreted unambiguously without improved power.

Finally, several other issues not directly related to treatment efficacy can help to differentiate the relative value of different MTs: 1) the breadth of applicability of the MT—What range of severity and characteristics of presenting problems can the therapy effectively treat? 2) its efficiency—What is the attrition rate of couples in therapy? How long does the therapy last? How difficult is the therapy to learn and to apply? and 3) its agreeability—How pleasant or unpleasant is the therapy for the clients and the therapists?

In summary, MTOC research has improved methodologically and has demonstrated effectiveness in the short term in comparison with WL control groups. The latter groups, however, have consistently demonstrated the absence of spontaneous remission. This evidence suggests that future MTOC studies should avoid creating WL control groups for ethical and humane reasons.

Efficiency may be a more important research issue than effectiveness at this point in time. Efficiency refers to customer satisfaction, acceptance of approach, completion of therapy, and long-term follow-up. A possible gold standard for MTOC research that would focus on efficiency rather than effectiveness and allow for new approaches to be evaluated without the need for control groups might include the following stipulations: 1) that 50% of the couples offered the MT accept the approach and complete the required course of therapy, 2) that 50% of the couples who accept and complete the MT demonstrate subjective and objective improvement on a measure of theoretical relevance to the MT approach, and 3) that 50% of these couples maintain improvement at one-year follow-up.

**Clinical Implications**
- No spontaneous remission of marital discord.
- No clear superiority for any specific type of marital therapy.
- Efficiency is an important concept for marital therapy.

**Limitations**
- Sample sizes are too small.
- Effects are modest.
- There were few investigators.

**References**

Résumé

Objectif : L’article examine d’une manière critique les problèmes théoriques et méthodologiques de la recherche sur les issues afin d’évaluer l’efficacité de la psychothérapie conjugale (PC). La psychothérapie conjugale comportementale (PCC), à orientation intellectuelle (PCOI), à orientation affective (TMOE) et par la compréhension de soi (PCCS) semblent efficaces.

Méthode : On a d’abord dépouillé la documentation grâce aux bases de données MEDLINE et au CD-ROM Psychlit, puis effectué une deuxième recherche à partir des citations dans les articles. Trois grandes catégories délimitant l’échantillon de clients, le type d’étude et les effets du traitement ont servi à déterminer la qualité des preuves illustrant l’efficacité de la PC.

Résultat : Aucune des 4 approches n’est plus efficace qu’une autre et la recherche ne fournit pas d’indication sur l’approche se prêtant le mieux à la résolution de certains problèmes conjugaux. Il semble néanmoins qu’à l’avenir, les chercheurs devraient éviter de recourir à des «groupes témoins tirés de la liste d’attente». Pareille méthode paraît contraire à la déontologie et à la compassion dans ce type de recherche.

Conclusion : Les auteurs suggèrent que l’efficacité du traitement est cliniquement plus appropriée à la recherche sur les issues. La meilleure preuve de «l’efficacité» d’un traitement est qu’il suscite des améliorations subjectives et objectives chez la moitié des couples admissibles. Pareille amélioration devrait encore être manifeste pour la moitié des couples à l’évaluation de suivi, un an plus tard. Une telle norme se comparera aux résultats sur l’efficacité des 4 approches données ici.