Acta Psychiatr Scand 2007: 115: 352–359
All rights reserved
DOI: 10.1111/j.1600-0447.2007.01007.x

Copyright © 2007 The Authors
Journal Compilation © 2007 Blackwell Munksgaard
ACTA PSYCHIATRICA SCANDINAVICA

Clinical Overview

Treating depression with the evidence-based psychotherapies: a critique of the evidence

Parker G, Fletcher K. Treating depression with the evidence-based psychotherapies: a critique of the evidence.

Objective: While Cognitive Behaviour Therapy (CBT) and Interpersonal Therapy (IPT) have been positioned as first-line evidence-based treatments for depression, we suggest that limitations to the 'evidence' deserve wider appreciation.
Method: A systematic literature search was undertaken, and limitations to the evidence base discussed.
Results: The review suggests that the specificity of CBT and IPT treatments for depression has yet to be demonstrated and details likely reasons.
Conclusion: The superiority of CBT and IPT may well be able to be demonstrated across defined rather than universal circumstances. To achieve this aim, outcome research should move away from testing treatments as if they have universal application for heterogeneous disorder categories. Findings have distinct implications for the clinical management of depressive disorders, and particularly in relation to the utility of psychotherapy.

G. Parker, K. Fletcher

School of Psychiatry, University of New South Wales, and Black Dog Institute, Prince of Wales Hospital, Sydney, Australia

Key words: psychotherapy; cognitive behaviour therapy; interpersonal psychotherapy; depression; clinical overview

Gordon Parker, Black Dog Institute, Prince of Wales Hospital, Randwick 2031, Sydney, Australia.
E-mail: g.parker@unsw.edu.au

Accepted for publication February 2, 2007

Clinical recommendations

- The identification of true treatment validity of CBT and IPT has likely been obscured by randomized controlled trial procedures and paradigms. The level 1 evidence base has been inappropriately constructed and analysed.
- The specific efficacy of CBT and IPT treatments for depression remains unestablished, and insufficient appreciation of their non-specific therapeutic impact exists.
- When tested as 'universal' therapies for heterogeneous depressive disorder categories, CBT and IPT efficacy is likely to be similar to other 'appropriate' psychotherapies.

Additional comments

- It may be possible to demonstrate the utility and treatment validity of CBT and IPT across defined circumstances, through identification of candidate disorders and aetiological factors that best reflect each therapy's theoretical rationale.
- Current clinical trial procedures present unsophisticated and limited models for actual clinical application. Thus, the transportability of evidence-based study efficacy estimates to any 'real world' effectiveness estimate cannot be assumed.
- The clinical implications of over-selling CBT and IPT as specific and superior treatments for depression impact on both patients and practitioners. Critical interpretation of the 'evidence' is therefore warranted for consideration by clinicians and researchers.

Introduction

The perceived relevance of psychotherapy for managing depression has varied considerably in Western regions. In the first half of the 20th century, it was the dominant general paradigm. Challenges appeared mid-century. For example, Eysenck (1) reviewed 24 studies involving > 7000

cases and concluded that 'They fail to prove that psychotherapy, Freudian or otherwise, facilitates the recovery of neurotic patients'. Observing that the data indicated that roughly two-thirds of 'neurotic patients' recover or improve markedly within 2 years of illness onset – whether treated by psychotherapy or not – Eysenck imputed spontaneous or natural resolution as the principal explanation. He noted the 'strong feelings of usefulness and therapeutic success which many psychiatrists and clinical psychologists hold', and argued for relying less on beliefs and more on facts. In this review, we suggest that, while the factual base to psychotherapy is now vast, beliefs continue to influence the *interpretation* of the evidence. As described in another context (2), the risk here is of 'evidence b(i)ased' extrapolation dominating interpretation of the facts.

While decades ago the psychotherapies were criticized for lacking specific benefits, theorists importantly drew attention to their non-specific ingredients. In a seminal paper, Frank (3) drew links between psychotherapy and 'healing' ingredients across culture and time, identifying four non-specific therapy factors. Firstly, an emotionally charged, confiding therapeutic relationship. Secondly, a healing setting. Thirdly, a rationale providing a plausible explanation for the symptoms and a logic for the recommended treatment procedure. Fourthly, a treatment procedure (or 'ritual') believed by both patient and therapist to be restorative. As Psychiatry has become more evidence based, acknowledgment of such non-specific ingredients has decreased.

During the 1950s and 1960s, antidepressant drugs were introduced, with the placebo-controlled trial becoming the key evaluative paradigm. A challenge was issued to those espousing psychotherapy to submit to the same evaluative rules. Numerous analyses of aggregated data sets were undertaken. For example, a meta-analysis (4) of 375 psychotherapy studies revealed an effect size of 0.68, indicating that the average psychotherapy client was better off than 75% of untreated controls. Another meta-analysis of 475 published and unpublished studies (5) derived an even larger effect size (of 0.85), favouring psychotherapy over control treatment. However, no distinct differences were found between differing psychotherapies. Luborsky and colleagues (6) reported a similar meta-analytic finding – and concluded that, as all analysed psychotherapies appeared to produce improvement in a high percentage of patients, common elements were operative.

During the last 2 decades, two psychotherapies – cognitive behaviour therapy (CBT) and interpersonal psychotherapy (IPT) – have been positioned as (i) 'specifically' effective treatments for depression; (ii) superior to 'other psychotherapies' and, in some evaluations; (iii) superior to antidepressant drugs.

Such advocacy is illustrated by the following commentaries. 'Cognitive therapy is superior to other forms of psychotherapy in the treatment of depression' (7). 'The evidence for the efficacy of cognitive behaviour therapy is considerable whereas there is little evidence for the more traditional psychotherapies. Because of the evidence base for it, cognitive behaviour therapy has advanced as the treatment of choice, leaving the other psychotherapies as "also rans"' (8). 'On a scale of effectiveness from "rubbish" to "pure gold", cognitive behaviour therapy is well towards the latter' (9).

Formalized guidelines reify such high status – usually by reference to their 'evidence base'. 'When considering individual psychological treatments for moderate, severe and treatment-resistant depression, the treatment of choice is CBT' (10). 'Cognitive behavioural therapy and interpersonal therapy have the best-documented effectiveness in the literature for the specific treatment of major depressive disorders' (11). 'Cognitive behavioural therapy and interpersonal therapy are as effective as antidepressant medications in mild-to-moderate depression and can be recommended as first-line therapies (level 1 evidence)' (12).

Aims of the study

In this review, we overview limitations to the 'evidence' that has allowed CBT and IPT to be positioned as superior treatments for depression. We also challenge the claim for their 'specificity', and suggest – as earlier demonstrated for psychotherapy generally, and increasingly recognized as contributing to the high rate of failed placebo-controlled antidepressant drug studies – that there has been insufficient appreciation of their non-specific therapeutic impact. While both CBT and IPT have been held to have short-term and long-term efficacy, as well as having demonstrable prophylactic potential in reducing relapse and recurrence, we focus on acute-phase studies – as it is here that their status is cemented in practice guidelines.

Material and methods

We examined the psychotherapy literature, respecting early findings that argued the need to delineate psychotherapy's specific and non-specific effects, and then focussing on relevant studies evaluating CBT and IPT.

Our overview of the literature indicated that, while CBT and IPT as 'Evidence Supported Therapies' (or ESTs) have high cachet status, and are commonly imputed as superior to other therapies, the supportive evidence base is less substantive than generally considered. We now detail supportive arguments for that interpretation.

Their high status emerges from their formal evaluation in efficacy studies that have – for better and for worse – respected antidepressant drug study paradigms. For better, in that the individual studies operate to the tenets of 'evidence based medicine', which accords 'level 1' status to evaluation emerging from multiple randomised placebo-controlled trials. For worse, as EST procedures and paradigms may have obscured identification of such psychotherapies' true treatment validity.

Both CBT and IPT have now been formally examined as acute-phase treatments of major depression by meta-analytic techniques. Gloaguen and colleagues (13) concluded that CBT was superior to no therapy, to all other psychotherapies apart from behaviour therapy, and to antidepressant drugs. However, when Wampold (14) reanalysed the same data set, after removing 'non-bona fide' therapies, CBT did not differ from 'other psychotherapies'. The only published meta-analysis of IPT (15) quantified only a non-significant trend for IPT to be superior to placebo in terms of remission rates.

In our specific reviews of the aforementioned CBT (16) and IPT (17) meta-analyses, we argued that any suggested superiority had been achieved by selection of inappropriate comparators (e.g. wait list assignment) or implausible control strategies, sometimes termed 'intent-to-fail' (18) or 'non-bona fide' therapies. Thus, the actual 'evidence' suggests that, with such ineffective control strategies, CBT and IPT are not superior to other psychotherapies. Issues emerging from our specific reviews of CBT (16) and IPT (17) encouraged this overview paper.

After closely considering the large evidence base for psychotherapy, Wampold (19) arrived at four conclusions. Firstly, psychotherapies intended to be therapeutic produce similar results. Secondly, evidence supporting specific effects of the psychotherapies is absent. Thirdly, and as a corollary, the evidence is consistent with improvement during psychotherapy more being a consequence of non-specific or common effects (e.g. therapist factors and therapeutic alliance) shared across the psychotherapies. Fourthly, allegiance of the therapist to the treatment being tested is strongly related to outcome, and can shape differences between comparator treatments.

Each point will be considered, with some requiring extension beyond CBT, IPT and depression domains.

Results

Treatments intended to be therapeutic produce similar results

In terms of suggested 'equipotency', we are not asserting that all 'treatments' are equal, noting Wampold's (20) observation that classification of treatments as bona fide or not is critical in considering 'treatment' efficacy. As observed by Westen and Bradley (21), 'We know of no published study in the last decade purported to demonstrate specific efficacy of a treatment that has ruled out the most parsimonious rival hypothesis: that something intended to be effective works better than something intended to be ineffective' (p. 267). Differential effects generally disappear when analyses are limited to therapies intended to be therapeutic. In addition to his reanalysis of the CBT meta-analysis, Wampold and colleagues (22) reviewed all 277 psychotherapy studies published from 1970 to 1995 comparing two or more treatments designed to be therapeutic. The preponderance of effect sizes were near zero, the aggregate 0.20 (a small effect size) and the frequency of large effect sizes was consistent with that expected by chance. Importantly, analyses showed that only 1% of the variance in outcome was due to specific (i.e. components integral to the psychotherapy) effects. Additionally, any 'treatment as usual' (TAU) comparator is a limited strategy in practice. As detailed elsewhere (23), TAU strategies tend to favour the active psychotherapy, as therapist enthusiasm, patient expectancy and other common factors are less likely to be activated in the TAU condition.

Evidence supporting specific effects to the psychotherapies is absent

Evidence of psychotherapy's benefits emerging from specific ingredients is lacking. Baskin and colleagues (24) demonstrated that well-designed treatments lacking specific ingredients produce benefits approaching EST psychotherapies (i.e. an effect size of 0.15). Wampold (25) emphasised two relevant process issues: psychological mechanisms do not mediate treatment effects; with benefits commonly attained before specific ingredients are delivered. As evidence of the former, Jacobsen and colleagues (26) undertook a study to determine what components of CBT contributed to its efficacy – testing (i) behavioural activation;

(ii) coping strategies for automatic thoughts and depressogenic schemas; and (iii) modification of core depressogenic schemas. Outcome at termination and follow-up did not differentiate the three presumed prototypic components, with the authors observing that dismantling of 'probably the most established psychological treatment in existence, failed to demonstrate that the components of the treatment were responsible for the benefit'. In reference to the latter point, Ilardi and Craighead (27) reported that, as most improvement during a course of CBT occurs before formal introduction of cognitive restructuring techniques, early improvement more reflects non-specific treatment factors rather than theorised cognitive mediation. Clearly, all four of Frank's non-specific therapy factors (noted earlier) are provided within recommended CBT and IPT frameworks. For example, the first few sessions of CBT tend to be characterized by forming a collaborative therapeutic alliance, while a treatment rationale is formally presented (28). Lambert (29) therefore suggested that early responders may be more resilient, better prepared, more motivated and more receptive to therapeutic influences, and that early response is more likely to reflect response to common factors than to specific interventions.

Several 'deconstruction' reviews and studies have considered the impact of designated specific ingredients. Wampold (25) argued that treatments designed for particular deficits (e.g. CBT for patients with cognitive deficits) are no more effective than treatments that have no such design objectives (e.g. IPT). Several studies have reported counter-intuitive findings in relation to CBT and IPT. For example, in the NIMH TDCRP study (30) of CBT, IPT, imipramine and clinical management, Sotsky and colleagues (31) found paradoxically that depressed patients with fewer dysfunctional attitudes (the suggested theoretical lever for CBT) had a superior response to CBT. In another TDCRP study analysis (32), it was concluded that neither therapy showed clear and consistent effects on measures related to their theoretical origins (i.e. cognitions for CBT, social adjustment for IPT). Examining transcripts from the TDCRP study, Ablon and colleagues (33) argued that both IPT and CBT *as actually practiced*, strongly resembled the ideal prototype for CBT. Thus, despite careful efforts at manualization and adherence checks, the program may have compared two cognitive therapies. A review of the evidence (34) revealed no support for specificity (i.e. that CBT will be preferentially beneficial to those with irrational thoughts and IPT for those with maladaptive relations).

Improvement during psychotherapy is a consequence of non-specific effects shared across the therapies

We now consider the proposition that improvement in depression during a course of CBT or IPT is more likely to reflect the impact of 'common' therapeutic factors rather than the therapy's 'specific' ingredients. For 'common', some might use the term placebo which, by its negative attribution, is inappropriate for many therapies – and particularly the psychotherapies – if Frank's delineation (3) of common (and intrinsically healing) ingredients is respected. From a review of empirical studies, Lambert (35) estimated that only 15% of improvement during psychotherapy was attributable to specific techniques (as against 30% attributable to the therapeutic relationship, 15% to expectancy effects, and 40% to client variables and extra-therapy factors). Wampold (25) suggested that common factors accounted for nine times more variability in outcome than specific ingredients, while his meta-analysis estimated that specific therapeutic effects accounted for only 8% of the variance.

'Therapeutic alliance' is one of the most frequently cited non-specific therapeutic factors contributing to successful psychotherapy (36). Krupnick and colleagues (37) estimated that therapeutic alliance accounted for 21% of the variance in improvement in the TDCRP study – compared with 2% for treatment differences. Furthermore, they found that ratings of video-taped therapeutic alliance measures had a significant (and essentially non-differential) effect on all four treatment arms (including CBT and IPT). For each unit of increased therapeutic alliance or patient contribution to it, the estimated odds of remission increased threefold.

Both patient and therapist factors contribute to any working alliance. Patient factors include severity of disturbance, motivation and expectations, capacity to relate, psychological mindedness, and application to therapy, and contribute – as noted earlier (35) – some 40% of the variance in improvement. In relation to therapist factors, Burns and Nolen-Hoeksema (38) demonstrated that the level of therapist empathy predicted response to CBT in those with depression. Kim and colleagues (39) quantified 8% of the variance in the TDCRP psychotherapy arms as attributable to the therapists – as against 0% to the particular treatment delivered. Intriguingly, and while individual studies variably show experienced or inexperienced therapists as superior, Lambert (29) suggested that little support exists for the view that outcome is advanced by necessarily having

experienced or highly trained therapists. Given that therapists have been shown to differ in clinical trials in terms of outcomes (despite therapists being trained and therapies manualised), Wampold (19) proposed that '*how* therapy is conducted is more important than *what* therapy is conducted' (p. 196).

Therapy allegiance is strongly related to outcome

Fourthly, we consider therapy allegiance as a determinant of therapeutic outcome. Wampold (25) noted that meta-analyses investigating allegiance have reported effect sizes ranging up to 0.65, in comparison with 0.20 for specific effects. A number of meta-analyses have examined allegiance effects on therapeutic outcome. An earlier review (5) reported that, in 88% of the 475 studies, investigators were biased in favour of a particular treatment type. The authors concluded 'where the allegiance was in favour of the therapy, the magnitude of effect was greatest. Where there was bias against therapy, the effect was least' (p. 120).

Allegiance is not only strongly associated with outcome but may equalize treatment differences when taken into account (40). A meta-analysis (41) of 58 studies of psychotherapy for treatment of depression, established that the therapeutic model of researchers was associated with more favourable results than the comparison model. It was concluded that treatment efficacy differences may be an artefact of therapy allegiance, estimating that almost one-third of the variance in effect sizes produced was due to this factor. In a review of eight comparative treatment meta-analyses, Luborsky and colleagues (42) estimated that allegiance accounted for over 69% of the variance in outcome. Westen and colleagues (18) estimated that, if the correlation of 0.85 found in the Luborsky *et al*. study was converted to a binomial effect size (43), the implication would be that 92.5% of the time, therapy allegiance alone could predict which treatment would be most successful.

Implications

All material presented to this point might suggest that we are cynical about the efficacy of the evidence-based psychotherapies under review. Rather, we wish to argue that their specific efficacy remains unestablished. We now consider limitations to the models used to test these psychotherapies in acute-phase studies.

While distinct limitations exist in the current testing and interpretation of RCTs of antidepressant drugs (44), reliance on the same paradigm is probably even less appropriate for testing CBT and IPT – and particularly what constitutes an appropriate control condition. The comparator therapy can be designed to control comparably for non-specific therapeutic ingredients such as those articulated by Frank (3), or be 'de-powered' by lacking any such components. The dilemma here is that the former risks disallowing the specific psychotherapy demonstrating any differential efficacy, while the latter risks contriving false superiority. In the latter instance, any suggested superiority of the 'specific' psychotherapy may have been achieved by the greater contribution of non-specific ingredients. Requesting a therapist to be inert (in so-called 'clinical management' arms) or to provide a 'psychotherapy' with only non-specific therapeutic ingredients is difficult for any therapist to implement.

Moreover, in comparison with a drug study, therapists delivering the 'active' psychotherapy will not only be aware that they are giving the therapy being tested but that they are doing something 'active' (24). Both factors are likely to increase therapist motivation and advance a range of positive non-specific therapeutic components. Therapists providing a more inert control 'therapy' risk compromising both the non-specific therapeutic components and the 'therapeutic alliance' – all of which, as detailed, link strongly to psychotherapy outcome.

Moving forward

If evidential support cannot be obtained from current trial evaluations, this does not disallow the possibility that CBT and IPT may be superior psychotherapies when operating across defined rather than universal circumstances. Paul (45) suggested that outcome research should be directed at the following: '*What* treatment, by *whom*, is most effective for *this* individual with *that* specific problem, and under *which* set of circumstances' (p. 111). Testing the utility and treatment validity of CBT and IPT might better proceed by weighting these components. This would require abandoning the current non-specific model which allows treatments (psychotherapies or drugs) to be tested as if they have universal application for non-specific disorder categories (e.g. major depression and dysthymia). Logically, if a non-specific treatment (i.e. viewed as having universal application) is tested against a non-specific disorder category (e.g. major depression), why would we not anticipate a non-specific result? For example, if a specifically effective asthma drug was provided to a sample of patients with major dyspnoea, it might appear ineffective if there were too few

subjects with asthma (the specific 'target' in the overall sample).

We earlier noted that the specific component to any psychotherapy contributes only slightly to outcome. If not addressed in study designs assuming universal application of the therapy, very large sample sizes would be required to confirm that a truly specific therapeutic intervention is efficacious, and certainly larger than those employed historically in CBT and IPT studies. Alternative strategies may be preferential. One would be to have advocates of CBT and IPT identify those candidate disorders and aetiological circumstances that might best reflect the theoretical rationale for each therapy. For example, for those with non-melancholic disorders, antecedent and ongoing (rather than mood state dependent) dysfunctional attitudes and cognitive schemas might provide the aetiology-treatment linked circumstance for demonstrating CBT efficacy. Alternately, a life pattern of poorly handled interpersonal crises contributing to depression might be a starting point for demonstrating IPT's specific efficacy. If differing treatments demonstrated specificity in factorial cell designs (treatment type examined against presence or absence of putative and non-putative aetiological factors), treatment validity of each psychotherapy could be claimed and the boundaries of each treatment progressively circumscribed to its ecological niche.

Nature of the evidence

We are not critical of evidence-based medicine (EBM) or evidenced-based psychiatry as a theoretical ideal. EBM ordering of 'evidence' assigns level 4 observational studies low in the hierarchy, recognizing the risk of interpretive error. Level 1 evidence (based on multiple controlled trials) is positioned as the ideal, and does not recognize interpretive error risk. We suggest, however, that in relation to ESTs, the level 1 evidence base has both been inappropriately constructed and analysed – and therefore misconstrued - before being reified in influential treatment guidelines. If reviewers (and meta-analytic researchers) exerted quality control by excluding inert control interventions (and only included psychotherapies with comparable non-specific ingredients), such a misconstruction would not have emerged. If it is to be respected, evidence-based psychiatry must transcend evidence that is filtered or shaped by the eye of the beholder or advocate.

Thus, we are not contesting that CBT and IPT are effective, but reprising an old refrain – that when tested as a 'universal' therapy for depression, their efficacy is likely to be comparable with other 'appropriate' psychotherapies, reflecting the dominant contribution of shared non-specific therapeutic factors. We acknowledge that CBT and IPT may nevertheless be specifically effective in some circumstances, but suggest that evidence of such specificity still needs to be demonstrated – presumably in more context-specific clinical syndromes that respect the intrinsic logic of the specific therapy.

The 'non-specific' factors integral to psychotherapy (and good medicine) currently evoke little interest from researchers. Stravynski and Greenberg (46) suggested that all models of psychotherapy may be 'equally unsound scientifically but they energize the therapists and provide useful fictions to activate the patients to lead somewhat more satisfactory lives'. This is an unnecessarily cynical and nihilistic view. All psychotherapies are and should be a mix of specific and non-specific therapeutic components, with the former providing the technical rationale, but each benefits from close and comparative evaluation. As Wampold (19) noted, 'The specific ingredients are necessary, but their status as the critical aspect is unwarranted' (p. 196).

Past criticisms about the ineffectiveness and non-specificity of psychotherapy have probably contributed to the ESTs being better marketed than measured. To obtain or regain credibility, there were theoretical advantages to imposing a definable structure and undertaking controlled scientific studies – if psychotherapy was to approach the credibility given to biologically weighted treatments. Perhaps as a consequence, the field sought structured psychotherapies that could be elevated to a status akin to drug treatments, with an evidence-based blessing. Regrettably, too much faith has been invested in pursuing a 'specificity' model rather than embracing a more pluralistic model.

Discussion

The baby does not need to be thrown out with the bathwater. The clinical rationale of CBT and IPT seemingly attracts more patients than being offered a set of generic 'therapy' sessions. Both CBT and IPT have been progressively advanced over several decades, with theoretical expositions, manuals and evaluation in numerous studies. The developers deserve credit for addressing the criticisms of Eysenck and others half a century ago, in seeking to introduce therapies with a theoretical rationale and in articulating specific technical (as against non-specific) components.

But, until proponents of CBT and IPT can reject the null hypothesis that treatments intended to be therapeutic are equally efficacious, there is no high level evidence base suggesting or identifying CBT and IPT as distinctively superior than other psychotherapies for major depression. Claiming that they are 'evidence based', implicating or stating their superiority, and formalizing such differential status in therapeutic guidelines is disingenuous at best and risks distorting the value of an evidence-based approach to Psychiatry. The clinical implications of over-selling CBT and IPT as specific and superior treatments for depression impact on both patients and practitioners. We need then to interpret 'the evidence' more critically.

Acknowledgements

Preparation of this paper was supported by an NHMRC Program Grant (222708) and an Infrastructure Grant from the Centre for Mental Health, NSW Department of Health.

References

1. Eysenck HJ. The effects of psychotherapy: an evaluation. J Consult Psychol 1952;**16**:319–324.
2. Melander H, Ahlqvist-Rastad J, Meijer G, Beerman B. Evidence b(i)ased medicine – selective reporting from studies sponsored by pharmaceutical industry: review of studies in new drug applications. BMJ 2003;**326**: 1171–1173.
3. Frank J. Common features of psychotherapy. Aust N Z J Psychiatry 1972;**6**:34–40.
4. Smith ML, Glass GV. Meta-analysis of psychotherapy outcome studies. Am Psychol 1977;**32**:752–760.
5. Smith ML, Glass GV, Miller TI. The benefits of psychotherapy. Baltimore, MD: Johns Hopkins University Press, 1980.
6. Luborsky L, Singer B, Luborsky L. Comparative studies of psychotherapies: is it true that "everyone has won and all must have prizes"? Arch Gen Psychiatry 1975;**32**: 995–1008.
7. Dobson KS. A meta-analysis of the efficacy of cognitive therapy for depression. J Consult Clin Psychol 1989;**57**:414–419.
8. Tarrier N. Commentary: yes, cognitive behaviour therapy may well be all you need. BMJ 2002;**324**:291–292.
9. Neighbour R. Commentary: benevolent skepticism is just what the doctor ordered. BMJ 2002;**324**:290–291.
10. National Institute for Clinical Excellence. Depression: management of depression in primary and secondary care. Clinical Guideline No. 23, London: Nice, 2004:23.
11. American Psychiatric Association. Practice Guidelines for the treatment of patients with major depressive disorder (revision). Am J Psychiatry 2000;**157**(Suppl.):1–45 (p11).
12. Kennedy SH, Lam RW, Morris B, Canadian Network for Mood and Anxiety Treatments (CANMAT) Depression Work Group. Clinical guidelines for depressive disorders. Summary of Recommendations relevant to family physicians. Can Fam Physician 2003;**49**:489–491 (p489).
13. Gloaguen V, Cottraux J, Cucherat M, Blackburn IM. A meta-analysis of the effects of cognitive therapy in depressed patients. J Affect Disord 1998;**49**:59–72.
14. Wampold BE, Minami T, Baskin TW, Tierney SC. A meta-(re)analysis of the effects of cognitive therapy versus 'other therapies' for depression. J Affect Disord 2002;**68**:159–165.
15. de Mello MF, de Jesus Mari J, Bacaltchuk J, Verdeli H, Neugebauer R. A systematic review of research findings on the efficacy of interpersonal therapy for depressive disorders. Eur Arch Psychiatry Clin Neurosci 2005;**225**:75–82.
16. Parker G, Roy K, Eyers K. Cognitive behaviour therapy for depression? Choose horses for courses. Am J Psychiatry 2003;**160**:825–834.
17. Parker G, Parker I, Brotchie H, Stuart S. Interpersonal psychotherapy for depression? The need to define its ecological niche. J Affect Disord 2006;**95**:1–11.
18. Westen D, Novotny CM, Thompson-Brenner H. The empirical status of empirically supported psychotherapies: assumptions, findings, and reporting in controlled clinical trials. Psychol Bull 2004;**130**:631–663.
19. Wampold BE. Establishing specificity in psychotherapy scientifically: design and evidence issues. Clin Psychol Sci Pract 2005;**12**:194–197.
20. Wampold BE. Methodological problems in identifying efficacious psychotherapies. Psychother Res 1997;**7**:21–43.
21. Westen D, Bradley R. Empirically-supported complexity. Re-thinking evidence-based practice in psychotherapy. Curr Dir Psychol Sci 2005;**14**:266–271.
22. Wampold BE, Mondin GW, Moody M, Stich F, Benson K, Ahn HN. A meta-analysis of outcome studies comparing bona fide psychotherapies: empirically, "all must have prizes". Psychol Bull 1997;**122**:203–215.
23. Westen D, Novotny CM, Thompson-Brenner H. EBP ≠ EST: reply to Crits-Cristoph et al. (2005) and Weisz et al. (2005). Psychol Bull 2005;**131**:427–433.
24. Baskin TW, Tierney SC, Minami T, Wampold BE. Establishing specificity in psychotherapy: a meta-analysis of structural equivalence of placebo controls. J Consult Clin Psychol 2003;**71**:973–979.
25. Wampold BE. The great psychotherapy debate: models, methods, and findings. Mahwah, NJ: Erlbaum, 2001.
26. Jacobsen NS, Dobson KS, Truax PA et al. A component analysis of cognitive-behavioral treatment for depression. J Consult Clin Psychol 1996;**64**:295–304.
27. Ilardi SS, Craighead WE. The role of non-specific factors in cognitive-behavioral therapy for depression. Clin Psychol Sci Pract 1994;**1**:138–156.
28. Beck AT, Rush AJ, Shaw BF, Emery BG. Cognitive therapy for depression. New York: Wiley, 1979.
29. Lambert MJ. Early response in psychotherapy: further evidence for the importance of common factors rather than "placebo effects". J Clin Psychol 2005;**61**:855–869.
30. Elkin I, Shea MT, Watkins JT et al. National Institute of Mental Health Treatment of Depression Collaborative Research Program: general effectiveness of treatments. Arch Gen Psychiatry 1989;**46**:971–982.
31. Sotsky SM, Glass DR, Shea MT et al. Patient Predictors of Response to Psychotherapy and Pharmacotherapy: findings in the NIMH Treatment of Depression Collaborative Research Program. Am J Psychiatry 1991; **148**:997–1008.
32. Imber SD, Pilkonis PA, Sotsky SM et al. Mode-specific effects among three treatments for depression. J Consult Clin Psychol 1990;**58**:352–359.
33. Ablon JS, Jones EE. Validity of controlled clinical trials of psychotherapy: findings from the NIMH Treatment of Depression Collaborative Research Program. Am J Psychiatry 2002;**159**:775–783.

34. Ahn H, Wampold BE. Where oh where are the specific ingredients: a meta-analysis of component studies in counselling and psychotherapy. J Counsel Psychol 2001;**48**:251–257.
35. Lambert MJ. Implications of outcome research for psychotherapy integration. In: Norcross JC, Goldstein MR, eds. Handbook of psychotherapy integration. New York: Basic Books, 1992: 94–129.
36. DeRubeis RJ, Brotman MA, Gibbons CJ. A conceptual and methodological analysis of the nonspecifics argument. Clin Psychol Sci Pract 2005;**12**:174–183.
37. Krupnick JL, Sotsky SM, Simmens S et al. The role of the therapeutic alliance in psychotherapy and pharmacotherapy outcome: findings in the National Institute of Mental Health Treatment of Depression Collaborative Research Program. J Consult Clin Psychol 1996;**64**:532–539.
38. Burns DD, Nolen-Hoeksema S. Therapeutic empathy and recovery from depression in cognitive-behavioral therapy: a structural equation model. J Consult Clin Psychol 1992;**60**:441–449.
39. Kim D, Wampold BE, Bolt DM. Therapist effects in psychotherapy: a random-effects modelling of the National Institute of Mental Health Treatment of Depression Collaborative Research Program data. Psychother Res 2006;**16**:161–172.
40. Luborsky LB, Barrett MS. What else materially influences what is represented and published as evidence?. In: Norcross JC, Beutler LE, Levant RF, eds. Evidence-based practices in mental health: debate and dialogue on the fundamental questions. Washington, DC: American Psychological Association Press, 2005:257–267.
41. Robinson LA, Berman JS, Neimeyer RA. Psychotherapy for the treatment of depression: a comprehensive review of controlled outcome research. PsycholBull 1990;**108**:30–49.
42. Luborsky L, Diguer L, Seligman DA et al. The researcher's own therapy allegiances: a "wild card" in comparisons of treatment efficacy. Clin Psychol Sci Pract 1999;**6**:95–106.
43. Rosenthal R. Meta-analytic procedures for social research. Newbury Park, CA: Sage, 1991.
44. Parker G, Anderson IM, Haddad P. Clinical trials of antidepressant medications are producing meaningless results. BMJ 2003;**183**:102–104.
45. Paul GL. Outcome research in psychotherapy. J Consult Psychol 1967;**31**:109–118.
46. Stravynski A, Greenberg D. The psychological management of depression. Acta Psychiatr Scand 1992;**85**:407–414.