Confounding placebo response in clinical trials: Analysing the little we know and current strategies to address the crisis

The crisis caused by increasing placebo response rates in efficacy studies in psychiatry has forced academics and drug companies to relook the science of placebo effects. Clinical response resulting from nonspecific factors (placebo effect) related to the conducting of clinical trials has grown significantly, with effect sizes of placebo groups more than doubling between 1980 and 2005.[1] For some reasons, poorly understood at this point, patients included in randomised controlled trials (RCTs) have become increasingly responsive to nonspecific therapeutic factors inherent in the clinical context in which trials are conducted. In depression studies a stage has now been reached where drug-placebo separation has become very difficult to achieve. Surprisingly, rigorous research of placebo effects has only gained momentum in the past decade.

Naturalistic studies such as CATIE,[2] STAR-D[3] and BALANCE[4] yielded important information about the comparative effectiveness of currently available treatments. However, these studies also clearly demonstrated a need for more efficacious drugs. Despite the recent introduction of comparative effectiveness methodologies, RCTs, at least for the time being, are still required for registration purposes and are still regarded as the highest level of evidence of efficacy in psychiatry research.

Unfortunately, the increasing number of failed RCTs in psychotropic drug development during the past decade, in combination with the recent recession, has led some companies to re-allocate resources to other less risky and less costly areas of drug discovery and development.[5]

Despite major advances in our understanding of mechanisms involved in psychiatric disorders, central nervous system (CNS) drug development has unfortunately not delivered the anticipated dramatic new developments over the past two or three decades. Intensive efforts to find more efficacious drugs with novel mechanisms of action notwithstanding, the rate of real innovation in terms of mechanism of action in all major drug classes has been very slow. The problem of failed studies is not new, but has been growing in magnitude for decades. A high placebo effect may be beneficial in clinical settings, but is one of the major causes of study failure and remains a fundamental issue that needs to be dealt with to ensure the future of drug development.[6]

What is the placebo response?

What is generally perceived as the placebo response is not necessarily response to placebo. Until recently (and still so in efficacy studies) placebo has been defined by its inert content and use as control in clinical trials. Placebo response has also been simplistically defined as the quantifiable improvement of symptoms in subjects in the placebo control group. In reality, placebo effect is the improvement resulting from a number of nonspecific factors (which may include the use of a placebo or not) related to the conducting of the trial.[7] It is the benefit to the patient accruing from simulated treatment and the surrounding clinical context. Therefore, placebo effect may account for a percentage of improvement observed in patients on antidepressants, both in research and in clinical settings.

From recent functional imaging studies, it becomes clear that placebo effects are genuine psychobiological events occurring in responders treated with either active drug or placebo. Mayberg *et al*.[8] demonstrated that the pattern of metabolic change observed in responders, regardless of treatment modality (antidepressant or placebo), overlapped to a large degree. Response in both groups was associated with an increased glucose metabolic rate in cortical brain regions (prefrontal, anterior cingulate, premotor, parietal) and decreased limbic-paralimbic (subgenual cingulate, para-hippocampus, thalamus) metabolic rates. Patients who responded to the active drugs, however, had additional decreased metabolic rates in subcortical and limbic regions (brainstem, striatum, anterior insula and hippocampus).

The increased placebo effect

Placebo response rates have increased from approximately 20% in the 1980s, to about 35% during the 1990s, to the current level of between 45% and 48% in depression studies.[6] After three recent important meta-analyses of efficacy of antidepressants in RCTs, the authors suggested that the overall advantage of antidepressants over placebo is trivial.[9-11] The press then widely reported that antidepressants do not work well and should therefore not be used. What was not conveyed was the fact that the incremental effect of antidepressants relative to placebos increases as the severity of depression increases. The effect of antidepressants

was robust in the severely ill patients (Hamilton Depression Rating Scale >23). The fact that antidepressants performed numerically better than placebo in 75% of the studies analysed by Kirsch *et al.*,[10] while placebo outperformed antidepressants numerically in only one study, was also not highlighted clearly enough.[6] One should also be circumspect about findings from the meta-analysis by Fournier *et al.*,[11] where only 6 out of 2 000 RCTs were included for analysis.[12]

If the prevalence of depression is taken into account, the 10 - 15% average drug-placebo difference in modern studies is significant. One should also remember that the studies included in these meta-analyses were not done in real-world situations. The significant placebo effects reported in RCTs have not been shown to be of similar magnitude in everyday clinical settings and therefore could not be generalised to real-world situations. Efficacy studies for other treatment modalities in depression (i.e. psychotherapy, repetitive transcranial magnetic stimulation, etc.) have exhibited similar difficulties in separating specific from nonspecific factors that determine treatment outcomes.

Variables affecting placebo response

Expectations, trial design and probability of placebo allocation

Clear evidence from laboratory studies has demonstrated the potent role of patient or clinician expectations in either poor or good clinical outcomes.[13] Expectancy by either patient or doctor is a major determinant of placebo effect, confirmed by many scientists working in various fields of medicine. The effect was even present when the clinician's negative expectations were not verbally communicated to the patient. The powerful effect of expectancy is further supported by frequent improvement of patients on suboptimal doses, dramatic early response, adverse effects experienced on placebo, or the huge variation in prescription habits of doctors driven by personal preference. The serious influence of expectancy also has major implications for patients being switched to generic drugs, about which patients and doctors often have strong personal preferences. For this reason, it is generally unwise to overrule a patient's stated preference for a non-generic, because the generic may prove ineffective or not be tolerated. More frequent administration, colour and size of placebo and familiarity with the brand all increase the likelihood of response to placebo. Unfortunately, placebo research often features a trickster quality which, for ethical reasons, is not duplicated in psychiatric settings.

A lower probability of placebo allocation (i.e. increased likelihood of being on active drug) has been shown to increase placebo response rates.[14] The mechanism for this is not clear, but it may act by influencing expectations of both patients and investigators or by influencing the rating style of investigators. The opposite is also true: an increased allocation of patients on placebo (i.e. a decreased probability of receiving active drug) decreases the placebo response rate. Simplicity of study design, limitation of number of study arms and limited assessment time are important strategies to limit sensitivity to nonspecific factors that could affect response.[15]

Importance of baseline ratings and quality of interview and raters

Appropriate patient selection is possibly the most important factor in CNS trial success.[6,15] Avoiding pitfalls relating to patient selection is regarded as the key to solving increased placebo response rates. A considerable body of research has implicated inappropriate patient selection as a major contributing factor to study failure. Strategies aimed at more careful selection of suitable patients are currently being investigated.[16]

The debate on the use of centralised (independent) versus on-site raters has not been resolved, as study findings have been inconsistent and limited in number. From the studies comparing the utility of central versus site-attached raters, it is only clear that central and on-site raters rate differently. Differences were pronounced during the early study phases (baseline ratings), but ratings converged progressively as the studies advanced, achieving very low inter-rater variability at study endpoints.[17,18] At present it is not clear who is more accurate or what could be driving the initial pronounced inter-rater variability at baseline, as neither the central raters nor the raters attached to the sites consistently rated higher or lower across different studies. Baseline score inflation by raters attached to the sites was therefore not found to be a confounding factor across the limited number of studies conducted. Significantly, the placebo response rates were lowest where the two groups had similar baseline ratings.

It is likely that in future additional centralised raters will be employed during early study phases. At this point cross-checking (by an independent rater) of subjects included in trials is seen as a probable solution. Inclusion of a run-in phase has also been implemented to avoid the problem of excessive baseline rating variability. Some current studies have also disconnected the inclusion process from response rating. Efficacy raters are not involved in selection of patients for inclusion in these studies.

Operational oversight with hands-on quality control should be a priority rather than the current emphasis on high turnover, often enforced by unreasonable timelines.

Poor raters or situations where rating skills do not match interviewing skills may also contribute towards increased placebo responses.[15] It is not clear whether placebo response rates for specific raters remain consistent across studies, as these data are often analysed in private and not shared among different sponsors. There seems to be conflicting evidence in this regard, with some investigators consistently showing lower placebo response rates and others varying response rates between studies.[19] Some experts are strongly proposing increased monitoring by the sponsor in the clinical setting of the rating process for single patients.

Inadequacy of current diagnostic categories

The current Diagnostic and Statistical Manual of Mental Disorders, 4th edition (DSM-IV) diagnostic categories have been disappointing and may be too broad for application in clinical research.[16] STAR-D, for instance, demonstrated a considerable difference in treatment response at level 2 in anxious and non-anxious depressed patients.[20] At present, studies also do not take factors such as improvement in patient environment (resolution of stress factors) into consideration when formulating inclusion criteria or when analysing data. The current emphasis on quantitative assessments also does not allow for exclusion of patients with understandable symptoms.[20]

The upside of placebo effects

Laboratory and clinical research creates the possibility of ethical use of placebo mechanisms during routine clinical care, and encourages the use of treatments that stimulate placebo effects.[7]

Conclusion

Despite considerable efforts, CNS drug development has been extraordinarily slow in all major classes. Drugs that work by essentially the same mechanisms of action as those developed in the 1950s and 1960s are still being used, and more efficacious treatments may not be found if the problem of confounding placebo effects in RCTs is not resolved. It is clear that the multiple dimensions of placebo effect have until recently been overly simplified to fit in with traditional research methodologies required for registration purposes. Until regulatory bodies are convinced to accept alternative methodologies, the 'placebo problem' should be analysed and solutions found to minimise response to nonspecific therapeutic factors. Funding bodies need to be convinced to sponsor research into alternative research methodologies in CNS investigations. The current body of available research shows that possible solutions do exist, but it is still too early to tell whether these strategies will yield meaningful results. In the meantime, clinicians should embrace nonspecifics contributing to treatment response, as this ensures improvement in many patients.

Paul J Pretorius
Lucas M van der Merwe
Ian B Westmore

Department of Psychiatry
University of the Free State
Bloemfontein

1. Rief W, Nestoriuc Y, Weiss S, Welzel E, Barsky AJ, Leucht S. Meta-analysis of placebo response in antidepressant trials. J Affect Disord 2009;118:1-8.
2. Lieberman JA, Stroup TS, Mc Evoy JP, *et al*. Effectiveness of antipsychotic drugs in patients with chronic schizophrenia. N Engl J Med 2005;353(12):1209-1214.
3. Trivedi MH, Rush JA, Wisniewski SR, *et al*. Evaluation of outcomes with citalopram for depression using measurement-based care in STAR-D: Implications for clinical practice. Am J Psychiatry 2006;163:28-40.
4. Geddes JR, Goodwin GM, Rendell J, *et al*. Lithium plus valproate combination therapy versus monotherapy for relapse prevention in bipolar I disorder (BALANCE): a randomised open-label trial. The Lancet Online, 23 December 2009. http://www.thelancet.com (accessed 7 July 2010).
5. Lieberman JA. Psychiatric drug development: Why so slow? http://www.medscape.com/viewarticle/721671?src=mp&spon=12&auc=6217PX (accessed 31 May 2010).
6. Thase ME. Unmet needs in antidepressant therapy: What is the current state of the field? Symposium 009, 163rd Annual Meeting of the American Psychiatric Association, New Orleans, 22 - 26 May 2010. http://www.sessions2view.com/apa_library_2010/viewer.php (accessed 7 July 2010).
7. Finniss DG, Kaptchuk TJ, Miller F, Benedetti F. Biological, clinical, and ethical advances in placebo effects. Lancet 2010;375:686-695.
8. Mayberg HS, Artura Silva J, Brannan SK, *et al*. The functional neuroanatomy of the placebo effect. Am J Psychiatry 2002;159(5):728-737.
9. Walsh BT, Seidman SN, Sysko R, Gould M. Placebo response in studies of major depression. JAMA 2002;287:1840-1847.
10. Kirsch I, Deacon BJ, Huedo-Medina TB, Scoboria A, Moore TJ, Johnson BT. Initial severity and antidepressant benefits: a meta-analysis of data submitted to the Food and Drug Administration. PLoS Med 2008;5(2):e45.
11. Fournier JC, De Rubeis RJ, Hollon SD, *et al*. Antidepressant drug effects and depression severity: A patient-level meta-analysis. JAMA 2010;303(1):47-53.
12. Trivedi MH. Effectiveness of treatment strategies for depression. Presidential Symposium 2, 163rd Annual Meeting of the American Psychiatric Association, New Orleans, 22 - 26 May 2010. http://www.sessions2view.com/apa_library_2010/viewer.php (accessed 7 July 2010).
13. Kradin R. The Placebo Response and the Power of Unconscious Healing. New York: Routledge Taylor & Francis Group, 2008.
14. Papakostis GI, Fava M. Does the probability of receiving placebo influence clinical trial outcome? A meta-regression of double-blind randomized clinical trials in MDD. European Neuropsychopharmacology 2009;19:34-40.
15. Szegedi A. Placebo response is not necessarily response to placebo: Factors inflating placebo response in CNS trials. Symposium 44, 163rd Annual Meeting of the American Psychiatric Association, New Orleans, 22 - 26 May 2010.
16. Targum SD, Pollack MH, Fava M. Redefining affective disorders: relevance for drug development. CNS Neuroscience and Therapeutics 2008;14(1):2-9.
17. Dunn J. Comparison of site-based versus centralized ratings in a study of generalized anxiety disorder. Symposium 44, 163rd Annual Meeting of the American Psychiatric Association, New Orleans, 22 - 26 May 2010.
18. Targum SD. Evaluation of centralized ratings in a clinical trial of major depressive disorder. Symposium 44, 163rd Annual Meeting of the American Psychiatric Association, New Orleans, 22 - 26 May 2010.
19. Kremer C. Does de-coupling of entry criteria improve outcomes in CNS trials? Symposium 44, 163rd Annual Meeting of the American Psychiatric Association, New Orleans, 22 - 26 May 2010.
20. Fava M. The use of independent assessments for patient eligibility for CNS trials. Symposium 44, 163rd Annual Meeting of the American Psychiatric Association, New Orleans, 22 - 26 May 2010.