Key Words: depression, antidepressants, clinical trials, clinical practice, placebo, placebo response

Placebo Response in Depression: A Perspective for Clinical Practice

By Arif Khan, Shirin Khan

ABSTRACT *~ Practicing clinicians appreciate that depression is not an easy disorder to treat and manage. Despite the plethora of new treatments—both pharmacological and non pharmacological—that has flooded the market in the past years, we are still nowhere close to obtaining full symptom relief for all patients and eradicating the morbidity and mortality associated with depression.*

In this context, recent methodological research, concentrating on the effectiveness of antidepressants has raised doubts about their therapeutic index. Because of obtuseness of the methodology and biased interpretations, we are submitting this perspective to clinicians so that they can appreciate some of the deficits of the recent research publications. For the practicing clinician, the best available data suggest that clinically depressed patients warrant treatment and the most robust available body of data (published and unpublished) would favor the use of antidepressants. Psychopharmacology Bulletin. *2008;41(3):91–98.*

INTRODUCTION

Depression is a chronic and fluctuating disorder. Symptoms and signs of depression, along with level of psychological distress and level of functioning, for most patients wax and wane over time. There is not a single profile of symptoms and signs, as evidenced by changing definition of depression by the American Psychiatric Association (Diagnostic and Statistical Manual of Mental Disorders versions II, III, IV) over the past fifty years.

There is no unifying concept on the etiology and pathophysiology of depression, or the acute treatment and chronic management of depressed patients. Experienced clinicians know that treating depression, albeit a simple disorder is not just reaching for the prescription pad and automatically knowing the best drug/treatment for each patient. Although there is a plethora of physical (Electric Convulsive Therapy (ECT) and Vagus Nerve Stimulation (VNS)), pharmacological (32 antidepressants in the United States) and psychological treatments (cognitive behavioral therapy, interpersonal therapy) for depression, managing each individual patient requires patience, humility, and wisdom (Trivedi et al.[1]).

Khan A, MD, Northwest Clinical Research Center, Bellevue, Washington, Department of Psychiatry and Behavioral Sciences, Duke University Medical Center, Durham, North Carolina. Khan S, BS, Northwest Clinical Research Center, Bellevue, Washington.

To whom correspondence should be addressed: Arif Khan, MD, Northwest Clinical Research Center, 1900—116th Ave NE, Bellevue, WA 98004, Phone: 425-453-0404; Fax: 425-453-1033; Email: akhan@nwcrc.net

Given this background, with the introduction of over a dozen new antidepressants in the past two decades, there has been a sense of hope and optimism in managing depressed patients. However, research in the past decade has shown that these newer agents are not necessarily better than earlier ones, but they however are easier to use, have fewer side effects, and are generally safer if taken in overdose.

In this context, two recent publications (Turner et al.[2] and Kirsch et al.[3]) with high media coverage have raised further concerns about the usefulness and the therapeutic index of our current antidepressants. Much of the data that they presented are well known to research psychiatrists, but may not have filtered to practicing clinicians.

In this perspective, we aim to provide a background of developmental hurdles for depression treatments, the nature of placebo response, review recent publications in the context of clinical trial research, and lastly highlight the role of placebo response for non-specific treatment factors in clinical practice.

CLINICAL TRIAL RESEARCH IN DEPRESSION

In 1939, it was easy to see that ECT could have dramatic effects on psychopathology. Although similar magnitudes of effects were first observed with insulin coma for schizophrenia, this treatment would not withstand scientific scrutiny and went away. The early imipramine trials (Ball et al,[4] and Kiloh et al[5]) reported if patients "had a good or worthwhile result or a poor result," from imipramine or placebo treatment, similar to the Clinical Global Impressions rating scale.

However, most of medical treatments, especially for those chronic illnesses that do not neatly fulfill Koch's principles, came under scrutiny in part due to the seminal paper by Beecher[6] in 1955. Further, the US Food and Drug Administration (FDA) came with its own in the 1960's in conjunction with scientists and clinicians and developed criteria to evaluate treatment efficacy (remember not effectiveness) and safety of new treatments, certified under Code of Federal Regulations 312.[7]

Basically, the primary role of the FDA is to ensure that a placebo could not be potentially sold to the public and that there are reasonable disclosures of known adverse events and safety concerns. Over the past few decades, this has translated into the requirement that to approve a new depression treatment, a well designed and conducted, multi-center, randomized, double blind, and placebo controlled clinical trial needs to show statistical superiority for the treatment compared to placebo. Further, the sponsoring company/agency needs to replicate the findings of the first trial, using a similarly designed clinical trial.

At first glance, this seems a simple and somewhat silly requirement. However, time has shown that this hurdle has not been overcome by at least 25 potential new antidepressants in the past two decades. In this context, it is important to note that there have been no applications to the FDA for any form of psychotherapy and applications with VNS or repetitive transcranial magnetic stimulation (rTMS) have met with mixed success. As a Principal Investigator of nearly one hundred antidepressant clinical trials, conducting this sort of trial is similar to being in a night time race with no lights (neither street lights nor your own lights), know that there are unmarked slippery oily patches, you are without a road map and you do not know where your main competition (placebo) is at.

Why is it so difficult to overcome placebo effect in clinical trials? Although no specific and coherent answer is available, for now we know that there are several factors that may be at play. First and foremost, depression is an insidious onset illness and can fluctuate from day to day. It is difficult to know what the universal symptoms are and how to measure them. A considerable number of patients either exaggerate or minimize their symptoms confounding researchers. This has been elegantly demonstrated by Greist et al.[8] showing that clinical researchers note a profile of symptoms signs and when patients are asked via telephone, the profile of symptoms and signs provided by the same patients does not match.

Further, we know that many factors can make a depressed person get well. For many, the arrival of spring is a boon, others a new job, others a new relationship. Paradoxically, many of these same factors can make symptoms worse. More important, a significant number of depressed patients come to treatment as they are starting to recover from their episode. Additionally, repeat measurements lead to a shift to the mean as described by Francis Galton in 1859.

In fact, Frank and Frank[9] have encapsulated placebo response as follows; 'a placebo pill, though pharmacologically inert, has great symbolic value and power as a conditioned stimulus. Placebo treated patients receive all the components of the treatment situation common to any treatment including a thorough evaluation, an explanation for distress, an expert healer, a plausible treatment, an expectation of improvement, an opportunity to verbalize their distress, as well as a healer's commitment, enthusiasm, and positive regard'.

WHAT DO EUROPEANS DO?

On the other hand, the focus of European regulators has been not so much on how well antidepressants show up against placebo in a head to head comparison, but rather show how therapeutic effects of new

antidepressants last over time Fortunately, this model makes antidepressants look better than they do in head to head comparisons (Geddes et al.[10]). So, in fact, if a new antidepressant beats placebo head on acutely, then this type of study actually helps evaluate the actual therapeutic index.

CAN WE TAME PLACEBO RESPONSE?

Given the fact that we understand few facts about placebo response, there has been no dearth of efforts or experts who have attempted to tame placebo response. Paradoxically, considerable effort was made to enhance the placebo effect as severe side effects of drugs such as amitriptyline could unblind the observers among clinical trials in the 1960s. Benzatropine, as well as some other hypnotics were used to "mask" the side effects of these antidepressants and thus prevent unblinding of raters.

The more common effort in recent years is to minimize placebo response. This effort has been bolstered by the data that response to antidepressants is relatively the same (40–50%) in over 100 antidepressant clinical trials, whereas placebo response can fluctuate from 10–50% (Khan et al.[11]). Simply put, if the magnitude of symptom reduction is more than 30% with placebo, the chances of the antidepressant being statistically superior to placebo are only 20%.

Although post-hoc analysis has shown that several factors such as severity of depressive symptoms at the time of entry into the trial (Khan et al.[12]), use of flexible dose schedule (Khan et al.[13]), which in part is linked to number of treatments arms, duration of depressive episode and sex distribution (Khan et al.[14]), in the sample may be related to the success of antidepressant trials, so far nobody has been able to translate it to be used prospectively.

The most vocal and active groups have focused on increasing reliability in measuring symptoms of depression ranging from better training of so called 'rater' by specially developed companies, use of phone to assess depression symptoms and signs, and centralized video tapping to name a few. Sad to say, as of now, there are no convincing data that these rituals can actually reduce placebo response.

Another technique used by many of the pharmaceutical companies and their staff is to cherry pick depressed patients by having 35–40 inclusion/exclusion criteria in the misguided belief that such selection will reduce placebo response. Such practices have yet to bear fruit.

NEW CONTROVERSIAL PUBLICATIONS

Erick Turner, MD was a FDA physician who evaluated new antidepressant applications before he moved to Portland to join the medical school faculty. In his position at the FDA, he could see that the data analyzed by

the FDA scientists did not exactly match what was published later. So, over the past several years, he put his efforts to measure the publication bias (Turner et al.[2]). Not surprisingly, there is a publication bias. This should be taken in the context that until recently, there has been no interest in academic circles and medical journals to publish results of failed trials, let alone any pharmaceutical company trials if not glowingly positive.

Interestingly, the publication bias is not restricted to pharmaceutical companies. Dr. Turner has provided us with excellent references of publication bias as seen in grantees of NIH and Canadian government (Dickersin Min,[15] Chan et al.[16]). So, publication bias is not limited to commercial enterprises alone.

Lastly, it is foolish to believe that pharmaceutical companies are anything other than commercial endeavors. Most people would consider it silly and not enforceable if FDA insisted that supermarkets should put the worst of tomatoes on top and not at the bottom of their display shelves/baskets. In this regard, caveat emptor (let the buyer beware) is the only reasonable fact that should not be forgotten.

On the other hand, Dr. Kirsch and his colleagues[3] have a long history of attempting to persuade physicians and scientists that antidepressants do not work. Part of this is may be related to their professional orientation, i.e., these researchers are psychologists, with a presumed belief in effectiveness of psychotherapy, regardless of available data.

Following are Specific Weaknesses in Kirsch et al.[3] Publication

First, Kirsch et al.[3] imply they are the first group of investigators to show such analysis and results. This is not true. We conducted identical analysis on a much larger sample of antidepressant clinical trials in 2002 (Khan et al.[12]) and found almost the same results. Interestingly, our analysis and report were focused on the magnitude of placebo response and factors associated with it. Further, they avoided analyzing a much larger set that is available now compared to 2002, for no apparent reason.

Secondly, they *imply,* never state, that placebo may be the treatment of choice for their ill-defined 'mildly depressed patients'. They make the statement that 'given these data, there seems little evidence to support the prescription of antidepressant medications to any but the most severely depressed patients, unless alternative treatments have failed to provide benefit'.

Many including some of the British Government agencies have taken this to mean that psychotherapy should become the treatment of choice for the 'mildly depressed patients', albeit data to support this conclusion do not exist and the authors cleverly avoid this point in their manuscript.

Third, Kirsch et al.[3] in their manuscript note that antidepressant's effects are unrelated to severity of symptoms. However, they note a mild

trend for the depressed patients with more symptoms not responding as well as to placebo compared to those depressed patients with fewer symptoms. Their interpretation that antidepressants should not be prescribed except for severely ill (as noted above) is disingenuous. The following analogy will clarify the lack of reasonable logic in their argument.

The tires from company A (antidepressants) are good for both good (less severely ill depressed patients) and bad weather (severely ill depressed patients) conditions. On the other hand, tires from company B (placebo/alternate treatment-psychotherapy) work well in good weather but badly in bad weather. Hence, it is not surprising that individuals' holding company B stock would say you should buy their products as weather is rarely that bad (also this is the more expensive product).

Fourth, we have reanalyzed Figure 4 of Kirsch et al.[3] manuscript, using their own paradigm in Table 1. The 'success rate' for showing 3 points or more of difference between antidepressant-placebo effects is very similar regardless of mean baseline total HAM-D scores.

Lastly, severity of depressive symptoms is not the only contributing factor in the outcome of antidepressant clinical trials. There are factors such as age, sex, number of treatment arms, dosing schedule to name a few. Kirsch et al.[3] have not taken any of these factors into account in their analysis and hence this meta-analysis is flawed.

IS PLACEBO RESPONSE A PROBLEM IN DEPRESSION TRIALS ALONE AND HIGH PLACEBO RESPONSE LIMITED TO DEPRESSION?

Besides the data from the pharmaceutical companies, the STAR*D study clearly demonstrated that clinical management of depression is more complicated than prescribing a single antidepressant for the majority of depressed patients (Trivedi et al.[1]). So, diagnosing and managing depression is a combined skill of appreciating pathology, safety profile, and a combination of drug use and all the psychotherapy skills we can muster.

TABLE 1

RE-ANALYSIS OF FIGURE 4 (KIRSCH ET AL.[3]) FOR ANTIDEPRESSANT TRIALS MEETING THE ⩾3 POINTS ON MEAN TOTAL HAM-D SCORES BETWEEN ANTIDEPRESSANTS AND PLACEBO*

	FREQUENCY AND PERCENTAGE RATIO
Trials with mean total baseline severity < 25 HAMD	6/14 (43)
Trials with mean total baseline severity ⩾ 25 HAMD	10/21 (48)

*This analysis is the best estimate from Figure 4 of Kirsch et al.[3]

So, until the day we have discovered the etiology and pathophysiology of depression and develop specific treatments, use of placebo is not likely to go away. Having said this, it is important to note that antidepressant clinical trial results, based on the state of the knowledge are more affected by placebo response than antidepressant response and results of all such trials should be viewed in this context.

All of the data suggest that the effect sizes for current antidepressants are between 0.25–0.30 (mild to moderate). Unless we somehow overcome the deficiency of etiology and pathology, serendipity is the only reasonable alternative for now. Not surprisingly, the effect size we see in depression is similar in other chronic illnesses such as hypertension, panic disorder, seizures, chronic pain, cancer and irritable bowel syndrome to name a few (Khan et al.[17]).

In this context, it is important to note that psychopharmacologists have routinely underpowered their trials, unlike many of our colleagues in other specialties. This lesson has been learnt by some of the pharmaceutical companies and they are appropriately powering their antidepressant trials (Khan et al.[18], Berman et al.[19], Thase et al.[20]).

CONCLUSION

It is important to note that acute head to head comparison of antidepressants to placebo is a difficult test for most depression treatments to pass. In fact, antidepressant trials using the long-term withdrawal model may reflect a more reasonable therapeutic index for depression treatments. However, eliminating the high bar of placebo may allow ineffective depression treatments to become acceptable.

Of the two recent publications, the paper from Dr. Turner and his colleagues reinforce the idea that there is a publication bias, worth paying attention to. On the other hand, the paper from Dr. Kirsch and his colleagues appears to be mired in weak methodology, internal inconsistencies and contradictions in what their analysis shows to what they conclude and finally potential partisan bias. At best, this can be taken as a publication bias of these investigators, as highlighted by Dr. Turner and his colleagues, this being a common problem (caveat emptor). We strongly urge Dr. Kirsch and his colleagues to conduct an analysis evaluating the efficacy/effectiveness of psychotherapy that is not affected by publication bias.

For the practicing clinician, the best available data suggest that clinically depressed patients warrant treatment and the most robust body of data would favor use of antidepressants. We do wish we will have soon have antidepressant with a better risk/benefit ratio.✤

DISCLOSURE

Funding for this manuscript was supported by Northwest Clinical Research Center. This manuscript has not previously been presented.

Dr. Khan, Principal Investigator for over 280 clinical trials sponsored by more than 57 pharmaceutical companies and 23 CROs, has done no consulting or speaking on their behalf.

REFERENCES

1. Trivedi M, Rush J, Wisniewski S 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.
2. Turner E, Matthews A, Linardatos E, Tell R, Rosenthal R. Selective publication of antidepressant trials and its influence on apparent efficacy. *N Engl J Med*. 2008;358(3):252–60.
3. Kirsch I, Deacon BJ, Huedo-Medina TB, et al. Initial severity and antidepressant benefits: A meta-analysis of data submitted to the Food and Drug Administration. *PLoS Med.* 2008;5(2):e45. doi:10.1371/journal. pmed.0050045.
4. Ball J, Kiloh L. A controlled trial of Imipramine in treatment of depressive states. *BMJ*. November 21, 1959:1059–1055.
5. Kiloh, L, Ball J. Depression treated with Imipramine: a follow-up study. *BMJ*. January 21, 1961; 168–171.
6. Beecher HK. The powerful placebo. *JAMA,* 1955 Dec 24;159(17):1602–6.
7. Code of Federal Regulations, Food and Drug Administration, 21 Parts 300 to 499. Available at: http://www.accessdata.fda.gov/scripts/cdrh/cfdocs/cfCFR/CFRSearch.cfm?CFRPart=312&showFR=1 Accessed April 19, 2008.
8. Greist JH, Mundt JC, Kobak L. Factors contributing to failed trials of new agents: can technology prevent some problems? *J Clin Psychiatry*. 2002;63(Suppl 2):8–13.
9. Frank JD, Frank JB. Persuasion and healing: a comparative study of psychotherapy. Baltimore: The John Hopkins Press; 1991.
10. Geddes JR, Carney SM, Davies C, Furukawa TA, Kupfer DJ, Frank E, Goodwin GM. Relapse prevention with antidepressant drug treatment in depressive disorders: a systematic review. *Lancet*. 2003; 361:653–661.
11. Khan A, Detke M, Khan S, Mallenkrodt C. Placebo response and antidepressant clinical trial outcome. *J Nerv Ment Dis*. 2003;19(4):211–218.
12. Khan A, Leventhal R, Khan S, Brown WA. Severity of depression and response to antidepressants and placebo: an analysis of the FDA database. *J Clin Psychopharmacol*. 2002;22(1):40–45.
13. Khan A, Khan SR, Walens G, Kolts R, Giller EL. Frequency of positive studies among fixed and flexible dose antidepressant clinical trials: an analysis of the Food and Drug Administration Summary Basis of Approval reports. *Neuropsychopharmacology*. 2003;28:552–557.
14. Khan A, Brodhead AE, Schwartz KA, Kolts RL, Brown WA. Sex differences in antidepressant response among outpatients participating in antidepressant clinical trials. *J Clin Psychopharmacology*. 2005;25(4):318–324.
15. Dickersin K, Min YI. NIH Clinical Trials and Publication Bias [article]. Online *J Curr Clin Trials*. 1993 Apr 28;1993(Doc No 50).
16. Chan AW, Krleza-Jeric K, Schmid I, Altman D. Outcome reporting bias in randomized trials funded by the Canadian Institutes of Health Research. *CMAJ* 2004;171(7):735–40.
17. Khan A, Bhat A. Placebo response in depression: clinical trials versus clinical practice. (Submitted 2008).
18. Khan A, Schwarts K. Suicide risk and symptom reduction in patients assigned to placebo in duloxetine and escitalopram clinical trials: analysis of the FDA summary basis of approval reports. *Ann Clin Psychiatry*. 2007;19(1):31–36.
19. Berman RM, Marcus RN, Swanink R, McQuade RD, Carson WH, Corey-Lisle PK, Khan A. The efficacy and safety of aripiprazole as adjunctive therapy in major depressive disorder: a multicenter, randomized, double-blind, placebo-controlled study. *J Clin Psychiatry*. 2007;68(6):843–853.
20. Thase ME, Macfadden W, Weisler RH, Chang W, Paulsson B, Khan A, Calabrese J, BOLDER II Study Group. Efficacy of quetiapine monotherapy in bipolar I and II depression: a double-blind, placebo-controlled study (The BOLDER II Study). *J Clin Psychopharmacol*. 2006;26(6):600–609.