Author's response to reviews

Title: The co-administration of quetiapine or placebo to cognitive-behavior therapy in treatment refractory depression; a preliminary trial

Authors:

Yves JA Chaput (yveschaput@bellnet.ca)
Annick Magnan (tobedetermined@bellnet.ca)
Alain Gendron (alaingendron@astrazeneca.ca)

Version: 4 Date: 22 July 2008

Author's response to reviews: see over
To the Editorial Staff,

BMC – Psychiatry,

MS: 8312126711678731

Montreal, July 22, 2008.

Thank you for reviewing, for the third time, my manuscript entitled ‘The co-administration of quetiapine or placebo to cognitive-behavior therapy in treatment refractory depression; a preliminary trial’ that I re-submitted for your consideration in BMC-Psychiatry on the May 2008.

I have endeavored to answer in as much detail as possible the many queries Dr Stewart brought forth in his third revision. They were quite interesting and pertinent and greatly appreciated. I have answered them in the following manner. In bold (usually) is Dr Stewart’s query and in regular type is my response.

Perhaps it is style, but I would have placed the paragraph introducing quetiapine ahead of the one about CBT. Again, because their design cannot address the efficacy of CBT, but can address possible efficacy of quetiapine, why lead with the treatment you cannot address? The manuscript has to introduce what is known about the utility of CBT for TRD. That is not my beef. Rather, it is that CBT is presented, especially in an earlier version of this manuscript as the answer for TRD, when this is not the case and this study does not make a case in that direction. Yes, CBT needs to be introduced to the reader, but not as the main thing the manuscript is about, or, indeed, something the manuscript can say anything definitive about.

This has been revised in the INTRODUCTION. The introduction of quetiapine is now prior to that of CBT.

More importantly, while they state a study aim their paper can address, they make no comment about why they used CBT, rather than just doing a drug vs. placebo study. Certainly, it would be nice to know whether CBT could help in TRD, but since they designed the study so they cannot address that question, what is CBT doing in it? We are not informed.

This comment is factually true given my results. The results being different however this might not be the case. For example, if the placebo/cbt arm had significantly improved with statistical separation from the quetiapine-cbt arm, then these results might suggest
that CBT is doing something and or that quetiapine does not really add anything (or is disruptive?). Indeed, if one analyses only the placebo/CBT patients that completed the trial (N = 5) then both primary rating scale scores and all secondary rating scale scores were show significant reductions with the exception of the quality of life scale, for all 5 patients. Indeed, they show reductions entirely comparable to those in the quetiapine/CBT group. Although I obviously cannot exclude the finding that more patients on placebo did not finish the trial than those on quetiapine and even so, my study design would nevertheless not permit a definitive answer.

Secondly, there was a very interesting point-counterpoint in this month’s Canadian Journal of Psychiatry concerning clinical trials and the placebo issue. The crux of both arguments was the ethical issue of placing patients (even of their own free will) in trials when a known (even if not terribly efficacious) treatment was available. When initially writing this protocol, prior to sponsors coming onboard, it was primarily a CBT trial. CBT has been shown to be efficacious in the treatment of depression in many past studies and as such the treatment was justified. I doubt that (years ago or even today) with the absence of a more traditional treatment like CBT that this protocol would have passed ethics. Especially on the basis of the state of the literature on quetiapine (even today, although my very small contribution might help). In addition, I really do not think I would have felt comfortable or at ease either writing or doing such a protocol. Even within the context of a trial our primary responsibility is to the patient, not the trial and randomizing stage II patients to placebo without any safeguards other that withdrawal would have made every one involved uncomfortable. One would most likely have to recruit stage V patients to justify such a protocol and that would not at all be an easy task, especially with recruitment. Even then, one might say that these are pre psychotic patients and quetiapine might improve them via an antipsychotic effect.

I think that rearranging the paragraphs as Dr Stewart suggested flows better and the reader can appreciate why CBT was administered in this trail. The aims section is now right beneath the abovre paragraph where it is mentioned that ‘This suggests that CBT may be of use in stage II TRD. Such a finding would not only be pertinent in TRD but also, in the treatment of severely depressed patients where medication may pose a health risk, such as during pregnancy or in the aged.’

Questions about prior treatments, refractoriness and stage definitions.

The maximum recommended dose problem. I have replaced ‘maximum recommend dose’ with the more generic ‘therapeutically recommended doses’ in the METHODS section, second paragraph, line 4. It is most likely a more appropriate term, given that the maximum recommended dose for some medications may change over the years.

The source documentation problem. I realize that the source documentation used was primarily available in Canada. The ‘Compendium of Pharmaceuticals and Specialties; the Canadian Drug Reference for heath professionals’ does have a web site, although it is by subscription and, at 225$ a peek, would most likely not at all interest most readers. I got
hold of a circa 2002 edition, which should be quite similar (dosage wise) to the 2001 edition that I used for the protocol.

I took some time to assess the more readily (and widely) available resources on this topic. A recent, natively American and readily available in electronic form resource is the ‘AHFS Dosing Companion’ Version 10.1.2/2007 by the American Society of Health System Pharmacists. The dosing recommendations are very similar to those of the CPS. Dr Stewart mentions ‘PDR’. I am not clear what this is but I googled the letters and came up with Physician’s Desk Reference. There appear to be 2 variants, a free web site and one where one needs to register. The dosing recommendations of the Physician’s Desk Reference via their free web site are quite similar to the AHFS and the CPS. As it is free and Dr Stewart recommends it, I have referenced it in the METHODS section. If this is not what Dr Stewart is referring to then I will eliminate this reference and reference the AHFS.

The three sources recommend;

CPS and AHFS = Effexor 225 mg max for moderate depression, 350 mg for severe depression. PDR = 225 mg max.
CPS and AHFS = Citalopram 40 mg per day max. The CPS does state that 60 mg can be used but does not typically show greater efficacy. PDR recommends a maximum of 40 mg.
CPS and AHFS and PDR = Sertraline 200 mg.
CPS and AHFS and PDR = Paroxetine 50 mg.
CPS and AHFS and PDR = Fluoxetine 80 mg.
CPS and AHFS (not in PDR) = Nefazodone 600 mg.
CPS and AHFS and PDR = Phenelzine 90 mg.
CPS and AHFS and PDR = Mirtazapine 45 mg.
CPS and AHFS = Amitriptyline 300 mg (PDR unclear, possibly 250 mg)

These values have now all been added to the footnote “f ” of table 1, where they are also referenced. As such, I have re written the paragraph in the METHODS section concerning the medication dosage. It now takes into account these values and I have gone back to the data for each patient to select that treatment which most closely followed these recommendations. As can be appreciated, of the 62 separate antidepressant trials (31 patients X 2 treatments) in only 6 was the dose (3 each for paroxetine 40 mg and sertraline 150 mg) below the highest dose recommended by these source documents, and not by much. At this point there should be no ambiguity as to dose.

Could I have gone higher still with each drug, the answer is obviously yes. Anything is possible I could have required 412.5 mg of effexor and no less than 60 mg of citalopram. However, given that there was a lithium augmentation phase (the antidepressant was at the highest dose during this phase) and there is a reasonable expectation of response at the antidepressant doses used (i.e, the STAR*D reference) then I think that my approach was reasonable.

They studied too few patients but ideally they would report whether the more
highly refractory patients were less likely to benefit.

This query was actually asked by Dr Stewart in the first review of the manuscript. The idea is great. I just don’t think I have the data to even begin to answer this question. I think Dr Stewart would agree that given the complexity of establishing a priori what is a ‘refractory depression’, going one step further to ‘highly’ refractory is a stretch. For instance, patient A, with the following pharmacological hx (Paxil 60, effexor 300, mirtazapine 45, citalopram 40, prior lithium augmentation, prior olanzapine augmentation…) was in the quetiapine/cbt arm, went to the end of the study and did poorly. Patient B had (paxil 30, serzone 600, effexor 300, sertraline 100, fluvoxamine 200, citalopram 50 and concomitant buspirone…) was in the placebo/cbt arm, went to end of study and actually did well. Perhaps both patients were fast metabolizers and these ‘po’ doses may not mean all that much.

Only 8 of 31 patients had 2 prior pharmacological trials. The others had 3 or more (the overall average being 3 +/- 1). Looking at it from the opposite point of view, how did those 8 patients with only 2 prior treatments do. Well, quite well actually. Overall, they showed some significant reductions in the primary rating scales. However, given that they were scattered among the three treatment groups (lithium and the two blinded ones) is was very difficult to speculate that number of treatments was correlated with.

Embedded in the Discussion we find that “some patients were at stage V.” This information and the number at each stage and/or number who previously received how many adequate trials belong in the Results.

Agreed, this information is contained within Table 1 of the RESULTS section. However, I have added a mention in the text in the RESULTS section, first paragraph. I have also slightly rephrased the sentence in the DISCUSSION. Thanks for pointing this out to me.

In his comments to me, the author pointed out that the cut-off doses used in his study were equal to or higher than those patients typically responded to in STAR*D. And, there is a similar comment in the text of the manuscript. I do not think this is the way to think about treatment refractoriness. Thus, just because 20 mg of fluoxetine is the commonest dose for people to improve does not mean I think patients who have received 20 mg/d of fluoxetine have had an “adequate trial”. Would Dr. Chaput say to someone unresponsive to 20 mg/d of fluoxetine, “Well, fluoxetine has not worked for you, let’s stop it and move to something else”? I doubt it, or, at least, hope not. At least clinically, Dr. Chaput does not consider 20 mgd of fluoxetine to be an adequate trial. So, the dose most people benefit from seems a poor way to define adequacy.

I fully agree with Dr Stweart’s above statements. I completely reworked the dose data in the METHODS and in Table 1 in the RESULTS section. It should be much clearer now exactly what patients got as a treatment, which were, at a minimum, what most
physicians would consider more than minimal therapeutic doses. I have also slightly reworked the STAR*D reference. I am using it as an adjunct, a reference point, simply stating that the doses that I used, in other people’s hands, have been shown to be adequate.

Statistical analysis still states alludes to patients being “randomized to CBT.” In fact, determination of who received CBT was not by random. Whether patients received quetiapine or placebo is what was randomly determined. Probably a left-over, but still misleading.

Sorry about that. Yes, it was a left-over. I ran some scenarios using the same 5 to 6 point drops in order to get the sample size for both post randomization groups. I used Systat rather than Sample power (I probably lost that CD years ago).

There must be a misprint on p. 13 as both significant secondary efficacy measures are listed as being the CGI-S.

Yes, one was the CGI-I.

 Appropriately, they give the mg/d and range of the patients assigned to quetiapine. However, the standard deviation is missing and the number of pills the patients were taking in the placebo group. They also should show, perhaps in parentheses, the number of pills the active group took with the statistical comparison of number of pills each group took. In my experience, patients taking placebo generally take more than those taking active, both because they are less likely to be getting better (at least if the drug is effective) and because they are less likely to have side effects that might prevent increases. Because this study had such a high drop-out rate in its placebo arm, they may have had the opposite effect, that is, the placebo arm took fewer pills because they never had a chance to increase the dose. Whatever, the reasons for any differences, they should show the data.

I would like to know is how many pills patients actually took. Dr. Chaput states in comments but not in the manuscript that these were identical for the quetiapine/CBT vs. the placebo/CBT subjects. This seems unlikely given the wide range of doses taken by the patients in the active drug group.

This query took quite some time to resolve as I had to go back to the storage bins to take out the study binder and the patient charts to double check all pill data. As stated in the text, the pill counts should be equal. However, I was referring to the average pill count per day for both groups. Pills were taken bid and, with the various doses in the blister packs, most dose augmentations could be achieved by taking 2 pills per day, although there were exceptions.
As such, the pill count per day in the study should have been the same or close to the same. This was actually the case. The placebo group took 2.1 +/- 0.7 pills per day whereas the quetiapine group took 1.9 +/- 0.6 pills per day. The difference is obviously not statistically significant.

Dr Stewart however is right when he states that the overall pill count (the amount of pills returned by all patients in the placebo group vs all patients in the quetiapine group) would differ substantially (much lower in the placebo group) as there were many placebo drop outs. This indeed was the case, although this *a posteriori* (or end of study) finding.

Note that in double checking all pill data I noticed that my coordinator had included a few taper doses into the spreadsheets, that slightly lowered the maximal dose of quetiapine. I recalculated that dose and corrected it in this version of the manuscript (a small but detectable 8 mg increase in the average quetiapine dose). I also calculated the placebo average, that was a bit over 200 mg per day. The difference between average daily doses was not significant.

The paragraph on page 14 in the METHODS section has been rewritten and now includes this data.

I do not object to the term “notable” as I agree that the high placebo drop-out rate certainly seems notable. However, why report this rate without informing the reader whether the proportions differ? Showing that the times in the study differ does not address whether the proportions differ. Also, does the fact that the placebo/CBT group had significantly less time in the study (i.e., less opportunity to improve) affect your interpretation of the statistics. I.e., was it more time or effective drug that got the quetiapine/CBT people better?

Here, unfortunately, I am not quite certain what Dr Stewart is referring to when he states ‘do the proportions differ’. As for the second part of the query, the answer is that it does affect the statistical results. This ‘notable’ a drop in patient numbers in the placebo/CBT group would naturally tend to shift the rating scale averages towards the absence of change. As stated above, when only placebo/CBT completers are examined there is a significant reduction in all rating scale scores except the quality of life scale.

I mention to this in the DISCUSSION section when it is stated that one possibility here is that quetiapine is acting as a non specific sedative helping patients make the transition from Drug to CBT. The absence of this in the placebo group most likely did not help.

I have slightly rearranged the paragraph in the RESULTS section as well as the in the DISCUSSION section.
Rates of various side effects are shown, but we are not informed whether any differ between groups.

This is now included in the last paragraph of the RESULTS section. Only somnolence showed statistical separation between groups.

We are informed how many patients “responded” to lithium, but not how many remitted. Nor are we informed (at least that I could find) how “response” was defined. Table 2 shows % improvement on a variety of ratings, but we are never told how many patients met response or remission criteria. The reader therefore has no sense at all whether the small per cent improvements represent most patients getting a little better or a few achieving remission. Since we strive to achieve remission, it is important to know whether any patients did. Thus, if quetiapine outperformed placebo on continuous measures but no one remitted, that seems less supportive of quetiapine’s use than it would be if some people’s depression completed went away and no ones did on placebo. While the comments to the reviewer acknowledge the importance of treating to remission, and this is reflected in the Introduction, as best I can tell this is the end of it. If the authors agree this is important, why do they not make this their major end point, or at least report how many subjects reached whatever their définition of remission might be.

Response has always been defined in the RESULTS section. It has been present (unmodified from the first version of the manuscript sent). However, given the length of some of the sections (METHODS especially) this information tended to get buried and it is understandable that a reader might not pick up on it.

Response is defined as “…a HRSD scores were reassessed at day 21 and patients with a ≥ 40% reduction (or a score < 18) were classified as responders and excluded” (page 8, second paragraph). I have included it also in the STATISTICAL analysis section on page 11.

As for the discussion on remission this study was never designed to assess it. The inclusion of this topic in the INTRODUCTION was at the reviewer’s request. That being said it was an interesting request and I agreed to it as it is a more modern concept.

However, presenting such data would require a Post Hoc analysis and probably a bad one at that. The problem with post hoc analysis is exactly that. You have to be lucky in the initial protocol design, as the analysis is usually not supported, which typically gives a type II error. If remission had been a primary measure (other than being silly and redefining it to suit my taste) study design would have been substantially different. For instance, if response had been defined as a HRSD of under 10, then the number of CBT sessions would have most likely been increased to a minimum of 16 or 18 and patient input as to dose would have gone out the window in favor of a more rigid, multiple arm dosing schedule (with a target dose for each arm, the placebo group being less numerous as a consequence). In other words, it would have been a big budget, N of 60 study, not an exploratory response-based study. At this point, I simply have to live with the design.
The last sentence of the first paragraph of the Discussion accurately raises questions about the utility of substituting quetiapine for unsuccessful antidepressant. However, doesn’t the failure of the placebo/CBT group to change at all also raise questions about the utility of CBT alone for this population? Seems to me, if this is where they want to raise issues about quetiapine’s efficacy as monotherapy, the efficacy of CBT as monotherapy should also be raised at the same time. If nothing else, to the extent that CBT is not doing anything, their argument that the study tells us nothing about quetiapine as monotherapy is undermined.

This paragraph has been slightly modified in the revised manuscript to more accurately qualify the ‘placebo/CBT’ group’s failure to respond. Even with this slight revision, the relative paucity of improvement with the placebo/CBT group is more fully explained in the paragraph below, which gives several hypotheses as to the modest improvement in the quetiapine/CBT group and the even poorer placebo/CBT group. This is also taken up in the next to last phrase in the DISCUSSION, prior to the conclusion.

At least as I understand what they did, they are comparing apples and oranges in the second paragraph of the Discussion. Maybe even apples and hammers. That is, the 30% they refer to in their study is the per cent change in rating scale score. The comparison 73% is the per cent Keller reported to have met his “response” criterion. As best I could tell, they did not report how many of their subjects met Keller’s “response” criterion. So, on what basis can the authors assert that their rates are “modest” relative to Keller. Perhaps they are, but not based on the information they give us.

I have revised this paragraph to more accurately reflect Keller’s data. Keller’s data has been more extensively dissected. I report that of the 73% response rate only about ½ had remission, the other half had a ‘satisfactory response’, which Keller defined as a HRSD of 15 or less, which is about what the response rate in my study amounted to. I also state that they administered 16 to 20 CBT-analysis system sessions, which is a nice tie in to the paragraph underneath, which goes into the much fewer sessions that we did, as remission was not an issue in my protocol. I could have gone into patient selection and underlying pathology, which was much more heterogeneous in the Keller study but did not.

They never speculate about the possible efficacy of quetiapine as monotherapy. Nor do they speculate about the implications of the lack of change in the placebo/CBT group for CBT monotherapy for TRD and whether that might inform our thoughts about using quetiapine as monotherapy vs. combining it with CBT. Finally, Dr. Chaput states in his comments to me that he has mentioned speculation about CBT’s poor showing. He certainly speculates that perhaps longer treatment would have been of benefit. However, he does not speculate that perhaps CBT is ineffective for this population.

I believe that many of the above changes made to the manuscript have addressed these
latter interesting and pertinent queries.

In the hope that this revision will meet with your satisfaction,

Yves Chaput, MD, FRCP(C), PhD