Author's response to reviews

Title: Anti-tumor necrosis factor alfa drugs in rheumatoid arthritis: systematic review and meta-analysis of efficacy and safety

Authors:

Alberto Alonso-Ruiz (aalonso@teleline.es)
Jose Ignacio Pijoan (joseignacio.pijoanzubizarreta@osakidetza.net)
Eukene Ansuategui (eaansuate@chdo.osakidetza.net)
Arantxa Urkaregi (arantxa.urkaregi@ehu.es)
Marcelo Calabozo (Marcelo.calabozoraluy@osakidetza.net)
Antonio Quintana (kfquloa@ehu.es)

Version: 2 Date: 17 January 2008

Author's response to reviews: see over
Dear Sir:

I am writing this letter in order to send you the paper “Anti tumor necrosis factor α drugs in rheumatoid arthritis: systematic review and metaanalysis of efficacy and safety” which has been modified following the recommendations and remarks made by the reviewers.

First of all you will find a list of the major amendments made to the manuscript. Then there is a list of answers and comments made in response to each remark made by each reviewer following the order of those remarks.

Yours faithfully

Marcelo Calabozo MD
**Modifications in the manuscript:**

We have somewhat shortened the abstract and included in it data sources.
Reference number 69 has been withdrawn.
We have corrected the misspellings and errors in Tables.
We have added in Table 1 the specific Jadad score for each selected trial.
We have made changes all throughout the Tables which include now only one decimal figure.
Leave all non-statistically significant NNTs out of the Tables.
Outcome frequencies (raw numbers) have been included in Table 4.

With the aim of polishing the English in our article we have had it corrected by a recommended “Manuscript Presentation Service” (www.biomedes.co.uk).

---

**Answer to Pierre Gevorek**

**MAJOR COMPULSORY REVISION**

It is true that the use of NNTs (and similarly of NNHs) as a summary effect measure in meta-analyses of clinical trials should be regarded with caution as they are strongly influenced, among other factors, by the particular mixture of clinical trials recruiting low, intermediate and high risk patients, different doses and duration of treatments, etc.

We can therefore apply this overall effect estimate only to individuals likely to have a risk close to the average risk of patients included in the meta-analysis at hand (*Egger M, Smith GD. Principles of and procedures for systematic reviews. In: Egger M, Smith GD, Altman DG, editors. Systematic Reviews in Health Care. 2nd ed. London: BMJ Publishing Group; 2001. p. 43-68*). The best way of applying the summary effect estimates from meta-analyses to individual patients or particular groups of patients is by using their individual/group baseline risk and calculating from the pooled relative risk reduction the treatment due outcome risk. From the treatment-associated absolute risk difference obtained is then straightforward to get the corresponding, baseline risk specific NNT (*Egger M, Smith GD. Principles of and procedures for systematic reviews.*
We have provided these pooled NNTs as a kind of effect estimate for average risk patients. In an attempt to minimise the presence of factors known to influence risks and therefore NNTs, we have selected rather homogeneous studies in terms of minimum follow-up, diagnostic criteria and have further made subgroup analyses accounting for several important clinical characteristics (resistance to previous DMARD, etc.). As the individual trials’ baseline risks can be estimated from the numbers provided in the article’s tables and figures (Table 2 provides detailed information including different outcomes and treatment durations), it is possible to obtain more individualised NNTs if required through the use of very simple calculus or freely available epidemiological calculators (see for instance http://medinformatics.uthscsa.edu/calculator/calc.shtml).

Despite these considerations pooled NNTs are today commonplace in the medical literature. As an example we give some recent references in the field of rheumatoid arthritis that use them:


In our paper we have included a sentence commenting on both clinical usefulness and pitfalls regarding the use of NNTs in meta-analyses (pp. 13-14).

As the reviewer rightly guesses, we have used the data provided by the authors in the published clinical trials. In some of these trials the authors explicitly acknowledge the use of **intention-to-treat approach to analysis**. In some others this strategy is not explicitly described as the one used in the study, but from the figures provided it seems that they have considered all the patients in the initially assigned treatment arm all throughout the study.
It is a well known fact that intention-to-treat approach to study design and analysis is neither completely nor adequately reported in many trials (Holly S, Campbell F. What is meant by intention to treat analysis? Survey of published randomised controlled trials. BMJ 1999;319:670-4). A more complete description of how to handle deviations from random allocation and missing values is being demanded but still there is a lack of uniformity and completeness in studies’ methods reports. We do agree with the reviewer that an analysis based on better information on the response pattern of individual patients could reinforce validity of efficacy and safety results. As we explain below, a systematic attempt to get extra information from trials’ promoters (pharmaceutical companies) had no success.

With regard to the **use of the NNH**, we do absolutely agree with reviewer’s comments and very much appreciate them.

The main author of this paper repeatedly contacted with representatives of the **pharmaceutical companies** promoting the selected clinical trials but he failed to obtain any extra information but the one already being published.

**MTC metanalysis use**: we have read with interest the recently published paper referred to by the reviewer and do agree that it seems a potentially interesting approach to the complex issue of indirect comparisons. There was one obvious reason for us not using it as this paper was published after our paper had been written and sent for review. On the other hand this new method uses some meta-regression techniques. Our meta-analysis, based on a not very large number of clinical trials and with apparently more within-trial variation as compared with between-trial variation in relevant aspects such as doses used does not seem to provide an optimal setting for the use of some of the proposed statistical techniques (e.g. random effects meta-regression). (Thompson SG, Higgins JPT. How should meta-regression analyses be undertaken and interpreted?. Statistics in Medicine 2002;21:1559-1573).

As we were also concerned about the lack of direct, head-to-head comparisons of effects among the commercialised anti-TNFα drugs, we had previously carried out and indirect

The results of this analysis are shown in the following table:

<table>
<thead>
<tr>
<th>Data from anti-TNF plus MTX versus MTX alone</th>
<th>ACR20 Odds Ratio (95% IC)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Infliximab versus etanercept</td>
<td>0.59 (0.18-1.91)</td>
</tr>
<tr>
<td>Etanercept versus adalimumab</td>
<td>1.28 (0.45-3.59)</td>
</tr>
<tr>
<td>Infliximab versus adalimumab</td>
<td>0.75 (0.35-1.61)</td>
</tr>
</tbody>
</table>

As it can be appreciated from the table there was no clear trend using this approach towards any specific anti-TNFα drug showing a differential effect size. Later on, and due to the rather large amount of data incorporated in the paper, we decided not to include this aspect of the meta-analysis.

MINOR ESSENTIAL REVISION
We have corrected the misspellings and errors in tables.

Reference number 69 (which was the same as number 9) has been withdrawn.

We have not included any reference to large observational studies for several reasons: there are several and diverse observational studies assessing uncontrolled effects of anti-TNFα drugs in RA in different clinical and health care practice settings and therefore there is wide variability in terms of baseline outcome risk, clinical severity, follow-up time, etc. To compare our results with those coming out of observational studies would doubtless increase length and complexity in an already long and complex article.

NNT/NNH ratio: some articles that incorporate this ratio as an estimate of the benefit-risk ratio:


In some of the referenced works the use of this ratio is suggested as a way to get insight into the benefit/risk ratio relationship.

Following reviewer’s instructions we have somewhat shortened the abstract and included in it data sources.

**Answer to Andrew Moore**

**MAJOR COMPULSORY REVISION**

Results of clinical trial’s quality assessment through the Jadad’s quality score are described at the end of the first paragraph in the Results section: “The methodological quality of the studies was moderate to high (3-5) except Bathon's trial (24) with a lower Jadad score of 2 because it neither mentioned nor explained whether treatment allocation was performed using a random procedure”.

We have added in Table 1 the specific Jadad score for each selected trial.

We have made changes all troughout the tables which include now only one decimal figure.
It is clear that statistical significance is not tantamount to **clinical relevance**. We have however presented statistically significant results. We have also made additional comments in the discussion section about the magnitude of estimated effect indexes.

Although there are several ways of reporting non-statistically significant NNTs (see Egger, Systematic reviews in health care), we do agree with the reviewer’s comment and **leave all non-statistically significant NNTs out of the tables**.

The commonly used **funnel plot** is also useful as a descriptive and exploratory tool for showing the potential existence of heterogeneity due to selection bias, different quality of the individual trials and other artefactual and real factors potentially affecting effect estimates (Glasziou PP, Sanders SL. Investigating causes of heterogeneity in systematic reviews. *Statistics in Medicine* 2002;21:1503-1511). In our study we could not perform an individual data meta-analysis, which could have minimised some of the limitations of our standard approach. Other factors irrespective of the potential existence of publication bias, however, can also influence the general shape of the funnel plot. Caution should be taken regarding the interpretation of asymmetry as other aspects such as the quality of individual trials can also be involved. If we perform the analysis limiting ourselves to trials with MTX-naïve patients, the previously observed asymmetry disappears as so it does the statistical significance [Begg and Mazumdar test (p = 0.462), Egger’s regression asymmetry test (p =0.506)].
A similar phenomenon is seen when analysis is focused on patients already resistant to MTX [Begg and Mazumdar test (p = 0.108), Egger’s regression asymmetry test (p =0.03)]

![Funnel Plot of Precision by Log risk ratio](image)

We make a comment in the Discussion section of the paper stating that there are several reasons apart from publication bias that can explain the finding of asymmetry in the funnel plot.

The outcome frequencies (raw numbers) have been included in table 4, so that percentages can be easily calculated. We have however kept in this table Q and I² statistics as they are currently commonly used and provide a flavour of observed statistical heterogeneity of results.

MINOR ESSENTIAL REVISIONS
With the aim of polishing the English in our article we have had it corrected by a recommended “Manuscript Presentation Service” (www.biomedes.co.uk).

Forest plots are traditional, very informative graphs showing effect estimates of individual trials included in a meta-analysis as well as information about precision of estimates, heterogeneity of estimates across trials, etc. we are hence keeping them as a way of conveying some important results of our meta-analysis.
L’Abbé plots: although we do agree with the reviewer that heterogeneity is always an issue when performing meta-analysis, there has been traditionally a lot of controversy regarding its potential causes and how to deal with it using sound, bias-free methodologies. These plots have long claimed to be a useful way of showing some sources of heterogeneity, such as the influence of baseline risk, most often estimated from the control group event rates (Song F. Exploring Heterogeneity in Meta-Analysis Is the L'Abbé Plot Useful? J Clin Epidemiol 1999;52(8): 725-730). However this method has been also shown to be sensitive to problems such as the regression to the mean effect, especially when some trial’s sample size is small (Sharp SJ, Thompson SG, Altmann DG. The relation between treatment benefit and underlying risk in meta-analysis. BMJ 1996;313:735-738.). We acknowledge that both types of plots (forest and L’Abbé) are useful tools to summarize results and to glance over the degree of variation between studies. Both also have limitations and can lead, if misunderstood, to flawed interpretations.