Author's response to reviews

Title: Modeling Dose-Response Relationships of the Effects of Fesoterodine in Patients With Overactive Bladder

Authors:

Linda Cardozo (lcardozo@compuserve.com)
Vik Khullar (vik.khullar@imperial.ac.uk)
Ahmed El-Tahtawy (ahmed.el-tahtawy@pfizer.com)
Zhonghong Guan (zhonghong.guan@pfizer.com)
Bimal Malhotra (bimal.k.malhotra@pfizer.com)
David Staskin (staskinatt@gmail.com)

Version: 2 Date: 29 March 2010

Author's response to reviews: see over
March 29, 2010

Rachel Neilan, MSc
Assistant Scientific Editor
BMC-series Journals
BioMed Central
Floor 6, 236 Gray's Inn Road
London, WC1X 8HL
UK

Dear Rachel:

On behalf of my co-authors, I am pleased to submit a revised version of our manuscript “Modeling Dose-Response Relationships of the Effects of Fesoterodine in Patients with Overactive Bladder.”

We have revised the manuscript to address the reviewer’s comments and we believe that this has strengthened our manuscript. Our responses to the comments are itemized below.

We hope you find this paper suitable for publication in BMC Urology.

Sincerely,

Linda Cardozo, MD
King’s College Hospital
Denmark Hill
London SE5 9RS
UK
Phone: +44 020 7935 2357
Fax: +44 020 7224 2797
E-mail: lcardozo@compuserve.com
Reviewer 1:
The authors reported a study evaluating the dose-response relationship in a pooled analyses of 2 phase II and 2 phase III study on Fesoterodine. The study finally demonstrated that dose-response relationship did exist.

The study is extremely complex and its methodology is impossible to understand for all the average readers of the Journal.

We agree that the methodology of this study is complex. However, as BMC Urology considers articles on “all aspects of the prevention, diagnosis, and management of urological disorders, as well as related molecular genetics, pathophysiology, and epidemiology,” it is likely that many readers will lack expertise in the methodology of most studies reported in BMC Urology. We feel that this paper is relevant to the readership of BMC Urology, as modeling and simulation in drug development is well accepted by drug regulatory agencies, and is becoming increasingly valuable to understand the clinical trial data to their fullest extent.

Introduction: in the second paragraph of the second page of the Background section, the statement “Therefore, mathematical models were developed to describe quantitative and predictive dose-response relationships of the effects of fesoterodine using a rich, subject level longitudinal data set from the phase II and III studies. This model-based dose-response characterization is more comprehensive because all available subject-level data obtained at each study visit after administration of 3 different dose levels (4 mg, 8 mg, or 12 mg) in the phase II and III trials were combined and analyzed.” can be deleted.

In the authors’ view, this text should be retained, particularly in light of the previous comment, as in very simple terms it explains the input used for modeling and the value of the output from modeling. We have revised the text to “Therefore, mathematical models were developed to describe quantitative and predictive dose-response relationships of the effects of fesoterodine. These models utilize individual subject-level, longitudinal data set from two Phase II and two Phase III studies. Compared with the traditional analysis of the end-of-treatment results on study-by-study basis, this model-based dose-response characterization is more comprehensive because all available subject-level data obtained at each study visit after administration of 3 different dose levels (4 mg, 8 mg, or 12 mg) in the phase II and III trials were combined and analyzed.”

In the Results section and in table 2, it should be clarified if the predicted dose-response relationship yielded statistically significance.

The following text have been added for clarification: “The values reported in table 2 represent the final model parameter estimates. The final model evaluations included Stepwise forward or backward comparisons. The selection is based on the likelihood ratio test, across multiple models, each expressing different
covariate-parameter combinations.\textsuperscript{1} The statistical significance is considered when there is a decrease in likelihood corresponding to a chi-square distribution with $\alpha = 0.01$ and degrees of freedom equal to the difference in the number of estimated parameters between any two models.”

The selection of postvoid residual urine as a parameter to evaluate the drug safety sounds a little bit inappropriate, due to the well known impact of anticholinergic drugs on the voiding phase of the micturition circle. Overall adverse event rate, withdrawal rates, or some specific complication rates would have likely been more appropriate. Text added: Typical anticholinergic effects like dry mouth and constipation were generally of mild to moderate severity; any dose-relationship of adverse events and associated discontinuation rates was easily apparent from their descriptive summaries (shown in Table below from the USPI), and further exploration of the tolerability data did not reveal any significant patient covariance. Therefore these AEs were not considered for model based analysis. Because post void residual urine volume (PVR) can be of potential safety concern, it was selected to fully understand the dose-response relationship and patient covariates that may influence the effect of fesoterodine treatment on PVR.

\textsuperscript{1} Karlsson et al. Pharmacokinetic Models for the Saturable Distribution of Paclitaxel. Drug Metab. Dispos. 1999. 27(10): 1220-1223.
Table 3 Adverse events with an incidence exceeding the placebo rate and reported by ≥1% of patients from double-blind, placebo-controlled Phase 3 trials of 12 weeks treatment duration

<table>
<thead>
<tr>
<th>System organ class/Preferred term</th>
<th>Placebo N=554 %</th>
<th>Toviaz 4mg/day N=554 %</th>
<th>Toviaz 8mg/day N=566 %</th>
</tr>
</thead>
<tbody>
<tr>
<td>Gastrointestinal disorders</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dry mouth</td>
<td>7.0</td>
<td>18.8</td>
<td>34.6</td>
</tr>
<tr>
<td>Constipation</td>
<td>2.0</td>
<td>4.2</td>
<td>6.0</td>
</tr>
<tr>
<td>Dyspepsia</td>
<td>0.5</td>
<td>1.6</td>
<td>2.3</td>
</tr>
<tr>
<td>Nausea</td>
<td>1.3</td>
<td>0.7</td>
<td>1.9</td>
</tr>
<tr>
<td>Abdominal pain upper</td>
<td>0.5</td>
<td>1.1</td>
<td>0.5</td>
</tr>
<tr>
<td>Infections</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Urinary tract infection</td>
<td>3.1</td>
<td>3.2</td>
<td>4.2</td>
</tr>
<tr>
<td>Upper respiratory tract infection</td>
<td>2.2</td>
<td>2.5</td>
<td>1.8</td>
</tr>
<tr>
<td>Eye disorders</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dry eyes</td>
<td>0.0</td>
<td>1.4</td>
<td>3.7</td>
</tr>
<tr>
<td>Renal and urinary disorders</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dysuria</td>
<td>0.7</td>
<td>1.3</td>
<td>1.6</td>
</tr>
<tr>
<td>Urinary retention</td>
<td>0.2</td>
<td>1.1</td>
<td>1.4</td>
</tr>
<tr>
<td>Respiratory disorders</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Cough</td>
<td>0.5</td>
<td>1.6</td>
<td>0.9</td>
</tr>
<tr>
<td>Dry Throat</td>
<td>0.4</td>
<td>0.9</td>
<td>2.3</td>
</tr>
<tr>
<td>General disorders</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Edema peripheral</td>
<td>0.7</td>
<td>0.7</td>
<td>1.2</td>
</tr>
<tr>
<td>Musculoskeletal disorders</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Back pain</td>
<td>0.4</td>
<td>2.0</td>
<td>0.9</td>
</tr>
<tr>
<td>Psychiatric disorders</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Insomnia</td>
<td>0.5</td>
<td>1.3</td>
<td>0.4</td>
</tr>
<tr>
<td>Investigations</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ALT increased</td>
<td>0.9</td>
<td>0.5</td>
<td>1.2</td>
</tr>
<tr>
<td>GGT increased</td>
<td>0.4</td>
<td>0.4</td>
<td>1.2</td>
</tr>
<tr>
<td>Skin disorders</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Rash</td>
<td>0.5</td>
<td>0.7</td>
<td>1.1</td>
</tr>
</tbody>
</table>

ALT=alanine aminotransferase, GGT=gamma glutamyltransferase

In the “Postvoid residual volume” paragraph, please, add odds ratios, confidence interval, and p values for the significant covariates.

Text Added: A step-wise selection of covariates was followed to select the significant covariates. The log-likelihood ratio test was used to compare the base model without any covariates to the model with an added or deleted covariate. When addition of a covariate results in a p-value less than 0.01, the covariate is significant. The addition of covariates is followed by step-wise deletion of covariates; except, the criteria for deletion if the change in p-value is not significant (p=0.001). The process is repeated until no further covariate is deleted from the model. The odds ratios are not
appropriate for this analysis since we are predicting mean response and not probabilities.

In the “Discussion” section, please assess the study limitations. Study limitations have been added to Discussion.
Reviewer 2:
Major Compulsory Revisions; no revisions exist.

Minor Essential Revisions; no revisions exist.

Discretionary Revisions; The Poisson distribution can be applied to systems with a large number of possible events, each of which is rare. Authors' should explain why micturition counts and UUI episode counts follow the Poisson distribution, also suitable models as above parameters. Editors should hear the opinion of the statistician about the validity of the statistical analysis.

Poisson distribution is a discrete probability distribution that expresses the probability of a number of events occurring in a fixed period of time if these events occur with a known average rate and independently of the time since the last event. In general, in our experience and based on literature, count data (UUI [for example for another OAB drug, darifenacin\(^2\)], seizures, panic attacks, etc) are best described by the Poisson distribution. Each count event has two possible outcomes: occurrence or no occurrence (Bernoulli trial). Counts are sums of Bernoulli trials. For large samples, Poisson distribution is used to approximate the Binomial distribution. One of the desirable properties of Poisson distribution is its constraints to be non-negative.

Reviewer 3:
One key problem seems to be properly described by the authors themselves: “Methods: A population analysis was conducted using pooled data from several studies.” I don’t think the term “several” is appropriate in a publication claiming to be scientific, particularly when it is very difficult to find out from the rest of the text what this “several” may stand for. There are 2 phase III trial with 2 doses only, and 2 phase II trials with 3 doses of which only 1 is published, the other just an abstract.

The word “several” has been deleted here, and we now specifically note in this paragraph, as well as elsewhere in the text and in Table 1, that data from 2 phase II trials and 2 phase III trials were used in this study.

In addition we learn: “… a dose-response relationship in the strictest (?) sense cannot be established with fewer than 3 dose levels. Therefore (?), mathematical models were developed to describe quantitative and predictive dose-response relationships of the effects of fesoterodine using a rich (?), subject-level longitudinal data set from the phase II and III (?) studies. This model-based dose-response characterization is more comprehensive (?) because all available subject-level data obtained at each study visit after administration of 3 different dose levels (4 mg, 8 mg, or 12 mg) in the phase II and III trials were combined and analyzed.” Further the authors state: “Data from 2514 subjects given placebo or fesoterodine 4 mg, 8 mg, or 12 mg in 2 phase II …and 2 phase III …. double-blind 8- or 12-week trials were used to develop the dose-response models”? This language is more obscuring than explaining. How many complete data sets are there? What does that mean when there are only complete 3 dose data published from a total of 186 or 225 pts? Or has a less than “strictest sense” been used? What does it mean that a model is "anchored" on less than 10% of the subjects' data?

We state ‘strictest sense’ because sometimes researchers claim dose-response relationship based on 2 dose levels when we very well know that there is always a straight line (curve) that connect 2 points.

The following text in the Introduction section has been revised to address the reviewer’s concern around clarity. “Therefore, mathematical models were developed to describe quantitative and predictive dose-response relationships of the effects of fesoterodine. These models utilize individual subject-level, longitudinal data set from two Phase II and two Phase III studies. Compared with the traditional analysis of the end-of-treatment results on study-by-study basis, this model-based dose-response characterization is more comprehensive because all available subject-level data obtained at each study visit after administration of 3 different dose levels (4 mg, 8 mg, or 12 mg) from a combined dataset across trials were analyzed.”

The breakdown of the number of subjects in each treatment arm in each trial is provided in Table 1. Data from all subjects in each study who were randomized to
either placebo or fesoterodine and who had at least one post-baseline visit were included in the present study. One of the fesoterodine studies included a tolterodine extended release arm as an active comparator; subjects randomized to this arm were not included in the present study.

The statement that “The 12-mg data were included in the models to anchor the dose-response relationship…” has been elaborated clarified as: “While the 12-mg dose is not an approved dose, the addition of data from a total of 225 patients who received the 12-mg dose in Phase 2 trials and provided a total 3 dose levels to strengthen confidence in the shape of the dose-response relationship” The fewer numbers of subjects at the 12 mg dose level are reflected in the somewhat wider 95% confidence interval widths around the model-predicted response at that dose, particularly for micturitions, MVV, and PVR (Figures 1, 3, and 4).

Similarly obscure, it is stated for PVR that only 0.5% exceeded a threshold of 200 ml. From the trends shown in Fig 4 it would be most appropriate to additionally clarify this percentage dosage dependent for men over 70 yrs, i.e. those where such changes seem to occur.

For patient above 70 years, there is an interaction with dose that was deemed significant. The predicted rates, along with confidence limits, are presented in Figure 4 to estimate what may be expected in specific subgroups by age or by gender.

Overall, when modeling any changes of inter-dependent parameters, here number of voids and leaks and MVV, one would expect some plausibility checks for these changes. If the number of voids decreases for placebo by -1.1, approximately -10%, then the MVV should increase by +10%, with diuresis being constant (in first approximation ignoring leakage), but the increase is only 9.7 ml, and a MVV of 100 ml is unlikely(?). Similarly non-plausible are the changes for verum. Thus, any comprehensive discussion of changes in number of voids, in leakage, and MVV, has to be tested modeled against a plausible volume balance, particularly when most changes are so small.

In our experience with clinical trials, mean changes in micturitions and MVV do not represent equal percentages of baseline values. For instance, in one of the phase II trials used in this analysis, the fesoterodine 8 mg group had baseline means of 11.9 for micturitions and 150.2 mL for MVV. LS mean change at week 12 was –1.88 for micturitions and 33.62 mL for MVV, which represent a 15.8% decrease in micturitions and an 21.8% increase in MVV (Chapple et al. Eur Urol 2007;52:1204-1212). In the other phase III trial used in this analysis, the fesoterodine 8 mg group had baseline means of 12.0 for micturitions and 156 mL for MVV. LS mean change at week 12 was –2.09 for micturitions and 33.6 mL for MVV, which represent a 17.4% decrease in micturitions and an 21.5% increase in MVV (Nitti et al. J Urol 2007;178:2488-2494). Thus, in our model we would not necessarily expect to see a direct
correspondence between percentage decrease in micturitions and percentage increase in MVV.
Reviewer 4:
The only one addition to the analysis that I would like to see is related to the homogeneity across study sites. It would be important to the reader to know any difference, if any, in baseline data, outcome and side effects along time.

The strength of these model-based analyses across studies lies in the fact that each subject’s baseline and other demographic characteristics, including site and study, are analyzed along with their treatment response (beneficial and adverse). While these sources of non-homogeneity are accounted for in the model, they may not be explicit in the final model presentation unless they are influential covariates of response.