Bridging from drug registration trials to meaningful clinical evidence: The case of schizophrenia

Mattila, T.K.

Citation for published version (APA):

General rights
It is not permitted to download or to forward/distribute the text or part of it without the consent of the author(s) and/or copyright holder(s), other than for strictly personal, individual use, unless the work is under an open content license (like Creative Commons).

Disclaimer/Complaints regulations
If you believe that digital publication of certain material infringes any of your rights or (privacy) interests, please let the Library know, stating your reasons. In case of a legitimate complaint, the Library will make the material inaccessible and/or remove it from the website. Please Ask the Library: https://uba.uva.nl/en/contact, or a letter to: Library of the University of Amsterdam, Secretariat, Singel 425, 1012 WP Amsterdam, The Netherlands. You will be contacted as soon as possible.
Chapter 2

Insomnia medication: Do published studies reflect the complete picture of efficacy and safety?
Chapter 2
Insomnia medication: Do published studies reflect the complete picture of efficacy and safety?

Abstract

Selective publication can have a deleterious effect on evidence based medicine, health policy decision making and treatment guidelines. Using the European Public Assessment Reports (EPARs) as reference, this study examined selective publication and selective reporting of efficacy and safety of insomnia medication. EPARs of with three insomnia medications were used to identify all clinical trials that were performed between 1998 and 2007 for the purpose of registration in the EU. The matching publication for each trial was searched through a systematic literature search. Accuracy of information in the publications was examined by comparison to the information in the EPARs. Only 55% of the trials with insomnia medications identified in EPARs were published. Positive trials were approximately two times more likely to be published. The lag time from study completion to publication was shorter for the positive compared to the negative trials. Sample size did not correlate with publication of negative trials. The meta-analysis of the effect size of insomnia medication was 1.6 times larger in the published data compared to the complete data. While the primary end points of the trials were reported reliably in the publications, remarkable inconsistencies were detected in the reporting of the secondary end points, methods, results and, especially safety. In conclusion, selective publication and reporting lead to an overestimation of efficacy and underestimation of safety of insomnia products. Authors of treatment guidelines should be aware of this bias. EPARs/FDA reviews provide a more unbiased view of the benefit-risk balance of insomnia and other medications and hence these documents should be consulted by e.g. authors of meta-analyses and of treatment guidelines.
Chapter 2

Introduction

Published clinical trials in peer reviewed journals form the basis of evidence for practising medicine. Potential bias in the published evidence could influence treatment choices and subsequently might have a deleterious impact on public health and on health care costs. The increasing use of meta-analyses and systematic reviews of published articles in formulating treatment guidelines has increased the interest in the possible consequences of biased clinical information.

There are two processes contributing to biased information about clinical trials (publication bias): Selective publication is defined as a tendency to submit or accept manuscripts for publication based on the direction or strength of the study findings. Selective reporting appears most commonly as biased outcome reporting and inaccurate reporting of study design and research.

Selective publication and reporting exist also among CNS medication trials (1). For example, a comparison of the US Food and Drug Administration (FDA) reviews with published data has revealed selective publication of anti-depressant trials (2): positive trials were three times more likely to be published than negative trials. Another example in the field of psychiatry is the case of lamotrigine for the indication bipolar disorder, where positive trials were 2.5 times more likely to be published than negative trials (3).

Attempts to study selective publication and reporting have been hindered by the limited access to unpublished trials. To overcome this problem, published trials have been compared to information obtained from for example study protocols sent to ethics committees (4,5), institutional review boards (6,7,8) and funding agencies (9). Regulatory authorities, because of their access to complete dossiers, seem to be a most suitable body for examining publication bias with respect to evidence pertaining to efficacy and safety of pharmaceutical products. Only a few studies have compared the full dossier of a product submitted for registration to published data (10,11). Until recently, publication of results from the complete dossier data was hampered by the confidentiality restrictions of the data. However now, European Public Assessment Reports (EPARs) offer a possibility to use the full dataset with the exclusion of commercially confidential data. EPARs are documents that the European Medicines Agency (EMA) publishes on the products assessed by the Committee for Medicinal Products for Human Use (CHMP). These documents present information on the grounds for granting a marketing authorization via the centralized procedure in the EU, summarizing the evidence submitted by the applicant and the scientific discussion at the CHMP. All commercially confidential data are deleted before the publication of an EPAR. Since 2004, the EPARs of products with a withdrawn or refused application are published at the EMA website as well.
The purpose of this paper is to investigate selective publication and reporting in the field of insomnia products using EPARs as reference. Insomnia products were chosen for this research, since the magnitude of effect of the same insomnia product is variable among trials (12,13) and therefore it was likely that negative trials would be found in the examined EPARs as well. To the best of our knowledge, this study is the first one to compare the information on the EPARs to published trials in order to investigate publication bias.

Experimental procedures

For the purpose of this study we retrieved all EPARs of medicinal products that fall under the ATC code N05C, ‘hypnotics and sedatives.’ Three products with different active substances were identified: melatonin (Circadin®, RAD Neurim Pharmaceuticals EEC Ltd., Berkshire, UK), zaleplon (Sonata/Zerene®, Meda AB, Solna, Sweden) and ramelteon (Rozerem®, Takeda Global Ltd., Lincolnshire, Illinois, US). Circadin® and Sonata® are authorized medicinal products; the ramelteon EU application was withdrawn by the applicant in September 2008.

From the EPARs we retrieved all randomized, double-blind, placebo and/or active controlled trials performed in patients, i.e. phases II and III trials. A trial was included in this study if the EPAR provided sufficient information to identify the matching publication and to make a comparison.

For each clinical trial the following information was extracted from the EPAR: objectives, study design, primary and secondary end points, treatment arms and study outcomes. The outcome on the predefined primary end point was used as a criterion for determining a trial as positive or negative. A trial was deemed positive if the test product was significantly better than placebo (in placebo controlled trials), or if the test product was non-inferior to the active comparator and better than placebo (in placebo and active controlled trials). In trials with several treatment arms (different doses and/or measurement points), the trial was deemed positive or negative based on the results for the dose(s) which finally received marketing authorization.

To find whether trials in the EPARs were published, a literature search was performed in PubMed, Embase, JSTOR, NICE, Clinical Evidence, Web of Science, Cochrane and PsycInfo. In addition, references in review articles and identified publications, related articles and lists of publications on company websites were examined. Only full article publications of trials in peer reviewed journals were considered as publications, thus abstracts or presentations at meetings were not included.

For trials where no publication was found, we tried to contact the primary investigator,
the sponsor or authors of other publications for the same product. Since the response rate was very modest and the obtained answers did not provide useful information, this is not presented in the paper.

In the subset of clinical trials which were published, we examined if the trial details presented in the EPARs were accurately reflected in the scientific publications. Comparisons were made with respect to the following factors: (1) primary end point, (2) secondary end points, (3) study design, including possible protocol changes, (4) analysis method, including analysed population, (5) numerical results with respect to the primary end point and the secondary end points, (6) post hoc additional analysis and (7) adverse events and safety issues.

In addition we performed a meta-analysis to compare the overall effect size of a product based on all the trials in the EPAR versus the published trials only. The number of subjects, least-squared mean differences in primary end point (sleep latency) result and standard error of the differences between placebo and the studied drug were extracted. In case of several dose arms, the data concerning the marketed dose was used in the analysis. The standard error was used to estimate standard deviation needed for the meta-analysis. Meta-analysis was carried out with Stata version 10.1.

Results

Selective publication

In total 34 trials were identified in the EPARs of the three insomnia products. Five trials (one trial with melatonin and four trials with zaleplon) were excluded because of insufficient information, thus this study examined 29 trials from the EPARs. Only 55% (16/29) of these trials were published. As Fig. 1 indicates, a considerably higher percentage of the positive trials was published compared to negative trials: 68% (13/19) versus 30% (3/10). The difference between the percentage of published positive and negative trials was marginally significant (p = 0.056, one-tailed Fischer’s exact test) with this rather small number of trials included. The publication percentages for individual products are presented in Fig. 1. From all the published studies, 81% (13/16) was positive and 19% (3/16) negative.
Insomnia medication: Do published studies reflect the complete picture of efficacy and safety?

Because of the limited data on the melatonin (Circadin®, 11 studies) and zaleplon (Sonata®, 9 studies) EPARs, comparing the effect size of a product based on all the trials versus the published trials was only possible with the data of ramelteon (nine trials in total of which five published). The effect size comparison was made with ramelteon 8 mg dose as it is the marketed dose in the US. As indicated in Fig. 2, the overall effect size based only on data from published trials was approximately 1.6 times larger than the effect size calculated on the data from all the trials. The overall weighted mean difference in sleep latency was 6.37 min (95% CI 4.45, 8.3) for all the trials compared to 10.1 min (95% CI 7.43, 12.78) for the published studies only. The overall weighted mean difference in sleep latency for the unpublished trials was 2.34 min (95% CI: -0.44, 5.12).

Figure 1.
Publication status and outcomes across all three insomnia products and per product.
The relationship of study size and publication was examined for 18 trials where the number of subjects (intention-to-treat analysis) was available from the EPAR (nine ramelteon trials and nine melatonin trials) (Fig. 3). The number of subjects in zaleplon studies, including the negative ones which were published, was not available in the EPAR. The results suggest that for negative trials, size did not predict publication since all negative studies were not published regardless of sample size. For positive trials, larger study size increased the likelihood for publication.

Figure 2.
Meta-analysis of ramelteon studies. Difference in improvement of sleep onset (minutes). Ramelteon 8 mg versus placebo, stratified by published/unpublished trials.
Further we investigated the average time lag from study completion to publication, which in our dataset was 3.4 years (Table 1). There was a slightly shorter lag time for the positive trials (3.3 years) compared to negative trials (4 years). The time of submission for publication was stated in 11 publications out of 16. For these studies, the time lag from study completion to submission for publication was on average 3 years, with a shorter time lag for positive (2.9 years) as compared to negative trials (3.5 years). The time from manuscript submission to publication was on average 0.6 years and there was no notable difference between the positive and negative studies (0.7 and 0.5 years, respectively).
Chapter 2

Table 1.
Time lag from study completion to submission and to publication for trials with three insomnia medications.

<table>
<thead>
<tr>
<th></th>
<th>Study completion to submission (years)*</th>
<th>Study submission to publication (years)*</th>
<th>Study completion to publication (years)†</th>
</tr>
</thead>
<tbody>
<tr>
<td>Positive studies</td>
<td>2.9</td>
<td>0.7</td>
<td>3.3</td>
</tr>
<tr>
<td>Negative studies</td>
<td>3.5</td>
<td>0.5</td>
<td>4.0</td>
</tr>
<tr>
<td>Total</td>
<td>3.0</td>
<td>0.6</td>
<td>3.4</td>
</tr>
</tbody>
</table>

* examined for 11 studies
† examined for 16 studies

Selective reporting
From the published 16 studies, we examined if the trial details presented in the EPARs were reported reliably in the scientific publications. One published study with melatonin was excluded due to insufficient data in the EPAR. Therefore selective reporting was examined in 15 publications in total. The summary of the identified differences is presented in Fig. 4 and described in more detail later.

Figure 4.
Identified reporting differences in insomnia trial publications compared to EPARs.
The primary endpoints were reported consistently with the EPARs. In general, the publications reported more secondary endpoints (in 12 publications out of 15, 80%) and post hoc or additional analysis (9/15, 60%), and gave more information concerning study design (8/15, 53%) and analysis method (9/15, 60%) as compared to EPARs. Additional reported information was mostly favourable to the test drug. There were also substantial differences in numerical results concerning both primary endpoint (in 2 publications out of 15, 13%) and secondary endpoints (2/15, 13%). In the two cases concerning primary endpoint, the results (of 4 test drug/placebo comparisons) had changed either from not significant to significant or the level of significance was higher in the publication compared to the EPAR. In the two cases concerning secondary endpoints the results were significant and in favour of the test drug whereas in the EPARs they were not significant.

The publications reported much less safety issues than the EPARs. Safety data was missing in 8 publications out of 15 (53%), including both positive and negative trials, and most of it was on adverse events, with one publication lacking any safety data. Comparison of safety data between a trial publication and EPAR is hampered by the fact that the EPAR presents pooled safety data derived from several studies. However safety issues which are crucial for insomnia medication, such as withdrawal effects, rebound insomnia and next day performance, are registered in every insomnia trial and thus could have been reported. Therefore lack of data on these issues in a publication was considered as missing data.

Although benefit–risk balance estimation was not defined as an examined detail for selective reporting a priori, it was noted that the publications often did not reflect the benefit–risk balance discussion in the EPAR.

Discussion

This study examined selective publication and selective reporting of insomnia trials by comparing publications in peer review journals to information provided in EPARs. While it is recognized that the EPAR and publication serve different purposes, here the EPARs are used as a reference to verify information in publications. To the best of our knowledge, this study is the first one to use EPARs—thus the complete dossier data—as reference.

The results of this study show selective publication of insomnia trials: only 55% of the placebo controlled clinical trials performed with three insomnia products recently submitted for registration in the EU were published. This percentage is lower than what
Chapter 2

was found for anti-depressant trials. Turner et al. (2) found that 69% of the trials with anti-depressants approved by the FDA between 1987 and 2004 were published. Two other studies comparing the clinical trial data published in articles with FDA reviews (of all types of medicinal products) (14,15), found that 43% and 77% of the trials were published, respectively.

Furthermore, the results indicate that the positive insomnia trials were approximately two times more likely to be published than the negative ones (68% of the positive versus 30% of the negative). This is in line with the findings of the study with anti-depressant trials, where the corresponding likelihood for publication was three times higher for positive trials (97% of the positive trials published compared to 33% of the negative trials) (2). For all types of medicinal products this likelihood is even higher ranging from three to five times higher likelihood for positive compared to negative trials to be published (14,15).

Possible factors that might explain the observed difference in publication of positive versus negative trials were investigated, i.e. trial size and lag time from study completion to publication. Our investigation did not confirm the widely accepted belief that small sample size studies are published less frequently, while large-scale studies are more likely to be published regardless of the findings (16,17). In our study, negative trials remained unpublished regardless of the sample size.

Within the sample of published studies, the average time lag between study completion and publication was 3.4 years. Positive studies were published slightly faster than the negative ones (3.3 compared to 4.0 years, respectively). In previous studies, this time lag has varied for positive trials from 4 to 5 years and for negative trials from 6 to 8 years (7,18). Since we could not identify a difference between negative and positive trials in the time lag from manuscript submission to publication, it is assumed that journal editors are not the primary source of delay of publication of negative trials. The longer time gap between study completion and submission of negative studies as compared to the positive ones (3.5 versus 2.9 years, respectively) might indicate that it either takes more time for investigators to submit manuscripts of negative studies or such manuscripts were previously submitted and rejected in other journals.

A meta-analysis of the ramelteon trials showed that the effect size calculated on data from published trials was larger than the effect size based on the data from all trials (10.1 min decrease in sleep latency vs. 6.37 min, respectively). Therefore it can be concluded that relying only on data from published trials leads to an overestimation of the efficacy of medicinal products.
Insomnia medication: Do published studies reflect the complete picture of efficacy and safety?

In addition to the degree of selective publication, we examined whether information included in the scientific publications was consistent with the information in the EPARs (selective reporting). Previous studies examining end point reporting have found evidence of omitting or incomplete reporting of results in publications (6,9,15).

This study showed that with respect to insomnia trials, the primary end point of the trials was reported reliably in all examined publications. In contrast, in 80% of the publications additional secondary end points that were not included in the EPARs were reported in the publication. Interestingly, most of the results concerning these secondary end points were not statistically significant, but in majority of cases showed a numerical advantage in favour of the test drug. A common statement in such publications was that although not statistically significant, the result showed “a trend towards improvement”.

In 13% of the publications there were discrepancies in the numerical results reported compared to those in the EPARs. In majority of the cases, the results were shifted towards significance in favour of the test drug. These discrepancies are not common in publications, but they might affect the conclusions regarding efficacy of a product and the results of meta-analyses. The size of inconsistency was remarkable as numerical results of important end points, such as sleep latency or total sleep time that were not significant in the EPARs were reported as significant in the publications.

The publications also included additional or post hoc analyses that were not presented in the EPARs. All these additional analyses were in favour of the test drug. This finding confirms the results of Al-Marzouki et al. (19), who reported that additional subgroup analyses not defined in the study protocol often appear in publications.

To the best of our knowledge, none of the previous studies on selective reporting have examined the reporting of safety in publications. Alarmingly, within our dataset, safety information was lacking in 67% of the publications. This is considered to be a serious drawback, since for a realistic overview of the benefit–risk balance it is essential that both positive effects (efficacy) and negative effects (adverse events) are presented. An accurate benefit–risk balance is essential not only for the marketing authorization decision making, but also for meta-analyses, the development of clinical guidelines and for the information available to health care professionals and patients in the public domain.

One limitation of our study is its limited scope: only one class of medication and three products were examined. Nevertheless, the results are consistent with the findings for the anti-depressant trials and for trials with all types of medicinal products.
In conclusion, our study has demonstrated that publication bias exists for insomnia trials and that the positive trials are two times more likely to be published than the negative ones. This selective publication resulted in an overestimation of the effect by a factor of 1.6. Furthermore, publications of insomnia trials provide more information on secondary end points and post hoc analyses favourable to the study drug than EPARs while in contrast, information on safety is almost entirely omitted.

An evident problem with basing clinical guidelines on published studies is that the results from one trial cannot be viewed within the context of the entire investigation program. Therefore it is difficult to get a clear picture of the drug’s benefits and risks by reading individual articles describing single studies. The EPARs and FDA reviews present all studies within the application package. However these documents are not often used as a reference in reviews. One potential remedy for this deficiency could be for journals to require the authors to cross-reference their article to all other clinical studies (and their publications) from the same study program and provide a reference to the EPAR or FDA review. This would facilitate considerably the perception of the true efficacy and safety of a product. Using the EPARs and FDA reviews as a source of data for meta-analyses should also be considered, provided that the uniformity of the EPARs is first improved.

The importance of unbiased effect estimates and unbiased assessment of the benefit–risk balance is widely recognized and therefore several attempts have been made to reduce publication bias. Regulatory authorities worldwide have also re-assessed their policy on data publication. In the US several requests for obtaining all data sent to the FDA have been made based on the freedom of information act (20). The European Commission plans to make public phases III and IV trial data on the EU database of clinical trials, EudraCT, in the near future. This database will be an important additional source of clinical trial data for researchers, health care professionals and the public.
Insomnia medication: Do published studies reflect the complete picture of efficacy and safety?

References

20. Vitry A, Berete V, Garattini S. Drug regulation and the need for greater transparency. RAJ Pharma 2009;March 1-4
Insomnia medication: Do published studies reflect the complete picture of efficacy and safety?

Published melatonin, zaleplon and ramelteon studies:

melatonin:


zaleplon:
Ancoli-Israel S, Walsh JK, Mangano RM, Fujimori M. Zaleplon, a novel nonbenzodiazepine hypnotic, effectively treats insomnia in elderly patients without causing rebound effects. Prim Care Comp J Clin 1999; Psychiatry 1:114-20


Walsh JK, Fry J, Erwin CW, Scharf M, Roth T, Vogel GW. Efficacy and tolerability of 14-day administration of zaleplon 5 mg and 10 mg for the treatment of primary insomnia Clin Drug Invest. 1998;16:347-354


ramelteon:

