

Commentary

Pooling, meta-analysis, and the evaluation of drug safety

Michel Lièvre*, Michel Cucherat and Alain Leizorovicz

Address: Service de Pharmacologie Clinique, Faculté de Médecine Laënnec, rue Guillaume. Paradin, BP 8071, 69376 Lyon Cedex 08, France

E-mail: Michel Lièvre* - ml@upcl.univ-lyon1.fr; Michel Cucherat - mc@upcl.univ-lyon1.fr; Alain Leizorovicz - al@upcl.univ-lyon1.fr

*Corresponding author

Published: 19 March 2002

Received: 16 January 2002

Current Controlled Trials in Cardiovascular Medicine 2002, **3**:6

Accepted: 19 March 2002

This article is available from: <http://cvm.controlled-trials.com/content/3/1/6>

© 2002 Lièvre et al; licensee BioMed Central Ltd. Verbatim copying and redistribution of this article are permitted in any medium for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: The "integrated safety report" of the drug registration files submitted to health authorities usually summarizes the rates of adverse events observed for a new drug, placebo or active control drugs by pooling the safety data across the trials. Pooling consists of adding the numbers of events observed in a given treatment group across the trials and dividing the results by the total number of patients included in this group. Because it considers treatment groups rather than studies, pooling ignores validity of the comparisons and is subject to a particular kind of bias, termed "Simpson's paradox." In contrast, meta-analysis and other stratified analyses are less susceptible to bias.

Methods: We use a hypothetical, but not atypical, application to demonstrate that the results of a meta-analysis can differ greatly from those obtained by pooling the same data. In our hypothetical model, a new drug is compared to 1) a placebo in 4 relatively small trials in patients at high risk for a certain adverse event and 2) an active reference drug in 2 larger trials of patients at low risk for this event.

Results: Using meta-analysis, the relative risk of experiencing the adverse event with the new drug was 1.78 (95% confidence interval [1.02; 3.12]) compared to placebo and 2.20 [0.76; 6.32] compared to active control. By pooling the data, the results were, respectively, 1.00 [0.59; 1.70] and 5.20 [2.07; 13.08].

Conclusions: Because these findings could mislead health authorities and doctors, regulatory agencies should require meta-analyses or stratified analyses of safety data in drug registration files.

Introduction

The drug registration files submitted to European and US health authorities present overall safety data in an "integrated safety report," which takes into account all the clinical trials that were performed during the development of the new product. In the standard integrated safety report, rates of adverse events observed with the new drug, placebo or active control drugs are calculated by pooling data across the trials. Pooling, the simplest and most naively

intuitive way of summarizing the information from several clinical trials, consists of adding the number of events observed in a given group, and dividing the results by the total number of patients included in this group. For example, the number of headaches in the new drug group is obtained by adding the numbers of headaches reported in all groups that received the new drug in clinical trials, regardless of control group. To obtain the pooled rate of headache, this number is divided by the total number of

patients who received the new drug. Similar pooling is done in the placebo and, if applicable, active control groups. The proportion of events is thus obtained for each treatment group, and the relative risk of developing the event is the ratio of the rate in the active group to the rate in the control group. Because pooling focuses on treatment groups rather than on studies, this approach does not consider the validity of the comparisons and, therefore, is subject to a bias termed "Simpson's paradox in probability" (also known in epidemiology as "confusion bias" [1,2]. A more satisfactory technique involves combining the results of each trial, expressed for example as relative risks. This latter technique is used in stratified analyses, and, particularly, in meta-analyses.

The results of a stratified analysis or a meta-analysis can differ greatly from those obtained by pooling the same data, as shown by the following example. This phenomenon may hamper the accurate evaluation of safety by regulatory agencies during the drug registration process. We propose that safety be assessed using meta-analytic techniques.

Methods

Ideally, we would use an example from original material submitted by pharmaceutical firms to regulatory agencies. Such material, however, is confidential and cannot be found in publications. Instead, we provide numbers from a hypothetical, but not atypical, application (part IV of international drug registration dossiers), where a new drug is compared to 1) a placebo in 4 relatively small trials of patients at high risk for a certain adverse event (referred to as the "event"), and 2) an active reference drug in 2 larger trials of patients at low risk for this event. All trials were of the same duration, and the risk of developing the event is higher in the treated group. Sample sizes and number of events are presented in Table 1.

The results of the trials can be pooled as described above, and can also be analyzed by meta-analytic techniques. In this example, the meta-analysis was performed using the logarithmic mean of the relative risk weighted by the inverse of its variance [3]. The calculations were performed using EasyMA software [4].

Results

Pooling the data yields the following proportions (no. events/no. participants): new drug group (studies 1–6): 45/3,460 = 1.30%; placebo group (studies 1–4): 19/1,460 = 1.30%; active control group (studies 5–6): 5/2,000 = 0.25%.

Table 2 presents the relative risk for developing the event for the new drug compared with placebo or active control (and the respective 95% confidence intervals). P-values were calculated by Fisher's exact test for pooled relative risks. The results of the meta-analysis are presented with the usual association and heterogeneity p-values.

The two methods give very different results. The pooled analysis shows the same risk for the event with the new drug and placebo, although the risk was, in fact, greater in the new drug than in the placebo group in studies 2, 3 and 4 and the same in study 1 (Table 1). The pooled risk of event in the new drug group was more than five times that in the active control group, although the difference was much less pronounced in studies 5 and 6. Conversely, the meta-analysis showed that the rate of developing the event was significantly increased in the new drug group compared with the placebo group, and the event rate was increased only 2.2 times in the new drug compared with the active control group, a non-significant increase. The results of the meta-analysis agree with the results of the individual trials. The non-significant p-values testing for heterogeneity imply that the data show no evidence of inconsistency across the different studies.

Table 1: Results of four hypothetical placebo-controlled (1–4) and two active treatment-controlled (5 and 6) studies: number of events/number of patients in each group.

Study	New drug	Placebo-controls	Active-treatment controls	Relative Risk
1	2/60	2/60	–	1.0
2	9/400	5/400	–	1.8
3	10/500	5/500	–	2.0
4	13/500	7/500	–	1.6
5	5/1000	–	2/1000	2.5
6	6/1000	–	3/1000	2.0

Table 2: Relative risks of events and 95% confidence intervals obtained by pooling data or performing a meta-analysis.

	Relative risk (pooling)	Relative risk (meta-analysis)
New drug versus placebo	1.00 [0.59; 1.704], p= 1.00	1.78 [1.02; 3.12] association p = 0.04 heterogeneity p = 0.94
New drug versus active control	5.20 [2.07; 13.08], p = 0.00004	2.20 [0.76; 6.32] association p = 0.14 heterogeneity p = 0.84

Discussion

Practical consequences

Simpson's paradox arises when validity of the comparisons is ignored, and when there is a large imbalance of a factor at the different levels of the variable of interest [1,2]. In the above example, Simpson's paradox arose in the pooling process and caused a discrepancy between the pooling and the meta-analysis results. This discrepancy occurred because the risks of the populations included in the placebo- and active drug-controlled studies differed greatly, and the comparison considered treatment groups rather than studies. A similar imbalance may arise if diseases or disease stages are different, or if patients are recruited from different settings (e.g., in hospital for placebo-controlled trials and from the community for active drug-controlled trials). It can also be the case when variable follow-up duration is not taken into account.

These data are hypothetical, yet they reflect a number of actual small or medium-sized placebo-controlled efficacy studies, as well as larger studies comparing new drugs to already existing competitors. These hypothetical data are not unusual; similar data arise commonly in practice.

Although for some time, pooling has been considered inadequate for the assessment of efficacy [5], it still represents the usual approach for presenting safety data in the integrated safety report of application files submitted by the pharmaceutical industry to regulatory agencies. Cases similar to our example may, therefore, be encountered and may mislead the experts who evaluate these applications. This is particularly true for relatively frequent events for which the usual pharmacovigilance approach—based on imputation of individual cases of rare events—is unable to detect an increased risk. In the above scenario, although the absolute risk was not negligible, the numbers of events were small in each trial and the increase in risk would have probably been overlooked at the trial level.

Drug agency experts may perform meta-analyses or stratified analyses on safety data; however, this is not common practice, perhaps because of force of habit and/or the en-

suing workload, given the very strict time frame in which application files must be submitted.

Recommendation: the integrated safety report should change

The usual way of summarizing safety data in drug application files submitted by the pharmaceutical industry to regulatory agencies is potentially misleading. Pooled safety information about new drugs may underestimate the risk of some adverse events and overestimate others. In some instances, it may lead to approval of a drug with an unacceptable safety profile. Such oversights harm both physicians and patients, since safety data (in contrast to efficacy data) are seldom available in published clinical trial reports.

The end-users of drugs cannot verify the appropriateness of the safety information given by the Summary of Product Characteristics. They must rely on the assessment of the regulatory agencies. Meta-analysis, or equivalently, stratifying by study, is the only approach that can reliably summarize safety data from several clinical trials because it focuses on studies rather than treatment groups, and maintains the validity of the comparisons. A limitation to meta-analysis is found when data are very sparse, for example, when a high proportion of zero events exist in trials. In this situation, meta-analysis is biased, although it remains the best method to summarize the data. Nonetheless, statistical power would be so low that results could not be interpreted and the pharmacovigilance approach would, instead, be appropriate. Another potential pitfall would be to include in the same meta-analysis all trials that have assessed the new drug, regardless of the control group. Such a meta-analysis would be invalid, since it may combine trials that are not comparable and relative risks that cannot be interpreted in the same way. Regulatory agencies should require meta-analyses or stratified analyses of safety data for adverse events that are common enough to be reported in nearly all the clinical trials in an application.

Authors' contributions

M Lièvre conceived the original idea for this paper and drafted the present article, A Leizorovicz and M Cucherat made substantial contributions to writing and editing the article. Guarantor: M Lièvre

Acknowledgements

The authors thank Alison Foote for her editorial assistance in the preparation of this paper.

References

1. Simpson EH: **The interpretation of interaction in contingency tables.** *J R Statist Soc B* 1951, **2**:238-241
2. Julious SA, Mullee MA: **Confounding and the Simpson's paradox.** *Br Med J* 1994, **309**:1480-1481
3. Whitehead A, Whitehead J: **A general parametric approach to the meta-analysis of randomized clinical trials.** *Stat Med* 1991, **10**:1665-1677
4. Cucherat M, Boissel JP, Leizorovicz A, Haugh M, Easy MA: **A program for the meta-analysis of clinical trials.** *Comput Method Program Biomed* 1997, **53**:187-190
5. Egger M, Davey Smith G, Phillips AN: **Meta-analysis: Principles and procedures.** *Br Med J* 1997, **315**:1533-1537

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMedcentral will be the most significant development for disseminating the results of biomedical research in our lifetime."

Paul Nurse, Director-General, Imperial Cancer Research Fund

Publish with **BMC** and your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours - you keep the copyright

Submit your manuscript here:

<http://www.biomedcentral.com/manuscript/>



BioMedcentral.com

editorial@biomedcentral.com