Cardiovascular safety of drugs not intended for cardiovascular use: need for a new conceptual basis for assessment and approval

Jeffrey S. Borer1*, Hubert Pouleur2, Eric Abadie3, Ferrenc Follath4, Janet Wittes5, Marc A. Pfeffer6, Bertram Pitt7, and Faiez Zannad8

1Division of Cardiovascular Pathophysiology, Weill Medical College of Cornell University, 47 East 88th Street, New York, NY 10128-1152, USA; 2Pfizer, Inc., New York, NY, USA; 3AFSSAPS, Paris, France; 4University Hospital, Zurich, Switzerland; 5Statistics Collaborative, Washington DC, USA; 6Brigham and Women’s Hospital, Boston, MA, USA; 7University of Michigan, Ann Arbor, MI, USA; and 8Hypertension and Preventive Cardiology Division, Department of Cardiovascular disease, Centre d’Investigations Cliniques INSERM-CHU, INSERM U684, Nancy, France

Received 4 December 2006; revised 24 April 2007; accepted 3 May 2007; online publish-ahead-of-print 5 July 2007

KEYWORDS
Cardiovascular diseases; Safety; Drug evaluation

Introduction

All drugs have multiple pharmacological actions, some beneficial and some detrimental. The perceived relation between benefit and risk, defined from pre-approval studies, forms the basis for drug approval and drug use. Pre-approval requirements must enable (i) delivery of effective and acceptably safe therapy to the public as quickly as is reasonable, (ii) information for prescribers sufficient to assure administration plans that will maximize benefit and minimize risk, and (iii) fairness to drug manufacturers to maintain incentives for drug development. The resulting process yields a body of information that, at drug approval, is unlikely to define all clinically important risks associated with drug use, especially those involving rare events that are difficult to detect after the relatively modest mandatory pre-approval drug exposure but may appear when a drug is in widespread clinical use. Indeed, pharmacological actions of a single drug can differ among functionally different organ systems; benefit for non-cardiovascular targets may be associated with cardiovascular adversity. However, cardiovascular testing is not a major focus of evaluation and development of non-cardiovascular drugs. Consequently, at approval, the relation between non-cardiovascular benefit and cardiovascular risk may not be adequately defined.

Deficiencies of pre-approval data are exemplified by issues emerging from the withdrawal of two non-steroidal anti-inflammatory drugs (NSAIDs), both cyclo-oxygenase (COX)-2 selective.1–4 These actions followed several publications associating NSAIDs with cardiovascular event rates higher than observed among non-drug users, and highlighting the effects of COX-2 selective agents, the most recently approved members of this drug group.5–11 These deficiencies are difficult to remedy within current regulatory algorithms, which create barriers to detecting relatively uncommon adverse events before approval. The problem is illustrated in Table 1, with absolute event rates taken from the Adenoma Prevention with Celecoxib (APC) trial (n = 2035 patients). A clear relative risk gradient was apparent between placebo and the two doses of celecoxib. However, the number of patients employed in this post-approval trial exceeded the total required patient exposure needed to meet pre-approval regulatory requirements. Current laws do not permit regulatory agencies to mandate such post-approval studies; current regulations limit mandatory pre-approval exposure to limit unfair development costs in the absence of a compelling evidentiary basis (see in what follows). Given the relatively modest absolute risks observed in APC, an a priori commitment from the manufacturer (revocable at any time under current laws) would have been needed to assure a post-approval trial with power adequate for likely detection of the relative risks actually observed. Thus, it is fair to question whether we need a
new conceptual basis for assessment and approval. This article provides some responsive suggestions, developed at a meeting of cardiovascular clinical trialists, biostatisticians, FDA and European Medicines Agency (EMEA) regulators, representatives of the National Institutes of Health (NIH), and pharmaceutical industry scientists in Versailles, France, in December 2005.

Background

The NSAIDs are not the first drugs for non-cardiovascular targets evidencing unexpected cardiovascular risk after market release. For example, in 1997, fenfluramine and dexfenfluramine were withdrawn when data suggested they could cause cardiac valvular disease.12-16 The following year, terfenadine was withdrawn because of association with fatal cardiac arrhythmias when it was administered with cytochrome p450 enzyme system inhibitors;17,18 2 years later, cisapride was withdrawn upon reports of cardiac arrhythmias and deaths.19

Thus, new approaches are needed to document cardiovascular effects of new pharmacological agents not intended for cardiovascular targets. However, potential pathophysiological links may not be evident between cardiovascular events and drugs prescribed for non-cardiac conditions, particularly when clinical trials for the primary disease targets are short-term or if the relevant populations already are at risk for cardiovascular disease because of age, systemic inflammatory diseases, or other conditions. Even in the pre-approval period, detection of cardiovascular events during rigorous clinical trials of ‘non-cardiovascular’ drugs generally relies on spontaneous reporting of adverse events by investigators not trained in cardiology, who may not recognize subtle cardiovascular complaints. Most importantly, as noted earlier, pre-approval patient exposure requirements provide scant statistical power to detect differences in rates of infrequent events between the new drug and comparators, especially if pre-treatment cardiovascular risk is low. In the absence of organized trials, post-approval detection depends on spontaneous cardiac event reporting by practitioners (largely non-cardiologists), a notoriously insensitive and inefficient approach.

Because of these considerations and experiences, the United States Food and Drug Administration (FDA) is likely to change its process for evaluating drug safety.20 Optimally, such change should define the type and magnitude of cardiovascular risks that must be detected before marketing to adequately protect public health. If pre-approval studies can assure that these important adverse events occur at less than a pre-specified rate, further information may not be necessary prior to approval,21 though post-approval data enhancement may be useful. Concomitantly, the scientific rigor required to establish the association between a new drug and cardiovascular risk must be established, i.e. when must the relevant data come from prospective, randomized clinical trials, when will registries (voluntary inclusion or mandated consecutive patient series) suffice, and when will retrospective prescription-based data sets be adequate.

### Classification of cardiovascular risk

The drug development process aims to define the benefits and risks of an investigational agent for a specific population. Results may be skewed by patient selection criteria in pre-approval clinical trials, which generally are designed to eliminate characteristics that might confound interpretation of the primary study outcome, e.g. childhood or advanced age, pregnancy, or patients with various comorbidities (thus optimizing likelihood of observing a drug effect if it exists). The resulting cohort can differ importantly from the population that will receive the drug after approval. In addition, the duration of observation in clinical trials commonly is shorter than the prolonged exposure in clinical settings. Regulatory agencies often require trial durations in part based on the presumption that, since spontaneous reports of major problems in long-term use generally have been rare in the past, mandatory prolongation of trials is not needed. However, as earlier, recent experience belies this presumption.

Prolonging pre-approval trials or increasing sample sizes to detect rare events during a relatively short interval have potentially important drawbacks. The costs of trials can increase dramatically with duration. However, patent life, the basis upon which drug manufacturers can recoup their development costs, is relatively short. Increasing the size of study cohorts may maintain study duration but also is associated with increasing costs of patient evaluation and of accelerated recruitment. Consequently, decisions to prolong and/or enlarge pre-approval trials to detect relatively infrequent potential adverse effects must be based on a clear determination of the importance to the public health of the increment in information that can be obtained, based on such factors as the number of patients likely to use the drug after approval, the potential for use in high-risk populations, and the biological plausibility of cardiovascular risks.

There is general consensus on the nature of major cardiovascular risks. These comprise three broad categories: cardiovascular death or major irreversible morbidity (e.g. non-fatal MI and stroke); debilitating but reversible cardiovascular symptoms or events (e.g. fluid retention sufficient to cause dyspnoea, palpitations, lightheadedness, syncope due to non-lethal arrhythmias, angina, transient ischaemic attacks); and pathophysiological characteristics that increase the likelihood of cardiovascular adverse events.

### Table 1  Sample sizes [as a function of power (type II error)] to detect cardiovascular risk in a randomized, controlled clinical trial, assuming risk in the placebo group is 0.33%/year, and each participant is followed for 1 year (two groups, equal n/group, type I error = 0.05; two-tailed test)21

<table>
<thead>
<tr>
<th>Ratio of event rate (treatment group:placeboa)</th>
<th>Power</th>
<th>80%</th>
<th>85%</th>
<th>90%</th>
<th>95%</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>995 500</td>
<td>1138 800</td>
<td>1332 700</td>
<td>1648 100</td>
<td></td>
</tr>
<tr>
<td>1.2</td>
<td>260 700</td>
<td>298 100</td>
<td>349 000</td>
<td>431 600</td>
<td></td>
</tr>
<tr>
<td>2.0</td>
<td>13 500</td>
<td>15 500</td>
<td>18 100</td>
<td>22 364</td>
<td></td>
</tr>
<tr>
<td>3.0</td>
<td>2 600</td>
<td>5 300</td>
<td>6 200</td>
<td>7 700</td>
<td></td>
</tr>
</tbody>
</table>

Estimation of sample sizes based on normal approximation to binomial.

aEvent rates in the APC trial, 0.67%/year (drug) and 0.33%/year (placebo), see text. Calculations per van Belle et al.21
(e.g., hypertension, accelerated thrombogenesis, asymptomatic arrhythmogenesis). \textsuperscript{22-24}

**Surrogate measures of risk**

Cardiovascular events can be determined only by direct clinical observation. However, potentially important pathophysiological changes can be measured objectively to predict events. The appropriateness of such pathophysiological surrogates for determining cardiovascular benefit is controversial. At present, for example, the FDA accepts only blood pressure reduction in hypertensive patients and LDL reduction in hypercholesterolemic patients as surrogates for cardiovascular clinical benefit. In assessing adverse effects, however, greater latitude is accepted. For example, both the FDA and the EMEA mandate measurement of the electrocardiographic QT interval during drug development, expecting that its substantial increase predicts potentially lethal arrhythmia. The International Conference of Harmonisation (ICH) documents E14 and S7B emphasize the need for thorough assessments of the QT interval during pre-clinical and clinical testing. \textsuperscript{25-27} supported by \textit{in vivo} cardiovascular safety studies and \textit{in vitro} electrophysiological studies, and generally to include a dedicated study specifically to assess drug effect on cardiac repolarization. \textsuperscript{26,28}

Inferences concerning adverse events from such data can be overcome only with trials that measure clinical events directly.

Nonetheless, surrogates are imperfect predictors of adversity. For example, most drugs associated in humans with the potentially lethal arrhythmia, torsades de pointes (TdP), inhibit the rapidly activating delayed rectifier potassium current in myocardial tissues. \textit{In vivo} or \textit{in vitro} studies that identify this activity may be useful for determining the potential for drug-induced TdP; however, not all drugs that cause QT prolongation act through the delayed rectifier, and not all will result in TdP. \textsuperscript{29} Also, clinically important QT abnormalities may not be apparent pre-approval, even though they may occur in some patients after marketing, possibly due to unpredictable drug–drug interactions that emerge only with multiple drug therapies that were precluded in pre-approval trials.

Safety surrogates may indicate risk that may not be actualized in all patients. Nonetheless, even an incompletely predictive surrogate may preclude approval if other therapies can provide acceptable benefit without the risk of concern. On the other hand, even a moderate risk might be acceptable if relatively greater benefit is perceived, though approval on this basis should lead to well-designed post-marketing studies to evaluate mortality and major morbidity. Thus, safety surrogates must be assessed case by case, and structured post-approval monitoring is irreplaceable for complete assessment of drug safety.

**Timing of risk assessment**

Detection of the potential for a cardiovascular safety problem in pre-clinical or early clinical studies should lead to detailed pre-approval definition of acceptable cardiovascular safety if a drug is to be approved for non-cardiovascular targets.

Pre-approval assessments of drug-associated risks commonly suffer from inadequate statistical power because of the relatively modest pre-approval exposure requirements imposed by regulatory agencies. Currently, post-approval assessments usually suffer from inadequate controls, from cohorts not representative of the population actually using the drug, and from inadequate power to identify adverse events. Brass et al.\textsuperscript{21} have proposed an approach to resolving the latter problem by basing safety evaluations not on point estimates but rather on the upper bounds of confidence intervals circumscribing the point estimate, allowing predefinition of the unacceptable risk level and accrual of sufficient events to provide power adequate to assess whether risk is within the pre-determined acceptable range. If the confidence limit is acceptable relative to the demonstrated benefit of therapy, approval may be justified to avoid delaying availability of effective therapy. The approval letter and/or label could state the need for post-marketing assessments for increasingly precise definition of cardiovascular risk. Patient exposure needed to generate the required events could be gathered by extending safety follow-up beyond that required for the primary endpoint in pre-approval trials, eliminating the need to design and conduct new studies or to enrol additional patients in pivotal trials. \textsuperscript{21}

**Acceptable mechanisms to evaluate cardiovascular risk**

Drug development should incorporate cardiovascular safety assessment beginning with pre-clinical/animal studies of drug effects on cardiac physiology/pharmacology. Phase 1 and 2 trials should evaluate pre-specified parameters including, but not limited to, coagulation profiles, blood pressure, blood lipid concentrations, left ventricular size and function, body weight, cardiac dimensions, QT intervals, and other ECG parameters and, to the extent feasible, interactions between drug effects and specific genomic or proteomic characteristics. Data generated from these studies provide the foundation upon which Phase 3 trials can be designed.

Most Phase 3 trials of ‘non-cardiac’ drugs should pre-specify measurement of cardiovascular events of interest, clearly stating the diagnostic criteria by which these events will be identified, and mandating regular measurement of variables to detect clinically silent events. In addition, the protocol should define a method for combining the adverse experiences from all trials in the development program to enhance power to identify problems if they exist.

Given the relative frequency of the non-cardiac outcome events in many of these trials, large sample sizes often are not required for demonstrating superior benefit of a new agent vs. placebo if a beneficial drug effect exists. Since the need to avoid confounding influences in these relatively small populations commonly results in relatively young patients with low cardiovascular risk, the capacity to detect cardiovascular adversity is inherently limited even if such events truly are associated with therapy. For example, when placebo-treated cohorts have annual event rates for the pre-specified cardiovascular safety composite as low as that found in the APC trial (earlier), population sizes (and costs) for clinical trials would need to be several orders of magnitude greater than typical pre-approval
exposure for non-cardiac drugs (e.g. terfenadine = 5000 patients, fenfluramine = 340, dexfenfluramine = 120011) if trials were designed a priori to achieve power sufficient to establish clinically meaningful cardiovascular risk that is 1.5 to two-fold greater than placebo (Table 1). Enlarging trials to identify statistically significant differences in cardiovascular risk between a new non-cardiovascular drug and a comparator is rarely feasible and generally would represent a wasteful expenditure of research resources. Non-inferiority trial design also can be employed to assess relative safety; regulators and investigators then would need to agree upon the magnitude of difference between adverse event rates caused by the new agent and those caused by an established drug (or by placebo) that would be ‘clinically unimportant’. As in superiority trials, however, large numbers of patients are needed when adverse events are relatively uncommon (Table 2).

However, the absolute magnitude of risk reliably excluded by trial data can be defined from registries/observational studies, less rigorous than randomized controlled trials for comparisons of effects of different therapeutic strategies, but perhaps better suited for providing credible absolute risk estimates. Actual detection of relatively rare events is likely to require a prohibitively large sample size (Table 3) but excluding a risk of unacceptable magnitude can be achieved with population sizes likely attainable in practice. As projected in Table 4, such populations could provide a point estimate of absolute risk, with narrower and more credible confidence intervals, than now is commonly available, enabling more confident comparison of benefit and risk than currently is possible.

Observational studies are less intricate and costly and, therefore, often more practical, than randomized trials for defining absolute risk. Also, compared with randomized trials that must employ rigorous exclusion criteria to minimize factors likely to confound data interpretation, observational studies of absolute risk do not necessarily require such exclusions and, in fact, may avoid them to mimic the ‘real world’ population at risk. Unfortunately, the strength of the observational study also is its weakness: the lack of a comparator and the potential for multiple confounders limits the breadth of acceptable conclusions and can lead to erroneous inferences, declaring a drug acceptably safe when it is not, or unacceptably harmful when it is safe. Sample sizes for observational studies depend on the pre-defined magnitude of acceptable risk and on decisions about confounder exclusions.

To use observational studies for safety assessment, the design must be prospective, with pre-specified definition of cardiovascular outcomes. Retrospective assessment may fail to provide documentation of variables of interest. When comparative risks are required, randomized trials may be the only acceptable design. The strength of evidence inferable from the study design must be considered when evaluating drug safety.

Large insurance databases are often used in observational studies to evaluate drug safety. These databases contain thousands of patient records with up to several years of experience.
follow-up. Adequately powered studies using these databases can be prospectively designed with strict procedures for data collection and verification. These databases can be excellent tools for evaluating the long-term safety of cardiovascular and non-cardiovascular drugs in much larger populations than would be feasible in prospective randomized trials, though inferences from them are limited by lack of information about doses actually used vs. those prescribed, and by imprecision or inaccuracy in diagnoses and other patient characteristics, since these generally are not entered directly or checked by treating physicians.

Regulatory impediments to post-marketing safety assessment

Currently, the FDA can request post-marketing studies, but can legally require them only for new drugs approved under Subpart H, granted on the basis of a surrogate endpoint or when approval carries restrictions to ensure safe use. The EMEA has parallel limitations. To improve post-marketing safety assessment, all regulatory agencies must be empowered to mandate adequately designed post-marketing studies as a condition of approval, with the capacity to withdraw approval if the sponsor does not perform the mandated studies. Drug approval should be viewed as the initial, not the final, step in defining appropriate drug use. The drug label developed at the time of approval should be subject to continual refinement in response to results of post-marketing studies. Post-marketing experience should not be primarily a basis for withdrawing marketing approval but rather a basis for continually improving the precision of information to prescribers/users about risks (and benefits) to be considered in employing a therapy. In the early post-approval period, a label could include a warning that “the cardiovascular risk has not been fully evaluated” with a numerical statement about the level of risk that cannot be excluded, to be removed when risk is more fully examined. The label could also indicate the date by which post-marketing evaluations are anticipated.

Other considerations

Dose-response relations for safety end-points may differ quantitatively from those for efficacy. In addition, the exposure time required to observe an adverse cardiovascular effect is not known; it is likely to be different for different drugs. Drug-related risks may vary with the pre-therapy cardiovascular risk of the patient. For example, a drug may have minimal absolute cardiovascular risk in a young patient without comorbid conditions, but it may be associated with far greater absolute risk in a 70-year-old with known coronary disease and diabetes. Thus, for non-cardiovascular drugs, it may be useful to perform at least one efficacy study in patients with relatively high cardiovascular risk if the drug will ultimately be used in this population; this approach requires careful safety monitoring during the trial. However, without empirical data, it cannot be known if findings from a high-risk population (either positive or negative) can be extrapolated to lower risk patients. Finally, all risks may not be evident even if the processes for assessing cardiovascular safety are revised.

Conclusion

Harms identified after drug approval during the past decade suggest the need for new approaches to evaluating potential cardiovascular risk in the context of non-cardiovascular therapy. Unfortunately, currently feasible strategies cannot guarantee detection and quantitation of all major adverse drug effects; biological variability will continue to provide uncertainty, which will be minimized only with pharmacogenomic and other novel initiatives. Nonetheless, in the present, we can and must improve upon efficiency in defining the risk:benefit relations of non-cardiovascular drugs. This will happen only if patient exposure, in number and duration, is increased in pre- and post-approval studies sufficiently to enable exclusion, with reasonable certainty, of unacceptable cardiovascular risks. Strategies to achieve this goal, outlined in this article, may be useful.

Appendix

The following individuals participated in the December 2005 8th Cardiovascular Clinical Trialsist Workshop:

E.A., MD; Kirkwood F. Adams, MD; Corine Bernaud, MD; J.B., MD; John Cleland, MD; Rory Collins, MBBS; Nicolas Danchin, MD; David DeMets, PhD; Ferenc Follath, MD, PhD; Nancy Geller, PhD; Mathieu Ghadafar, MD; Sidney Goldstein, MD; David Gordon, MD, PhD, MPH; Peter Held, MD; H.M. James Hung, PhD; Desmond Julian, MD; Bridget-Anne Kirwan, PhD; Alain Leizorovicz, MD; Richard Lewis, PhD; Raymond Lipicky, MD; Alice Mascette, MD; M.A.P., MD, PhD; B.P., MD; Stuart Pocock, BA, MSc, PhD; Philip Poole Wilson, MD; H.P., MD, PhD; Edmond Roland, MD; Denise Simons-Morton, MD, PhD; Scott Solomon, MD; Christian Torg-Pedersen, MD; J.W., PhD; F.Z., MD, PhD.

Conflict of interest: The symposium was supported by an unrestricted educational grant from Pfizer, Inc., New York, NY, USA. J.S.B. receives an honorarium as member of a Data and Safety Monitoring Committee and a blinded Events Adjudication Committee for two Pfizer-sponsored trials of non-cardiovascular drugs. He is also a member of the Executive Committee for PRECISION, a comparison of several non-steroidal anti-inflammatory drugs, funded by a grant from Pfizer to the Cleveland Clinic Foundation; J.S.B. receives no remuneration for this work. H.P. is a Pfizer employee, and holds stock in Pfizer. E.A. has no relations to report. F.F. participates on a cardiovascular drugs advisory board on for Pfizer Switzerland. J.W. has no relations to report. M.A.P. reports having received honoraria from Pfizer; and served as a member of the Data Safety Monitoring Board for the National Cancer Institute Contract no. N01-CN-95015, Prevention of Sporadic Colorectal Adenomas with Celecoxib, which was funded by a grant from the Strang Cancer Prevention Center. B.P. is a consultant to Pfizer. F.Z. receives research support from Pfizer.

References


Appendix
Cardiovascular safety of non-cardiovascular drugs


