Is Randomization to Placebo Safe? Risk in Placebo-Controlled Angina Trials: Angina Risk Meta-Analysis

Stephen P. Glasser, James Willard, Albert Defelice, John Lawrence, James Hung, Evelyn Obot, John Girton, Shari Targum, Douglas Throckmorton, Dennis Mangano, Raymond J. Lipicky

University of Alabama at Birmingham, Birmingham, Ala., Food and Drug Administration, Rockville, Md., National Institutes of Health, Bethesda, Md., Ischemia Research and Education Foundation, San Francisco, Calif., Lipicky LLC, North Potomac, Md., and 6030 N. Moon Crest Dr., Tucson, Ariz., USA

Key Words
Angina • Arrhythmia • Congestive heart failure • Placebo • Stroke • Transient ischemic attack • Ventricular tachycardia

Abstract
Objective: It was the aim of this study to document the risks of symptomatic patients with angina in placebo-controlled, anti-anginal drug development trials in which symptom-limited exercise testing was used as the primary endpoint. Patients and Methods: The original case report forms submitted to the United States Food and Drug Administration in support of approval of new or supplemental new drug applications between 1973 and 2001 were identified and subjected to a by-patient meta-analysis, utilizing both a maximum likelihood analysis and classical Mantel-Haenszel methods. Results: There were 63 placebo-controlled, clinical trials that randomized 10,865 patients, with 1,047 patient-years of observation time. The trials involved 21 different chemical entities from 4 different drug classes. The relative risk (RR) for withdrawal (placebo compared to drug-treated patients) was not increased [RR = 0.92, 95% confidence interval (CI) 0.78–1.08; p = 0.28]. Of interest, a RR of 0.54 (95% CI 0.26–1.04; p < 0.068) for irreversible harm (a combination of cerebrovascular accidents, myocardial infarction and death) and a RR of 0.89 (95% CI 0.61–1.30; p = 0.56) for serious cardiovascular events (myocardial infarction, congestive heart failure, cerebrovascular accidents) both non-statistically significantly favored being randomized to placebo. Conclusions: For the development of current or future drugs for the treatment of angina, there is no obvious contraindication to the use of placebo controls and exercise tolerance testing.

Introduction

Anti-anginal drug development (intended for symptom relief) has usually involved obtaining data from at least 2 placebo-controlled clinical trials that demonstrated decreased angina as assessed by exercise tolerance testing (usually by showing an increase in exercise time before developing angina) as well as anti-ischemic effects (usually by showing an increase in exercise time before developing ST segment depression). In such trials, there have been ethical and safety concerns over the use of pla-
cebo as a control (as opposed to conducting a non-inferiority trial against some approved drug), but very little published data [1–10].

Because of the legal requirement that complete case report forms (CRFs) for any dropout, for any reason, from any trial be submitted at the time of a new drug application (NDA) or supplemental new drug application (SNDA) to the United States Food and Drug Administration (FDA), the FDA is in a unique position to provide data from such placebo-controlled trials. Recognizing the importance of making unique data available, in 1991, we published a by-patient meta-analysis from the then available NDA/SNDA database that involved only 2,921 patients [11]. Subsequently, we undertook and published a meta-analysis evaluating the risks of placebo-controlled antihypertensive trials during which we evolved better methodologies for data collection and analyses [12]. In 2001, since the NDA/SNDA database had expanded to 10,865 randomized patients, we felt that an update of the 1991 analysis was appropriate.

Since 2001, there has been one nitroglycerine spray that was NDA submitted and approved. That trial involved only 30 patients and is not contained in this meta-analysis. Ranolizine was approved in 2006, but the original NDA submission which contained all of the randomized, placebo-controlled, exercise trials supporting ranolizine approval was available prior to 2001 and is included in this meta-analysis. There have been no other NDAs or SNandas submitted to the FDA. Thus, although 10 years have elapsed since the date of the last NDA analyzed, this meta-analysis as of 2011 contains 99.725% of all currently (August 2011) available data.

Methods

Screening for Relevant Trials

We identified all NDAs or SNDAs involving the treatment of angina pectoris, and within those applications, identified only randomized, double-blind, placebo-controlled trials. Because these data had been submitted to the FDA only for the purpose of regulatory approval, written permission to include all trials in this meta-analysis was obtained from each individual sponsor of each application, all of whom complied.

Definition of Dropout – Classification Criteria and Combinations of Primary Events

We predefined 12 categories to which patients withdrawn from trials were assigned, similar to the procedure used in an examination of placebo-controlled trials of antihypertensive drugs, but tailored for studies of chronic stable angina (table 1) [12]. Three combinations of events were also prospectively defined: (1) total serious events, (death, unstable angina, arrhythmia, congestive heart failure (CHF), cerebrovascular accident (CVA), myocardial infarction (MI), transient ischemic attack (TIA)), (2) serious cardiovascular events (unstable angina, arrhythmia, CHF, MI), and (3) serious neurological events (CVA and TIA).

Three retrospective analyses of combinations of events were also performed: all dropouts (except for administrative and therapeutic failure), serious cardiovascular events (MI, CHF, CVA) without unstable angina, and a combination of events that included only documented events of irreversible harm (defined as death, MI, CVA).

Review of CRFs

Under FDA regulations (21 CFR 314.50), a NDA or SNDA must include complete CRFs (for every study visit) for every patient who died or left the study prematurely because of an adverse event. If the patient dropped out of a trial, the investigator usually indicates why the patient was withdrawn.

We retrieved all CRFs (for every study visit) for all reported deaths and all dropouts (for any reason) from all of the trials included in our study. For crossover trials, because of potential carryover effects from the first period, we retrieved the CRFs only for the first period of the crossover. When CRFs were not present in the archived NDA or SNDA submission, duplicates were obtained from the drug sponsor.

The details of the analytic methodology have been reported previously [12]. Briefly, two blinded, non-physician health professionals independently entered data from the CRFs into their own database via a web-based form. From the CRFs, each of the 2 reviewers (by patient and by study) separately generated a 77-field Excel spreadsheet record that included the patient’s demographics, blood pressures at baseline and, at last routine visit, any adverse events observed, and whether the dropout had been associated with admission to an emergency room (ER) or a hospital. The reviewers independently assigned the patient’s primary reason (only one per patient) for withdrawal to 1 of the 12 categories as listed in table 1. Reviewers were instructed that if ≥ 2 categories seemed equally applicable, the more serious category (e.g. death > MI > unstable angina > therapeutic failure; in some cases, it was a judgment call left up to the individual reviewer) should be chosen as the primary reason. Some convergence was given to the assessment that had been made by the investigator, but reviewers were free to assign a different category when it seemed to be more consistent with the other data in the remainder of that patient’s CRFs, or in the definitions provided in the protocol.

A committee of 3 medical reviewers resolved differences between the 2 reviewers for the primary reason for dropping out, by re-reviewing all CRFs. If the 3 medical reviewers could not reach consensus, the primary reason for dropping out was decided by a 9-physician dispute resolution executive committee. Other discrepancies were resolved by the data manager (J.W.) by going back to the CRFs and examining where the discrepancy occurred (e.g. birth dates, cholesterol values, blood pressures). Discrepancies were found by Excel macros developed for reconciling the data from both reviewers.

Statistical Analysis

This is a by-patient (not by-study) meta-analysis, and our primary measurement estimation was the relative risk (RR, defined...
as placebo/investigational drug) of dropout for any cause. Thus, RRs >1 favor investigational drug (i.e. more events occurred in the placebo group) and RRs <1 favor placebo (i.e. more events occurred in the investigational drug group).

For RR, p values were determined without correction for multiple comparisons. If Bonferroni corrections were applicable, a p value equivalent to a standard p of 0.05 would need an uncorrected p of about $2 \times 10^{-3}$, and that of a standard p of 0.01 would need an uncorrected p of about $4 \times 10^{-4}$.

We anticipated that most events of importance would be of low incidence and would not occur at all in many of the included trials. Assumption of an asymptotic normal model might then have led to meaningless confidence intervals (CIs) and other anomalies [13], so our primary methodology was the maximum-likelihood method stratified by study. For completeness, we also used the conventional method of Mantel and Haenszel [14] (and found similar results compared to the maximum-likelihood method).

Table 1. The 12 predefined endpoints

<table>
<thead>
<tr>
<th>Endpoint</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unstable angina pectoris</td>
<td>If the patient progresses from stable, exertional angina to a diagnosis of unstable angina</td>
</tr>
<tr>
<td>Arrhythmia</td>
<td>Any arrhythmia other than ventricular tachycardia that led to dropout. Non-sinus bradycardia/tachycardia, as well as 1st to 3rd degree ativoventricular block are classified as arrhythmia. Unspecified or sinus bradycardia/tachycardia or 'sustained tachycardia' are classified as other cardiovascular dropout.</td>
</tr>
<tr>
<td>CVA</td>
<td>Neurological changes (e.g. limb weakness or paralysis, aphasia, blindness) that are permanent as a result of either cerebral ischemia, infarction or hemorrhage. If sufficient evidence exists that neurological changes are permanent, we may classify as a CVA even if the investigator did not (see also TIA).</td>
</tr>
<tr>
<td>CHF</td>
<td>Typical signs/symptoms of either left- (e.g. pulmonary congestion) or right-sided (e.g. venous congestion) failure. All cases must be labeled CHF (either new onset or decompensation) by the investigator.</td>
</tr>
<tr>
<td>Death</td>
<td>While the subject was actively enrolled in the trial, or within 24 h of discontinuation or completion of the trial.</td>
</tr>
<tr>
<td>MI</td>
<td>Transmural and non-Q-wave infarctions. If a dropout occurs at the time of hospitalization to 'rule out' infarction and insufficient information is available to determine the outcome of hospitalization, the event is classified as an MI.</td>
</tr>
<tr>
<td>Other</td>
<td>Administrative, non-medical reasons for dropout.</td>
</tr>
<tr>
<td>Other adverse events</td>
<td>Non-cardiovascular events. Included here is chest pain that is not described as anginal.</td>
</tr>
<tr>
<td>Other cardiovascular events</td>
<td>Any cardiovascular symptoms or signs, other than the above (e.g. orthostatic hypotension, sinus bradycardia/tachycardia, dependent edema, cardiomegaly, asymptomatic electrocardiographic changes) leading to dropout.</td>
</tr>
<tr>
<td>Therapeutic failure</td>
<td>If dropped by the investigator because of unresponsive angina. Note the difference from unstable angina pectoris.</td>
</tr>
<tr>
<td>TIA</td>
<td>Neurological changes (e.g. limb weakness or paralysis, aphasia, blindness) that are temporary, with no residual impairment and resolution well within 48 h. If a dropout occurs at the time of hospitalization with a presumptive diagnosis of TIA or CVA and insufficient information is available to determine the outcome of hospitalization, the event is classified as a TIA.</td>
</tr>
<tr>
<td>VT</td>
<td>Episodes of ventricular arrhythmia with at least three consecutive ventricular ectopic beats.</td>
</tr>
</tbody>
</table>

These represent the 12 prospective categories for the primary cause for dropout from any of the trials reviewed. Each patient had only one primary cause for dropout. There were a total of only 14 VT events, so this category was merged for the final analysis resulting in 11 prospective event categories.

After analysis of individual events, we conducted several analyses that aggregated events in some combination, including those prospectively defined in the protocol. Absolute event rates were calculated as the number of events per 1,000 patient-years for each group, and the 95% CIs for the difference in rates (placebo/drug) were calculated assuming a constant risk across protocols and time of observation.

Other aggregate statistics were evaluated using paired comparisons and one-way analysis of variance; standard descriptive statistics were used as appropriate. Except for the ad hoc spreadsheet software noted above, we used standard software packages, including SAS and JMP (SAS Institute, Cary, N.C., USA), Access (Microsoft Corporation, Redmond, Wash., USA), S-Plus (Insightful Software, Seattle, Wash., USA) and Stata version 10.1.
Results

Eligible Trials

Review of the FDA databases found 63 trials from a total of 21 different applications that met our inclusion criteria. There were 10,865 patients randomized in the trials (3,624 assigned to placebo treatment and 7,241 to investigational drug treatment, with 1,047 patient-years of observation time, i.e. 373 patient-years for placebo and 674 patient-years for drug). The mean duration of the trials was 38 days. Twenty-one different chemical entities from 4 drug classes were represented, including β-blockers (n = 7), calcium channel blockers (n = 9), nitrates (n = 4) and ranolazine (n = 3 trials). There were only 4 protocols that allowed for background therapy with a β-blocker or nitrate, and this involved 293 patients in the investigational arm and 283 patients in the placebo arm. Due to the small numbers, these were not separately analyzed.

Assignment and Collapse of Categories

Because of discrepant primary event categorization by the primary reviewers, approximately 7% of the dropout CRFs were referred to the dispute resolution executive committee. After un-blinding, we noted that only 14 ventricular tachycardia (VT) events had been recorded, so this category was merged into the arrhythmia category (34 events in total, 20 arrhythmias and 14 VT).

Dropout Rates

Table 2 shows the results for each prospectively defined event, the event incidences, the RR (placebo/drug) for each event, absolute event rates, and absolute event rate differences (placebo minus drug).

There were 606 dropouts (199 of 3,624 randomized to placebo and 407 of 7,241 randomized to drug; RR = 0.92, 95% CI 0.78–1.08) by maximum-likelihood calculations (nominal p = 0.28), and an absolute excess of 70/1,000 patient-years in the drug group (placebo minus drug). As would be expected, therapeutic failure was more common...
in the placebo group (RR = 2.86, 95% CI 1.82–4.76; p = 10^{-5}). The single other category similarly favoring drug was arrhythmia/VT (RR = 1.61, 95% CI 0.79–3.23; p = 0.19), but this was not statistically significant. Major statistically significant findings favoring being randomized to placebo were: other adverse events (RR = 0.56, 95% CI 0.40–0.74; p = 0.00092) and other cardiac events (RR = 0.31, 95% CI 0.13–0.62; p = 0.014). The combined endpoint of all but therapeutic failure and administrative dropouts (RR = 0.63, 95% CI 0.50–0.79) was also statistically significant (p = 0.00005). Of the 607 dropouts, 157 (26%) had unscheduled hospitalizations or ER visits which showed a trend (RR = 0.81; nominal p = 0.22) but was not statistically significant in favor of being randomized to placebo.

The mean age of the patients at the time of randomization was 63 years (median 63.8), 79% were male, 88% white, 5.7% black, 1.5% Hispanic, 1.7% Asian, 0.3% Indian (American) and 0.3% other, while 2.5% were unknown. The mean sitting cuff blood pressure was 137 and 82 mm Hg, systolic and diastolic, respectively. The median time to dropout was 17 days (mean 22.9), with no significant differences across drug classes or dropout categories (data not shown). Additionally, the RR and the time to dropout did not noticeably vary with age, sex, ethnicity or study year (data not shown).

A description of selected individual event categories and combined event analyses follows.

**Therapeutic Failure and Unstable Angina**

Therapeutic failure was expected to be more common in the placebo group, and this indeed was seen. Of the 76 dropouts classified as therapeutic failure, 26 were on active drug, 50 on placebo (RR = 2.86, 95% CI 1.82–4.76; p = 10^{-5}). There were 5 hospitalizations for therapeutic failure (7%), all in the placebo group. In 3 of these 5 hospitalizations, although increasing angina is described, 1 patient also had uncontrolled diabetes, another bradycardia and severe lethargy, and a third experienced a syncopal episode. Thus, only 2 of the 5 were hospitalized mainly because of angina alone. On the other hand, unstable angina (in contrast to therapeutic failure) would be expected to be of major concern for physicians, implying a possible change in the state of the patient, with an enhanced risk of MI; this, in contrast to therapeutic failure, more often leads to hospitalization. Of the 71 dropouts classified as unstable angina, 49 on active drug and 22 on placebo (RR = 0.86, 95% CI 0.51–1.41; p = 0.56, a point estimate that non-statistically significantly favors being randomized to placebo), there were 57 hospitalizations or ER visits (80% of the unstable angina population which non-statistically significantly slightly favored being randomized to placebo). There were only 2 hospitalized patients in this category that had symptoms or signs other than increased frequency or severity of angina as the reason for being hospitalized.

Thus, although, as expected, there were clearly fewer therapeutic failures on active drug, unstable angina did not show any statistically significant difference and, in fact, showed a trend favoring being randomized to placebo.

**Other Adverse Events**

Overall, the other adverse events category was the most frequent event type (183 events, 30% of all dropouts). The RR of 0.56 (95% CI 0.40–0.79; p = 0.00092) is not surprising, since drugs used to treat angina are known to have adverse effects. Twenty-three percent of patients in this
category had an unscheduled visit (hospital or ER). The adverse effects observed in this category were severe enough to not only cause discontinuation from study but, in 1 out of 4 cases, required an unscheduled medical visit.

Most patients in this category had more than 1 descriptive event recorded. For example, in a sample of the first 46 patients in this category, headache was most common, and in decreasing order, asthenia and lethargy, dizziness, laboratory abnormalities, skin rash, drowsiness, sexual dysfunction, as well as a variety of other events were observed.

**Administrative Failure**

The second most frequent reason for dropping out was administrative, and there was no statistically significant difference (RR 1.02, 95% CI 0.7–1.45; p = 0.91) in terms of being randomized to placebo or drug. The reasons for dropout in this category were protocol deviations, loss to follow-up, withdrawal of consent, non-compliance with dosing or scheduled study visits, relocation, transportation impediments, intercurrent illness, non-cardiovascular surgery, or emergent physical limitations from accidents. Of the 134 patients in this category, there were only 6 hospitalizations, all for intercurrent illness, and 2 ER visits, both for intercurrent illness.

**Other Cardiac Adverse Events**

Of the 52 patients with cardiac adverse events (other than angina and MI), 8 were on placebo and 44 on active drug treatment, 27% had more than 1 event (e.g. hypotension and dizziness) and the RR was 0.31 (95% CI 0.13–0.62; p = 0.0014). These events were postural hypotension (1 placebo, 3 drug), bradycardia and/or palpitations (1 placebo, 13 drug), hypotension (1 placebo, 12 drug), non-specific electrocardiography changes, and non-diagnostic chest pain (1 placebo, 5 drug), pitting edema of extremities (0 placebo, 7 drug), dizziness (1 placebo, 17 drug), syncope (1 placebo, 1 drug, both patients were hospitalized), and fatigue and shortness of breath (5 placebo, 9 drug). Additionally, there were unscheduled visits in the drug group. Overall, increases in the other cardiac events were predominantly due to edema, hypotension and bradycardia, and these were a common cause of unscheduled visits (too few for statistical analysis). There was 1 patient from the drug group who had severe bradycardia with periods of sinus arrest and resultant cardiac arrest requiring cardiopulmonary resuscitation and ultimately pacemaker therapy. In retrospect, he might better have been included in the arrhythmia category but was coded under other cardiac events.

**Arrhythmia**

There were 34 (15 placebo, 19 drug) arrhythmic events (RR = 1.61, 95% CI 0.79–2.23; p = 0.19). Of the 20 arrhythmic events, 11 were atrial fibrillation/flutter (7 placebo, 4 drug) resulting in 5 unscheduled visits (4 hospitalizations). Of the 14 VT patients (6 drug, 8 placebo), there were 4 unscheduled visits.

**Combinations of Events of Interest**

As previously stated, we prospectively categorized three combinations of events: total serious events (death, unstable angina, arrhythmias, CHF, CVA, MI and TIA), serious cardiovascular events (unstable angina, arrhythmia, CHF, MI), and serious neurological events (CVA, TIA). There were insufficient serious neurological events (CVA, TIA) to do a statistical analysis, but all 5 events occurred in patients on drug (2 CVAs and 3 TIAs).

As shown in table 2, all combinations of events, whether prospectively or retrospectively defined, gave point estimates that favored being randomized to placebo, but were not or were only marginally statistically significant. Notably, eliminating only administrative and therapeutic failure (a retrospective combined endpoint but one that includes most events – 396 events, 65% of all events), there was a highly statistically significant (p = 0.00005) difference, with an absolute net excess in the drug group of 171 per 1,000 patient-years. Serious cardiovascular events (unstable angina, MI, CHF or CVA) were non-statistically significantly more common in the drug than in the placebo groups (RR = 0.89, 95% CI 0.61–1.30; p = 0.56). We also retrospectively combined dropouts we labeled as ‘irreversible harm’, and there was a trend favoring placebo (death, stroke or MI; RR = 0.54, 95% CI 0.26–1.04; p = 0.068).

**ER/Hospitalizations**

There were 157 ER/hospital visits, with an approximately equal frequency in drug and placebo groups (1.4% placebo, 1.5% drug), with a RR (maximum likelihood) favoring placebo (RR = 0.81, 95% CI 0.57–1.14; p = 0.22).

**Discussion**

Although the data presented herein represent all of the deaths and dropouts reported over 38 years of anti-anginal drug development, there were only 606 such events (199 placebo, 407 drug), and 210 of these events were administrative or therapeutic failure dropouts. That leaves
only 396 events of interest, and all of those events are represented in the combined endpoint designated ‘all endpoints except administrative and therapeutic failure’. The RR for this combined endpoint statistically significantly favored being randomized to placebo (RR = 0.63, 95% CI 0.50–0.79; nominal p = 0.00005; a p value of approximately 0.001 as a Bonferroni-corrected p value, see Methods); the absolute risk was 171 patients/1,000 patient-years less than dropping out when randomized to drug (–171, –244 to –99).

All point estimates of RR (except for administrative, therapeutic failure, and arrhythmia plus VT) was <1.0, favoring being randomized to placebo. Irreversible harm (all-cause death, non-fatal MI and non-fatal stroke), including a total of 46 events (35 drug and 11 placebo), although small in number, had an upper 95% confidence bound of only 1.04. These data provide no convincing evidence that being randomized to placebo was harmful. Since 2001, only one placebo-controlled trial using exercise tolerance as an endpoint (involving only 30 patients) has been submitted as a NDA or SNDA to the FDA. Angina drug development has historically been devoted to symptom relief and improvement of ischemia, and the latest ACC/AHA guidelines [15] as well as ESC guidelines [16] are consistent with this. As such, the FDA has almost invariably required two trials for drug approval, each showing relief of symptoms and improvement of ischemia. Placebo-controlled symptom limited exercise tolerance trials were among the most frequently suggested metric as a development program was evolving.

It will be at least several decades before the database evaluated by our current report becomes substantially larger. Thus, for the next several decades, our current study represents the only data that will be available for making data-dependent decisions about the risks of placebo in these kinds of placebo-controlled trials. Interestingly, except for the current study and our initial report [11], both from an FDA database, there have been no reports related to the safety of performing placebo-controlled trials in patients with stable, symptomatic angina. Similarly, the safety of placebo-controlled antihypertensive trials has had sparse reporting [1, 10, 12], the largest also being from an FDA database [12]. In these analyses, angina (1,047 patient-years of follow-up) and hypertension (12,657 patient-years of follow-up) short-term trials using placebo for purposes of evaluating a new therapy is not apparently harmful. These data should prove useful for institutional review boards, ethicists and investigators for future trials.

Limitations

Publication bias is one of the most important considerations in performing a meta-analysis. This meta-analysis, publication bias is minimal since it is unlikely that there are any (certainly no more than a few) instances where a drug receiving an approval was discontinued before a NDA was submitted, of which the authors are unaware. In any case, were this circumstance to occur, it is likely that this would favor placebo since the reason for failure to develop the drug would likely have been due to the ineffectiveness or the adverse event profile of the drug. Ascertainment bias was not evident in our analysis except for the one recognized instance of an arrhythmic event, which was classified in the other cardiac event category. Also, we did not assess dropout rates during the washout periods of these trials. In the 1991 article, the washout dropout rate was 2.8% but, as was pointed out at that time, since there is no comparator during the washout period, one can draw no conclusions other than that the rate is seemingly low. Finally, we have no data on whether the short-term absence of treatment might affect future cardiovascular events. However, there is no suggestion that the anti-anginal treatment of chronic stable angina impacts cardiovascular events beyond symptom control.

In conclusion, although a larger sample size, different categories and more advanced analytic techniques were utilized, the results are similar to that reported in 1991. Thus, within the limitations of this meta-analysis, there were no safety issues related to patients with chronic stable angina pectoris being enrolled in the placebo arm of short-term clinical trials.

Acknowledgements

This study was enabled by (1) support from the personnel of the Ischemia Research Foundation (Dennis Mangano), (2) by the original idea by Stephen Glasser, which has led to a previous publication related to angina trials, and (3) by the entire staff and document control support of the Division of Cardiovascular and Renal Products, FDA.

Conflict of Interest

No conflicts of interest are reported.
References