Medical Decision Making with Incomplete Evidence—Choosing a Platelet Glycoprotein IIbIIIa Receptor Inhibitor for Percutaneous Coronary Interventions

James M. Brophy, MD, PhD, FRCP(c), FACC, Lawrence Joseph, PhD

Background. Medical decision making must often be performed despite incomplete evidence. An example is the choice of a glycoprotein IIb/IIIa (GP2b3a) inhibitor, a class of potent antplatelet medications, as adjunctive therapy during percutaneous coronary interventions (PCIs). GP2b3a inhibitor efficacy in reducing adverse outcomes has been well documented with multiple placebo-controlled randomized trials, but there is a paucity of comparative data about their individual equivalency. Substantial cost differentials are also present between the drugs.

Methods. A systematic review of the literature was performed to identify all randomized placebo-controlled trials of GP2b3a inhibitors as adjunctive therapy for PCI. Three complimentary methods were used to assist in decision making regarding drug equivalency. First, the data from the single direct comparative trial are analyzed from a Bayesian perspective. Next, prior information from other GP2b3a inhibitor trials in similar but not identical patient populations is incorporated. In the 3rd method, indirect comparisons of GP2b3a inhibitors are carried out using a hierarchical meta-analytic model of the placebo-controlled trials identified by the systematic review.

Results. A total of 12 randomized trials were identified involving 3 agents (abciximab, eptifibatide, tirofiban), but only 1 involved a direct comparison of 2 drugs (abciximab v. tirofiban). In contradiction to the original publication, the authors’ Bayesian analysis both without (method 1) and with (method 2) the inclusion of some prior information suggests a reasonable probability of equivalency. The indirect comparisons from all randomized placebo-controlled trials (method 3) also failed to provide support for superiority of any agent over the others.

Conclusion. Decision making with incomplete evidence is a difficult but frequently occurring medical dilemma. The authors propose 3 methods that may elucidate the process and illustrate them in the context of the choice of GP2b3a inhibitor for adjunctive therapy during PCI. Further data may or may not eventually lead to a different conclusion, but based on the evidence available to date, the authors’ 3 methods suggest clinical equivalency between GP2b3a inhibitors, in contrast to the initial conclusions from the single comparative randomized trial.

Key words: percutaneous coronary interventions; platelet glycoprotein IIbIIIa receptor inhibitor; decision making; Bayesian analysis.

Received 23 April 2004 from the Division of Cardiology, Department of Medicine, and Department of Epidemiology and Biostatistics, McGill University, Montréal, Québec. Revision accepted for publication 27 September 2004.

Address correspondence to James M. Brophy, Division of Cardiology, Royal Victoria Hospital, 687 Pine Avenue West Room R4.12, Montréal, Québec H2L 4M1, Canada; james.brophy@mcgill.ca.

DOI: 10.1177/0272989X05275156
Ischemic heart disease remains the leading cause of patient mortality and morbidity in Western countries. Percutaneous coronary interventions (PCIs) have become a widely accepted therapy for the symptoms of ischemic heart disease, with more than 600,000 angioplasties performed annually in the United States. PCI is a generally safe technique with low rates of mortality and morbidity, but uncontrolled plaque rupture may expose underlying plaque debris, stimulating platelet aggregation and thrombosis. This may mediate the complications associated with interventional procedures, including death and myocardial infarction (MI). Three GP2b3a inhibitors (abciximab, eptifibatide, and tirofiban) are commercially available, and although approved indications as well as individual properties are slightly different, all will suppress platelet aggregation by at least 80% at therapeutic levels. The majority of PCI procedures are now done under protection of these medications, but there has been only 1 direct comparative randomized trial. A narrow interpretation of “evidence-based medicine” might base the choice of agent on these results, but existing practice guidelines have not made specific recommendations as to the choice of drug, perhaps implicitly recognizing the paucity of data. Nevertheless, clinicians and health care managers must make decisions as to whether meaningful health benefits exist between the agents, thereby possibly justifying differences in acquisition costs.

As an illustrative example of decision making with incomplete evidence and as an extension of previous work using indirect comparisons to circumvent a lack of direct comparative trials, we present several different methods to assess the equivalency of GP2b3a inhibitors as adjunctive therapy for PCI.

METHODS

Literature Review

Following the “Users’ Guides to the Medical Literature,” our 1st step, after framing the question, was to find all available evidence. We performed a systematic electronic search for all randomized controlled trials of GP2b3a inhibitors as adjunctive therapy for PCI. Using PUBMED and the key words randomized controlled trial, angioplasty, glycoprotein, and inhibition, we identified 186 articles published up to 31 March 2004. Trials had to report the meaningful clinical end points of death and MI. After identification of all pertinent information and summarization using a hierarchical model, we applied the following 3 techniques to assist in data interpretation and decision making.

Method 1—Objective Bayesian Analysis

In method 1, an objective Bayesian analysis, the data from the only comparative trial are examined from a Bayesian perspective, but without incorporating any of the information available outside of this single trial. Standard statistical analyses of randomized clinical trials, including the original analysis reported for this trial, fail to provide a direct estimate of the probability of treatment superiority, the probability that a clinically meaningful difference exists, or the probability of clinical equivalence, for any given clinical cutpoint. A Bayesian analysis, whether or not formally incorporating prior beliefs based on previous trials or other evidence, permits the calculation of these important and highly clinically relevant probabilities, allowing for more lucid decision making. In this 1st analysis, we assumed a noninformative prior so that the final probabilities are determined almost exclusively from the observed data of the comparative trial. We used a simple beta-binomial model. As the sample size in the trial was large, a Haldane or beta(0,0) prior distribution was used for the 2 probabilities of interest, that is, the probability of a clinical event (death or nonfatal MI at 30 d) with either drug, and a binomial likelihood represented the information contained in the observed comparative data. Probability differences were derived by forming the difference between the 2 beta posterior densities, and probabilities of clinical interest were calculated as areas under the curve of this density.

Method 2—Bayesian Analysis

Incorporating Prior Information

In method 2, we further develop our Bayesian approach by including information from past trials in the form of a prior distribution. In particular, we use indirect evidence from placebo-controlled trials of GP2b3a inhibitor in acute coronary syndrome patients, a similar, but not identical, patient population to the PCI trials. This furnishes separate prior beliefs for the effectiveness of each of the treatments used in these trials compared to placebo. Taking the difference of the 2 treatment effects (i.e., tirofiban compared to placebo and abciximab compared to placebo) provides an indirect estimate of the tirofiban to abciximab difference that is of main interest. A normal distribution is then fit
to this estimate, which forms our prior distribution. These prior beliefs are then updated by the direct comparative data from the direct comparative (TARGET) trial via normal distribution updating, approximating the binomial likelihood function with a normal density. Because the sample sizes for all of these trials are quite large, normal approximations fit very well. With smaller trials, exact methods can be used. Because the prior information is derived from indirect evidence and on a population not identical to that in the target trial, one can argue that the prior evidence should be discounted. At 1 extreme, one would completely discount all previous studies, leading to an analysis identical to the objective Bayesian approach (method 1) described above. At the other extreme, there would be no discounting. We performed a sensitivity analysis whereby the degree of discounting of the prior information from the indirect studies is varied, leading to a range of posterior distributions from which probabilities could be calculated.

Method 3—Indirect Comparisons via Hierarchical Bayesian Meta-Analysis

In the 3rd method, we again perform an indirect comparison in which GP2b3a inhibitors are compared from the systematic review of the placebo-controlled trials using a Bayesian hierarchical meta-analytic model. At the 1st level, each subject in each arm of each trial is assumed to follow a binomial distribution, with separate probabilities of events for each arm of each trial. At the 2nd level, logarithms of the odds ratios among each study are assumed to follow a normal distribution. Separate meta-analytic models are run for all trials of abciximab versus placebo, and for the trials of eptifibatide or tirofiban versus placebo, so that the end result of the meta-analyses is 3 posterior distributions, each summarizing the effect of drug versus placebo in all available trials. Taking the ratio of these posterior distributions then produces a ratio of these odds ratios. Values larger than 1 indicate superiority of abciximab compared to eptifibatide or tirofiban.

RESULTS

Available Evidence

The systematic literature search found 11 randomized placebo-controlled trials of GP2b3a inhibitors as adjunctive therapy to PCI, 8 with abciximab, 2 with eptifibatide, and 1 with tirofiban. Despite the high-risk populations studied in these 11 trials, death rates have been remarkably low for both those receiving GP2b3a drugs (93 deaths in 10,421 patients, 0.9%) and placebo (103 deaths in 8124 patients, 1.3%). Comparing death rates in GP2b3a drugs to placebo results in a wide confidence interval for the odds ratio (OR 0.761, 95% CI [0.546, 1.07]), indicating an inconclusive result. Although not including any head-to-head comparisons of the 3 drugs under consideration here, these trials are of high quality with little evidence of selection, performance, or attribution biases. Therefore, indirect comparisons can be formed as described in method 3 above.

The systematic literature search identified only 1 head-to-head double-blinded randomized trial comparing 2 GP2b3a inhibitors in the setting of modern PCI with coronary stenting. This TARGET trial was designed to demonstrate the noninferiority of tirofiban as compared with abciximab. The end point (composite of death or nonfatal MI at 30 d) occurred more frequently among the 2398 patients in the tirofiban group than among the 2411 patients in the abciximab group (7.2% v. 5.7%; hazard ratio [HR] 1.26; 1-sided 95% CI 1.51) and consequently failed to meet the prespecified limit for noninferiority of tirofiban. The equivalency requirement was an upper bound of the 95% CI of the hazard ratio for the comparison of tirofiban with abciximab <1.47, consistent with the preservation of a difference of at least 50% in the effect of abciximab as compared with that of placebo observed in the EPISTENT trial. The protocol permitted retesting the data, and the superiority of abciximab over tirofiban was concluded based on a 2-sided P value of 0.04. Given the borderline strength of this evidence, it is prudent to reexamine this conclusion using more informative methods and additional sources of related data.

An indirect source of evidence to assess equivalency comes from the placebo-controlled randomized trials of GP2b3a inhibitors for patients with acute coronary syndromes and where initial direct passage to the catheterization laboratory was not planned. For example, the PRISM+ trial evaluated tirofiban in 1915 patients with acute coronary syndromes, and the GUSTO IV trial investigated abciximab in 7800 similar patients. Similar clinical presentations (63% of TARGET patients had an acute coronary syndrome), mean age, percentage of diabetic patients, and the same standard outcomes recorded in the 3 trials suggest that it may not be unreasonable to assume that the patient groups are sufficiently similar to combine at least some of the earlier data (prior belief) with the TARGET results (as described in method 2 above). However, this is not a triv-
ial assumption, as drug dosages, duration of therapy, and the percentage undergoing PCI were not standardized across the studies.

Method 1—Objective Bayesian Analysis

With a diffuse prior distribution, our Bayesian analysis of the TARGET data of course gives a very similar estimate of the difference between the 2 drugs as reported in the original report, 1.5%, fewer deaths or nonfatal MIs with abciximab compared to tirofiban (95% CI 0.1%, 3.0%). Based on TARGET data alone, there is a 98% probability that abciximab has fewer combined clinical endpoints, deaths or nonfatal MIs, than tirofiban. However, our approach is also able to provide direct probability statements about clinically interesting thresholds, which aid in decision making. The probability that the difference in composite endpoints exceeds an absolute value of 1% is 0.798, or about 80%. Phrased differently, even when ignoring data from other trials, the data from TARGET remain compatible with a 20% probability that any difference in combined endpoints is less than 1% (Figure 1). For mortality, the same objective Bayesian methods again reveal an unchanged point estimate compared to the original report but, importantly, virtually excludes a mortality difference greater than 1% (probability < 0.01%) (see Table 1).

Method 2—Bayesian Analysis Incorporating Prior Information

Acute coronary syndrome patients represent a similar, but not identical, profile to those in the PCI trials and have also been extensively studied in placebo-controlled randomized trials with different GP2b3a inhibitors. In the PRISM+ and GUSTO IV trials, active treated patients had exactly equal incidences of death or nonfatal MI at 30 d of 8.6%. Combining this prior information with the TARGET data presented above shifts the probability density curve toward zero. If all previous data are considered (i.e., no discounting of previous data), the probability of abciximab being superior to tirofiban (by a more than 1% difference in the combined endpoint) falls to only 33.3% (see Table 1 and Figure 2). Even incorporation of only 20% of this prior knowledge results in a reduction of the probability of abciximab’s superiority to 69.1%.

Method 3—Indirect Comparisons via Hierarchical Bayesian Meta-Analysis

Using a Bayesian hierarchical model, the mean odds ratio for the reduction of the composite endpoint of death, MI, or repeat revascularization for the 8 studies comparing abciximab to placebo as adjunctive therapy

| Table 1 Probability of Equivalence between Tirofiban and Abciximab |
|----------------|----------------|----------------|----------------|
| Mortality Study | 0 | 1 | 2 | 3 |
| TARGET alone    | 0.327 | 0.999 | 1 | 1 |
| TARGET + 20% prior | 0.313 | 1 | 1 | 1 |
| TARGET + 50% prior | 0.317 | 1 | 1 | 1 |
| TARGET + 100% prior | 0.351 | 1 | 1 | 1 |
| Mortality & Nonfatal Myocardial Infarctions Study |
| TARGET alone |
| TARGET + 20% prior |
| TARGET + 50% prior |
| TARGET + 100% prior |
| Threshold for Clinical Difference (%) |
| 0.015 | 0.225 | 0.744 | 0.981 |
| 0.027 | 0.309 | 0.826 | 0.992 |
| 0.055 | 0.448 | 0.911 | 0.998 |
| 0.139 | 0.667 | 0.976 | 1 |
to PCI was 0.56 (95% CI 0.43, 0.70). The odds ratio for the comparison of other GP2b3a inhibitors to placebo was 0.79 (95% CI 0.56, 1.09). The ratio of the benefit from the abciximab/placebo studies to that observed from the other GP2b3a inhibitors/placebo studies was OR = 0.72 (95% CI 0.47, 1.06), with the wide CI indicating an inconclusive result. Similarly, when comparing the abciximab result to the separate eptifibatide or tirofiban trials, the ratio of benefits was not statistically different from unity (OR 0.75; 95% CI 0.44, 1.22; and OR 0.71, 95% CI 0.31, 1.38, respectively).

DISCUSSION

Adjunctive therapy with GP2b3a inhibitors is used in the majority of the more than 600,000 angioplasties performed annually in the United States. Three agents are commercially available, and it is clinically and economically important to determine if the drugs are clinically equivalent. The only comparative trial concluded that abciximab was superior to tirofiban at 30 d follow-up, and it appears unlikely that other randomized trials will be forthcoming. Therefore, this article, which presents 3 complimentary, supplemental approaches to assist in the decision-making process regarding drug equivalency in the face of this less than ideal evidence, is most pertinent. In contrast to the conclusions from the comparative trial, our analyses suggest at best weak evidence of any meaningful differences between the GP2b3a inhibitors from the data available to date.

Innovative means of using other data to complement and enhance the results of randomized trials are being increasingly examined. Contemporary statistical theory suggests more informed decisions might be reached by attempting to incorporate prior information with the results of the most recent randomized clinical trial. This is hazardous terrain, and to avoid erroneous conclusions, careful attention must be paid to study populations, treatments received, and outcome measures. Recognizing that medical decisions regarding resource allocation must occasionally be made in the absence of perfect information coming from sufficient numbers of well-designed clinical trials, indirect comparisons for estimating the efficacy of competing interventions are being increasingly proposed and appear to provide useful and valid supplemental information provided the studies are sufficiently similar with good internal and external validity.

Our Bayesian perspective permitted the assessment of the probability of a clinically meaningful difference between the 2 drugs going beyond the standard statistical methods of simplistic testing of point null hypotheses, which are of questionable clinical utility. We recognize that the choice of a clinically meaningful difference is value laden and somewhat arbitrary, although there is an unofficial consensus among the cardiovascular community that a 1% decrease in mortality is clinically important. In the present example, mortality differences are much smaller, and for the sake of argument, we will propose that a less stringent 1% difference of the composite endpoint be considered as clinically important. Our analysis shows a 22.5% probability of clinical equivalence under this condition. The application of this Bayesian analysis to the original study once again demonstrates how marginally statistically significant P values may become confused with clinical significance and lead to overestimations of the strength of the data in the minds of many clinicians and decision makers.

Incorporating prior evidence about these drugs with the TARGET results may further enlighten our understanding of their relative efficacy. The inclusion of prior knowledge shifts this curve to the left, and the probability of abciximab’s superiority over tirofiban for the combined endpoint of death and MI exceeding 1% falls substantially (Figure 1, Table 1). Even incorporation of only 20% of this prior knowledge results in a more than 30% probability of clinical equivalency be-
between the drugs. Importantly, this conclusion of clinical equivalency is relatively insensitive to the amount of prior information considered.

Table 1 shows how this probability of equivalency varies as a function of the outcome considered, the prior information, and the level of clinical equivalence chosen. Because the prognostic significance of a peri-procedural micro-infarct is not completely resolved, others may well argue that clinical equivalence is present when the difference in composite endpoints is < 2%. Under such a definition, the probability of superiority of abciximab is less than 25%.

Because none of the 3 methods employed in this evaluation are definitive, it is reassuring that all are concordant in their conclusion of no strong proof for the superiority of abciximab as adjunctive therapy for PCI. Astute clinicians may have possibly reached a similar conclusion by careful assessment of the TARGET data, as there were no differences in mortality (0.5% v. 0.4%, 95% CI –0.3%, 0.5%) or urgent revascularizations (0.8% v. 0.7%, 95% CI –0.4%, 0.6%). Moreover, these clinicians would be concerned by the uncertainty surrounding the prognostic implications of the increase in nonfatal MI observed with tirofiban (6.9% v. 5.4%, 95% CI 0.1%, 2.9%), as at least some were likely peri-procedural micro-infarcts. Before reaching a definitive conclusion, prudent clinicians and health care evaluators would also probably like confirmation of the durability of any benefits beyond the 30-d follow-up period of the original trial.

Longer term follow-up data from the TARGET study are now available and may be offered as a reasonable gold standard by which to evaluate our methods. At 6 mo, the early benefits of abciximab were attenuated and tirofiban provided a similar level of overall protection as abciximab against the composite endpoint (HR 1.04, 95% CI 0.90–1.21; P = 0.591), further supporting our conclusions based on the 30-d results.

Our methods suggest only weak evidence for clinically important differences between the different GP2b3a inhibitors and highlight the fragility of basing medical decisions on conclusions from a single isolated randomized clinical trial, even if well performed. Applying Bayesian reasoning to the original data, incorporating even small amounts of outside information and indirect comparisons based on the totality of evidence from placebo-controlled trials may be helpful in reaching reasonable health care decisions. The methods proposed in this article cannot be guaranteed to produce the “right” answer in all cases but do provide a reasonable structure for decision making in clinical situations with incomplete evidence.

ACKNOWLEDGMENTS

Drs Brophy and Joseph are funded researchers from Les Fonds de la Recherche en Santé du Québec (FRSQ) and the Canadian Institutes of Health Research (CIHR), respectively.

REFERENCES