Frequentist or Bayesian: Coronary artery bypass grafting offers advantages over percutaneous coronary intervention in left main coronary disease

Derrick Y. Tam, MD, PhD,a Jan O. Friedrich, MD, PhD,b Rakesh C. Arora, MD, PhD,c,d and Bobby Yanagawa, MD, PhDa

You flip a coin and ask what is the probability that the coin lands heads up. The Bayesian would say, on this toss, it’s 50%. The frequentist would say, it is either 100% heads or 0% heads, there is no in between. With this in mind, we delve into how both Bayesian and frequentist statistics are interpreted differently yet may give us the same conclusion when it comes to optimal revascularization of left main coronary disease.

Early work by Yusuf and colleagues demonstrated the superiority of coronary artery bypass grafting (CABG) over optimal medical therapy for severe left main coronary artery disease. The role of percutaneous coronary intervention (PCI) for left main coronary artery disease has been the subject of several randomized clinical trials (RCTs) in the past 2 decades. This remains a highly controversial area with substantive differences in recommendations from the major societal guidelines. In the European Society of Cardiology/European Association for Cardio-Thoracic Surgery Guidelines on Myocardial Revascularization, left main PCI received a Class I indication (ie, treatment is associated with benefit) in those with low coronary lesion complexity and Class IIa recommendation (ie, treatment is likely of benefit) for intermediate complexity. In contrast, in the American College of Cardiology/American Heart Association Guidelines on Coronary Artery Revascularization, a IIa recommendation is suggested for left main PCI. Results from the recently published 5-year outcomes of the Evaluation of XIENCE versus Coronary Artery Bypass Surgery for Effectiveness of Left Main Revascularization (EXCEL) trial demonstrated that all-cause mortality was greater with PCI compared with CABG in the secondary analysis. This new finding was somewhat inconsistent with previous prospective trial data and highlighted the need for a patient-level meta-analysis of available RCT evidence to better understand the association between left main revascularization strategy and late outcomes. The purpose of this review is to summarize the findings of the recently published meta-analysis on this topic; to provide cardiovascular surgeons with a primer on both frequentist and Bayesian statistics; and to provide knowledge to help their patients interpret the findings of this study to make clinical decisions and treatment strategies.

From the Division of Cardiac Surgery, Department of Surgery, and Critical Care and Medicine Departments, St Michael’s Hospital, University of Toronto, Toronto, Ontario; Department of Surgery, Section of Cardiac Surgery, Max Rady College of Medicine, University of Manitoba, Winnipeg, Manitoba; and Cardiac Sciences Program, St Boniface Hospital, Winnipeg, Manitoba, Canada.

Received for publication Feb 7, 2022; revisions received April 19, 2022; accepted for publication April 30, 2022; available ahead of print May 14, 2022.
Address for reprints: Bobby Yanagawa, MD, PhD, Department of Surgery & Department of Pharmacology and Toxicology, St Michael’s Hospital, 30 Bond St, 8th Floor, Bond Wing, Toronto, Ontario, M5B 1W8, Canada (E-mail: bobby.yanagawa@unityhealth.to).

J Thorac Cardiovasc Surg 2023;166:136-40 0022-5223/$36.00
Copyright © 2022 by The American Association for Thoracic Surgery
https://doi.org/10.1016/j.jtcvs.2022.04.045
end points included rate of spontaneous myocardial infarction (MI), procedural MI (defined by protocol or universal definition), stroke, or repeat revascularization. Overall, authors included 4 RCTs that met inclusion criteria: Premier of Randomized Comparison of Bypass Surgery versus Angioplasty Using Sirolimus-Eluting Stent in Patients with Left Main Coronary Artery Disease (PRECOMBAT), Synergy between PCI with Taxus and Cardiac Surgery trial (SYNTAX), the previously mentioned EXCEL, and Nordic–Baltic–British Left Main Revascularisation Study (NOBLE), with 4394 pooled patients. Patients in the PCI and CABG arms shared similar baseline characteristics, with a median age of 66 years, mostly male (76%), with similar European System for Cardiac Operative Risk Evaluation (3.0) and distribution of SYNTAX scores. For the primary outcome, the Kaplan–Meier estimate of 5-year all-cause mortality was 11.2% and 10.2% for patients undergoing PCI and CABG, respectively. This difference was not statistically different; the hazard ratio from the Cox-proportional hazard model that accounted for trials as a random effect was 1.10 (95% confidence interval [CI], 0.91-1.32, \( P = .33 \)), with an absolute risk difference of 0.9% (95% CI, -0.9-2.8). In contrast, in the Bayesian analysis, with “non-informative” priors (ie, no prior assumption of superiority of either treatment), there was an 85.7% probability that the absolute risk difference for mortality was >0% and greater in PCI than CABG (similar to the ~85% percentage of the area of the Bell curve [normal] distribution of the hazard ratio [HR] being >1 in Figure 1, A, and absolute risk ratio being >0%, indicating greater mortality of PCI compared with CABG). There was a 49.1% probability that this difference was at least 1.0% over 5 years (since with a mean estimate +0.9% just under half of the Bell curve probability distribution lies above +1.0%; see Figure 1, B). Importantly, there were also differences in secondary outcomes in the frequentist analysis; spontaneous MI over 5 years was greater in the PCI group (HR, 2.35; 95% CI, 1.71-3.23) along with repeat revascularization (HR, 1.78; 95% CI, 1.51-2.10) with no difference in stroke (HR, 0.84; 95% CI, 0.59-1.21). Procedural MI as defined by protocol was greater with CABG (4.7% vs 3.2%, \( P = .013 \)) but not different when the universal definition of MI was used (3.2% vs 2.3%, \( P = .15 \)). Given all this information, interpretation of this study requires at least a cursory understanding of the differences between frequentist and Bayesian statistics (Table 1).

While the biomedical literature has relied on traditional frequentist statistics, the use of Bayesian statistics has increased significantly in the literature. While the concept of the \( P \) values and their use in frequentist statistics dates back to the 1920s, Bayes’ theorem actually predated the use of \( P \) values by almost 150 years.7 However, the use of Bayesian approaches has only recently become more

![A](image1.png)  
**FIGURE 1.** Density plots for the primary outcomes from the frequentist analysis by Sabatine and colleagues.6 A, The hazard ratio (HR) for mortality shown with the black line (1.1) and 95% confidence interval (0.91, 1.32) in the dotted black lines. A total of 85.7% of the curve is shaded, representing the 85.7% probability that mortality was >0% (ie, The HR was >1.0) in the Bayesian analysis. B, The risk difference for mortality between PCI and CABG represented by the black line (+0.9%) with the 95% confidence interval (−0.9%, +2.8%) in the dotted black lines. A total of 49.1% of the curve is shaded, representing the 49.1% probability in the Bayesian analysis that the risk difference was >1%. PCI, Percutaneous coronary intervention; CABG, coronary artery bypass grafting.
TABLE 1. Comparison of frequentist hypothesis testing versus Bayesian analysis

<table>
<thead>
<tr>
<th>Hypothesis formulation</th>
<th>Frequentist hypothesis testing</th>
<th>Bayesian analysis</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Express the null and alternative hypothesis.</strong> In this case, the H₀ is that there is no difference in all-cause mortality between PCI and CABG for treating left main disease. The alternative hypothesis is that there is a difference between PCI and CABG. Determine the alpha or threshold for rejecting the null hypothesis. The threshold in this analysis was 0.05.</td>
<td>Choose the prior distribution for the analysis; this can be based on expert opinion, previous RCTs or observational studies, or biological or physiological basis. The prior distributions help specify what is known previously about differences between the 2 treatments. The prior may be informative or noninformative (ie, no assumption of superiority of either treatment), depending on what information is available. In this analysis, a noninformative prior was chosen.</td>
<td></td>
</tr>
</tbody>
</table>

| Choosing the correct statistical model | Calculate the test value from the data and compare this value with the predetermined threshold. In this case, the test value yielded a P = .33, which is above the prespecified threshold of 0.05. The null hypothesis is retained in this analysis. | From the observed data and the prior distribution, the posterior distribution is calculated using simulation software. |

| Calculation of the test values and interpretation of the results | Evaluation and interpretation of the posterior distribution. In this case, the probability of greater mortality with PCI over CABG and the probability for what these absolute differences are. The posterior distribution found that there was a >85% probability that there is difference in mortality that was greater in PCI than CABG corresponding to ~85% of the Bell curve (normal) distribution lying above HR >1 or absolute risk difference >0%. It also estimated a 49% probability that the absolute risk increase was at least 1%, corresponding to a mean estimate of +0.9% absolute risk increase or just under 50% of the Bell curve (normal) probability distribution being >1%. |

PCI, Percutaneous coronary intervention; CABG, coronary artery bypass grafting; RCTs, randomized clinical trials; HR, hazard ratio.
TABLE 2. Glossary of statistical terms

<table>
<thead>
<tr>
<th>Statistical term</th>
<th>Definition</th>
</tr>
</thead>
<tbody>
<tr>
<td>Confidence interval</td>
<td>In frequentist statistics, 95% confidence interval, which should be interpreted as when an experiment is repeated infinitely and under the exact conditions, 95% of the estimated intervals would contain the true value of the parameter, that said, the true value of the parameter is not known but rather what value the parameter can take with repeated sampling</td>
</tr>
<tr>
<td>Credible interval</td>
<td>Bayesian statistics provide a 95% credible interval; this is interpreted as a 95% probability that the true value lies within this range.</td>
</tr>
<tr>
<td>Informative prior</td>
<td>Informative priors are typically expressed as a parameter with a distribution. The distribution shape from an informative prior depends on how strong or weak the prior is. Informative priors may be obtained from previous experiments, meta-analyses, or even expert opinion. The informative prior is combined with probabilities generated from the new data to inform the posterior distribution. Use of different priors can be examined in a sensitivity analyses to understand the robustness of the results.</td>
</tr>
<tr>
<td>Noninformative prior</td>
<td>Prior knowledge is typically expressed as a uniform distribution and contains no information as the name implies. This means that the posterior distribution only reflects the probabilities from the new data and as such, may have results that converge with those from a frequentist analysis</td>
</tr>
<tr>
<td>Posterior distribution</td>
<td>Results after integrating prior knowledge with probability from the new data.</td>
</tr>
<tr>
<td>Null hypothesis</td>
<td>The null hypothesis is part of frequentist statistics and usually expressed as, &quot;there is no difference between treatment A and treatment B&quot;</td>
</tr>
<tr>
<td>Alternative hypothesis</td>
<td>The alternative hypothesis is part of frequentist statistics and is often the corollary of the Null hypothesis, and can be expressed as: &quot;there is a difference between treatment A and treatment B&quot;</td>
</tr>
<tr>
<td>P value</td>
<td>Used in frequentist statistics, the P value represents the probability of the null hypothesis occurring completely by chance (or random)</td>
</tr>
</tbody>
</table>

Referring back to the coin toss example, once the coin has landed, it can only be either heads or tails; similarly, from the frequentist perspective, the results can either be significant or not significant based on the prespecified threshold. Part of the reason for this strict interpretation relates to some of the concerns of relying on P values to interpret significance versus nonsignificance. One such pitfall is the concept of multiplicity and multiple hypothesis testing, which increases the chance for Type I error (a false positive) when P values are not adjusted for testing for multiple hypotheses.10 Recently, some journals have advocated for the presentation of the CI to provide an estimate of the margin of error rather than the P value alone. However, there exists a mathematical relationship between P values and CIs, and this quandary continues to be debated, even at the Journal. For readers who are further interested in this topic, we would suggest the following 3 articles by Dr Park, Dr Blackstone, and Drs Freames and Tam that were recently published here in the Journal.9-11

In contrast, Bayesian statistics does not perform hypothesis testing in the same way as frequentist statistics. In frequentist analyses, prior data are not explicitly used. Although some prior data may be used to help calculate sample size and power in frequentist statistics, prior data are not directly used in the final analysis. In contrast, for Bayesian analyses, prior data may be explicitly used, known as the “informative” prior and combined with the new data to form the posterior distribution. This is the crux of Bayes’ theorem and its mathematical relationship as described previously, where we take pretest probability (the prior) and integrate it with the data from the new study to derive a posttest probability (the posterior distribution). However, if there is not much information known previously, the prior may be noninformative (ie, no prior assumption of superiority of either treatment), which means that the posterior distribution only reflects the probabilities from the new data and, as such, yield results that converge with those from a frequentist analysis, which was the case in this study by Sabatine and colleagues.6 In contrast to frequentist statistics, we are provided with the probability that the parameter of interest is actually different between groups. Bayesian statistics provide a 95% credible interval; this is interpreted as 95% probability that the true value lies within this range—this is in contrast to the frequentist statistic of 95% CI, which should be interpreted as when an experiment is repeated infinitely and under the exact conditions, 95% of the estimated intervals would contain the true value of the parameter, that said, the true value of the parameter is not known but rather what value the parameter can take with repeated sampling. We compare and contrast these relevant concepts in frequentist and Bayesian analysis in Table 1 and how they apply to the meta-analysis by Sabatine and associates.6
In the meta-analysis by Sabatine and colleagues, the frequentist analysis showed that there was no difference in all-cause mortality; there was not enough evidence to reject the null hypothesis that PCI and CABG were not different (ie, the P value was not less than .05). However, this does not mean CABG and PCI are equivalent, as this study was not a noninferiority analysis, nor was a power calculation provided. Failure to reject the null hypothesis could be the result of an underpowered study: for example, the lower and upper bounds of the 95% CI of the 0.9% absolute increase ranged from a 0.9% decrease to a 2.8% increase in 5-year mortality, which does not rule out either significant benefit or even more severe harm of PCI versus CABG. However, the Bayesian analysis of this study provides an intuitive interpretation of results that we can easily provide to patients, which is that there is an 85% chance that CABG will help you live longer compared with stenting and that there is a ~50% probability that the absolute increase in mortality is at least 1%—this difference in longevity is likely small but still greater with CABG. In addition, CABG offers protection against future MIs and need for further revascularization procedures. However, we must interpret these data in the context of the limitations of the study, namely that these were selected patients with anatomy suitable for both PCI and CABG, at low surgical risk. In addition, some short-term outcomes favored PCI: although overall stroke risk was not different; it was significantly lower for PCI in the first year (P < .002), and mortality tended lower for PCI initially, with mortality benefits for CABG not beginning to manifest themselves until after the first year. Furthermore, in subgroup analysis, CABG seemed more clearly beneficial in patients not presenting with acute coronary syndrome and possibly in patients ≥65 years old. Consequently, evaluation by a heart team, consisting of both interventional cardiologists and cardiac surgeons, remains critical to individualize decision-making, particularly for patients who are high surgical risk or in those who have complex coronary anatomy. In high-risk cases, where perioperative mortality or morbidity is high, the risk of surgery may outweigh the late mortality benefit of CABG. However, in patients at low surgical risk with low complexity left main disease, a clear and thorough discussion of the risks and benefits of both revascularization procedures must be provided to allow patients to make informed choices. Where life expectancy is reasonable, CABG offers better event-free survival and a greater probability for improved overall survival. The use of individualized visual decision aids may help improve patients’ ability to comprehend and understand what these risks and benefits entail. Nonetheless, as surgeons and clinicians, it is prudent on us to understand both the frequentist and Bayesian interpretation of a study and to deliver these interpretations in a clear and understandable manner.

Conflict of Interest Statement
The authors reported no conflicts of interest.

The Journal policy requires editors and reviewers to disclose conflicts of interest and to decline handling or reviewing manuscripts for which they may have a conflict of interest. The editors and reviewers of this article have no conflicts of interest.

References

Key Words: CABG, PCI, left main disease, revascularization, Bayesian, frequentist