Incorporating Patient Preferences *via* Bayesian Decision Analysis

Shomesh E. Chaudhuri[1] and Andrew W. Lo[1,2,3,4,5,6]

The regulatory process for market authorization of medical diagnostic and therapeutic products is fraught with ethical dilemmas that regulators outside the medical industry do not face. The consequences of approving an ineffective therapy with potentially dangerous side effects (a "Type I error" or false positive) must be weighed against not approving a safe and effective therapy (a "Type II error" or false negative) that could help ease the burden of disease for many patients. Regulators must strike the proper balance by considering multiple factors, including scientific merit, clinical evidence from randomized control trials, the burden of disease, the current standard of care and alternatives, and patient preferences. How these factors are—and should be—weighed is not always clear, which only encourages criticism by whichever stakeholder group disagrees with the decision.

Given the complexity of biomedicine and the consequences of both types of errors, regulators must exercise discretion in making their decisions. However, such flexibility can be made more transparent and systematic by applying Bayesian decision analysis to the regulatory approval process, as described in a series of studies (1–6). The benefits can be best understood by contrasting it with traditional hypothesis testing in which a desired Type I error rate, say 5%, is chosen and the statistical significance of the clinical evidence is evaluated using this threshold. If results are inconsistent with the null hypothesis of no efficacy at a significance level, or P value <0.05, then the null hypothesis is rejected and, in our context, the therapy is approved.

The question raised and answered by Bayesian decision analysis is "why 0.05?" For fatal diseases with no existing treatments, patients may be willing to accept a much higher false-positive rate, especially if it yields a lower false-negative rate, as is often the case. For example, suppose the conventional 0.05 Type I error is associated with a Type II error of 0.25. A patient with glioblastoma who has exhausted the standard of care may be comfortable with a Type I error of 0.20 if it is associated with a Type II error of 0.10. Given that such patients have no other recourse for this terminal illness, the relative importance of false positives and negatives should reflect their circumstances. To do so, a new regulatory approval threshold

can be computed by explicitly minimizing the expected loss to patients due to both Type I and II errors, where the expected loss is the weighted sum of the measured effect of false positives and false negatives, weighted by their probabilities.

Bayesian decision analysis does require more information than the traditional approach—the losses under both types of errors must be specified, and in some cases, these losses may be difficult to gauge. However, several metrics have been developed for this purpose, including survey tools designed by patient advocacy groups to measure the preferences of their constituents. For example, through the collaboration of members of the Michael J. Fox Foundation, it was determined that patients with late-stage Parkinson's disease were significantly more tolerant of false positives in exchange for greater and more timely access to risky experimental therapies, such as deep brain stimulation devices (6). Similar patient preference surveys can be used to measure the degree to which patients with kidney failure would be willing to bear the risk of switching from in-center dialysis to a yet to be approved wearable KRT, which can then serve as one of several inputs for determining the optimal regulatory decision. The Bayesian decision analysis framework provides the flexibility needed to incorporate these burden of disease and patient preference considerations in regulatory decisions.

The simplest form of the framework uses benefit-risk and time preferences to estimate the value lost from the patient's perspective of making an incorrect or delayed approval decision. For example, for the case of an incorrect approval, the new therapy is assumed to provide no benefits relative to the standard of care, but does contain additional risk in the form of adverse side effects and missed opportunities to be treated by more effective therapies. The estimated annual discount rate for patients with Parkinson's disease ranges from 14.5% to 32.7%, increasing as a function of both age and disease severity (6). However, an incorrect failure to approve results is the missed opportunity to receive a therapy that is more effective than the standard of care. Finally, even in the case of a correct approval, value can be lost due to the time it takes to reach a regulatory decision.

Multiplying these costs by their probabilities and summing across the various scenarios yields the expected loss of an incorrect or delayed regulatory

[1]QLS Advisors, LLC, Cambridge, Massachusetts
[2]Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Massachusetts
[3]Laboratory for Financial Engineering, Massachusetts Institute of Technology, Cambridge, Massachusetts
[4]Computer Science and Artificial Intelligence Laboratory, Massachusetts Institute of Technology, Cambridge, Massachusetts
[5]Department of Electrical Engineering and Computer Science, Massachusetts Institute of Technology, Cambridge, Massachusetts
[6]Santa Fe Institute, Santa Fe, New Mexico

Correspondence:
Dr. Andrew W. Lo, Massachusetts Institute of Technology Sloan School of Management, 100 Main Street, E62-618, Cambridge, MA 02142. Email: alo-admin@mit.edu
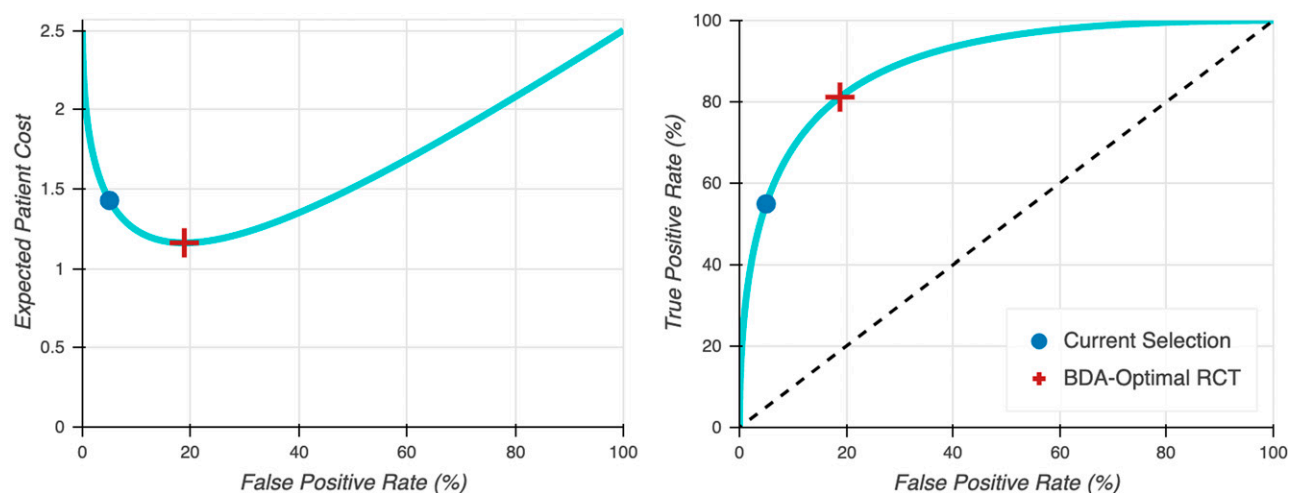
Figure 1. | **Optimal size and power for two-armed balanced fixed sample clinical trial according to Bayesian decision analysis (BDA) in which the expected loss of false positives and negatives is jointly minimized.** The left figure displays the expected cost to patients as a function of the false positive rate (the BDA optimal value minimizes this expected cost), and the right figure displays the true positive rate (or power) as a function of the false positive rate. For an interactive version of these graphs, see the Supplemental Material.

decision. The Bayesian decision analysis framework jointly determines the optimal number of study participants and statistical significance threshold such that the expected harm of these downside scenarios is minimized. Figure 1 illustrates how the optimal statistical significance threshold changes as a function of Type I and II error costs.

For patients who had previously received deep brain stimulation treatment, the optimized significance levels were between 0.04 and 0.10, similar to or greater than the traditional value of 0.05; conversely, patients never having received deep brain stimulation were less comfortable with the risks, and hence, their optimal significance levels ranged from 0.002 to 0.044 (6). In both populations, optimal significance levels increased with the severity of their symptoms. Additional examples of the Bayesian decision analysis framework applied to disease areas such as obesity, oncology, and coronavirus disease 2019 for both fixed sample and adaptive clinical trials can be found in refs. 2–5. These examples show that the optimal Type I error is sometimes more conservative than 0.05; hence, this framework does not necessarily imply lower approval standards, something physicians would likely oppose (7).

The most challenging practical issue in implementing this framework is the consequences of a larger number of false positives. This can be addressed by creating a temporary license to market "speculative" therapies that expires after a short period (say, 2 years) (8). During this period, the licensee is required to collect and share data on the performance of its therapy, and if the results are positive, the license converts to a standard approval; otherwise, the therapy is withdrawn upon expiration. Regulators should have the right to terminate the temporary license at any time in response to adverse events or significantly negative data. Such licenses would greatly accelerate the pace of therapeutic development for many underserved medical needs—including alternatives to dialysis for kidney failure—without limiting regulatory flexibility.

Flexibility is particularly important because any system can be gamed, leading to unintended outcomes; hence, no single interest group should be allowed to exercise undue influence in this process. Therefore, regulators must, and do, apply discretion, judgment, and a wealth of experience in their review process. Nevertheless, a systematic, rational, transparent, reproducible, and practical framework in which regulators' decisions can be clearly understood by and communicated to all stakeholders while explicitly incorporating their feedback may still have value. Such a framework is consistent with the Food and Drug Administration's 2009 guidance on its use of patient-reported outcomes to support labeling claims, as well as its 2017 guidance on using real-world evidence for medical device approval decisions. This would also help satisfy the agency's mandate under the 21st Century Cures Act to incorporate patient preferences and real-world data explicitly into its approval process.

Supplemental Material

This article contains the following supplemental material online at http://cjasn.asnjournals.org/lookup/suppl/doi:10.2215/CJN.12110720/-/DCSupplemental.

Supplemental Material. An interactive version of Figure 1.

References

1. Isakov L, Lo AW, Montazerhodjat V: Is the FDA too conservative or too aggressive?: A Bayesian decision analysis of clinical trial design. *J Econom* 211: 117–136, 2019
2. Montazerhodjat V, Chaudhuri SE, Sargent DJ, Lo AW: Use of Bayesian decision analysis to minimize harm in patient-centered randomized clinical trials in oncology. *JAMA Oncol* 3: e170123, 2017
3. Chaudhuri SE, Ho MP, Irony T, Sheldon M, Lo AW: Patient-centered clinical trials. *Drug Discov Today* 2: 395–401, 2018
4. Chaudhuri SE, Lo AW, Xiao D, Xu Q: Bayesian adaptive clinical trials for anti-infective therapeutics during epidemic outbreak [published online ahead of print May 14, 2020]. *Harvard Data Sci Rev* 10.1162/99608f92.7656c213
5. Chaudhuri SE, Lo AW: Bayesian adaptive patient-centered clinical trials. 2019
6. Chaudhuri SE, Hauber B, Mange B, Zhou M, Ho M, Saha A, Caldwell B, Benz HL, Ruiz J, Christopher S, Bardot D, Sheehan M, Donnelly A, McLaughlin L, Gwinn K, Sheldon M, Lo AW: Use of Bayesian decision analysis to maximize value in patient-centered randomized clinical trials in Parkinson's disease. 2020
7. Kesselheim AS, Woloshin S, Lu Z, Tessema FA, Ross KM, Schwartz LM: Physicians' perspectives on FDA approval standards and off-label drug marketing. *JAMA Intern Med* 179: 707–709, 2019
8. Lo AW: Discussion: New directions for the FDA in the 21st century. *Biostatistics* 18: 404–407, 2017

Published online ahead of print. Publication date available at www.cjasn.org.