p-Values vs. Patient Values: An Analytic Perspective

Andrew W. Lo, MIT
September 13, 2018

ASA Biopharma Regulatory-Industry Statistics Workshop

MIT Laboratory for Financial Engineering
FDA Approvals Involve Trade-offs

- Goal: minimize errors by setting the threshold for approval (typically a “p-value” of 5%)
- But there’s a trade-off between these errors
- “Greatest good for the greatest number” J. Bentham

<table>
<thead>
<tr>
<th></th>
<th>Approve</th>
<th>Reject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ineffective Therapy</td>
<td>Type I Error (False Positive), $\alpha$</td>
<td>Correct</td>
</tr>
<tr>
<td>Effective Therapy</td>
<td>Correct</td>
<td>Type II Error (False Negative), $\beta$</td>
</tr>
</tbody>
</table>
Why 5%??

“...If one in twenty does not seem high enough odds, we may, if we prefer it, draw the line at one in fifty or one in a hundred. Personally, the writer prefers to set a low standard of significance at the 5 per cent point, and ignore entirely all results which fails to reach this level. A scientific fact should be regarded as experimentally established only if a properly designed experiment rarely fails to give this level of significance...”

Why 5%??

The ASA’s Statement on p-Values: Context, Process, and Purpose

In February 2014, George Cobb, Professor Emeritus of Mathematics and Statistics at Mount Holyoke College, posed these questions to an ASA discussion forum:

Q: Why do so many colleges and grad schools teach $p = 0.05$?
A: Because that’s still what the scientific community and journal editors use.

Q: Why do so many people still use $p = 0.05$?
A: Because that’s what they were taught in college or grad school.

2014) and a statement on risk-limiting post-election audits (American Statistical Association 2010). However, these were truly policy-related statements. The VAM statement addressed a key educational policy issue, acknowledging the complexity of the issues involved, citing limitations of VAMs as effective performance models, and urging that they be developed and interpreted with the involvement of statisticians. The statement on election auditing was also in response to a major but specific
“...better that ten guilty persons escape than that one innocent suffer” — Blackstone (1765)
Incorporating Patient Preferences

Guidance for Industry & FDA Staff (2012)

“FDA recognizes that patient tolerance for risk and a patient-centric assessment of risk may reveal reasonable patients who are willing to tolerate a very high level of risk to achieve a probable benefit, especially if that benefit results in an improvement in quality of life.”


“How the FDA plans to use relevant patient experience data and related information when evaluating the risks and benefits of a drug.
A New Approach

What If We Try To Reduce The Average Loss?

\[
\text{Loss(False Positive)} \times \text{Prob(False Positive}; p) \\
+ \text{Loss(False Negative)} \times \text{Prob(False Negative}; p) \\
\text{Average Loss}(p)
\]

- Given the losses from false positives and negatives, choose \( p \) to minimize the average loss \( \Rightarrow \) BDA
Is the FDA Too Conservative and Too Aggressive?: A Bayesian Analysis of Clinical Trials
Leah Jackson, Andrew W. Lo, and Yung Meng Lo
This Draft: 28 November 2018

Abstract

Is the Bayesian approach to regulatory drug development more efficient than the traditional approach? In this article, we consider a simple model in which the FDA requires two trials for approval of a new drug. We show that in this setting, the Bayesian approach is more efficient than the traditional approach. In particular, we show that the Bayesian approach requires fewer trials on average than the traditional approach.

Keywords: Clinical Trial Design; Drug-Approval Process; Adaptive Design.

Bayesian Adaptive Patient-Centered Clinical Trials

Shomesh E. Chaudhuri, Martin P. Ho, Tolba Irvany, Murray Sheldon, and Andrew W. Lo

Introduction

The regulatory approval process for new medical therapies involves a series of decisions that a patient may be better served by a different therapy. In this context, the decision to use a particular therapy may be influenced by the patient's prior medical history, their current health status, and their preferences. In addition, the decision to use a particular therapy may be influenced by the patient's prior medical history, their current health status, and their preferences. In this section, we describe a model of a patient's decision to use a particular therapy.

Bayesian Decision Analysis

In this section, we describe a model of a patient's decision to use a particular therapy. The model is based on the assumption that the patient is aware of the potential benefits and risks of the therapy and is willing to take the risk if the benefits outweigh the risks. In this model, the patient's decision to use a particular therapy is influenced by the patient's prior medical history, their current health status, and their preferences.

Slide 8
Bayesian Decision Analysis

Collaborators:

- Dawn Bardot, Heather Benz, Brittany Caldwell, Shomesh Chaudhuri, Stephanie Christopher, Katrina Gwin, Brett Hauber, Martin Ho, Telba Irony, Leah Isakov, Brennan Mange, Lauren McLaughlin, Vahid Montazerhodjat, Kyle Myers, John Ruiz, Annie Saha, Dan Sargent, Murray Sheldon, Mo Zhou

- FDA (CDRH), MDIC, MJF Foundation, RTI Health Solutions, MIT LFE
What About the Value of Life?
What About the Value of Life?

$9,100,000.00
What About the Value of Life?

NEJM Aug 2014

Updating Cost-Effectiveness — The Curious Resilience of the $50,000-per-QALY Threshold

Peter J. Neumann, Sc.D., Joshua T. Cohen, Ph.D., and Milton C. Weinstein, Ph.D.

If one had to select a single threshold outside the context of an explicit resource constraint or opportunity cost, we suggest using either $100,000 or $150,000.

- Compare new drug’s price to QALYs
- How much incremental value is it providing?
BDA-optimal decision minimizes expected cost
# Measuring “Size” and “Importance”

<table>
<thead>
<tr>
<th>Disease Name</th>
<th>Prevalence (Thousands)</th>
<th>Severity</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ischemic heart disease</td>
<td>8,895.61</td>
<td>0.12</td>
</tr>
<tr>
<td>Lung cancer</td>
<td>289.87</td>
<td>0.45</td>
</tr>
<tr>
<td>Ischemic stroke</td>
<td>3,932.33</td>
<td>0.15</td>
</tr>
<tr>
<td>Hemorrhagic/other non-ischemic stroke</td>
<td>949.33</td>
<td>0.16</td>
</tr>
<tr>
<td>Chronic obstructive pulmonary disease</td>
<td>32,372.11</td>
<td>0.06</td>
</tr>
<tr>
<td>Diabetes</td>
<td>23,694.90</td>
<td>0.05</td>
</tr>
<tr>
<td>Cirrhosis of the liver</td>
<td>78.37</td>
<td>0.49</td>
</tr>
<tr>
<td>Alzheimer’s disease</td>
<td>5,145.03</td>
<td>0.18</td>
</tr>
<tr>
<td>Colorectal cancer</td>
<td>798.90</td>
<td>0.15</td>
</tr>
<tr>
<td>Pneumococcal pneumonia</td>
<td>84.14</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Source: Isakov, Montazerhodjat, Lo (2015)
# Measuring “Size” and “Importance”

<table>
<thead>
<tr>
<th>Post-Trial</th>
<th>In-Trial</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\hat{H} = H_0$</td>
<td>$\hat{H} = H_1$</td>
</tr>
<tr>
<td>$H = H_0$</td>
<td>0</td>
</tr>
<tr>
<td>$H = H_1$</td>
<td>$C_2$</td>
</tr>
</tbody>
</table>
Application to Alliance Cancer Trials

Source: Montazerhodjat, Chaudhuri, Sargent, Lo (JAMA Onc. 2017)
Results for Parkinson’s Device

Subgroups ordered from low to high risk tolerance

Mild symptoms, no DBS

Severe symptoms, DBS

© 2018 by Andrew W. Lo
All Rights Reserved
Proposal: “Right-to-Try” License:

- Two-year license, no off-label use, strict monitoring and data collection/sharing requirements (paid by company)
- Can be revoked any time during two-year period
- At the end of two years, either license expires or it converts to regular approval
- Similar to adaptive trials, but much greater economic incentives for bipoharma industry
Qualifications

- How to choose parameters? (ODAC, but with patients)
- Whose preferences should be reflected? (patients!)
- Potential backlash from toxicities and side effects?
- Ethical considerations

But these issues already exist in one form or another for current methods (e.g., eteplirsen); BDA provides a more systematic, transparent, objective, repeatable, rational framework for addressing them.
Conclusion

“We should care about patient values as well as $p$-values.”

– Donald Berry, MD Anderson
Thank You!