Bayesian inference for 21st century drug development and approval

Traditional statistical hypothesis testing methods have been the mainstay of global regulatory agencies for decades. Dr Stephen Ruberg of Analytix Thinking, argues that a Bayesian approach, combining current data with prior knowledge, offers advantages over traditional methods. He reasons that these quantifiable probability assertions from Bayesian approaches are much more beneficial to both doctors and patients. Regulators should also find them more useful, with explicit evaluations of benefits and risks supporting decisions regarding new drug approval.

For decades, traditional statistical hypothesis testing methods have been the mainstay of global regulatory agencies, such as the US Food and Drug Administration (FDA). These methods inform decisions regarding the effectiveness of new pharmaceutical treatments seeking new drug approval. The research and development of a new drug takes many years and is underpinned by numerous preclinical experiments and clinical trials. During this time, scientific knowledge is advancing; innovations from molecular biology or clinical medicine may emerge relating to the new treatment's mechanism of action (the particular process through which a drug produces its effect) or the specific disease state.

Dr Stephen Ruberg, President of Analytix Thinking, and his collaborators suggest that this accumulation of knowledge and data is better suited to a Bayesian statistical approach. This approach formally summarises existing knowledge and data to describe the efficacy and safety of the new treatment. This existing evidence, quantified in what is known as a prior, is updated with results from new research experiments and clinical data, creating posterior probabilities for both the treatment's effectiveness and its safety outcomes. Ruberg and his team advocate that the decision makers, who range from regulatory agencies to the patient receiving the treatment, would find the ability to estimate the probability 'that a drug works' and 'that a drug is safe' highly desirable. Furthermore, these probabilities then become the prior probabilities for the next investigation (see Figure 1) as the process of drug development continues. Despite these benefits, however, the drug development process is still heavily reliant on the traditional hypothesis testing approach, known as the frequentist paradigm, where trials are treated as separate and distinct evidentiary entities.

TRADITIONAL STATISTICAL HYPOTHESIS TESTING

Traditional statistical hypothesis testing is analogous with the mathematical concept of proof by contradiction; we begin with an assumption and then work through all the logical steps until a result is obtained that is obviously false, therefore invalidating the initial assumption.

When traditional statistical hypothesis testing is carried out during a clinical trial of a new treatment, the starting point is the assumption that there is no effect from the treatment whatsoever. This is the null hypothesis. The clinical trial is then carried out and the results from the new treatment are compared to either a placebo or another established treatment. Data is collected on patients taking part in the trials and the effect of the new treatment is measured and compared with that of the placebo or established treatment. For instance, when the average responses of patients on each of the treatments are compared, if the averages are similar, then the new treatment is no better than the established treatment. If, however, there is a larger, positive difference between the average responses, then the new treatment is better than the established one.

THE P-VALUE

Traditionally, averages are compared using a p-value. The p-value is calculated using an appropriate probability distribution and has a value between 0 and 1. This is a standardised measure of how far the data lie from the null hypothesis. A small p-value indicates that the null hypothesis is unlikely to be true, so there may be a difference between the two treatments. Usually, if the p-value is less than 0.05 the null hypothesis is rejected, suggesting that there is a treatment effect.

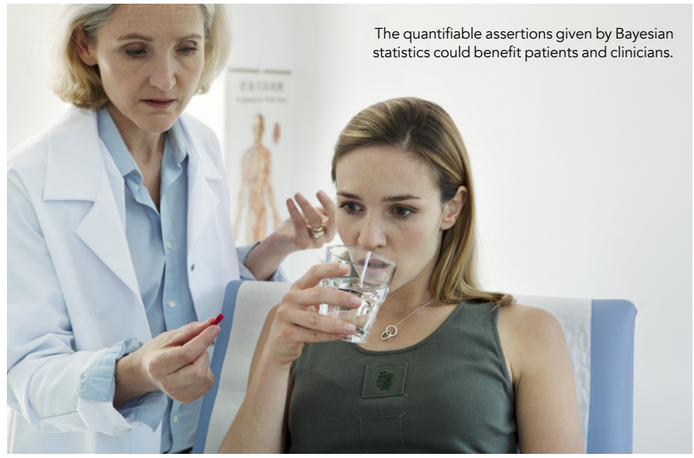
Dr Ruberg explains that, "while such quantitative approaches have helped bring greater rigor to the decision-making process for the approval of new drug/biologic treatments, they have shortcomings as well."

BAYESIAN METHODS

Numerous advances in high performance computing and Bayesian statistical theory mean that new approaches are now available; these techniques can handle more data and information, via more complex analysis, in order to ascertain the effectiveness of new pharmaceutical treatments. These methods are based on Bayes' Theorem, and collectively are known as Bayesian statistics, which calculates the probability of an outcome by combining prior knowledge with data from a current experiment. When a clinical trial is carried out, the statistician can combine current data with prior knowledge of the hypothesis resulting in an updated probability or belief. This is known as the posterior probability of the treatment effect and offers quantifiable statements such as 'the probability that treatment X is 30% better than treatment Y is 0.90' (Figure 2 overleaf). Dr Ruberg believes that this offers advantages over traditional methods, and that these quantifiable assertions are much more beneficial to both doctors and patients. Regulators should also find them more useful as they give explicit evaluations of benefit and risk, supporting decisions regarding the approval of new treatments.

Countless new treatments and drugs fail to make it through clinical development. This 'wastage' adds to the cost of new medicines. There is evidence to suggest that the frequentist paradigm, where trials are treated individually, and each requires a p-value of less than 0.05, is somewhat to blame. No one actually knows how

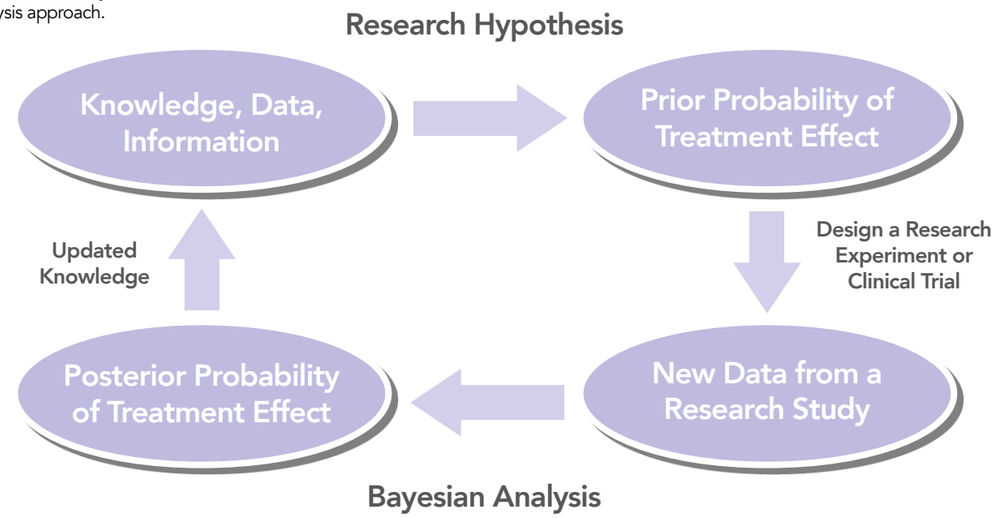
This accumulation of knowledge and data is better suited to a Bayesian approach.



The quantifiable assertions given by Bayesian statistics could benefit patients and clinicians.

Image Point Fr/Shutterstock.com

Figure 1
The cycle of scientific discovery aligns with the Bayesian statistical analysis approach.



many effective drug treatments have been abandoned because they 'failed' clinical trials with p-values of 0.05 or more.

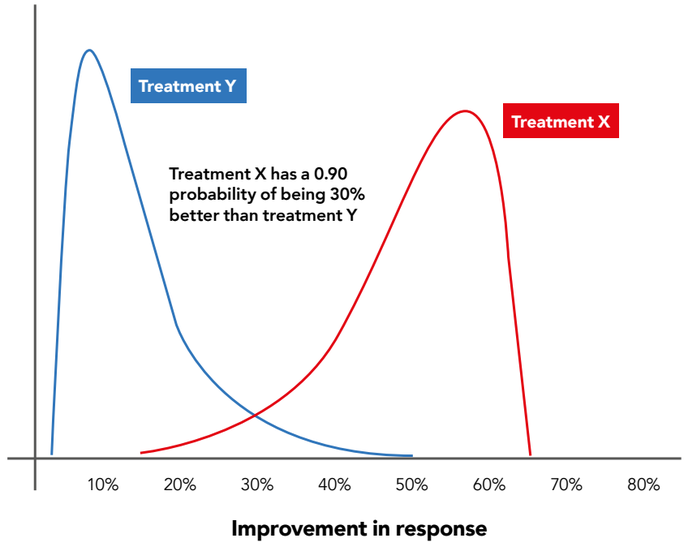
PHASES OF CLINICAL TRIALS

A clinical trial is only carried out when there is already evidence to suggest that a new drug treatment could improve patient care. Prior to the clinical trials, tests are carried out and evaluated in laboratory research that is *in vitro* (literally "*in glass*") or *in vivo* (i.e. animal research). These tests assess various features of the medicine prior to testing on human volunteers, i.e. clinical trials. Such experiments can be a useful basis for creating a prior probability about the new treatment's efficacy or safety. Clinical trials are performed in successive stages, called phases. There are usually four – Phase 1 to Phase 4, and with each phase, trials become larger, longer and more complex (and more expensive) to understand fully the benefits and risks of a new treatment. Dr Ruberg's research has a particular focus on Phase 2 and 3 trials. In general, Phase 2 studies are the initial studies in patients and assess what doses might be most likely to have sufficient benefit without significant adverse effects. Phase 3 trials are known as 'confirmatory' or 'definitive' trials and are the basis for whether a treatment is approved for marketing by a regulatory agency. It is the accumulation of data over time that is leveraged by Bayesian methods.

A BAYESIAN APPROACH

The probability-based inference deployed in Bayesian methods has recently been shown to be useful in the drug development process. Dr Ruberg and his team propose a Bayesian approach that uses data from other trials as the prior knowledge for answering the question, "Is this new treatment effective and safe?". The outcome of the Phase 1 trial becomes the prior probability for Phase 2 and together with the Phase 2 data, they generate an updated probability, the Phase 2 posterior. Likewise, the Phase 2 posterior probability becomes the prior probability for Phase 3 and combined with the Phase 3 data produce an updated probability, the Phase 3 posterior, which is effectively the probability that the new drug is effective and safe.

Figure 2: Posterior distributions of treatment effect



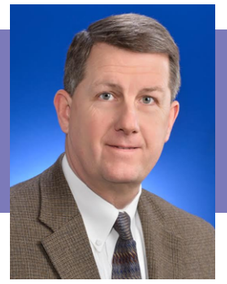
The research team have demonstrated the value of a Bayesian approach with data from real clinical trials for a new biologic (a drug that is derived from living organisms or contains components of living organisms e.g. vaccines or allergenics). They have shown that using a traditional frequentist approach, the analysis produces negative or ambiguous results from two Phase 3 clinical trials. Using the same two clinical trials and a Bayesian analysis demonstrates the unequivocal effectiveness of the new biologic treatment.

Much of their research has been devoted to the analysis of efficacy data as this is where hypothesis testing and p-values currently play a particularly dominant role. However, from their work to date, the research team believe the Bayesian approach is an appropriate methodology for the safety evaluation of a new treatment as well.

Ruberg argues that a Bayesian model that synthesises data from preclinical research and previous clinical trials as a 'prior' for a Phase 3 trial should be employed to generate probability statements that support the understanding of the magnitude of the actual treatment effect. As Dr Ruberg says, "Such a Bayesian paradigm provides a promising framework for improving statistical inference and regulatory decision making."

Such a Bayesian paradigm provides a promising framework for improving statistical inference and regulatory decision making.

The utility of the Bayesian approach requires a greater level of sophistication when compared with the frequentist methods, and therefore demands more thoughtful implementation. A prior probability or belief in the treatment efficacy is also required. Nevertheless, when the health of the population is at stake, it is imperative that the best possible synthesis of all available data is employed to make informed decisions.



Behind the Research

Dr Stephen Ruberg

E: analytixthinking@gmail.com **T:** +1 317 506 5686 **W:** www.analytixthinking.blog

Research Objectives

Dr Ruberg's work examines the statistical methods used to determine drug efficacy.

Detail

11121 Bentgrass Court
Indianapolis, IN 46236
USA

Bio

Dr Ruberg received a BA in Mathematics, MS in Statistics and a PhD in Biostatistics. He worked in the pharmaceutical industry for 37 years applying statistics to all areas of pharmaceutical research and development. Dr Ruberg was a Distinguished Research Fellow at Eli Lilly & Company before he retired and formed his own consulting company, Analytix Thinking. He is a Fellow of the American Statistical Association and an elected Fellow of the International Statistics Institute.

Collaborators

- Margaret Gamalo
- Karen Price
- Haoda Fu
- Carl Peck
- Frank Harrell (www.fharrell.com)
- J. Jack Lee

References

Ruberg, S.J., Harrell, F.E., Gamalo-Siebers, M., LaVange, L., Lee, J.J., Price, K. & Peck, C. (2019). Inference and Decision Making for 21st-Century Drug Development and Approval, *The American Statistician*, [online] 73:sup1, 319-327, Available at: URL doi.org/10.1080/00031305.2019.1566091 [Accessed 26 June 2019].

Personal Response

How do you intend to persuade the regulatory agencies to adopt the Bayesian paradigm?

"The US FDA already allows for the use of Bayesian methods when evaluating and approving medical devices. When developing new therapeutic treatments, there are some special situations – like the study of treatments for rare diseases – where there is some willingness to consider Bayesian approaches. The same or similar statements apply to selected other regulatory agencies outside the US. As our knowledge and experience with Bayesian approaches grow, there is a greater opportunity to shift from traditional frequentist hypothesis testing to Bayesian estimates of a treatment's effectiveness and safety."

$$P\left(\begin{matrix}\text{Treatment}\\ \text{works}\end{matrix}\middle|\begin{matrix}\text{Observed}\\ \text{Data}\end{matrix}\right)=\frac{P\left(\begin{matrix}\text{Observed}\\ \text{Data}\end{matrix}\middle|\begin{matrix}\text{Treatment}\\ \text{works}\end{matrix}\right)P\left(\begin{matrix}\text{Treatment}\\ \text{works}\end{matrix}\right)}{P\left(\begin{matrix}\text{Observed}\\ \text{Data}\end{matrix}\middle|\begin{matrix}\text{Treatment}\\ \text{works}\end{matrix}\right)P\left(\begin{matrix}\text{Treatment}\\ \text{works}\end{matrix}\right)+P\left(\begin{matrix}\text{Observed}\\ \text{Data}\end{matrix}\middle|\begin{matrix}\text{Treatment}\\ \text{doesn't work}\end{matrix}\right)P\left(\begin{matrix}\text{Treatment}\\ \text{doesn't work}\end{matrix}\right)}$$

Bayes' Theorem