Published in final edited form as:

Stat Biopharm Res. 2012 January 1; 4(3): 293–300. doi:10.1080/19466315.2012.698933.

Assessing Similarity to Existing Drugs to Decide Whether to Continue Drug Development

Guoguang Julie Ma,

Gilead Sciences Inc., 333 Lakeside Drive, Foster City, CA 94404

(JulieGuoguang.Ma@gilead.com)

Eric Chi,

Amgen Inc., One Amgen Center Drive, Thousand Oaks, CA 91320

Joseph G. Ibrahim, and

University of North Carolina, 301-305 East Cameron Avenue Chapel Hill, NC 27514

Robert A. Parker

University of Michigan, Ann Arbor, MI 48109

Abstract

Developing a drug requires large investments, over many years, with dramatic increases in development costs at later stages. Thus, one wants to make a No Go decision on a compound early, unless evidence continues to suggest that the project will ultimately be successful, so that resources can be focused on the most promising compounds to benefit patients. Instead of predicting the probability of success of a Phase III study, our approach to this decision uses the Phase II study results to assess similarity of the novel compound to existing drugs that are classified by different decision categories, such as a clear Go decision (e.g., a clearly effective drug), a (unfortunately common) Not Sure decision (e.g., a potentially useful but not outstanding drug), and a clear No Go decision (e.g., a clearly not effective drug). We describe how this modeling can be done using both individual and binary endpoints and how results can be combined for several different endpoints. Potential extensions of the method are also discussed.

Keywords

Bayesian methods; Clinical trial; Medical decision making; Prior elicitation

1. Introduction

Consider the problem of a decision maker with a drug portfolio deciding whether to proceed with a substantial investment for Phase III clinical trials for a compound. Ideally, this decision would be made such that the expected value of proceeding with the specific compound maximizes the entire value of the drug portfolio given all the other potential uses of the funds, especially to invest in other compounds at earlier stages of development. Such a calculation for this compound alone requires other factors that usually include an assessment of the probability of technical success in the Phase III program, meaning both a clinically important and statistically significant effect for specific endpoints in at least one,

and normally two clinical trials, as well as an assessment of the likelihood of approval of the drug based on its benefit–risk profile (Spiegelhalter, Abrams, and Myles 2004).

Clearly, this is a difficult decision problem that involves the true (unknown) benefit of the compound, often for different endpoints than those already studied. For example, long-term endpoints may be required for the regulatory filing rather than the short-term surrogate endpoints used in the Phase II study. In addition, often different populations of patients will be included in the Phase III studies. Thus, even assessing the probability of success in a single Phase III study from the Phase II study results is a major challenge (Chan et al. 2008).

An alternative approach is to assess whether the compound seems to be similar to other drugs that have been approved in the past for similar indications, or whether it appears similar to compounds that ultimately did not become approved or were commercially unsuccessful. Therefore, in this approach, we are estimating the probability that the experimental compound is like the other compounds. This changes the decision process from estimating the probability of success of the proposed Phase III study to classifying the experimental compound as similar to one of several existing compounds. One implication of this alternative approach is that the specifics of the Phase III study design are not taken into consideration in the decision process. We pursue this approach in this article.

This article is organized as follows. Section 2 introduces the basic issues and notation, and describes the basic model for binary endpoints. Section 3 provides an approach to combining results from multiple endpoints and/or multiple studies. Section 4 provides an illustration with hypothetical and real data combining results across three endpoints, and Section 5 discusses potential extensions to the approach. The final section provides a brief discussion of the potential advantages and disadvantages of the approach.

2. Notation and Basic Models

2.1 Assumptions

We assume that we have results from a randomized, parallel-group Phase II study comparing the compound of interest against a control arm. We assume binary outcomes in most of this article, and discuss additional extensions in Section 5. We assume further that the decision maker has available data for one or more other products for the same or a similar indication, so that there is at least some information on what was and was not considered sufficient benefit to justify a drug approval in the indication of interest. This prior information will usually consist of publications or presentations describing the benefit for other compounds previously studied, and possibly regulatory guidance or recommendations on what would be considered minimally adequate evidence for a submission. We assume also that the decision maker can formulate the circumstances in which they would be uncomfortable in making a decision based on the information from the Phase II study; that is, when they would require additional information before making a decision. Thus, we have changed the problem from one of determining the probability of success to one of classification, determining the probability that the results for the compound come from the class of successful drugs, of indeterminate information (or a marginal drug), or of not successful drugs.

In summary, the following priors will be constructed, based on which the posterior probability will be computed given the result from the Phase II study:

- *Drug class prior:* this is the prior distribution of a drug that leads to a clear Go decision (e.g., a clearly effective drug), a Not Sure decision (e.g., a potentially useful but not outstanding drug), and a clear No Go decision (e.g., a clearly not effective drug).

- *Compound class prior*: this is a prior that describes how likely is it that the compound belongs to each of the three drug classes in decision makers' minds.

Sometimes, a decision is based on multiple endpoints in a clinical trial, in which case it is necessary to have an *endpoint importance prior* to indicate how important each of the multiple endpoints is relative to the decision. This prior is also determined by decision makers. The extension to multiple binary endpoints is discussed in Section 3.

2.2 Drug Class Prior

The benefit of the compound is the difference in the response between treatment groups. Thus, the natural metric for the decision maker is the treatment effect ($\widehat{\delta}$), rather than the individual response rates in the two groups. We use $\widehat{\delta}$ to classify the results of our Phase II study as similar to one or another of the approved drugs. We do not allow for the possibility of a negative $\widehat{\delta}$ in our model since there would be no interest in pursuing a drug with an unfavorable effect.

Three drug class priors are considered: a Go drug class (G), a Not Sure drug class (I), and a No Go drug class (N). Since the true treatment effect δ ranges from 0 to 1, we define the distribution of δ for each of the three drug classes as a beta distribution, $\text{Beta}(a_i, b_i)$, $i = G, I, \text{ or } N$, where a_i and b_i are obtained based on the results of what our clinical colleagues considered the characteristic study for the drug underlying the decision class, or based on a meta-analysis of multiple independent studies.

2.3 Compound Class Prior

It is likely that the decision maker has some prior beliefs about the compound's utility, and that these views would tend to influence the decision. For example, the decision maker may feel before a formal analysis that this specific compound is strong and therefore have an optimistic prior with larger weight on the Go decision (d_G), and relatively less weight on the other two decisions (d_I, d_N). Alternatively, the decision maker may be very pessimistic about the compound, or be relatively neutral and weight the three outcomes equally before having the results of the Phase II studies, with the constraint that

$$d_G + d_I + d_N = 1.$$

Since the compound class prior is how strongly the decision maker feels about the compound before having the data from the current study, it can be based on any approach the decision maker wants to use and sometimes can incorporate various types of "extra" information. There is no right or wrong prior, and it can be different for different decision makers.

2.4 Phase II Study Result

For a binary endpoint, the data available would be the number of subjects having a response, r_A and r_C (active and control, respectively) from n_A and n_C participants on study. The outcome is the proportion responding, $\widehat{p}_A = r_A/n_A$ and $\widehat{p}_C = r_C/n_C$, and the effect of active treatment is measured by $\widehat{\delta} = \widehat{p}_A - \widehat{p}_C$. We assume throughout that benefit of the new compound compared to the control is coded positively, that is, $0 < \widehat{\delta} \leq 1$.

2.5 Transformation of the Observed Data for Analysis

Since we are modeling the probability that the observed difference, $\widehat{\delta}$, comes from one of the three underlying beta distributions, we need to convert the results of the original Phase II

study into equivalent results for a conceptual experiment in which we would have observed the outcome $\widehat{\delta}$. The conceptual experiment would be a single arm study where the observed outcome is the difference between the outcome with treatment and with control. Thus, the results of the Phase II study ($r_A/n_A, r_C/n_C$) need to be converted into an equivalent set of results ($\widehat{r}_\delta/\widehat{n}_\delta$) as if the outcome ($\widehat{\delta}$) was observed from a single arm study where the values of $\widehat{r}_\delta$ responses from $\widehat{n}_\delta$ observations should have the same mean and asymptotic standard error as the results from the parallel-group study. To do this, we first solve for the equivalent number of observations $\widehat{n}_\delta$ given the observed proportion difference so that the variance is the same as the observed data; that is,

$$\widehat{n}_\delta = \widehat{\delta} (1 - \widehat{\delta}) / (\widehat{p}_A (1 - \widehat{p}_A) n_A^{-1} + \widehat{p}_C (1 - \widehat{p}_C) n_C^{-1}),$$

and then solve for

$$\widehat{r}_\delta = \widehat{\delta} \widehat{n}_\delta.$$

This transformation will fail if $\widehat{\delta}$ is negative, but in that case there would be little interest in pursuing the result. Let $\tilde{r}$ and $\tilde{n}$ be the nearest integers to $\widehat{r}_\delta$ and $\widehat{n}_\delta$, the results of the conceptual experiment with a single arm being used in place of our original Phase II parallel-group study. For example, suppose $r_A = 110$, $r_C = 50$, and $n_A = n_C = 200$, then $\widehat{\delta} = 0.3$, $\tilde{n} = 97$, and $\tilde{r} = 29$.

In reality, even though a study is planned to recruit a specific number of subjects, often the actual number enrolled and evaluable differs slightly from the target. Thus, strictly speaking, both n_A and n_C in the original experiment are random variables, but this complication is ignored in most analyses. Similar to this conventional assumption, we consider that the number of observations in our conceptual trial, $\tilde{n}$, is preplanned, so that only $\tilde{r}$ is a random variable.

2.6 Modeling for a Single Binary Endpoint

Given the assumptions above, for the decision distribution i , the probability of observing $\tilde{r}$ responses from $\tilde{n}$ (fixed) trials follows a beta-binomial distribution, that is,

$$f(\tilde{r}|\tilde{n}, p) = \binom{\tilde{n}}{\tilde{r}} p^{\tilde{r}} (1-p)^{\tilde{n}-\tilde{r}}, \text{ where } p \text{ is a random variable that follows a beta distribution}$$

for decision prior i , $f(p|a_i, b_i) = \frac{p^{a_i-1} (1-p)^{b_i-1}}{B(a_i, b_i)}$. Thus,

$$\begin{aligned} P(\tilde{r}|i, a_i, b_i, \tilde{n}) &= \int_0^1 f(\tilde{r}|\tilde{n}, p) f(p|a_i, b_i) dp \\ &= \binom{\tilde{n}}{\tilde{r}} \frac{B(\tilde{r}+a_i, \tilde{n}-\tilde{r}+b_i)}{B(a_i, b_i)}. \end{aligned}$$

Then the probability that the results arise from distribution i is given by:

$$P(i|\tilde{r}, \tilde{n}, a_i, b_i, d) = \frac{d_i P(\tilde{r}|i, a_i, b_i, \tilde{n})}{\sum_i d_i P(\tilde{r}|i, a_i, b_i, \tilde{n})}. \quad (1)$$

Using the beta-binomial gives a single unique value for the result. However, we recommend that these posterior probabilities be evaluated using Markov chain Monte Carlo (MCMC) methods (Gilks, Richardson, and Spiegelhalter 1995) via WinBUGS, or SAS, for example. Using simulations allows additional information to be provided to decision makers, which becomes important when the results are close to the Go/No Go boundary. For example, in addition to knowing what the average result is, decision makers might want to know what fraction of simulations is above the boundary. Furthermore, simulations can easily be extended for multiple endpoints as discussed in the next section and other extensions described in Section 5.

3. Extensions to Multiple Binary Endpoints

Assume that we now have data and priors available on m endpoints ($m > 1$). Although an active treatment is likely to improve results on multiple different endpoints, while an inactive treatment is likely to have minimal impact on all endpoints, we ignore at this stage the problem of potential correlation between the endpoints. We discuss an extension to the analysis allowing for correlation in Section 5.

It is unlikely that all endpoints would be equally important in the decision to proceed with drug development. As an illustration, a strong signal in the primary clinical endpoint used for regulatory decisions would be far more important to the decision than would a patient-reported health-related quality-of-life outcome.

There are at least two potential approaches for summarizing the results of the different endpoints. The simplest approach would be a weighted average of the probabilities separately estimated from decision i for endpoint j , denoted by $P_j(i)$. Unlike Equation (1), $P_j(i)$ is calculated without incorporating the decision maker's weights, since the weights should be incorporated only once in the overall decision in Equation (3). With an obvious extension of notation, therefore, $P_j(i)$ is given by

$$P_j(i) = \frac{P(\tilde{r}_j | i, a_{ij}, b_{ij}, \tilde{n}_j)}{\sum_i P(\tilde{r}_j | i, a_{ij}, b_{ij}, \tilde{n}_j)}, \quad (2)$$

where each $P(\tilde{r}_j | i, a_{ij}, b_{ij}, \tilde{n}_j)$ is a random draw from the beta-binomial distribution for decision i and endpoint j .

Let $e_j, j = 1 \dots, m$ denote the relative weights for endpoint j with a constraint $\sum_j e_j = 1$. The weighted average of the results, incorporating the decision maker's weights, would then give the overall probability that the drug belongs to class i by

$$P(i) = \frac{d_i \sum_j e_j P_j(i)}{\sum_i (d_i \sum_j e_j P_j(i))}. \quad (3)$$

Again, a single unique value could be calculated if desired, using the beta-binomial formula, but again this would seem to neglect the potential for variability in the results.

An alternative approach would be based on the likelihood for each endpoint. We do this in the MCMC framework as follows. We draw a random sample from the corresponding beta distribution for decision i and endpoint j denoting the k th set of such draws by $\theta_{i,j,k}$. We calculate, for each of the observed results ($\tilde{r}_j$ events among $\tilde{n}_j$ observations), the overall likelihood kernel of the results for this draw,

$$l_k(i) = \prod_j \left(\theta_{i,j,k}^{\tilde{r}_j} (1 - \theta_{i,j,k})^{\tilde{n}_j - \tilde{r}_j} \right)^{e_j} \quad (4)$$

The probability that the set of results come from decision i for that specific draw would then be given by

$$P_k(i) = \frac{d_i l_k(i)}{\sum_i d_i l_k(i)}, \quad (5)$$

and the overall result would be the summary of these results.

4. Example

We applied the proposed method to a Phase II dose ranging study where patients were equally randomized to a placebo group or one of three dose groups. The analysis was focused on the comparison of three binary efficacy endpoints between the highest dose group and the placebo group.

4.1 Drug Class Prior

Three drug classes were denoted as G-like (a Go decision), I-like (a Not Sure decision), and N-like (a No Go decision). The drug class prior was constructed based on the following steps: (1) A review was conducted for historical studies that our clinicians would consider representative for the underlying decision classes (see Section 2.2). In this example, about two to six studies were reviewed for each drug class and a typical study was chosen to represent the corresponding drug class. (2) The observed proportion difference from the study chosen for each drug class was then converted into a result from an equivalent single arm study. Suppose the response rates from a study for the N-like drug class are $r_A = 95$ and $r_C = 55$ and the sample sizes are $n_A = 250$ and $n_C = 251$ for active and control groups, respectively. Using the same method as described in Section 2.4, we obtain $\hat{\delta} = 0.16$ and $\hat{n} = 0.16$. (3) The parameters of the beta distribution were calculated. For the N-like drug class, the parameters are $a = \hat{\delta} \times \hat{n} + 1 = 14$ and $b = \hat{n} - a + 1 = 70$ for the beta distribution.

Figure 1 shows a set of decision distributions for each endpoint by three curves: green for a Go decision; orange for a Not Sure decision; and red for a No Go decision. These colors are used throughout the figures. As is typical in our experience, there is considerable overlap between adjacent decisions, and even modest overlap between the Go and the No Go decision. This emphasizes the importance of using multiple criteria for making the decision.

4.2 Compound Class Prior

The decision makers may have different prior beliefs about how likely it is that the investigational drug would fall in each drug class, in which case sensitivity analyses should be performed as shown in Table 1. As a primary analysis, the key decision maker's beliefs are used, $d_G = 0.2$, $d_I = 0.6$, and $d_N = 0.2$; that is, a much stronger belief on a I-like drug. In a sensitivity analysis, different beliefs (neutral, optimistic, or skeptical) were also used and results are presented in Table 1.

4.3 Endpoint Importance Prior

The endpoint importance was unanimous among decision makers that there was a modest differential weighting of the three endpoints, with the second endpoint considered the most important with 50% of the total weight, the first endpoint being counted as 30%, and the last

endpoint being counted as 20%. The Win-BUGS code used for the calculations is attached in the Appendix. Each simulation was run using 20,000 MCMC iterations.

4.4 Results and Decision Making

Figures 2 and 3 show two sets of hypothetical results. Figure 2 intends to show a difficult decision situation where it is very unclear how strong the data are for the Go decision. For each endpoint, the results are close to a 50:50 decision, reflecting results close to the mid-point of overlap between the Go and the Not Sure decisions shown in Figure 1. Since there is no strong signal for any of the three endpoints, both the weighted average and weighted likelihood approaches reflect that the overall results are not clear. Figure 3 shows a much clearer result for the individual endpoints. In both cases, the weighted likelihood approach increases the magnitude of the difference between the two choices compared to the averaging approach, emphasizing the consistency of the results in the three endpoints.

Figure 4 shows the real results of the Phase II study. The drug is similar to N-like drugs with a probability of 87% or higher for endpoints 1 and 2, while the probability is about 83% for endpoint 3. The similarity to G-like drugs is as low as 1%. The overall estimates, after taking into account the weighting of the three endpoints, show clearly that this would be a No Go decision.

As a robustness check of the primary analysis, we also assume a neutral opinion, an optimistic prior, and a skeptical prior, in addition to the key decision maker's belief. Table 1 displays the weighted likelihood estimates of the probabilities based on various compound class priors. Prior opinions about the drug do change the results, but even very strong prior beliefs about the drug do not overwhelm the data. In conclusion, it is unlikely that the compound is as effective as a G- or I-like drug and clearly is an N-like (No Go) drug.

5. Extensions

We briefly mention several potential extensions of the method, without providing specific details of any of them. The most obvious extension would be to other types of endpoints; for example, normally distributed endpoints and time-to-event data. These types of endpoints can easily be modeled using standard approaches for Bayesian data analysis available in existing software, and present no additional problems to the methods presented above.

Another obvious extension would be when there are, as usually occurs, multiple preliminary and Phase II studies, usually in somewhat different indications or in different patient populations, possibly with different compound posology. The problem becomes how relevant the results in these other studies would be for the current decision, and the decision maker must provide information on this. Often the results of these other studies are ultimately considered of modest value for the decision. Depending on the specifics of the situation, it may be necessary to analyze the results using priors for the different decisions specific to each study, since the success criteria may well be substantially different across different indications and populations. Results of these analyses could then be presented separately, as a type of sensitivity analysis, or combined with explicit weights based on the decision maker's assessment of the importance of the individual studies, similar to the approach illustrated with the weighted likelihood approach across multiple endpoints shown in Equation (5). A third approach would be to assume an underlying hierarchical model, with each of the specific prior studies a realization from the hyperdistribution of the decision classes. One would then use the different study results to estimate the parameters of the hyperdistribution, and then use the hyperdistribution as the observed data to estimate the probabilities of the various decision classes (Chen and Ibrahim 2006).

A third area for extension would be the common case where there are multiple potential studies available on other treatments that could be used for priors for the different decisions. There are several ways that this can be approached as mentioned above in the case of using results from more than one preliminary study. We recommend that the results be analyzed separately for each of these individual studies to assess the consistency of the decision. Obviously, if very different decisions would be made depending on which prior study is used (e.g., clear evidence for a Go decision with one study, and a No Go decision with another), then this is critical information for the decision maker.

A fourth area for extension would explicitly model the unreliability of the results in the specific Phase II study itself. For example, the distribution of the true underlying probability of success for a binomial endpoint based solely on the Phase II study is given by a $\text{Beta}(r_{\delta} + 1, n_{\delta} - r_{\delta} + 1)$ distribution. One could potentially sample from this underlying distribution, denoting the k th sample by Θ_k , sample $r_{\delta,k}$ from a binomial distribution with parameters Θ_k and n_{δ} , and then estimate the probability of class membership using Equations (1) or (4) and (5), based on this sample.

Finally, there is the extension to correlations among the different endpoints. This raises several issues. First, when simulating results from the prior distributions, results of one endpoint would need to be incorporated in the generation of a sample from the second prior. This assumes that there is a causal model; for example, that endpoint 1 affects endpoint 2, rather than solely a correlation, so that one knows how to model the sampling. Second, it becomes less clear what weights are actually being applied to the different endpoints. In particular, the weight attached to the second endpoint provides additional weight on the first endpoint assuming a positive association. Determining the actual weight in this case would be challenging. Expanding associations to more than two endpoints would be even more challenging.

6. Discussion

We have outlined a basic approach for deciding whether to proceed to a Phase III program, based on the chance that the treatment being developed resembles existing treatments. We have framed the problem as classifying the Phase II results as similar to a successful treatment (a “Go” decision), an unsuccessful treatment (a “No Go” decision), or as one in which further information is required (based on similarity to a “marginal” product). The approach can handle single or multiple endpoints of various kinds; we give formulas for binary outcomes, but the approach can easily be extended to other endpoints and combinations of endpoints as well. The approach can explicitly incorporate the relative importance of different endpoints, as well as the decision maker’s prior opinions about whether the treatment would be successful or not. The approach described here has provided useful guidance to the decision-making process.

There are several advantages to our approach. First and foremost, it forces clear thinking about how to make the decision. Most importantly, it forces decision makers to clearly identify what success looks like for the treatment. In addition, it focuses attention on the importance of different endpoints in this decision. This is particularly important as there is sometimes a tendency to define success post hoc, after the results are available. Our approach encourages systematic background data collection, so that priors can be formed for the different drug classes. By assessing the results using different potential priors for the decision classes, it is possible to assess the sensitivity of the results to assumptions regarding the definition of success. Furthermore, our approach allows different decision makers to use different priors, to weight endpoints differently, and to have different prior probabilities as

to the likely success of the treatment. This allows the impact of these different assumptions to be explicitly recognized.

Our approach is fundamentally different from an approach in which the results of the Phase II study are used to predict the success of the Phase III study. To do this requires assumptions about the nature of the relationship between the Phase II endpoint and the Phase III endpoint required for regulatory approval and about the impact of patient characteristics on both the Phase II and Phase III endpoints, among other things. Such an approach attempts to answer the question “will the Phase III study be successful?” In contrast, our approach answers the question “does this treatment look like other successful treatments?” Although definitely not the same question, the information can be viewed as a necessary, but not sufficient, condition for success.

Our approach requires extensive data collection, which can be difficult, time consuming, and expensive. It could be hard to find the relevant information on other compounds, especially if the Phase II study involves short-term surrogate endpoints for decision making, rather than longer-term endpoints used in Phase III trials, which is often the primary information available in the published literature. In addition, the data abstraction/reduction may on occasion reduce the problem to one or at most a very small number of dimensions, which would give these endpoints more importance than they might deserve. Although we have focused on combining efficacy endpoints, the approach can be extended easily to incorporate different dimensions, such as targeted safety signals when there is an issue with other drugs having the same mechanism of action, for example.

Finally, it is important to realize that by quantifying the classification problem, we may well lead decision makers to focus on the posterior probabilities from our approach rather than in the broader decision itself. Thus, it is important when presenting results from this approach that the robustness and variability of estimates be emphasized at least as much as the estimates themselves. Given the magnitude and expense of a Phase III program, it is essential to emphasize that these results, although quantified, should be only one aspect of the overall decision.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

References

- Chan JK, Ueda SM, Sugiyama VE, Stave CD, Shin JY, Monk BJ, Sikic BI, Osann K, Kapp DS. Analysis of Phase II Studies on Targeted Agents and Subsequent Phase III Trials: What are the Predictors for Success? *Journal of Clinical Oncology*. 2008; 26:1511–1518. [293]. [PubMed: 18285603]
- Chen MH, Ibrahim JG. The Relationship Between the Power Prior and Hierarchical Models. *Bayesian Analysis*. 2006; 1:551–574. [298].
- Gilks, WR.; Richardson, S.; Spiegelhalter, DJ. *Markov Chain Monte Carlo in Practice: Interdisciplinary Statistics*. Chapman & Hall; Boca Raton, FL: 1995. [295]
- Spiegelhalter, DJ.; Abrams, KR.; Myles, JP. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. Wiley; New York: 2004. [293]

Figure 1.
Sample endpoint decision criteria.

Figure 2.
Hypothetical data—moderate effect.

Figure 3.
Hypothetical data—strong effect.

Figure 4.
Real data—minimal effect.

Table 1

Probabilities based on different prior beliefs: weighted likelihood real data—minimal effect

	A-like No Go decision	B-like Not Sure decision	C-like Go decision
Key decision makers belief 20%, 60%, 20%	0.908	0.090	0.002
Neutral belief: 33%, 33%, 33%	0.960	0.038	0.003
Optimistic belief: 5%, 20%, 75%	0.861	0.109	0.029
Skeptical belief: 75%, 20%, 5%	0.988	0.011	0.000