A Bayesian Approach to Sample Size Estimation and the Decision to Continue Program Development in Intervention Research

Ding-Geng Chen *University of North Carolina at Chapel Hill, University of Pretoria*
Mark W. Fraser *University of North Carolina at Chapel Hill*

ABSTRACT ***Objective:*** In intervention research, the decision to continue developing a new program or treatment is dependent on both the change-inducing potential of a new strategy (i.e., its effect size) and the methods used to measure change, including the size of samples. This article describes a Bayesian approach to determining sample sizes in the sequential development of interventions. ***Description:*** Because sample sizes are related to the likelihood of detecting program effects, large samples are preferred. But in the design and development process that characterizes intervention research, smaller scale studies are usually required to justify more costly, larger scale studies. We present 4 scenarios designed to address common but complex questions regarding sample-size determination and the risk of observing misleading (e.g., false-positive) findings. From a Bayesian perspective, this article describes the use of decision rules composed of different target probabilities and prespecified effect sizes. Monte-Carlo simulations are used to demonstrate a Bayesian approach—which tends to require smaller samples than the classical frequentist approach—in the development of interventions from one study to the next.

KEYWORDS: intervention research, research design, Bayesian, sample size, Monte-Carlo simulation

doi: 10.1086/693433

The determination of sample sizes is a crucial element in the development of new interventions over a series of studies. Precision in research design is critical because the early stages of intervention research are rarely so well-funded as to afford recruiting large samples. Without proper upfront consideration, a poorly specified design can lead to results that are unreliable, difficult to interpret, or not statistically robust. Indeed, a poorly specified design could produce misleading findings that result in a premature decision to stop program development, or, alternatively, a false-positive decision that leads to unwarranted further development.

Journal of the Society for Social Work and Research, Volume 8, Number 3. 2334-2315/2017/0803-0007/$10.00.

In determining sample sizes in intervention research, a Bayesian framework offers certain advantages. Differing from the more common frequentist perspective on probability, a Bayesian perspective incorporates data from prior studies into the design and development process. The Bayesian approach tends to produce studies that are relatively smaller in size but adequately powered, thus decreasing costs while preserving the reliability and robustness of the statistical findings. Focusing on the design of small-scale trials and sample sizes, this paper describes and demonstrates sample-size determination from a Bayesian perspective in intervention research.

The Intervention Research Process

Early in the development process, small effectiveness trials are often designed to compare the outcomes of a novel intervention with the outcomes observed in a routine services condition (Fraser, Richman, Galinsky, & Day, 2009; Galinsky, Fraser, Day, & Richman, 2013). Typically, the research aims to demonstrate the superior effectiveness of the novel treatment, but the research might alternatively aim to show that the novel treatment is at least equivalent to and not inferior to routine services. Additionally, a researcher might aim to demonstrate that the novel treatment offers advantages over usual services, such as reduced costs or equal effectiveness with decreased duration of treatment.

A common starting point in the development of a new intervention is to construct a null and an alternative hypothesis as follows:

- H_0 (null hypothesis): The new intervention and control conditions have the same effect on the outcome of interest; and
- H*a* (alternative one-tailed hypothesis): The new intervention is more effective than the control on the outcome of interest.

Data are collected during the intervention study and analyzed at the trial's end to assess the strength of evidence in support of the null hypothesis with a *p* value. If the *p* value is sufficiently small, the null hypothesis is rejected in favor of the alternative hypothesis, and the researcher can conclude that the novel treatment is more effective than the reference control. This is often called the frequentist perspective, and it dominates intervention research (Chen & Fraser, 2017; Chen & Peace, 2011).

In designing intervention studies, statistical power is calculated as a safeguard to make certain that an adequate number of study participants is recruited to control the probability of a Type I error (i.e., incorrectly rejecting the H_0 or accepting a false-positive finding) to a certain level (α) and to ensure that a sufficiently high probability exists of correctly rejecting a false H_0 (i.e., power). Using the frequentist approach, the sample size is calculated under the assumptions that the magnitude

of the difference between the new intervention or treatment (T), when compared with the usual services control (C), is a known, prespecified value, and the variability of the outcome of interest is also known. In practice, this information is often estimated from previous or similarly designed studies.

This frequentist approach works well when the assumption about the magnitude of the difference between the novel treatment and the reference control is correct, as is the assumption about the variability of the outcome of interest. However, the information on which these assumptions rest is often insufficient, especially in the early stage of conceptualizing a new intervention. Indeed, intervention researchers are sometimes required to make best guesses.

Rather than setting up a study without sufficient evidence, a prudent researcher designs initial studies to assess the possibility that the new intervention might have some benefit, or that it might have a magnitude of benefit that the researcher has specified. However, such questions are difficult to address in the classical frequentist statistical approach because these questions require reliable data for calculations of sample size and statistical power. Given that such reliable data are often unavailable in the early stages of intervention development, we consider in this paper an alternative: the Bayesian approach.

A Bayesian approach can be used to develop and answer critical questions early in the intervention research process, including questions such as "What sample sizes are required to assess a targeted possibility that the new intervention has a benefit over the reference control?" and "What sample sizes are required to assess a targeted possibility that the magnitude of a benefit over the reference control is at least a certain amount?" These two questions are fundamental to the sequential design/redesign process that makes up intervention research and repeatedly asks, "To warrant further development, what level of confidence is required that the new intervention is more effective?"

Corresponding to this question, this article addresses the following two questions:

Question 1: What sample sizes are required so that the researcher has sufficient confidence that the novel intervention (T) has some benefit over the reference control (C)—that is, the probability that (T) is better than (C), which can then be written as $\Pr(T > C)$ or $\Pr(T-C > 0)$?

Question 2: What sample sizes are required so that the researcher has sufficient confidence to demonstrate that the novel treatment (T) meets a prespecified benefit (δ) over the reference control (C)—that is, that (T) is better than (C) by the prespecified δ, which can be then be written as $\Pr(T-C > \delta)$?

Question 2 is more stringent than Question 1 because the intervention benefit must meet a positive prespecified value (δ). Whereas Question 1 is intended to evaluate the probability that the new treatment (T) is better than the conventional control (C), Question 2 is intended to evaluate the probability that the new treatment

(T) is superior to the conventional control (C) by a specified positive value (δ). Notice that in the Bayesian paradigm, the probabilities $\Pr(T-C > 0)$ and $\Pr(T-C > \delta)$ are more intuitive, and these replace the typical Type I error, Type II error, and significance level in the frequentist paradigm. Building on extensive theoretical developments in sample-size determination using Bayesian models in medical clinical trials (e.g., Brutti, De Santis, & Gubbiotti, 2008; Gajewski & Mayo, 2006; Mayo & Gajewski, 2004; Whitehead, Valdés-Márquez, Johnson, & Graham, 2008), this article briefly describes Bayesian methods in intervention research (see also, Chen & Fraser, 2017) and then demonstrates a Bayesian approach to research design using Monte-Carlo simulations. (For a more detailed explanation of simulations, see Chen & Chen, 2017).

Bayesian Framework for Intervention Research

Bayesian methods can be used in intervention research to answer questions 1 and 2. For researchers familiar with the R program, a review of Bayesian methods for intervention research can be found in Chen and Fraser (2017) and in Chen and Peace (2011), both of which provide programming details for calculations. For researchers familiar with SAS, Bayesian methods are described in Chen, Peace, and Zhang (2017), which presents details for step-by-step implementation using R or SAS for analysis with Bayesian methods. A comprehensive discussion on the use of Bayesian methods in clinical trials can be found in Barry, Carlin, Lee, and Muller (2010). The core Bayesian approach involves the following three steps:

1. Based on existing information or belief, identify a prior distribution (i.e., a probability distribution) regarding the effect size of a novel treatment.
2. Construct the data likelihood from the new data collected to explore the effect size of the new intervention.
3. Construct the posterior distribution by multiplying the constructed data likelihood (data collected) from Step 2 and the prior distribution in Step 1.

Once the posterior distribution is obtained (Step 3), the researcher can use this value as an updated prior distribution in Step 1 to repeat the process, gathering additional data in Step 2 (e.g., by including more participants in the study or conducting a new study) and updating the prior information with each repetition of the steps. This successive prior distribution updating process is a key feature of a Bayesian perspective on intervention research.

Sometimes no prior information is available, or the available information is based on an untested application of a new construct. Although a researcher might believe a new intervention represents an advance over conventional practice, that belief likely has low value in the selection of an effect size on which to power a preliminary study. This situation commonly occurs in the early stages of interven-

tion research or when an intervention is being applied to an entirely new population or setting. In a data-poor situation, no reliable prior information is available to construct a prior distribution.

How to proceed? The simple Bayesian solution is to start from ignorance—the equivalent of a flat prior distribution—where no information supports a reason to believe that the novel treatment will be better than the reference control. In other words, all possible values for the treatment effect are considered equally plausible. In this circumstance, data are collected and used to test how this *noninformative prior* changes based on the new information. In this case, the posterior distribution is entirely dependent on the observed data and not on the prior distribution.

Simulation Design on Sample Size and Decision-Making

Suppose we have developed a novel intervention to be compared to a usual-services control, and we want to show that this novel intervention is not only promising but also warrants further development. Because this is the new intervention's first effectiveness trial, funders will not want to invest in a large study. Rather, consistent with a design approach in which an intervention is developed in a series of small studies, we want to explore the possible effectiveness using a small-scale trial. If a reasonable probability exists that the novel intervention provides benefit, then further design and development would be warranted.

We constructed Monte-Carlo simulation studies within the Bayesian framework to assess the potential impact of different sample sizes on our decision-making process (i.e., the decision to continue development). The simulation study starts from a noninformative prior where we assume no reliable prior information is available.

Based on ordered effect size (ES) categories as recommended in Cohen (1988; see Table 1), we designed the following four scenarios:

- Scenario 1 is designed to address the first question regarding what sample sizes are required for each condition (i.e., arm) to continue developing the new intervention, based on a decision rule that the novel intervention has at least a targeted 90% probability of observing a minimal benefit (defined as ES > 0) over the reference control.

Table 1
General Guideline for Effect Sizes

Effect size	< 0.1	0.1–0.3	0.3–0.5	> 0.5
Intervention effect	Trivial	Small	Moderate	Large

- Scenario 2 is intended to answer the second question regarding what sample sizes are required for each condition to continue developing the new intervention, based on a decision rule that the intervention has at least a targeted 90% probability of observing a prespecified ES greater than 0.3 (i.e., a small to moderate ES, as defined in Table 1).
- Scenario 3 is intended to further answer the second question on what sample sizes are required for each arm to continue developing the new intervention, based on a decision rule that the intervention has at least a targeted 90% probability of observing a prespecified ES greater than 0.5 (i.e., a moderate to large ES, as defined in Table 1).
- Scenario 4 is designed to further examine the impact of different target confidence probabilities (i.e., for deciding whether to continue developing the novel intervention) under a fixed sample size of 30 subjects per arm. A small sample size is selected because small samples are often used early in the design and development process.

These four scenarios are summarized in Table 2.

Simulation Settings for Scenario 1

Suppose we decide that if a targeted posterior probability of at least 90% is observed, then we will continue developing the intervention. Suggesting that the novel treatment is better than the routine services control is equivalent to saying that the posterior probability that the true ES is ≥ 0 is at least 90%. Based on this decision rule and using the simulations, we then choose an appropriate sample size for the intervention study. In addition, we also want to avoid continuing to develop the intervention if the new treatment is no better than the control (i.e., a false-positive outcome), where the ES < 0.

For this purpose, the simulation process can be carried out in the following five steps.

Table 2
Summary of Four Simulation Scenarios and Decision Rules for Continuing Intervention Development

In each scenario, continue development if the posterior probability of the effect size (ES) . . .

	Scenario 1	Scenario 2	Scenario 3	Scenario 4
Decision rule	ES > 0 is at least 90%	ES > 0.3 is at least 90%	ES > 0.5 is at least 90%	ES > 0 is at least 60%, 70%, 80%, 90%
Sample size	20–100 per arm	20–100 per arm	20–100 per arm	Fixed at 30 per arm

Step 1: For a true standard deviation of $\sigma = 10$, create a list of values for a treatment mean (μ_T) from -2 to 10 by 1-unit intervals, and set the mean for control to $\mu_C = 0$ without loss of generality. With this setup, the values for $\mu_T \leq 0$ (i.e., -2, -1, 0) indicate the new treatment (T) either performs worse than or not better than the usual services control (C). The rest of the values for $\mu_T > 0$ (i.e., 1 to 10 by an interval of 1, such as 1, 2, 3, 4, 5, 6, 7, 8, 9, and 10) indicate the new treatment performs better than the control. Based on this setup, the values for the ES are then from -0.2 to 1 by an interval of 0.1 (i.e., -0.2, -0.1, 0, 0.1, 0.2, 0.3, 0.4, 0.5, 0.6, 0.7, 0.8, 0.9, 1).

Step 2: Create a series of sample sizes (n) from 20 to 100 by an interval of 20 (i.e., 20, 40, 60, 80, and 100).

Step 3: For each sample-size selection in Step 2, simulate a set of samples from both the new treatment (T) and conventional control (C) based on the normal distribution with the specific value of ES from Step 1.

Step 4: Based on the samples from Step 3, calculate the posterior probability of $\Pr(T-C > 0)$ and evaluate whether the posterior probability is greater than the targeted probability of 90%.

Step 5: Simulate steps 3 and 4 many times (N; we used $N = 100{,}000$) to calculate how often the decision rule of $\Pr(T-C > 0) > 90\%$ is satisfied among the N simulations, which would be the probability to continue the new intervention.

Simulation Settings for Scenarios 2 and 3

The steps for scenarios 2 and 3 are similar to those for Scenario 1, with the only change needed at Step 4.

The following change is made at Step 4 (Scenario 2): Based on the samples from Step 3, calculate the posterior probability of $\Pr(T-C > 0.3)$ and evaluate whether it is greater than the targeted probability of 90%.

The following change is made at Step 4 (Scenario 3): Based on the samples from Step 3, calculate the posterior probability of $\Pr(T-C > 0.5)$ and evaluate whether it is greater than the targeted probability of 90%.

Simulation Settings for Scenario 4

Scenario 4 was designed to examine the impact of different target probabilities under a fixed small sample size of 30 participants per arm. We believe that the initial tests of all new interventions should routinely include exploring the probability for continuing development of the proposed intervention under a specified small sample size. For this purpose, the simulation process for Scenario 4 can be carried out in the following five steps.

Step 1: It is the same as Scenario 1, Step 1.

Step 2: With a sample size of 30 subjects per arm, create a series of target probabilities of 0.6, 0.7, 0.8, and 0.9.

Step 3: For sample size $N = 30$, simulate a set of samples from both the new treatment (T) and conventional control (C), based on the normal distribution with the specific ES value from Step 1.

Step 4: Using the generated sample from Step 3, calculate the posterior probability of Pr(ES > 0) and evaluate whether it is greater than the target probability from each ES value in Step 2.

Step 5: Simulate steps 3 and 4 many times (N; we used N =100,000) to calculate how often the decision rule of Pr(ES > 0) > the values of target probability is satisfied among the N simulations, which would be the probability to continue developing the new intervention.

Results

Table 3 summarizes the simulation results for scenarios 1, 2, and 3 for the combinations of different ESs and sample sizes. From this table, the probability to continue development of the new intervention generally increases as the ESs increase, from a small value close to 0 to a large value close to 1 for all the sample sizes. For example, in Scenario 1, the probability increases from 3.1% to 97% for a sample size of 20 subjects per arm; the probability increases from 1.5% to 99.9% for a sample size of 40 subjects per arm; the probability increases from 0.9% to 100% for a sample size of 80 subjects per arm; and the probability increases from 0.4% to 100% for a sample size of 100 subjects per arm. As shown in Table 3, with a sample size of 20 subjects per arm, if the new treatment has an ES of 0.3, then we would continue development 37.6% of the time. However, if the new treatment has an ES of 0.5, then we would continue development 62% of the time. If we increase the sample size to 80 subjects per arm and the new treatment has an ES of 0.3 or 0.5, then we would continue development 73.3% or 97% of the time, respectively.

If the new treatment is no better than the conventional control (i.e., ES = 0), a 10% probability of a false positive exists; that is, we would continue development even though the new treatment is not effective. If the new intervention is worse than the control (i.e., ES < 0), a small chance exists of making an incorrect decision to continue development, although this probability decreases with larger sample sizes.

To address the more stringent criterion of Question 2 (i.e., treatment benefit must meet a positive prespecified value, δ), Scenario 2 and Scenario 3 seek to determine the probability that the novel treatment (T) has a small/moderate ES greater than the reference control defined as ES > 0.3 (i.e., Scenario 2) and a moderate/large ES defined as ES > 0.5 (i.e., Scenario 3) across different sample-size specifications, ranging from 20 to 100 subjects per arm. As shown in Table 3, moving from Scenario 1 through Scenario 3, the probabilities to continue development of the new intervention are sequentially smaller. Specifically, as shown for the small/moderate ES of 0.3 on Table 3, the probabilities to continue intervention development for

Table 3

Simulated Probabilities for Continuing Development of a New Intervention Under Different Combinations of Effect Size and Sample Size in Three Scenarios

	Scenario 1					Scenario 2					Scenario 3				
	Sample Size					Sample Size					Sample Size				
ES	20	40	60	80	100	20	40	60	80	100	20	40	60	80	100
−0.2	0.031	0.015	0.009	0.005	0.004	0.015	0.007	0.004	0.002	0.002	0.010	0.004	0.002	0.001	0.001
−0.1	0.058	0.044	0.033	0.030	0.024	0.032	0.022	0.017	0.014	0.011	0.020	0.014	0.010	0.009	0.006
0.0	0.105	0.104	0.101	0.101	0.100	0.061	0.060	0.058	0.060	0.057	0.042	0.040	0.038	0.040	0.038
0.1	0.172	0.207	0.235	0.260	0.282	0.106	0.133	0.153	0.174	0.195	0.075	0.095	0.112	0.128	0.145
0.2	0.263	0.349	0.429	0.492	0.555	0.180	0.253	0.317	0.379	0.434	0.134	0.194	0.249	0.307	0.356
0.3	**0.376**	**0.528**	**0.642**	**0.733**	**0.801**	**0.269**	**0.411**	**0.528**	**0.622**	**0.706**	**0.210**	**0.336**	**0.449**	**0.544**	**0.635**
0.4	0.498	0.697	0.816	0.894	0.939	0.380	0.586	0.728	0.827	0.895	0.310	0.508	0.659	0.773	0.853
0.5	**0.620**	**0.830**	**0.928**	**0.970**	**0.987**	**0.501**	**0.744**	**0.877**	**0.944**	**0.975**	**0.426**	**0.675**	**0.831**	**0.918**	**0.961**
0.6	0.732	0.919	0.979	0.994	0.998	0.627	0.862	0.954	0.987	0.996	0.551	0.813	0.932	0.978	0.993
0.7	0.824	0.967	0.994	0.999	1.000	0.736	0.937	0.988	0.998	1.000	0.668	0.909	0.980	0.996	0.999
0.8	0.891	0.989	0.999	1.000	1.000	0.828	0.976	0.997	1.000	1.000	0.773	0.963	0.995	0.999	1.000
0.9	0.941	0.996	1.000	1.000	1.000	0.894	0.992	1.000	1.000	1.000	0.854	0.987	0.999	1.000	1.000
1.0	0.970	0.999	1.000	1.000	1.000	0.941	0.998	1.000	1.000	1.000	0.915	0.996	1.000	1.000	1.000

Note. ES = effect size. The bold rows for ESs 0.3 and 0.5 correspond to small/moderate and moderate/large ESs, respectively.

Scenario 1 are 0.376, 0.528, 0.642, 0.733, and 0.801; for Scenario 2, the probabilities decrease to 0.269, 0.411, 0.528, 0.622, and 0.706; and the probabilities decrease further for Scenario 3 to 0.210, 0.336, 0.449, 0.544, and 0.635. Similarly, as shown in Table 3 in the row for the moderate/large ES of 0.5, the probabilities to continue intervention development for Scenario 1 are 0.620, 0.830, 0.928, 0.970, 0.987; for Scenario 2, the probabilities decrease to 0.501, 0.744, 0.877, 0.944, and 0.975; and the probabilities decrease further for Scenario 3 to 0.426, 0.675, 0.831, 0.918, and 0.961.

Simulation results showing the probability of continuing intervention development for Scenario 1 are illustrated graphically in Figure 1. The horizontal axis in this figure represents the ES, with values greater than 0 indicating a benefit associated with the new intervention and values less than 0 indicating a detrimental effect of the new intervention. The three vertical dashed lines on the horizontal axis represent no real effect (ES = 0), a small/moderate effect (ES = 0.3), and a moderate/large effect (ES = 0.5). The vertical axis represents the probability of the decision to continue development of the new intervention. The various colored dashed lines and the solid line on the figure represent different sample sizes, showing the benefits of increasing the sample size.

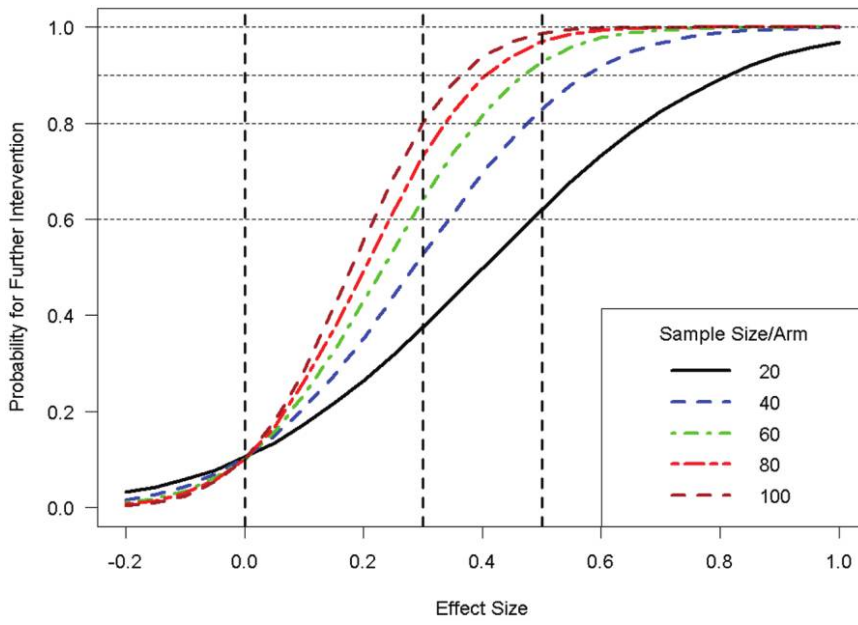

Figure 1. Simulation results showing the probability for continuing development with different specifications of sample sizes and effect sizes under Scenario 1.

Scenario 4 differs from the first three scenarios in that the aim of Scenario 4 is to examine different specifications of the target values (60%, 70%, 80%, and 90%) on the probability to continue development of the new intervention using a fixed small sample size of 30 participants per arm. This scenario reflects real-world situations in which available resources, such as research budgets and time, are limited and only a small sample can be afforded. In this situation, the target value of 90% would have to be lowered. In other words, the cost of a smaller sample is increased risk for a false-positive finding and making an incorrect decision to continue development.

Table 4 summarizes the results from the simulation study for Scenario 4, and Figure 2 presents a graphic illustration of the simulation results for Scenario 4. As can be observed from Table 4 and Figure 2, when the target value is lowered from 90% to 80%, then to 70%, and then to a less stringent 60%, the probabilities increase for continuing intervention development. For example, if we choose 60% as a target value (the second column in Table 4 and top solid line in Figure 2), the probability for continuing development increases from 0.4 to 1, with a corresponding positive intervention ES ranging from 0 to 1. Specifically, when a new intervention has a small/moderate ES of 0.3 and a moderate/large ES of 0.5, we would continue devel-

Table 4

Simulated Probabilities for Continuing Development of a New Intervention Under Different Combinations of Target Probabilities and Effect Sizes in Scenario 4

	Target Probability			
ES	60%	70%	80%	90%
−0.2	0.152	0.099	0.055	0.021
−0.1	0.264	0.184	0.112	0.048
0.0	0.400	0.301	0.201	0.103
0.1	0.552	0.446	0.325	0.188
0.2	0.701	0.596	0.476	0.311
0.3	**0.817**	**0.739**	**0.627**	**0.457**
0.4	0.902	0.849	0.762	0.611
0.5	**0.953**	**0.920**	**0.861**	**0.744**
0.6	0.981	0.965	0.931	0.851
0.7	0.993	0.986	0.968	0.922
0.8	0.998	0.995	0.988	0.965
0.9	0.999	0.998	0.996	0.986
1.0	1.000	1.000	0.999	0.995

Note. ES = effect size. Sample size is fixed at $N = 30$ per arm. The bold rows for ESs 0.3 and 0.5 correspond to small/moderate and moderate/large ESs, respectively.

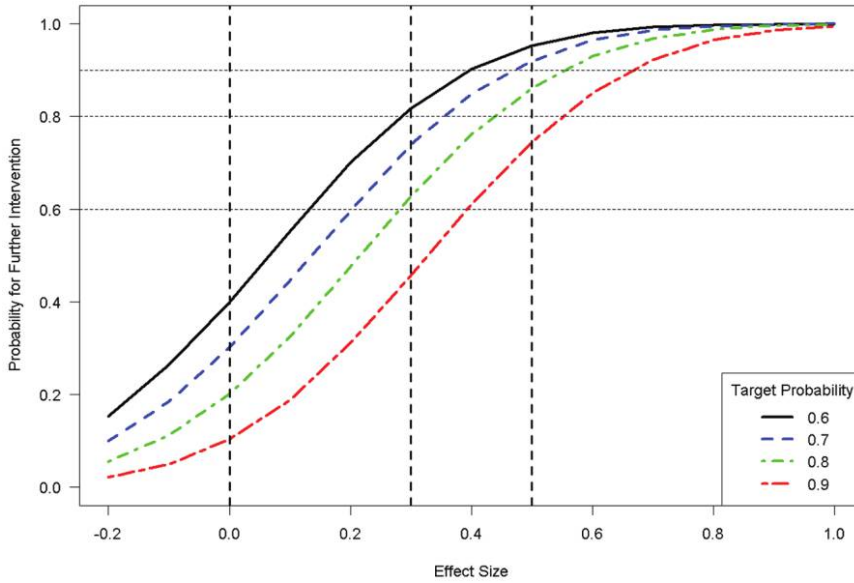

Figure 2. Simulation results showing the probability for continuing development with different specifications of target confidence probabilities and effect sizes under Scenario 4 with sample size fixed at $N = 30$ per arm.

opment more than 81.7% or 95.3% of the time, respectively. However, if the benefit of the new intervention is no better than the usual services control (ES = 0), then there is a 40% probability that we would continue development even though the new intervention is not effective. Because 60% is a far less stringent target, this situation has an inherently high risk of a false-positive recommendation.

Discussion

In intervention research, a Bayesian approach is useful when, as a part of a sequence of studies intended to gradually refine a new social or health program, projecting an expected ES in calculating power is difficult because prior information is of uncertain value. Bayesian methods enable intervention researchers to design intervention studies in the context of differentially informative prior data.

Using Monte-Carlo simulations, we presented four scenarios to show the applicability of the Bayesian approach with continuous and normally distributed data with a noninformative prior distribution. That is, the shape of the prior distribution was flat, and there was no contribution of a prior intervention to the new study. However, the results can be easily extended to an informative prior distribution if reliable data are available. With the use of a noninformative prior distribution, researchers can design intervention studies and evaluate results within a Bayesian framework without searching for prior data that are potentially inappropriate or even misleading.

Intervention research is developmental in nature, often beginning with a new construct or mediating mechanisms, the creation of program activities and other content (e.g., recruitment and training materials), and small pilot studies. To power these studies under a frequentist perspective, researchers must make assumptions based on scant information. Bayesian methods hold the potential to be more informative and forgiving in the sense that these methods provide a technique for repeatedly reassessing the state of the evidence and incorporating new findings—in the form of revised prior distributions—in decision-making processes.

Author Notes

Ding-Geng Chen, PhD, is the Wallace H. Kuralt Distinguished Professor and Director of Statistical Development and Consultation at University of North Carolina at Chapel Hill and an extraordinary professor at University of Pretoria, South Africa.

Mark W. Fraser, PhD, is the John A. Tate Distinguished Professor for Children in Need in the School of Social Work at the University of North Carolina at Chapel Hill.

Correspondence regarding this article should be directed to Ding-Geng Chen via e-mail to dinchen@email.unc.edu

Acknowledgments

We thank Ms. Diane C. Wyant for providing helpful comments on the draft of this paper.

References

Barry, S. M., Carlin, B. P., Lee, J. J., & Muller, P. (2010). *Bayesian adaptive methods for clinical trials*. Boca Raton, FL: Chapman & Hall/CRC Biostatistics Series.

Brutti, P., De Santis, F., & Gubbiotti, S. (2008). Robust Bayesian sample size determination in clinical trials. *Statistics in Medicine*, *27*(13), 2290–2306. doi:10.1002/sim.3175

Chen, D. G., & Chen, J. D. (2017). *Monte-Carlo simulation-based statistical modeling*. New York, NY: Springer/ICSA Book Series in Statistics.

Chen, D. G., & Fraser, M. W. (2017) A Bayesian Perspective on Intervention Research: Using Prior Information in the Development of Social and Health Programs. *Journal of the Society for Social Work and Research*, *8*. Advance online publication. doi:10.1086/693432

Chen, D. G., & Peace, K. E. (2011). *Clinical trial data analysis using R*. Boca Raton, FL: Chapman & Hall/CRC Biostatistics Series.

Chen, D. G., Peace, K. E., & Zhang, P. (2017). *Clinical trial data analysis using R and SAS*. Boca Raton, FL: Chapman & Hall/CRC Biostatistics Series.

Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Mahwah, NJ: Lawrence Erlbaum.

Fraser, M. W., Richman, J. M., Galinsky, M. J., & Day, S. H. (2009). *Intervention research: Developing social programs*. New York, NY: Oxford University Press.

Gajewski, B. J., & Mayo, M. S. (2006). Bayesian sample size calculations in phase II clinical trials using a mixture of informative priors. *Statistics in Medicine*, *25*(15), 2554–2566. doi:10.1002/sim.2450

Galinsky, M. J., Fraser, M. W., Day, S. H., & Richman, J. R. (2013). A primer for the design of practice manuals: Four stages of development. *Research on Social Work Practice*, *23*, 219–228. https://doi.org/10.1177/1049731512468957

Mayo, M. S., & Gajewski, B. J. (2004). Bayesian sample size calculations in phase II clinical trials using informative conjugate priors. *Controlled Clinical Trials*, *25*(2), 157–167. doi:10.1016/j.cct.2003.11.006

Whitehead, J., Valdés-Márquez, E., Johnson, P., & Graham, G. (2008). Bayesian sample size for exploratory clinical trials incorporating historical data. *Statistics in Medicine*, *27*(13), 2307–2327. doi:10.1002/sim.3140

Manuscript submitted: January 26, 2017
Revision submitted: March 2, 2017
Accepted: March 7, 2017
Electronically published: July 12, 2017