A Bayesian Perspective on Intervention Research: Using Prior Information in the Development of Social and Health Programs

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*:** By presenting a simulation study that compares Bayesian and classical frequentist approaches to research design, this paper describes and demonstrates a Bayesian perspective on intervention research. ***Method*:** Using hypothetical pilot-study data where an effect size of 0.2 had been observed, we designed a 2-arm trial intended to compare an intervention with a control condition (e.g., usual services). We determined the trial sample size by a power analysis with a Type I error probability of 2.5% (1-sided) at 80% power. Following a Monte-Carlo computational algorithm, we simulated 1 million outcomes for this study and then compared the performance of the Bayesian perspective with the performance of the frequentist analytic perspective. Treatment effectiveness was assessed using a frequentist *t*-test and an empirical Bayesian *t*-test. Statistical power was calculated as the criterion for comparison of the 2 approaches to analysis. ***Results*:** In the simulations, the classical frequentist *t*-test yielded 80% power as designed. However, the Bayesian approach yielded 92% power. ***Conclusion*:** Holding sample size constant, a Bayesian analytic approach can improve power in intervention research. A Bayesian approach may also permit smaller samples holding power constant. Using a Bayesian analytic perspective could reduce design demands in the developmental experimentation that typifies intervention research.

KEYWORDS: intervention research, *t*-test, Bayesian, prior distribution, posterior distribution, statistical power, Monte-Carlo simulation

doi: 10.1086/693432

Interventions are purposefully implemented change strategies. They range from brief communication techniques, such as those intended to promote client motivation in substance-abuse treatment, to multiyear prevention programs, such as those intended to improve the academic achievement and life-course outcomes of elementary school children. Intervention research is the systematic design and development of these change strategies. Intervention research tends to follow a design and evaluation process in which activities build on prior information over time.

Journal of the Society for Social Work and Research, Volume 8, Number 3. 2334-2315/2017/0803-0006/$10.00.

Typically, these activities include five steps: Step 1, develop problem and program theories; Step 2, design program materials; Step 3, refine program components in a series of pilot tests and experiments; Step 4, test effectiveness in large trials in practice settings; and Step 5, disseminate program findings and materials (e.g., Fraser & Galinsky, 2010; Fraser, Richman, Galinsky, & Day, 2009). The process is iterative and nonlinear in the sense that researchers often find themselves refining and testing new materials while they run effectiveness trials and disseminate findings from earlier phases of program development.

Intervention researchers use design processes that rely on prior information (e.g., Charles, Gorman-Smith, & Jones, 2016; Wu, Fraser, Guo, Day, & Galinsky, 2016). For example, in Step 1 (develop problem and program theories), a theory of the problem is usually developed from previous research on the causes and correlates of the problem an intervention is intended to address. A problem theory is then used to identify malleable mediators and to specify those mediators in a theory of change, which describes how an intervention is expected to work. Together, problem theory and change theory comprise *program theory*—that is, the latent causal argument of an intervention. This argument is sometimes called a program's *deep structure* in the sense that the program theory specifies a risk process that intervention researchers hope to change with program activities that create new opportunities, build knowledge, strengthen skills, and change environmental conditions (Resnicow, Soler, Braithwaite, Ahluwalia, & Butler, 2000).

In intervention research, program materials are developed, tested, and revised over many studies; these materials specify program activities intended to alter risk processes (e.g., Bender et al., 2015; Kim, Oesterle, Hawkins, & Shapiro, 2015; Schwinn, Hopkins, & Schinke, 2016). Activities can range from a direct focus on risk factors to a direct focus on protective factors that might reduce risk exposure (e.g., creating an after-school mentoring program with academic and recreational activities that promote health knowledge, improve study skills, and reduce opportunities for association with delinquent peers). Programs often build in complexity over time, and one study informs the next. New information sometimes leads to major program adaptations to address sociocultural risk factors that operate in a new population. For example, an acculturation and settlement intervention for arriving immigrants who have escaped a war zone may need new content on trauma recovery and resources. Although the deep structure of an intervention might change little, program activities and processes—so-called *surface structures*—are often adapted to new settings and populations. This development of a new intervention takes place through revision of program materials based on qualitative and quantitative data collected in a series of related studies (Galinsky, Fraser, Day, & Richman, 2013).

Although prior information informs successive iterations of program materials in intervention research, prior information is rarely considered in data analysis per se. To be sure, information from prior studies is incorporated in power estimates

(e.g., specification of an expected effect size based on prior research). However, information from prior studies does not condition analyses per se. But it might.

This article describes a Bayesian perspective on intervention research. Bayesian methods make use of prior information in analyses. In particular, the use of prior information is incorporated in the estimation of means and standard errors in statistical data analysis. From a research-design perspective, Bayesian methods have the potential to improve power and, assuming a constant significance level (α), reduce required sample sizes in pilot tests. If smaller samples could be used, the cost of pilot tests might be reduced, which in turn could reduce the design demands of intervention research.

Method

From Bayes' Theorem to Bayesian Modeling

The Bayesian analytic perspective originated from Bayes' theorem (Bayes & Price, 1763) developed by Thomas Bayes (1702–1761). Based on the principle of conditional probability, the theorem is simple, yet elegant. Given two events (A and B) with $P(B) > 0$, the conditional probability of A given B is defined as the ratio of the joint probability of A and B, and the probability of B—that is, $P(A \mid B) = P(AB)/P(B)$. Similarly, with $P(A) > 0$, the conditional probability of B given A is defined as the ratio of the joint probability of A and B and the probability of A—that is, $P(B \mid A) = P(AB)/P(A)$. Based on these two conditional probabilities, we can easily obtain:

$$P(B \mid A) = P(B)P(A \mid B)/P(A). \tag{1}$$

This equation exactly denotes Bayes' theorem that the "posterior" probability, $P(B|A)$, is proportional to the prior probability, $P(B)$, and a conditional probability for event A given event B, $P(A|B)$, with proportional constant, $1/P(A)$. To be exact, the posterior probability, $P(B|A)$, is the ratio of $P(B)P(A|B)/P(A)$.

To illustrate use of Bayes' theorem, suppose that B denotes the event of a patient having breast cancer and A denotes the event of the patient having a positive mammogram. From Equation 1, the conditional probability $P(B|A)$ of the patient having breast cancer (B) given the patient has a positive mammogram (A) can be calculated from (prior) knowledge of the unconditional probabilities $P(A)$ and $P(B)$, as well as the conditional probability of $P(A|B)$. $P(A)$ could be estimated as the proportion of patients with a positive mammogram; $P(B)$ could be estimated as the proportion of patients having breast cancer; and $P(A|B)$ could be estimated as the proportion of patients having breast cancer and positive mammogram results. Technically, Bayes' theorem expresses the posterior probability for a hypothesis (B) of having breast cancer after a positive mammogram (A) is observed—in terms of the prior probabilities of B and A, and the probability of A given B. For instance, suppose a

mammogram is 95% accurate in detecting breast cancer among patients with known breast cancer—that is, the *sensitivity* of the mammogram, $P(A|B) = 0.95$—and is 99% accurate in failing to detect breast cancer among patients not having breast cancer—that is, the *specificity* of the mammogram, $P(A^c|B^c) = 0.99$—where A^c and B^c denote the compliment of A and B. Further, suppose 1% of all women will have breast cancer—that is, the *prevalence* of breast cancer, $P(B) = 0.01$. Bayes' theorem enables researchers to calculate the probability that a patient has breast cancer, given the mammogram was positive—that is, the *precision* of the mammogram, $P(B|A)$. For example,

$$P(B|A) = \frac{P(B)P(B|A)}{P(A)} = \frac{P(B)P(B|A)}{P(A|B)P(B) + P(A|B^c)P(B^c)}$$
$$= \frac{0.95 \times 0.01}{0.95 \times 0.01 + (1 - 0.99) \times (1 - 0.01)} = 0.45$$

—a surprisingly small probability that demonstrates why good medicine requires further diagnostics.

The researcher can also calculate the probability that a patient does not have breast cancer given the mammogram was negative—that is, the *negative predictive probability* of the mammogram, $P(B^c|A^c)$, such as,

$$P(B^c|A^c) = \frac{P(A^c|B^c)P(B^c)}{P(A^c|B^c)P(B^c) + P(A^c|B)P(B)}$$
$$= \frac{0.99 \times (1 - 0.01)}{0.99 \times (1 - 0.01) + (1 - 0.95) \times 0.01} = 0.999.$$

Applied to intervention research, if A represents observed data (D) and B is described in terms of a hypothetical intervention effect parameter (θ), then the probability, $P(A|B)$, is the likelihood function $L(\theta) = L(\theta|D)$, and $P(B) = P(\theta)$ is the prior distribution about the parameter (θ). In this setting Bayes' theorem becomes

$$P(\theta \mid D) = P(\theta)P(D \mid \theta)/P(D), \tag{2}$$

where θ stands for any intervention parameter of interest whose probability can be affected by data; D corresponds to data yet to be observed and therefore not used in contributing to the prior probability; $P(\theta)$, the prior probability distribution, is the probability of intervention parameter θ before D is observed, which is the prior belief about how likely different parameters are; $P(D \mid \theta)$ is the probability of observing D given intervention parameter θ (also known as the data likelihood that is used for the classical frequentist statistical modeling); $P(D)$ is the marginal likelihood as the integration of $P(D \mid \theta)$ over all θ, which is then a constant and unrelated to θ. $P(\theta \mid D)$ is the posterior probability, which is the probability of intervention pa-

rameter θ after D is observed; this posterior probability is the probability of various hypotheses about θ given the observed data incorporated as prior knowledge. This *posterior probability function* is central in Bayesian modeling and is used for estimation of the intervention effect after the data are collected.

Equation 2 shows that in Bayesian modeling, the posterior probability of a hypothesis is determined by a combination of the prior belief of the likeliness of a hypothesis and the compatibility of the observed data with the hypothesis (likelihood). This likelihood means that the posterior probability is proportional to the prior information multiplied by the likelihood function (as seen in Equation 2).

For modeling purposes, the Bayesian paradigm starts with the *prior distribution* of the parameter, $P(\theta)$, where θ typically represents the intervention effect size. After new data (D) are collected, Bayes' theorem is applied to derive an updated distribution, which is called the *posterior distribution* of the parameter, $P(\theta \mid D)$, and is used to make statistical inferences about the intervention parameter (θ). The prior distribution can be updated as more data are cumulated, and a new posterior distribution can be derived for updating statistical inferences. In this manner, the Bayesian approach follows the general scientific principle of using cumulative information or knowledge to make inferences.

The advantage of a Bayesian perspective is that it provides a way to combine new data with prior information through the application of Bayes' theorem. Prior information might come from pilot tests, other research, and theory. A design and development approach to intervention research makes use of prior knowledge to formulate a program theory and relies on new information from pilot tests to refine intervention protocols. Bayes' theorem can be applied iteratively: After observing some data, the resulting posterior probability can be treated as a prior probability, and a new posterior probability can be computed from new data. This iterative process allows for Bayesian principles to be applied to various kinds of data, whether viewed all at once or over time. The literature on Bayesian modeling is vast but largely unapplied in social work intervention research (e.g., Berger, 1985; Carlin & Louis, 2008; Chen, Peace, & Zhang, 2017; Gelman et al., 2013). In the next section, we illustrate the implementation of a Bayesian perspective in intervention research.

Pilot Intervention Study

Developing an intervention involves a series of pilot studies. Although large effectiveness trials typically garner all the media attention, intervention research begins with conceptualization and small controlled studies, often with both qualitative and quantitative data collection. After the initial design phase of program materials, implementation of pilot studies typically starts with a power analysis based on information from previous research. To estimate a sample size, a researcher sets power (usually at .80) and posits an expected effect size and a standard deviation.

Suppose a researcher had conducted an initial pilot study to compare the effects (denoted by θ) of a new Intervention 2 (i.e., treatment) to a usual services Intervention 1 (i.e., control). The researcher would use a small sample (n_0) from each condition. This pilot study produced a treatment effect (such as a difference in outcomes between intervention and control conditions) of θ_0 and the corresponding standard deviation (τ). This pilot study information would serve as prior information, denoted as

$$P(\theta) = N(\theta_0, \tau^2). \tag{3}$$

This prior information is then used for statistical power analysis to determine the sample size for subsequent intervention studies. Using this prior information, the researcher can perform a power analysis with the probability of Type I error (α) controlled, such as at $\alpha = 2.5\%$ (one-sided) with a power of 80% (or 90%), the sample size can be determined as n_1.

Statistical Analysis With *t*-Test: The Frequentist Perspective

From the power analysis, a new pilot study with sample size n_1 is designed, and data are collected for both treatment and control conditions. The classical *t*-test can be used to test the treatment effect (θ) from the collected data with the associated t statistic defined as

$$t = \frac{\hat{\theta}}{\hat{\sigma}}, \tag{4}$$

where $\hat{\theta}$ is the estimated treatment effect and $\hat{\sigma}$ is the standard error. The typical estimate for $\hat{\theta}$ is the observed mean difference between conditions with $\hat{\sigma}$ as the pooled standard error. In statistics, this *t*-test is classified as part of the *frequentist paradigm*; in this paradigm, a conclusion is drawn based on only the current intervention study. (For a discussion of the frequentist paradigm, see Levy, 2016.) That is, in the analysis, the researcher ignores the prior information from pilot studies (Box, 1987). Ignoring this prior information can reduce power, as we show in the remainder of this article.

Bayesian Perspective on Intervention Research

Rather than ignoring prior information, the Bayesian approach to intervention research incorporates prior information from new data distributions based on Bayes' theorem in Equation 1. Information from pilot tests is used to formulate a posterior distribution. This posterior distribution is then incorporated in the inferential process. In this sense, a Bayesian approach to intervention research analyzes current trial data by drawing on information from previous trials. The Bayesian perspective provides a sequential quantitative method for estimating outcomes in newly obtained data by making use of the previous understanding of effects.

Specifically, the researcher can denote the data distribution, $P(D|\theta)$—that is, likelihood function $L(D|\theta)$—from the new intervention study to estimate an intervention effect as

$$D|\theta \sim P(D|\theta) = N(\hat{\theta}, \hat{\sigma}^2). \quad (5)$$

Using Bayes' theorem, the Bayesian paradigm is then characterized by combining this data likelihood—$P(D|\theta)$ from Equation 5—with the prior distribution—$P(\theta)$ from Equation 3—to construct the Bayesian posterior distribution, $P(\theta|D)$. It has been shown (e.g., Chen et al., 2017; Gelman et al., 2013) that

$$P(\theta|D) \propto P(D|\theta)P(\theta) = L(D|\theta)P(\theta) = N(\hat{\mu}_P, \hat{\sigma}_P^2). \quad (6)$$

Central in Bayesian modeling, this combined posterior distribution then incorporates information from both the new data collection (D) and the prior information. The combined distribution is used for further statistical inference to test for an intervention effect. In this posterior distribution, $\hat{\mu}_P = w\hat{\theta} + (1 - w)\theta_0$ is the posterior mean calculated as a weighted mean of the prior treatment effect (θ_0) and the treatment effect from the current study ($\hat{\theta}$). The weight in the posterior mean is defined as

$$w = \frac{\frac{n_1}{\hat{\sigma}^2}}{\frac{1}{\tau^2} + \frac{n_1}{\hat{\sigma}^2}},$$

which depends on the prior variance and data variance. The posterior variance is

$$\hat{\sigma}_P^2 = \frac{1}{\frac{1}{\tau^2} + \frac{n_1}{\hat{\sigma}^2}},$$

which incorporates variances from the prior variance (τ^2) and the data variance from the current intervention study ($\hat{\sigma}^2$). It can be proven that this posterior variance is less than the current data variance (i.e., $\hat{\sigma}_P^2 < \hat{\sigma}^2$) regardless of any values of the prior variance (τ^2), which indicates that the Bayesian estimate (i.e., posterior mean, $\hat{\mu}_P$) of the intervention effect has shrunk in comparison to the classical frequentist estimate of the treatment effect, where the variance is estimated by $\hat{\sigma}^2$. The posterior mean ($\hat{\mu}_P$) and the posterior variance ($\hat{\sigma}_P^2$) are used to calculate the empirical Bayesian t ratio between $\hat{\mu}_P$ and $\hat{\sigma}_P$ for statistical inference.

Note that the only situation in which the posterior variance of $\hat{\sigma}_P^2$ approaches the current data variance for the treatment effect (i.e., $\hat{\sigma}^2/n_1$) is when the prior variance (τ^2) is infinite, as seen from the formulation in the posterior variance of

$$\hat{\sigma}_P^2 = \frac{1}{\frac{1}{\tau^2} + \frac{n_1}{\hat{\sigma}^2}}.$$

In this case, the first pilot study provides no information to the current study. This scenario is called a *noninformative prior* in Bayesian modeling. If a noninformative

prior is used, the Bayesian posterior distribution approaches the data likelihood or frequentist solution. Another application of the noninformative prior involves cases in which prior information on the intervention effect might be biased or irrelevant (e.g., because of substantial changes to the intervention or application to an entirely new population). In this situation, the prior variance (τ^2) can be chosen to be infinitely large to diminish the influence of prior information on the Bayesian inference.

Monte-Carlo Simulation Study

We designed a Monte-Carlo simulation to illustrate this Bayesian perspective and compare outcomes with the classical frequentist perspective. For this simulation, suppose that the researcher would like to test an intervention effect size of 2 or greater (i.e., $H_0 : \theta = 0$ vs. $H_a : \theta = 2 > 0$). In addition, suppose the researcher had a pilot study with a relatively small sample size ($n_0 = 20$) from each condition (i.e., treatment and usual services). This pilot study produced a treatment effect of $\theta_0 = 2$ and standard deviation $\sigma_0 = 10$, with a standardized treatment effect of $\theta_0/\sigma_0 = 2/10 = 0.2$. With these data, a *t*-test can be performed, which produces a *p* value of 0.265. This value means that the researcher does not have enough evidence to reject the null hypothesis and argue that the new intervention is significantly better than the conventional intervention. However, the observed effect size of 0.2 is promising; therefore, the researcher might want to develop a new study to further test the new intervention. Using the pilot study data, the researcher performs a power analysis with the probability of Type I error controlled at 2.5% (one-sided) with a power of 0.80. From this analysis, the researcher estimates the sample size for a new trial to be $n_1 = 393$ per arm. We used this information to simulate the new study.

Using the sample size of 393 per arm, we simulated 1 million intervention studies to evaluate the performance of a Bayesian versus a frequentist perspective on intervention research. Each simulation used the *t*-test and was constructed using the following steps:

- Step 1: To mimic the predefined treatment effect of $\theta = 2$, randomly generate n_1 subjects from the control arm with a mean of 0 and standard deviation of 10, and n_1 samples for the treatment arm with a mean of 2 and standard deviation of 10.
- Step 2 (frequentist perspective): Using the samples from both conditions, the traditional two-sample *t*-test in Equation 4 is estimated to test the H_0, and the associated *p* value is recorded.
- Step 3 (Bayesian perspective): From a Bayesian perspective, the samples from the current study are used to construct a data likelihood distribution, as shown in Equation 5. The posterior distribution incorporates the prior

data and is constructed based on Equation 6. A p value was obtained from this posterior distribution to test the H_0.

We ran these three steps 1 million times, simulating 1 million separate intervention studies. The statistical power can then be calculated as the proportion of the total number of simulations with p value $< \alpha$ (i.e., reject H_0) from the 1 million simulations.

Results

Distributions for Means

For each of the simulated studies with sample size $n_1 = 393$, we can calculate the means for each arm, where the true mean for the control arm is 0 and the true mean for the treatment arm is 2. The 1 million simulations produce 1 million means for each arm; the distributions of the means are illustrated in Figure 1. The distributions for both arms are normally distributed with the center at 0 and 2 with 95% confidence intervals of [–0.99, 0.99] for the control arm and [1.01, 2.99] for the treatment arm. These confidence intervals indicate that the simulations worked as expected.

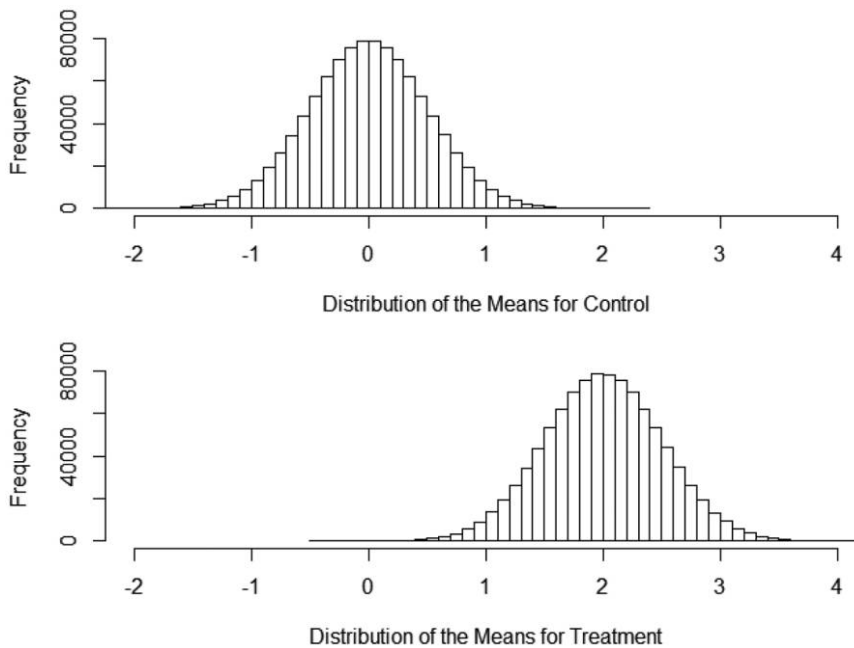

Figure 1. Distributions of the means for the control and treatment arms obtained from 1 million simulations.

Distributions for Standard Errors

For each of the simulated studies, we calculated the standard error (*SE*) for the treatment effect (i.e., the mean difference) as shown in Equation 4, and we calculated the *SE* for the Bayesian model as shown in Equation 6. In the classical frequentist model, the *SE* for the mean difference can be calculated as the square-root of the pooled variance ($\hat{\sigma}^2$), which is

$$\hat{\sigma} = \frac{10}{\sqrt{393/2}} = 0.713,$$

corresponding to the setups in the simulation studies. In the Bayesian model, the *SE* (σ_P) can be calculated as in Equation 6 by incorporating the prior study and data from the current study. This calculation produces a standard error of 0.504, corresponding to the setup in this simulation study.

In general, because the Bayesian model incorporates more information, the Bayesian *SE* should be smaller than the *SE* in the classical frequentist model. Shown in Figure 2, the means of these 1 million *SE*s were 0.713 for the classical frequentist perspective and 0.504 for the Bayesian perspective. Also shown in Figure 2, the 95% confidence intervals (typically called *credible intervals* in the Bayesian perspective) were [0.678, 0.749] for the classical frequentist model and [0.479, 0.529] for the Bayesian model.

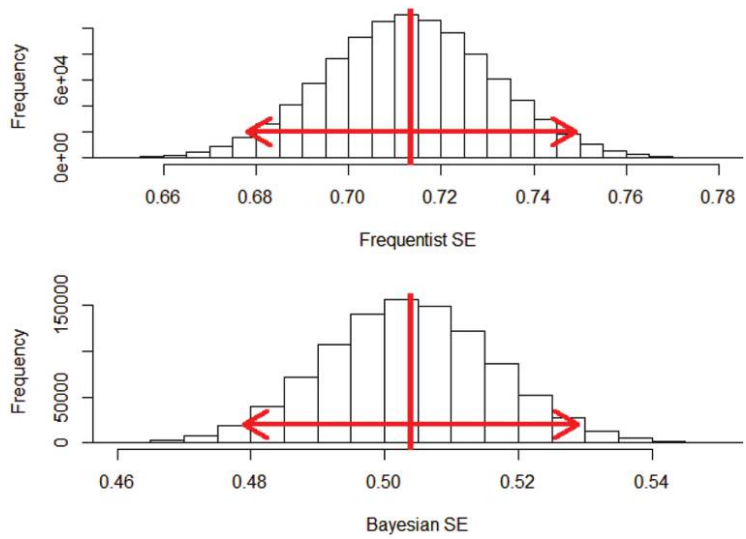

Figure 2. Distributions of standard errors (*SE*) for both frequentist and Bayesian models obtained from 1 million simulations. The vertical red lines signify the means, and the horizontal red lines represent 95% confidence intervals.

Statistical Power: The Bayesian or the Frequentist Perspective?

As demonstrated in the simulation, the Bayesian *SE* is smaller than the frequentist *SE* because the Bayesian model incorporates information from the prior study, whereas the frequentist *SE* ignores the prior study information. With a smaller *SE*, the statistical power of the Bayesian approach is higher. That is, holding the sample sizes and mean differences constant, we are more likely to detect a significant difference when using a Bayesian perspective.

For the 1 million simulated studies, we can track the number of *t*-tests that have a *p* value less than the one-sided prespecified α—the significance level used to calculate power. For the classical frequentist *t*-test, the 1 million simulations contain 799,510 significant studies. This value confirms the power to be 80%, which is consistent with the study design of 80% power. For the Bayesian model, however, the 1 million simulations contain 921,838 significant studies, yielding a power of 92%—higher than the classical frequentist approach.

An Illustrative Study

Because Bayesian modeling is more powerful than the classical frequentist *t*-test, the Bayesian analysis is more likely to detect significant differences. We illustrate this concept with a simulated study for which the control arm and treatment arm data distributions are shown in Figure 3. As seen in the figure, the true mean for the control arm is 0 and the true mean for the treatment arm is 2.

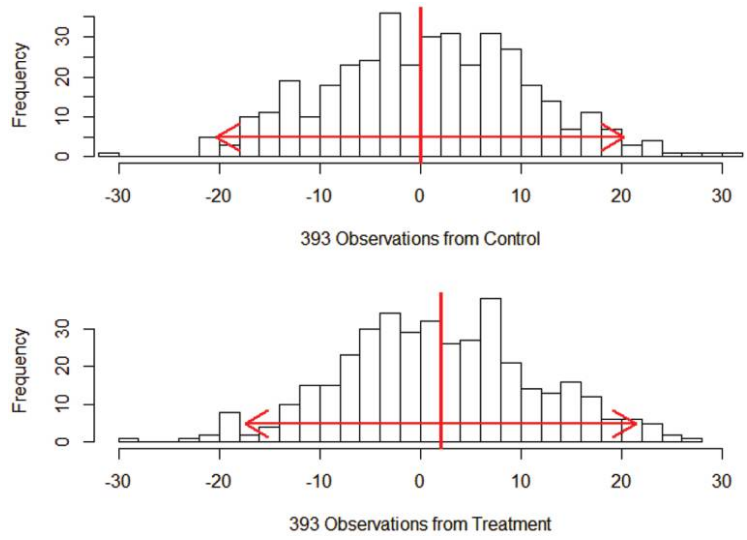

Figure 3. Data distribution for a simulated study with 393 observations. The vertical red lines signify the means, and the horizontal arrowed red lines represent 95% confidence intervals.

Using the frequentist and Bayesian approaches, we come to different conclusions. With the 393 observations per arm, the estimated means for both the control arm and the treatment arm are 0.781 and 1.689, respectively, which results in an estimated treatment effect of 0.908 and an associated *SE* of 0.708. In a classical frequentist *t*-test, these values produce a *t* statistic of 1.282 and a one-sided *p* value of 0.100. The mean difference would not be statistically significant.

Using a Bayesian perspective and incorporating the prior information of a treatment effect of $\theta_0 = 2$ and standard deviation of $\tau = 10$, the estimated posterior mean treatment difference is 0.911, and the standard error is 0.500, which yields an empirical Bayes *t* ratio of 1.820 and a one-sided *p* value of 0.035. The use of prior information yielded a significant treatment effect.

Discussion

In this article, we described a Bayesian perspective on intervention research, comparing the Bayesian perspective to the frequentist perspective that dominates analysis in social work and the social sciences. Intervention research is characterized by the sequential development of new programs. That is, at the beginning of the design phase, intervention research is deeply rooted in the replication of pilot studies that rely, in turn, on meta-analyses and carefully constructed conceptual models. As interventions are refined over time, researchers hope that subsequent pilot studies will support and extend prior results. However, most studies use prior information only in the conceptualization of a program theory and in estimates of statistical power. Trials are regarded as independent, and prior information is not used to condition means and *SE*s. In Bayesian analysis, however, prior information is used to condition means and *SE*s. We demonstrated that the Bayesian perspective provides more power in the design and development processes that comprise intervention research.

We illustrated the Bayesian perspective using a simple case of a two-arm comparison. However, Bayesian applications can be extended to include multiple treatment comparisons with the analysis of variance, treatment comparisons with covariates in regression, adjustments for clustering in multilevel modeling, and complex parameter estimation in structural equation modeling (e.g., Holtmann, Koch, Lochner, & Eid, 2016). For the more complicated designs, advanced Bayesian modeling with Markov-Chain Monte-Carlo estimation can be used. Computation software is available in SAS and R (e.g., Brooks, Gelman, Jones, & Meng, 2011; Chen et al., 2017; Gelman et al., 2013).

When to Consider a Bayesian Approach

The Bayesian perspective has been developed extensively in statistics, but it has rarely been applied to intervention research where new programs are developed through sequenced studies. A key conceptual question in the use of Bayesian meth-

ods in the design of new interventions requires consideration: To what extent can prior information be used? Typically, interventions change over the course of development. Interventions have deep structures that focus on risk processes believed to mediate risk exposure and subsequent negative developmental or life-course outcomes. In the same vein, interventions also have surface characteristics, which tend to be program activities and processes that can be tailored to improve treatment uptake and adherence in alternative settings.

One consideration in the use of a Bayesian perspective is the relevance of prior information to program structures. Unlike in drug trials in which doses can be carefully controlled, the dosage of social interventions often has great variability. Through the development process, researchers test programs of different lengths; they create programs with different features and content; they use intervention agents with different backgrounds, such as the provision of intervention content by social workers versus teachers; and they alter the way content is presented. Given the iterative development of interventions—usually through multiple pilot tests—it is not clear when the use of prior information can be justified. At a minimum, use of prior information would appear justifiable when the deep structure remains unchanged across trials. To be sure, no harm is done when using noninformative priors because the Bayesian model resolves to the frequentist model when prior information has no value. However, research on the goodness of prior information is still emerging and caution is warranted, especially when program structure has low resemblance from one trial to the next (e.g., see Holtmann et al., 2016).

Limitations

Simulations are inherently limited by the way they are parameterized. In this study, we based findings on informative prior information, including an explicit treatment effect and an adequate sample size. Simulations offer the advantage of known parameters, and therefore, the precision of competing models, such as the frequentist and Bayesian approaches, can be compared. However, simulations can be developed in infinite ways. Future research could, for example, estimate power under varying conditions, such as small samples and weak or misleading prior information.

A key component of conducting Bayesian analysis is the specification of a prior distribution for intervention parameters. This prior information can be either informative or noninformative. The prior information is typically drawn from previous research; however, it is quite possible to use prior belief, which can be subjective. In the case of a subjectively determined prior, the posterior distribution could be misleading, which would lead to incorrect statistical conclusions. A key feature of intervention research is the iterative development of knowledge based on prior research. We caution researchers who are interested Bayesian methods to carefully choose prior information based on systematically derived information. Interven-

tion research is a sequential process based on knowledge and understanding that accumulates over many types of studies. If based on prior research, the prior is likely to be informative.

To be sure, there may be times when no or little useful information exists to set a prior distribution. In circumstances of uncertainty, researchers can use noninformative priors that are loosely based on prior research. Priors are selected deliberately to minimize constraints on the posterior distribution parameters. In this situation, the prior can be set with a large prior variance. This approach suggests that not all prior information is totally noninformative. For example, if insufficient evidence exists about the distribution of an intervention effect, researchers can set the prior as a normal distribution with a mean of 0 and a large variance. It is not unusual to set the variance as high as 1,000, or higher. By using such a large variance for the prior distribution, the researcher acknowledges the lack of credible information regarding the posterior distribution. Hence, the posterior distribution is left largely unaffected by the prior distribution and relies primarily on observed data to obtain the estimate for an intervention effect. When prior information is subjectively determined or is of questionable value, a noninformative prior with its variance set to be infinite forces reliance on the observed data.

In the situation when uncertainty exists about prior distributions, the researcher should also conduct sensitivity analyses to determine the influence of different specifications of the priors. If the influence is not large, we would conclude that the Bayesian model is robust to different specifications of priors. Otherwise, a reasonable prior should be chosen for robust Bayesian analysis. In analysis, priors can be determined within a range of accumulated knowledge.

Conclusion

Although Bayesian methods are scarcely used in social work research, they offer potential advantages over the commonly used frequentist perspective. For intervention researchers, Bayesian approaches provide a logically cohesive framework to permit updating analyses with new information. This feature amounts to a mechanism to adjust analyses based on the level of uncertainty, which tends to decline in intervention research as more information is cumulated through pilot studies. To be sure, the selection of priors and the use of small samples are challenging issues in Bayesian analysis. However, within the frequentist perspective, the need to improve point estimates and growing dissatisfaction with p values are perhaps more challenging. The Bayesian approach provides an alternative inferential method that, like intervention research itself, incorporates prior information in the knowledge development process (Berry, Carlin, Lee, & Muller, 2011; Oldehinkel, 2016). Rather than rejecting null hypotheses, Bayesian analyses permit testing alternative theories about the data, and, holding other factors constant, Bayesian analyses appear to provide intervention researchers with more statistical power.

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

This paper was presented at the 2017 Annual Meeting of the Society for Social Work and Research in New Orleans, LA. We thank Dr. Shuyen Ho and Ms. Diane C. Wyant for providing helpful comments on the draft of this paper.

References

Bayes, T., & Price, R. (1763). An essay towards solving a problem in the doctrine of chance. By the late Rev. Mr. Bayes, communicated by Mr. Price, in a letter to John Canton, A. M. F. R. S. *Philosophical Transactions of the Royal Society of London*, *53*, 370–418. doi:10.1098/rstl.1763.0053

Bender, K., Begun, S., DePrince, A., Haffejee, B., Brown, S., Hathaway, J., & Schau, N. (2015). Mindfulness intervention with homeless youth. *Journal of the Society for Social Work and Research*, *6*, 491–513. doi:10.1086/684107

Berger, J. O. (1985). *Statistical decision theory and Bayesian analysis* (2nd ed.). New York, NY: Springer-Verlag Press.

Berry, S. M., Carlin, B. P., Lee, J. J., & Muller, P. (2011). *Bayesian adaptive methods for clinical trials*. Boca Raton, FL: CRC Press.

Box, J. F. (1987). Guinness, Gosset, Fisher, and small samples. *Statistical Science*, *2*(1), 45–52. http://www.jstor.org/stable/2245613

Brooks, S., Gelman, A., Jones, L., & Meng, X. (2011). *Handbook of Markov chain Monte Carlo*. Boca Raton, FL: CRC Press.

Carlin, B. P., & Louis, T. A. (2008). *Bayesian methods for data analysis* (3rd ed.). Boca Raton, FL: CRC Press.

Charles, P., Gorman-Smith, D., & Jones, A. (2016). Designing an intervention to promote child development among fathers with antisocial behavior. *Research on Social Work Practice*, *26*, 20–27. doi.org/10.1177/1049731515580440

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

Fraser, M. W., & Galinsky, M. J. (2010). Steps in intervention research: Designing and developing social programs. *Research on Social Work Practice*, *20*, 459–466. https://doi.org/10.1177/1049731509358424

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

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

Gelman, A., Carlin, J. B., Stern, H. S., Dunson, D. B., Vehtari, A., & Rubin, D. B. (2013). *Bayesian data analysis* (3rd ed.). Boca Raton, FL: CRC Press.

Holtmann, J., Koch, T., Lochner, K., & Eid, M. (2016). A comparison of ML, WLSMV, and Bayesian methods for multilevel structural equation models in small samples: A simulation study. *Multivariate Behavioral Research*, *51*, 661–680. https://doi.org/10.1080/00273171.2016.1208074

Kim, B. K. E., Oesterle, S., Hawkins, J. D., & Shapiro, V. B. (2015). Assessing sustained effects of Communities That Care on youth protective factors. *Journal of the Society for Social Work and Research*, *6*, 565–589. doi:10.1086/684163

Levy, R. (2016). Advances in Bayesian modeling in educational research. *Educational Psychologist*, *51*, 368–380. https://doi.org/10.1080/00461520.2016.1207540

Oldehinkel, A. J. (2016). Bayesian benefits for child psychology and psychiatry researchers. *Journal of Child Psychology and Psychiatry*, *57*, 985–987. https://doi.org/10.1111/jcpp.12619

Resnicow, K., Soler, R., Braithwaite, R. L., Ahluwalia, J. S., & Butler, J. (2000). Cultural sensitivity in substance use prevention. *Journal of Community Psychology*, *28*, 271–290. https://doi.org/10.1002/(SICI)1520-6629(200005)28:3<271::AID-JCOP4>3.0.CO;2-I

Schwinn, T. M., Hopkins, J. E., & Schinke, S. P. (2016). Developing a web-based intervention to prevention drug use among adolescent girls. *Research on Social Work Practice*, *26*, 8–13. doi.org/10.1177/1049731515579204

Wu, F., Fraser, M. W., Guo, S., Day, S. H., & Galinsky, M. J. (2016). Strengthening the social information-processing skills of children in China: A controlled pilot test of the Let's Be Friends program. *Research on Social Work Practice*, *26*(5), 525–538. doi:10.1177/1049731514556995

Manuscript submitted: December 6, 2016
Revision submitted: January 10, 2017
Accepted: January 18, 2017
Electronically published: July 12, 2017