



The influence of baseline marijuana use on treatment of cocaine dependence: application of an informative-priors Bayesian approach

Charles Green^{1*}, Joy Schmitz¹, Jan Lindsay¹, Claudia Pedroza¹, Scott Lane¹, Rob Agnelli², Kimberley Kjome¹ and F. Gerard Moeller¹

¹ University of Texas Health Sciences Center at Houston, Houston, TX, USA

² SAS, Inc., Cary, NC, USA

Edited by:

Thomas Kosten, Baylor College of Medicine, USA

Reviewed by:

Thomas Kosten, Baylor College of Medicine, USA

Thomas Newton, Baylor College of Medicine, USA

*Correspondence:

Charles Green, University of Texas Medical School at Houston, 6431 Fannin, MSB 2.106, Houston, TX 77030, USA.

e-mail: charles.green@uth.tmc.edu

Background: Marijuana use is prevalent among patients with cocaine dependence and often non-exclusionary in clinical trials of potential cocaine medications. The dual-focus of this study was to (1) examine the moderating effect of baseline marijuana use on response to treatment with levodopa/carbidopa for cocaine dependence; and (2) apply an informative-priors, Bayesian approach for estimating the probability of a subgroup-by-treatment interaction effect. **Method:** A secondary data analysis of two previously published, double-blind, randomized controlled trials provided complete data for the historical (Study 1: $N = 64$ placebo), and current (Study 2: $N = 113$) data sets. Negative binomial regression evaluated Treatment Effectiveness Scores (TES) as a function of medication condition (levodopa/carbidopa, placebo), baseline marijuana use (days in past 30), and their interaction. **Results:** Bayesian analysis indicated that there was a 96% chance that baseline marijuana use predicts differential response to treatment with levodopa/carbidopa. Simple effects indicated that among participants receiving levodopa/carbidopa the probability that baseline marijuana confers harm in terms of reducing TES was 0.981; whereas the probability that marijuana confers harm within the placebo condition was 0.163. For every additional day of marijuana use reported at baseline, participants in the levodopa/carbidopa condition demonstrated a 5.4% decrease in TES; while participants in the placebo condition demonstrated a 4.9% increase in TES. **Conclusion:** The potential moderating effect of marijuana on cocaine treatment response should be considered in future trial designs. Applying Bayesian subgroup analysis proved informative in characterizing this patient-treatment interaction effect.

Keywords: cocaine, marijuana, treatment response, Bayesian, subgroup analysis

INTRODUCTION

Multiple substance use is common in cocaine patients, making it a challenge to obtain samples of “pure” or singly dependent subjects for clinical trials research. The response has been to broaden the eligible subject pool by allowing concurrent use or abuse of other substances, although recent articles have debated the degree of acceptable heterogeneity for pharmacotherapy efficacy research (O’Brien and Lynch, 2003; Rounsaville et al., 2003). Moreover, when patients who use multiple drugs are included in a cocaine clinical trial, the investigation needs to take into account this increased variability. Not all clinical trials have consistently reported the impact of concomitant substance use on response to cocaine treatment.

Marijuana use is especially prevalent among patients with cocaine dependence (Miller et al., 1990a,b; Aharonovich et al., 2005, 2006). In an earlier report our research group described some interesting characteristics associated with marijuana-using individuals presenting for medication treatment trials (Lindsay et al., 2009). Among the large percentage (46.4%) of concurrent

marijuana users seeking treatment for cocaine dependence, those who used marijuana frequently (i.e., 10 or more days over the past 30) showed a profile of greater clinical impairment, exemplified by more severe addiction severity scores, and heavier patterns of drug use, than those who used marijuana occasionally or not at all.

The extent to which marijuana-using cocaine-dependent patients fare worse in treatment compared to their non-marijuana-using counterparts is unclear. Aharonovich et al. (2006) examined the impact of continued cannabis use during methylphenidate treatment of comorbid cocaine dependence and ADHD and found that moderate or intermittent use of cannabis was associated with better retention in treatment than heavy use or abstinence. The effect of cannabis use on cocaine abstinence rates was not significant, contrary to previous findings (Aharonovich et al., 2005). Other studies have failed to find an effect of marijuana use on retention or cocaine treatment outcomes (Budney et al., 1996; Higgins et al., 2003). In a recent retrospective analysis of data from three clinical trials of contingency management, Alessi et al.

(2011) found that pretreatment marijuana use did not dampen the benefit of this behavioral treatment in reducing cocaine use. In sum, the small literature shows mixed results on the impact of cannabis use on cocaine treatment outcome, underscoring the need to better understand this potential interaction, especially as it relates to the development of effective medications for treatment of cocaine dependence.

The prognostic significance of baseline marijuana use on treatment outcome is the focus of the present study. Having reported positive results from earlier trials of levodopa pharmacotherapy for cocaine dependence (Mooney et al., 2007; Schmitz et al., 2008), we carried out a secondary analysis to determine whether levodopa treatment effects varied as a function of marijuana use at baseline. Observations by Lindsay et al. (2009) led us to expect that marijuana use would moderate the effect of levodopa by reducing overall treatment effectiveness. Specifically, we hypothesized that within the subgroup of cocaine subjects having high levels of baseline marijuana use, treatment with levodopa pharmacotherapy would be less effective.

The second aim of this study is to demonstrate a more appropriate, statistical method for conducting secondary analysis of subgroup treatment effects. Analytically, evaluation of subgroup effects entails testing the interaction of treatment and some baseline, subgroup measurement. A large literature has criticized the use of traditional, Frequentist approaches as inadequate due to their often low power, and dichotomous evaluation of the evidence (i.e., significance testing; Dixon and Simon, 1991; Simon et al., 1996; Simon and Freedman, 1997; Simon, 2002; Green et al., 2009). For instance, Brookes et al. (2001) showed that a trial with 80% power to detect a main effect of treatment possessed only 29% power to detect an interaction effect as large as the main effect. Positing an interaction effect matching the magnitude of a main effect reflects an extremely optimistic scenario. Stated another way, if the interaction was as large as the anticipated main effect, the sample size would still require quadrupling in order to maintain 80% power to detect the interaction. The same literature, criticizing conventional Frequentist methods for analyzing treatment-by-subgroup interaction effects, has recommended Bayesian methods as an alternative.

Bayesian analytical methods avoid issues of reduced power and significance testing by evaluating the probability that the alternative (H_A) hypothesis exists, given the observed data, and any prior evidence for H_A . These probabilities can be meaningfully refined, unlike p values that cannot, by using informative prior probability distributions (for more extensive review of the limitations of p values, and the advantages of Bayesian statistical approaches see Dixon and Simon, 1991; Schervish, 1996; Goodman, 1999, 2005; Berry, 2006). Thus, the researcher is able to attach a probability value to the parameter of interest, in this case, the interaction effect, e.g., “given the observed data, the probability that an interaction of at least magnitude X exists is Y .” The Bayesian approach provides a natural analysis for the present situation, in which we sought to incorporate information from a previous study to improve precision in estimating the interaction parameter. Referred to as the Bayesian *power prior* approach, this paper demonstrates how to pool historical data from an earlier randomized clinical trial with

current data, and produce more precise conclusions regarding the hypothesis of interest, i.e., that heterogeneity in response to treatment is a function of baseline marijuana use.

MATERIALS AND METHODS

SAMPLE

The present study analyzed data obtained from two completed levodopa trials for cocaine dependence. In Study 1, cocaine-dependent subjects enrolled in a 9-week, randomized, double-blind, placebo-controlled trial comparing placebo to 400/100 mg and 800/200 mg levodopa/carbidopa in a sustained-release preparation (Sinemet CR). In Study 2, cocaine-dependent subjects received levodopa/carbidopa (800/200 mg) or placebo delivered in combination with behavioral therapy interventions of varying intensity. Study 1 established the safety, tolerability, and feasibility of levodopa therapy in cocaine outpatients, while Study 2 demonstrated support for use of levodopa pharmacotherapy with behavioral contingency management. Details of each trial design, sample composition, and main findings are presented elsewhere (Mooney et al., 2007; Schmitz et al., 2008). The samples used here are subsets of the overall trial data, selected due to their comparability in dosing for levodopa (800/200 mg), and behavioral intervention (cognitive-behavioral therapy without contingency management). To demonstrate the proposed Bayesian analysis, Study 1 ($N = 64$ complete observations) formed the historical dataset and Study 2 ($N = 113$ complete observations) provided the current dataset.

All participants provided written informed consent. The Committee for the Protection of Human Subjects (CPHS) of the University of Texas Medical School, Houston reviewed and approved the research, consent, and all study materials.

MEASURES

Moderator variable

Baseline level of marijuana use was defined according to self-reported number of days of use over the past 30 using the Addiction Severity Index (ASI; McLellan et al., 1980). In addition to self-reported marijuana use, urine specimens were screened for Δ -9-THC, the primary active metabolite of marijuana. The concordance rate between self-report and urinalysis testing was acceptable (76%).

Treatment response variable

The Treatment Effectiveness Score (TES; Ling et al., 1997) was used as an outcome indicator of treatment response. The TES is calculated by assigning one point for each cocaine-negative urine sample (cutoff < 300 ng/ml), and no points for positive or missing samples. Both studies required three urine samples per week. Studies 1 and 2 were 9 and 12 weeks respectively permitting TES values that ranged from 0–27 to 0–36 negative urines.

ANALYSIS

Bayesian statistical reasoning

Frequentist and Bayesian statistical reasoning comprise the principal modes of statistical reasoning, taking distinct but symbiotic approaches to uncertainty/probability. Frequentist reasoning defines probability as the frequency of an event in the limit of

a series of infinite, repeated trials. Often illustrated using a large number of repeated coin flips, counting the number of heads, or tails, long run frequency counts provide an estimate of the fixed parameter that governs the performance of the coin. Bayesian reasoning defines probability as a judgment regarding the likelihood of an event (Hájek, 2003). Often referred to as a subjective or personalist probability, Bayesian reasoning is best operationalized in terms of the proportion of a fixed amount of money an observer would bet on a specific outcome (Hájek, 2003). Statistically, the Frequentist approach models the data as random, and the parameter as fixed and unknown. In contrast the Bayesian approach models a parameter as unknown and random, and the observed data as fixed (Lucke, 2004). Frequentist reasoning indirectly evaluates the alternative hypothesis (H_A) by rejecting or failing to reject the null hypothesis (H_0). This permits statements such as “given that the null hypothesis is true, the probability of observing data this extreme or more extreme is Z .” Indirect evaluation of the H_A by testing H_0 means that no probability valuation is attached to directly to possible values of H_A . While Frequentist Confidence Intervals represent an attempt at providing an index of H_A they are properly interpreted as those interval which have a 95% chance of capturing the true parameter estimate and say nothing about the differential plausibility of the values they include. They cannot, because in Frequentist statistics the true parameter estimate is fixed (i.e., it has a probability of one): while the true parameter is fixed, the confidence interval moves with each new sample. Bayesian reasoning, defines the governing parameter of interest (in this case the interaction coefficient) as random, and therefore permits probability statements regarding what the value of that parameter might be. The product of a Bayesian analysis is the posterior distribution, which indicates the differential probability that the parameter of interest takes various values. Since the relative probability of various values being the true parameter within this interval is defined by the shape of the posterior distribution, it is possible to comment on the probability that H_A takes on some value or range of values. This means that, given the observed data and a formally articulated prior distribution representing the anticipated value and uncertainty for the parameter of interest, Bayesian modeling permits direct quantification of evidence for the alternative hypothesis (H_A). The Bayesian approach permits statements such as “given the observed data, the probability that an interaction of at least magnitude X exists is Y .” Under certain conditions (e.g., vague priors) the values of the estimates from Bayesian and Frequentist methods are often quite similar, however their interpretation is quite different: Bayesian inference directly addresses the alternative hypothesis, while Frequentist inference does so indirectly by rejecting or failing to reject the null hypothesis. In making this distinction the current paper is not arguing that one approach is inherently correct and the other not. Rather, we agree with Kendall (1949) that each of these different notions of probability ultimately cannot stand without the other. More recently, Wijeysondera et al. (2009) pointing out that the approaches answer different albeit complementary, questions.

Bayesian posterior probabilities

In this paper we present a Bayesian analyses as an alternative approach to assessing the association between baseline marijuana

use and treatment effectiveness. In brief, Bayesian reasoning holds that inferences about a hypothesis should be encapsulated in a probability distribution, given the observed data. This distribution, known as the posterior probability distribution, summarizes evidence for the parameter value as the product of previous evidence (prior distribution), and any newly gathered data. Bayes' Theorem (Eq. 1) expresses this relation:

$$p(\theta|\text{data}) = \frac{p(\text{data}|\theta) p(\theta)}{p(\text{data})} \quad (1)$$

where $p(\text{data}|\theta)$ is the likelihood and $p(\theta)$ is the prior distribution. The denominator $p(\text{data})$ functions as a scaling coefficient and is often omitted from the equation to give:

$$p(\theta|\text{data}) \propto p(\text{data}|\theta) p(\theta) \quad (2)$$

Specifically, the probability that the parameter takes on some value (or range of values) is proportional to the product of the observed data (i.e., likelihood) and prior evidence. Stated in prose, this reads, “The probability (i.e., “ p ”) of the parameter value θ given (i.e., “ I ”) the data is proportional to (i.e., “ \propto ”) the probability of the data given the parameter value θ [i.e., “ $p(\text{data}|\theta)$ ”] multiplied by the prior probability of θ [i.e., “ $p(\theta)$ ”].

Bayesian prior distributions

The prior distribution is a mathematical formalization of existing evidence for a parameter value before observing new data (Gill, 2002). In the absence of historical evidence, the prior distribution is based on subjective judgment, representing varying degrees of skepticism regarding the pre-existing evidence (i.e., enthusiastic, neutral, or skeptical). When access to data from previous studies is available, it is scientifically reasonable and statistically advantageous to incorporate this historical information into the prior distribution. However this requires investigators to evaluate the comparability or *exchangeability* of the historical and current data. Ibrahim and Chen advocate the use of *power priors* (Chen et al., 2000; Ibrahim and Chen, 2000; Ibrahim et al., 2003; Chen and Ibrahim, 2006) to produce posterior distributions that incorporate historical data sets. Implementation of such methods permits the evaluation of the sensitivity of posterior estimates to assumptions regarding the comparability of the constituent samples. Assuming that such comparability exists, combining the information existing in more than one sample may have the benefit of improving the precision of the resulting estimates.

Bayesian power priors

Study 1 provided the existing, historical data set (D_0 of size n_0) and Study 2 provided the current data set (D_1 of size n_1). We assumed an initial prior distribution (i.e., before observing D_0) represented by $p(\theta)$. Incorporation of D_0 into the estimation of the posterior distribution based on D_1 takes the following form:

$$p(\theta|D_1) \propto p(D_1|\theta) p(D_0|\theta)^{a_0} p(\theta) \quad (3)$$

Where $p(D_0|\theta)^{a_0} p(\theta)$ represents the power prior which incorporates the historical data from D_0 , raised to the power a_0 which

is restricted to $0 \leq a_0 \leq 1$, and the initial prior [i.e., $p(\theta)$]. Conceptually, values of a_0 range from $a_0 = 0$, in which D_0 is fully discounted or excluded from the model, to $a_0 = 1$ in which D_0 carries the same weight as D_1 . In essence a_0 provides a means of regulating the amount of information that D_0 contributes to the analysis of D_1 . Two reasons for adjusting the weight of the historical data are: (1) discounting very large historical data sets so that they do not overwhelm the information in a smaller, more current data set, and (2) accounting for inter-sample heterogeneity which may result from differences in the experimental protocol, sampling, etc. (Chen and Ibrahim, 2006; De Santis, 2006). While empirical estimation of a_0 from the data might reflect a weighting that minimizes the loss of information across the two samples (Ibrahim et al., 2003; Chen and Ibrahim, 2006) such an approach is problematic (Neuenschwander et al., 2009), leading to extremely low estimates of a_0 even for identical data sets (Neelon and O'Malley, 2010). Moreover, an attempt to remedy this (Duan et al., 2005) leads to improved but low estimates of a_0 (Neelon and O'Malley, 2010). We adopt a "conditional power priors" (Neelon and O'Malley, 2010, p. 2) approach, following Spiegelhalter et al. (2004) and evaluating the full range of values for a_0 to understand the degree to which discounting the historical data influences inferences based on the resulting posterior distribution.

Model specification

Negative binomial regression evaluated TES as a function of treatment condition (levodopa/carbidopa, placebo), baseline marijuana use (number of days in the past 30), and their interaction. Unlike Poisson regression, negative binomial regression accounts for over-dispersion in the count data characterizing both the historical and current data sets. Following Eq. 3, initial prior distributions [i.e., $p(\theta)$] took the form $\sim \text{Normal}(0, \text{var} = 1 \times 10^6)$ for regression coefficients (i.e., normally distributed with a mean of zero and a variance of 1,000,000; the prior is centered on the null hypothesis and expresses substantial uncertainty). While SAS 9.3 specifies a $\sim \text{Gamma}(0.001, 0.001)$ prior as a default for the dispersion term (Gamma functions describe a family of distribution on the positive real number line. The parameters values 0.001 and 0.001 refer to as the α and β parameters respectively. These two quantities capture the shape (i.e., α) and rate (i.e., β) of the specified distribution. The inverse of the rate (i.e., $1/\beta$) is called the scale parameter. The two parameters convey the shape and the spread or dispersion of the distribution. In, perhaps more familiar terms, the mean of a Gamma distributed variable is α/β while its variance is α/β^2). There has been some concern that such priors may not be appropriate (see Gelman, 2006 for a discussion in the context of normal distributed data). As such we follow DiPrete et al. (2011) as well as Zheng et al. (2006). Reasoning that the dispersion parameter might take values ranging $(0-\infty)$, these authors specify a $\sim \text{Uniform}(0,1)$ prior on the inverse of the dispersion coefficient [a $\sim \text{Uniform}(0,1)$ distribution spans the range zero through one with all values having an equal probability of occurring (i.e., it is a straight line)]. SAS v. 9.3 code for doing so is relatively straightforward and available from the corresponding author. These vague, neutral prior distributions acknowledge a relative state of ignorance regarding

parameter values prior to observing either the historical or current data sets. Again, following Eq. 3, the likelihood incorporating the historical data (i.e., $p(D_0|\theta)^{a_0}$) required specification of the term a_0 . Following Spiegelhalter et al. (2004) we evaluated use of the historical data at values ranging from $a_0 = 0$ to $a_0 = 1$. Estimates of parameter values for the interaction term are the appropriate indices for evaluating heterogeneity in response to treatment as a function of baseline marijuana use. As in Poisson regression, exponentiated parameter estimates and intervals correspond to risk ratios (R.R.) and Credible Intervals (C.B.I.) respectively.

Computational software

Analyses utilized Proc Markov Chain Monte Carlo (MCMC; SAS v. 9.3; SAS Institute Inc, 2010) which provides a flexible computing environment for the MCMC simulations for estimating Bayesian posterior distributions. An example applying Proc MCMC to a relatively straightforward use of power priors in estimating a binomial proportion using historical and current data is provided in SAS Documentation (SAS Institute Inc, 2011). Examples of statistical code and salient output for the current analysis are available from the corresponding author.

RESULTS

Baseline marijuana use characteristics of pharmacotherapy groups from their respective studies are provided in **Table 1**.

Table 2 displays Bayesian estimates for $a_0 = 0$ and $a_0 = 1$ as well as frequentist estimates of the interaction term. Estimates of the interaction coefficient for $a_0 = 0$ and $a_0 = 1$ and $a_0 = 1$ are R.R. = 0.887 (95% C.B.I. 0.781–1.022) and R.R. = 0.899 (95% C.B.I. 0.808–1.016) respectively. That is, relative to patients in the placebo condition, participants receiving levodopa demonstrated a decrease in TES by a factor of 0.887 (i.e., 11.3%) or 0.899 (i.e., 10.1%) for the two most extreme weights that might be attached to the historical evidence (i.e., $a_0 = 0$ and $a_0 = 1$ respectively). Note that the point estimates are quite similar and that the 95% Credible Interval for $a_0 = 0$ entirely contains the corresponding interval for $a_0 = 1$. The broader credible interval for $a_0 = 0$ which encompass the interval for $a_0 = 1$ reflects the decrease in uncertainty that occurs as the historical data receives greater weight in the analysis; a reasonable result given that inclusion of more information from the historical data permits greater precision in the resulting estimates. Inspection of **Figures 1** and **2** show that, at different levels of a_0 , the value of the parameter estimate for the interaction, as well as the probability that increased baseline

Table 1 | Baseline marijuana use levels.

Variable	Placebo Mean, SD (N)	I-DOPA Mean, SD (N)
HISTORICAL DATA		
Baseline marijuana use (days in past 30)	3.41, 7.68 (29)	1.97, 5.67 (35)
Baseline marijuana use (years of use)	9.17, 10.06 (29)	11.03, 9.96 (34)
CURRENT DATA		
Baseline marijuana use (days in past 30)	3.28, 7.23 (58)	3.24, 6.94 (55)
Baseline marijuana use (years of use)	13.43, 10.71 (58)	10.55, 9.21 (55)

Table 2 | Exponentiated parameter estimates and 95% credible intervals for $a_0 = 0$ and $a_0 = 1$.

	Risk ratio	95% Credible limits	
		95% Lower limit	95% Upper limit
NEGATIVE BINOMIAL REGRESSION FOR TES			
$a_0 = 0$			
Intercept	3.158	1.895	5.694
Medication	3.346	1.483	7.650
Marijuana	1.062	1.000	1.160
Medication*marijuana	0.887	0.781	1.022
Dispersion	28.566	11.451	107.651
$a_0 = 1$			
Intercept	3.376	2.184	5.552
Medication	2.376	1.217	4.646
Marijuana	1.050	0.999	1.124
Medication*marijuana	0.899	0.808	1.016
Dispersion	38.695	16.805	118.570
FREQUENTIST MAXIMUM LIKELIHOOD ESTIMATES FOR THE INTERACTION OF MEDICATION AND MARIJUANA			
Sample	Risk ratio estimate	95% Confidence interval	Likelihood ratio statistic
Historical sample	0.856	0.687–1.066	$\chi(1) = 1.77, p < 0.184$
Current sample	0.887	0.787–1.001	$\chi(1) = 3.12, p < 0.078$
Combined samples	0.899	0.810–0.998	$\chi(1) = 3.30, p < 0.069$

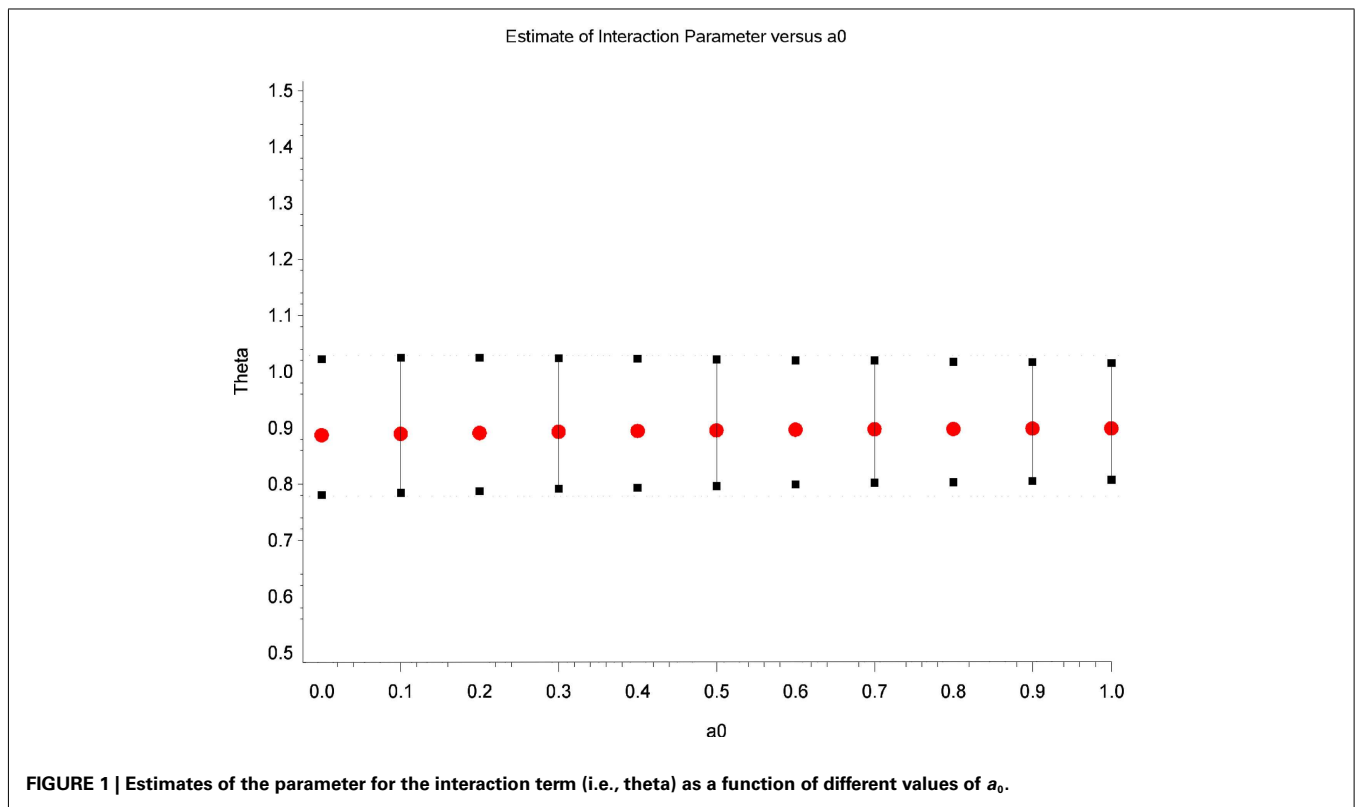


FIGURE 1 | Estimates of the parameter for the interaction term (i.e., theta) as a function of different values of a_0 .

marijuana use predicts decreased TES scores in the levodopa condition, remains relatively constant. Inspection of **Figure 3** shows the magnitude of the decrease in the Credible Interval range and

the variance of the posterior distribution as a percentage of its width at $a_0 = 0$, as a function of altering a_0 . In the current case, the range of the Credible Interval decreases by approximately

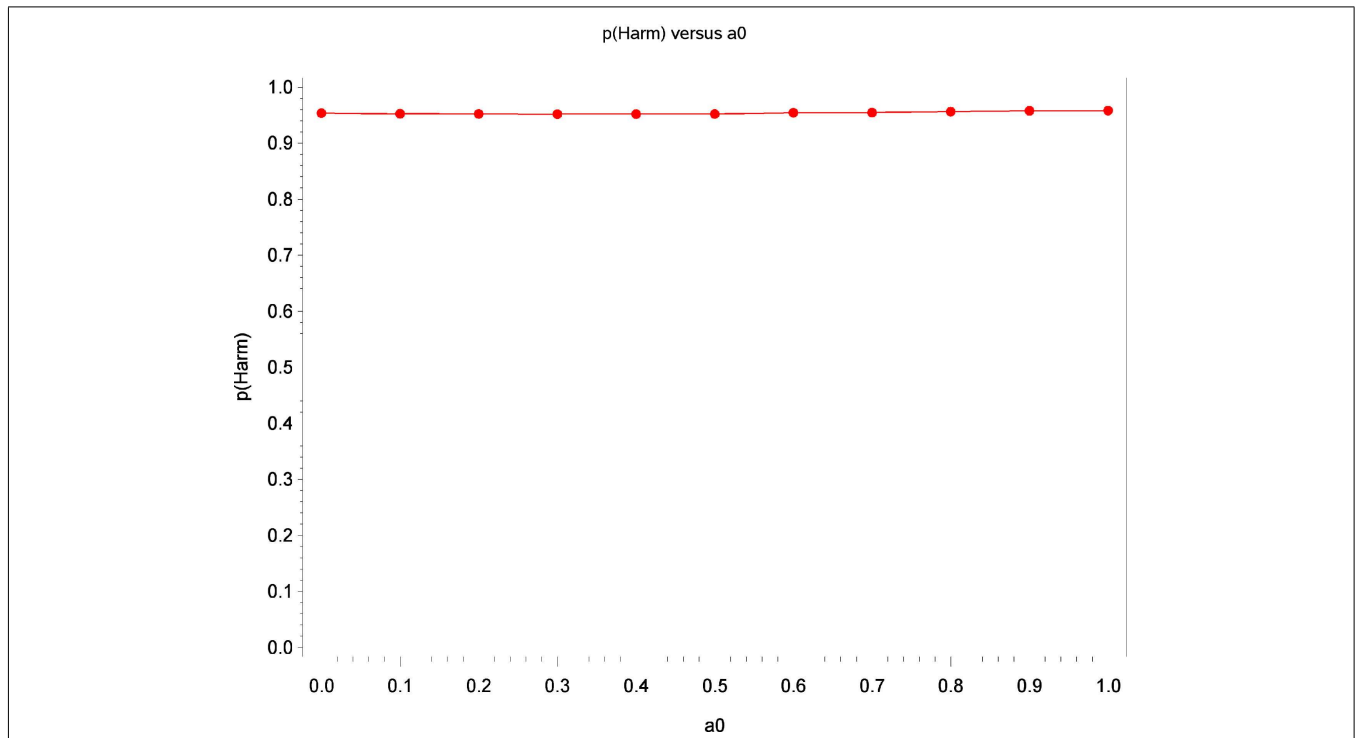


FIGURE 2 | Estimates of the probability that increased baseline marijuana use predicts decreased TES scores in the levodopa treatment as a function of a_0 .

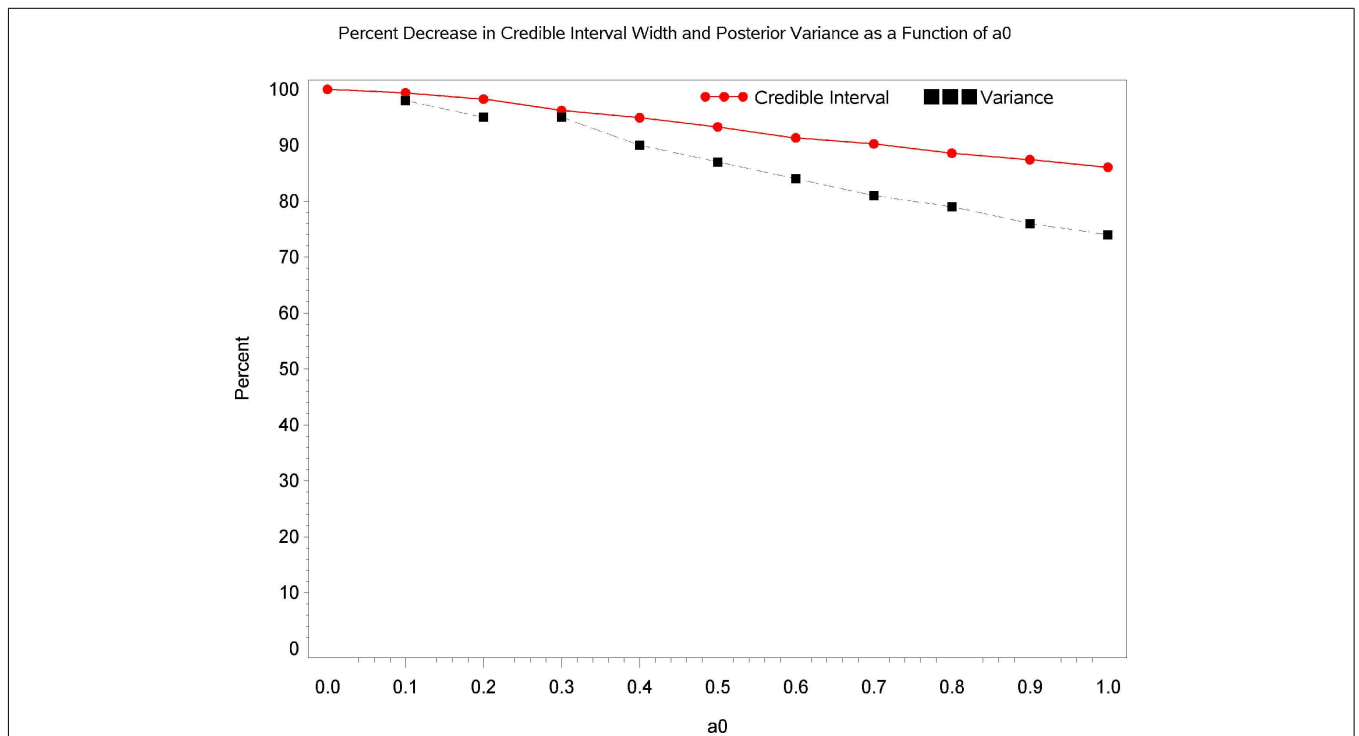
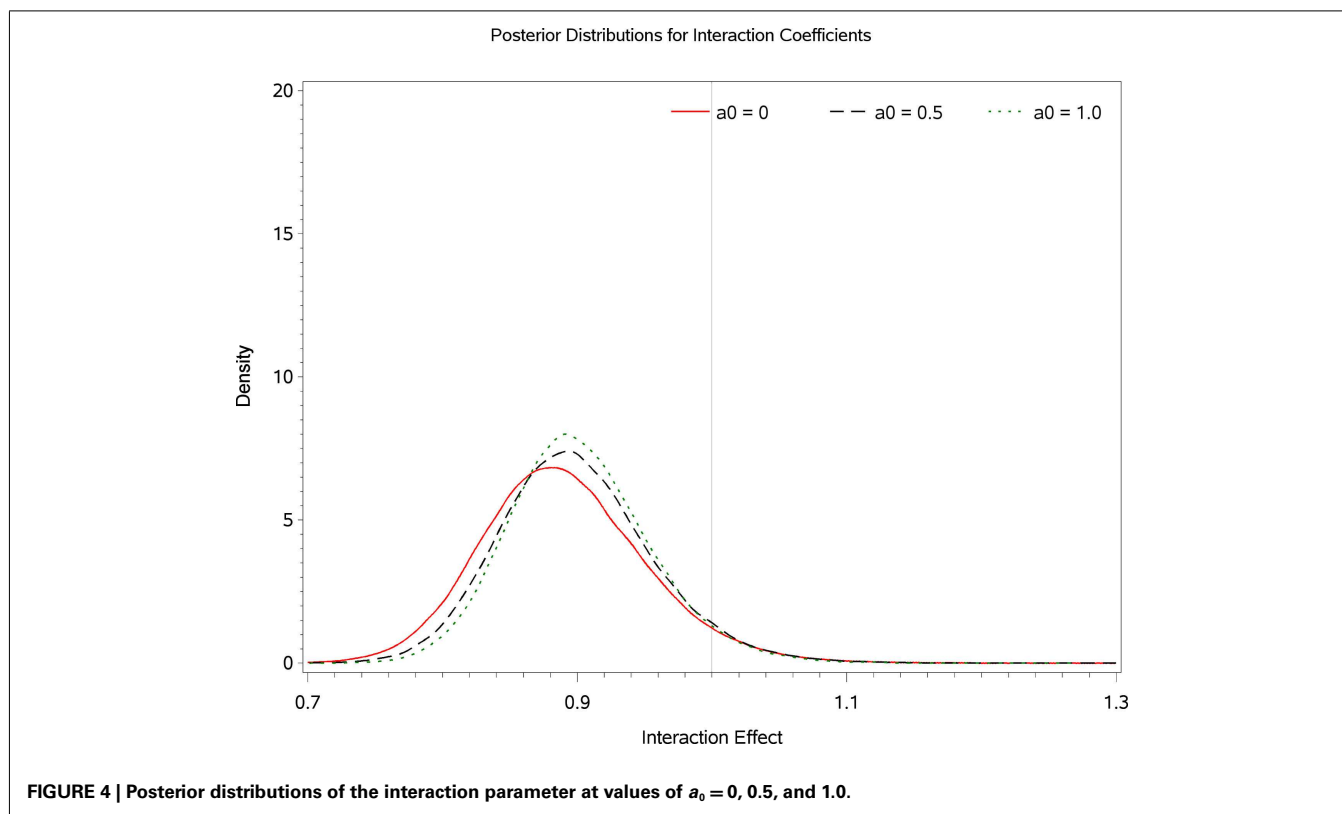


FIGURE 3 | Change in the precision of the parameter estimate as a function of a_0 . The red line shows the magnitude of the decrease in the Credible Interval range as a percentage of its width at $a_0 = 0$, as a

function of altering a_0 . The blue line shows the magnitude of the decrease in the variance of the posterior distribution as a function of altering a_0 .



14% while the variance of the posterior distribution of the exponentiated coefficient decreases by almost 26%. Finally, **Figure 4** shows the posterior distributions of the parameter value for the interaction coefficient. The distributions, displayed for $a_0 = 0, 0.5,$ and 1.0 , appear to be quite similar. Given the failure to demonstrate that alterations in a_0 result in substantial changes to the estimated value of the interaction coefficient, coupled with the increase in precision for the estimate, it is reasonable to give the historical data a weight comparable to the current data (i.e., $a_0 = 1$).

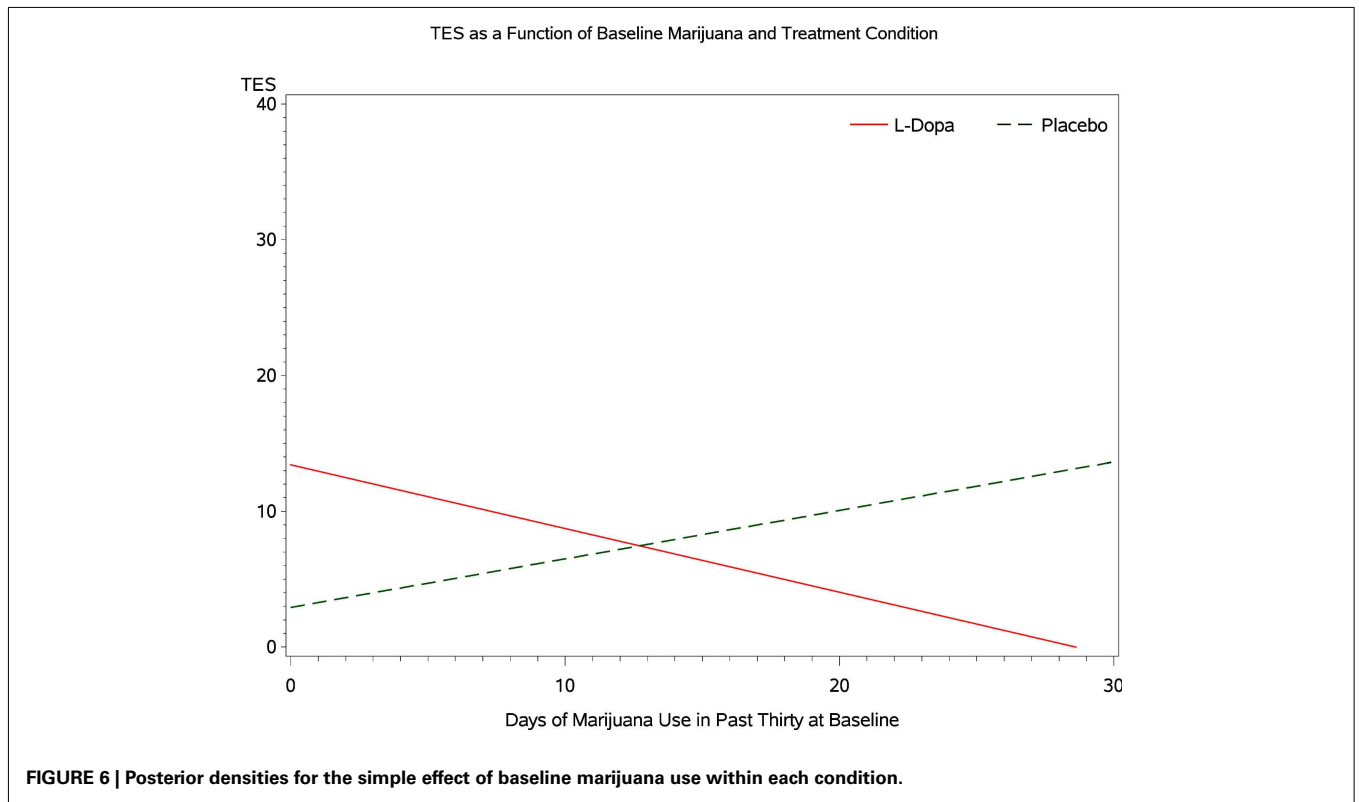
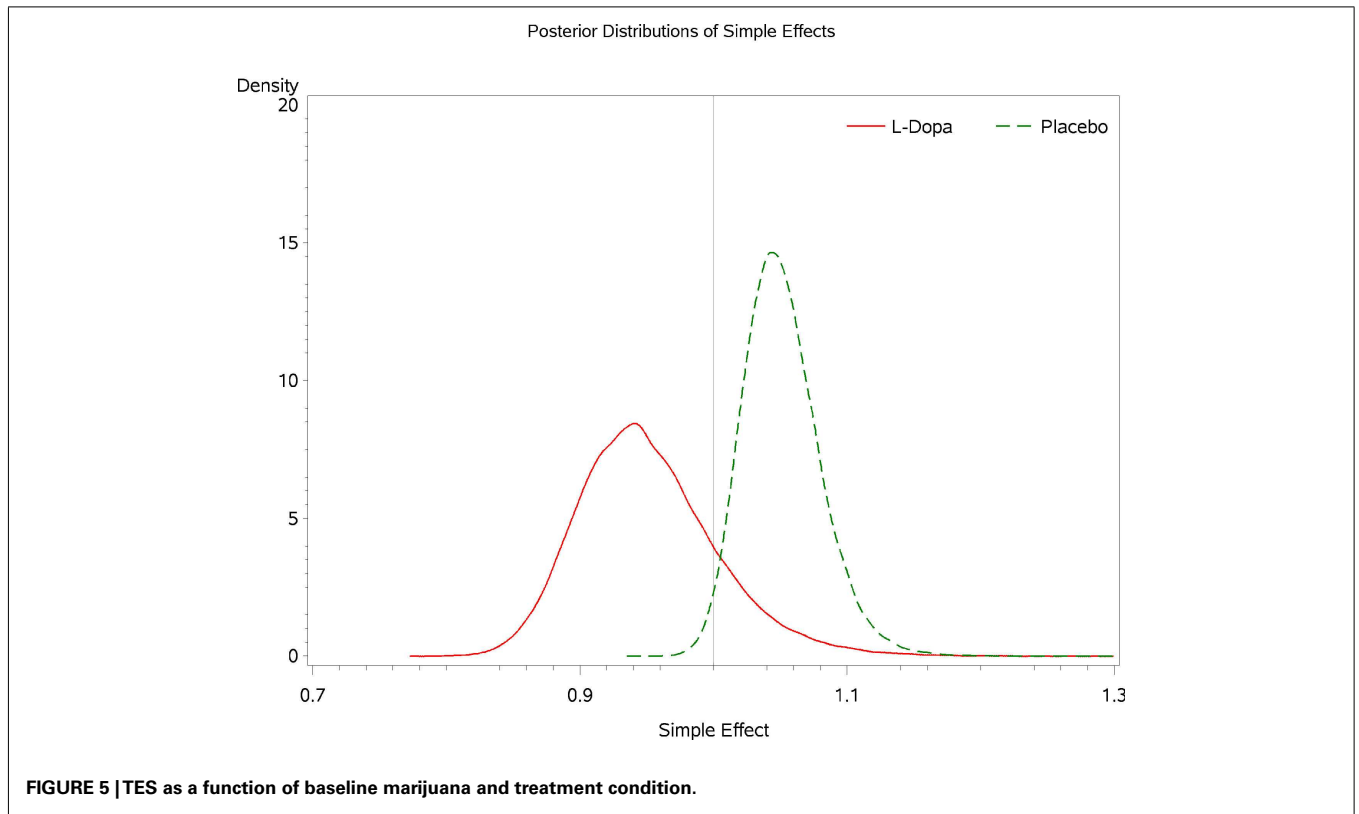
From a substantive perspective, using $a_0 = 1$, the probability that baseline marijuana use predicts differential response to treatment with levodopa/carbidopa is 0.96 (i.e., where θ is less than or equal to a R.R. of 1.0; **Figure 4**). Based on the Frequentist Likelihood Ratio Test statistics presented in **Table 1**, neither the historical, current nor the combined data sets lead to rejection of the null hypothesis.

Inspecting the simple effects of baseline marijuana use indicates that among participants receiving levodopa/carbidopa the probability that baseline marijuana confers harm in terms of treatment outcome is 0.981 (**Figure 5**). The probability that marijuana confers harm within the placebo condition is 0.163 (**Figure 5**). For every additional day of marijuana use reported at baseline, participants in the levodopa/carbidopa condition demonstrate a 5.4% decrease in TES (R.R. 0.946, 95% C.B.I. 0.864–1.069; **Figure 6**). For every additional day of marijuana use reported at baseline, participants in the placebo condition demonstrate a 4.9% increase in TES (R.R. 1.049, 95% C.B.I. 1.002–1.115; **Figure 6**).

DISCUSSION

Subgroup analyses are informative for characterizing heterogeneity in response to treatment for cocaine dependence. Inadequate power minimizes the usefulness of conventional, Frequentist tests of the salient interaction terms. As such Bayesian methods provide a more appropriate, probabilistic measure of the evidence for subgroup effects. Improved precision for estimates of subgroup effects can result from incorporation of historical data into the analysis of a current data set. Doing so raises questions regarding the comparability of the data sets, and the weight subsequently accorded the historical data. Power priors offer a mechanism for formalizing the degree to which historical data is incorporated into the analysis of a newer data set.

In the current context this type of analysis points, with high probability, to the existence of a subgroup effect of baseline marijuana use on response to treatment for cocaine dependence with levodopa/carbidopa. Specifically, higher marijuana use predicted lower treatment effectiveness (cocaine-negative urines) in the levodopa/carbidopa condition, but not in the placebo condition. It is possible that higher marijuana use at baseline operates as a proxy for cocaine severity that has been shown to distinguish subpopulations in terms of treatment response. Repeating our analysis with baseline cocaine use as a covariate, however, failed to alter any substantive conclusions. To the extent that frequent marijuana use at baseline continued during treatment, this pattern of concomitant use may have influenced levodopa's efficacy via providing competing drug reinforcement, perhaps counteracting the putative dopamine-restoring effects



of levodopa. Although generally a well-tolerated medication in cocaine-dependent patients, levodopa may interact with marijuana to produce less tolerable effects and thus reduce compliance and efficacy. Based on these findings, future research should examine how marijuana use *during* treatment interacts with cocaine use in moderating the effects of levodopa treatment: (i) chronic marijuana use may produce changes in cannabinoid-1 and/or dopamine receptor availability which may alter the effects of levodopa in cocaine-dependent subjects; (ii) intoxication following acute marijuana use may alter the probability of concurrent cocaine use in the levodopa condition relative to placebo via drug–drug interactions not yet fully understood.

Limitations of the current study include the *post hoc*/secondary nature of the data analysis. A large literature has discussed the problem of *post hoc* subgroup analyses (Adams, 1998; Assmann et al., 2000; Pocock et al., 2002; Cook et al., 2004; Rothwell, 2005).

REFERENCES

- Adams, K. F. (1998). Post hoc subgroup analysis and the truth of a clinical trial. *Am. Heart J.* 136, 753–758.
- Aharonovich, E., Garawi, F., Bisaga, A., Brooks, D., Raby, W. N., Rubin, E., et al. (2006). Concurrent cannabis use during treatment for comorbid ADHD and cocaine dependence: effects on outcome. *Am. J. Drug Alcohol Abuse* 32, 629–635.
- Aharonovich, E., Liu, X., Samet, S., Nunes, E., Waxman, R., and Hasin, D. (2005). Postdischarge cannabis use and its relationship to cocaine, alcohol, and heroin use: a prospective study. *Am. J. Psychiatry* 162, 1507–1514.
- Alessi, S., Rash, C., and Petry, N. (2011). Contingency management is efficacious, and improves outcomes in cocaine patients with pretreatment marijuana use. *Drug Alcohol Depend.* 118, 62–67.
- Assmann, S. F., Pocock, S. J., Enos, L. E., and Kasten, L. E. (2000). Subgroup analysis and other (mis)uses of baseline data in clinical trials. *Lancet* 355, 1064–1069.
- Berry, D. (2006). Bayesian clinical trials. *Nat. Rev. Drug Discov.* 5, 27–36.
- Brookes, S. T., Whitely, E., Egger, M., Smith, G. D., Mulheran, P. A., and Peters, T. J. (2001). Subgroup analyses in randomized controlled trials: quantifying the risks of false-positives and false negatives. *Health Technol. Assess.* 5, 1–4.
- Budney, A. J., Higgins, S. T., and Wong, C. J. (1996). Marijuana use, and treatment outcome in cocaine-dependent patients. *Exp. Clin. Psychopharmacol.* 4, 7.
- Chen, M.-H., and Ibrahim, J. G. (2006). The relationship between the power prior and hierarchical models. *Bayesian Anal.* 1, 551–574.
- Chen, M.-H., Ibrahim, J. G., and Shao, Q.-M. (2000). Power prior distributions for generalized linear models. *J. Stat. Plan. Inference* 84, 121–137.
- Cook, D. I., Gebski, V. J., and Keech, A. C. (2004). Subgroup analysis in clinical trials. *Med. J. Aust.* 180, 289–291.
- De Santis, F. (2006). Power priors and their use in clinical trials. *Am. Stat.* 60, 122–129.
- DiPrete, T. A., Gelman, A., McCormick, T., Teitler, J., and Zheng, T. (2011). Segregation in social networks based on acquaintanceship and trust. *Am. J. Sociol.* 116, 1234–1283.
- Dixon, D. O., and Simon, R. (1991). Bayesian subset analysis. *Biometrics* 47, 871–881.
- Duan, Y., Ye, K., and Smith, E. P. (2005). Evaluating water quality using power priors to incorporate historical information. *Environmetrics* 17, 95–106.
- Gelman, A. (2006). Prior distributions for the variance parameters in hierarchical models. *Bayesian Anal.* 1, 515–533.
- Gill, J. (2002). *Bayesian Methods: A Social and Behavioral Sciences Approach*. Boca Raton: Chapman & Hall/CRC.
- Goodman, S. (1999). Toward evidence-based medical statistics. 1: the P value fallacy. *Ann. Intern. Med.* 130, 995–1004.
- Goodman, S. M. (2005). Introduction to Bayesian methods I: measuring the strength of evidence. *Clin. Trials* 2, 282–290.
- Green, C., Moeller, F., Schmitz, J., Lucke, J., Lane, S., Swann, A., et al. (2009). Evaluation of heterogeneity in pharmacotherapy trials for drug dependence: a Bayesian approach. *Am. J. Drug Alcohol Abuse* 35, 95–102.
- Hájek, A. (2003). “Interpretations of probability,” in *The Stanford Encyclopedia of Philosophy*, Summer Edn, ed. E. N. Zalta. Available at: <http://plato.stanford.edu/archives/sum2003/entries/probability-interpret>
- Higgins, S., Sigmon, S., Wong, C., Heil, S. H., Badger, G. J., Donham, R., et al. (2003). Community reinforcement therapy for cocaine-dependent outpatients. *Arch. Gen. Psychiatry* 60, 1043–1052.
- Ibrahim, J. G., and Chen, M.-H. (2000). Power prior distributions for regression models. *Stat. Sci.* 15, 46–60.
- Ibrahim, J. G., Chen, M.-H., and Sinha, D. (2003). On optimality properties of the power prior. *J. Am. Stat. Assoc.* 98, 204–213.
- Kendall, M. G. (1949). On the reconciliation of theories of probability. *Biometrika* 36, 101–116.
- Lindsay, J. A., Stotts, A. L., Green, C. E., Herin, D. V., and Schmitz, J. M. (2009). Cocaine dependence and concurrent marijuana use: a comparison of clinical characteristics. *Am. J. Drug Alcohol Abuse* 35, 193–198.
- Ling, W., Shoptaw, S., Wesson, D., Rawson, R. A., Compton, M., and Klett, C. J. (1997). Treatment effectiveness score as an outcome measure in clinical trials. *NIDA Res. Monogr.* 175, 208–220.
- Lucke, J. E. (2004). Fall-prevention programs for the elderly: a Bayesian secondary meta-analysis. *Can. J. Nurs. Res.* 36, 48–64.
- McLellan, A. T., Luborsky, L., Woody, G. E., and O’Brien, C. P. (1980). An improved diagnostic evaluation instrument for substance abuse patients. The addiction severity index. *J. Nerv. Ment. Dis.* 168, 26–33.
- Miller, N. S., Gold, M. S., and Klahr, A. L. (1990a). The diagnosis of alcohol and cannabis dependence (addiction) in cocaine dependence (addiction). *Int. J. Addict.* 25, 735–744.
- Miller, N. S., Klahr, A. L., Gold, M. S., Sweeney, K., Cocores, J. A., and Sweeney, D. R. (1990b). Cannabis diagnosis of patients receiving treatment for cocaine dependence. *J. Subst. Abuse* 2, 107–111.
- Mooney, M. E., Schmitz, J. M., Moeller, F. G., and Grabowski, J. (2007). Safety, tolerability and efficacy of levodopa-carbidopa treatment for cocaine dependence: two double-blind, randomized, clinical trials. *Drug Alcohol Depend.* 88, 214–223.
- Neelon, B., and O’Malley, A. J. (2010). Bayesian analysis using power priors with application to pediatric quality of care. *J. Biom. Biostat.* 1, 103.
- Neuenschwander, B., Branson, M., and Spiegelhalter, D. J. (2009). A note on the power prior. *Stat. Med.* 28, 3562–3566.
- O’Brien, C. P., and Lynch, K. G. (2003). Can we design and replicate clinical trials with a multiple drug focus. *Drug Alcohol Depend.* 70, 135–137.

- Pocock, S. J., Assmann, S. E., Enos, L. E., and Kasten, L. E. (2002). Subgroup analysis, covariate adjustment, and baseline comparisons in clinical trial reporting: current practice and problems. *Stat. Med.* 21, 2917–2930.
- Rothwell, P. M. (2005). Treating individuals 2: subgroup analysis in randomized controlled trials: importance, indications, and interpretations. *Lancet* 365, 176–186.
- Rounsaville, B. J., Petry, N. M., and Carroll, K. M. (2003). Single versus multiple drug focus in substance abuse clinical trials research. *Drug Alcohol Depend.* 70, 117–125.
- SAS Institute Inc. (2010). *SAS/STAT® Software*. Cary, NC: SAS Institute Inc.
- SAS Institute Inc. (2011). *Bayesian Binomial Model with Power Prior Using the MCMC Procedure*. Available at: <http://support.sas.com/rnd/app/examples/stat/BayesPop/bayespop.pdf>
- Schervish, M. J. (1996). P values: what they are and what they are not. *Am. Stat.* 50, 203–206.
- Schmitz, J. M., Mooney, M. E., Moeller, F. G., Stotts, A. L., Green, C., and Grabowski, J. (2008). Levodopa pharmacotherapy for cocaine dependence: choosing the optimal behavioral therapy platform. *Drug Alcohol Depend.* 94, 142–150.
- Simon, R. (2002). Bayesian subset analysis: application to studying treatment-by gender interactions. *Stat. Med.* 21, 2909–2916.
- Simon, R., Dixon, D. O., and Freidlin, B. (1996). “Bayesian subset analysis of a clinical trial for the treatment of HIV infections,” in *Bayesian Biostatistics (Statistics: a Series of Textbooks and Monographs)*, eds D. Berry and D. Stangl (New York: Marcel Dekker), 555–576.
- Simon, R., and Freedman, L. S. (1997). Bayesian design and analysis of two X two factorial clinical trials. *Biometrics* 53, 456–464.
- Spiegelhalter, D. J., Abrams, K. R., and Myles, J. P. (2004). *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. West Sussex: John Wiley & Sons Ltd.
- Wijeyesundera, D. N., Austin, P. C., Hux, J. E., Beattie, W. S., and Laupacis, A. (2009). Bayesian statistical inference enhances the interpretation of contemporary randomized controlled trials. *J. Clin. Epidemiol.* 62, 13–21.
- Zheng, T., Salganik, M. J., and Gelman, A. (2006). How many people do you know in prison? Using overdispersion in count data to estimate social structure in networks. *J. Am. Stat. Assoc.* 101, 409–423.
- that could be construed as a potential conflict of interest.

Received: 30 July 2012; accepted: 02 October 2012; published online: 30 October 2012.

Citation: Green C, Schmitz J, Lindsay J, Pedroza C, Lane S, Agnelli R, Kjome K and Moeller FG (2012) The influence of baseline marijuana use on treatment of cocaine dependence: application of an informative-priors Bayesian approach. *Front. Psychiatry* 3:92. doi: 10.3389/fpsy.2012.00092

This article was submitted to *Frontiers in Addictive Disorders and Behavioral Dyscontrol*, a specialty of *Frontiers in Psychiatry*.

Copyright © 2012 Green, Schmitz, Lindsay, Pedroza, Lane, Agnelli, Kjome and Moeller. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in other forums, provided the original authors and source are credited and subject to any copyright notices concerning any third-party graphics etc.

Conflict of Interest Statement: Robert Agnelli who was instrumental in programming the SAS statistical code is an employee of SAS, Inc., Cary, NC, USA. The other authors conducted the current research in the absence of any commercial or financial relationships