



HHS Public Access

Author manuscript

Am J Drug Alcohol Abuse. Author manuscript; available in PMC 2015 May 13.

Published in final edited form as:

Am J Drug Alcohol Abuse. 2009 ; 35(2): 95–102. doi:10.1080/00952990802647503.

Evaluation of Heterogeneity in Pharmacotherapy Trials for Drug Dependence: A Bayesian Approach

C. E. Green, Ph.D.,

Center for Clinical Research & Evidence-Based Medicine, University of Texas, Houston, Texas, USA

F. G. Moeller, M.D.,

Department of Psychiatry and Behavioral Sciences, University of Texas, Houston, Texas, USA

J. M. Schmitz, Ph.D.,

Department of Psychiatry and Behavioral Sciences, University of Texas, Houston, Texas, USA

J. F. Lucke, Ph.D.,

Center for Clinical Research & Evidence-Based Medicine, University of Texas, Houston, Texas, USA

S. D. Lane, Ph.D.,

Department of Psychiatry and Behavioral Sciences, University of Texas, Houston, Texas, USA

A. C. Swann, M.D.,

Department of Psychiatry and Behavioral Sciences, University of Texas, Houston, Texas, USA

R. E. Lasky, Ph.D., and

Center for Clinical Research & Evidence-Based Medicine, University of Texas, Houston, Texas, USA

J. P. Carbonari, Ed.D.

Department of Psychology, University of Houston, Houston, Texas, USA

Abstract

Aims—Difficulty identifying effective pharmacotherapies for cocaine dependence has led to suggestions that subgroup differences may account for some of the heterogeneity in treatment response. Well-attested methodological difficulties associated with these analyses recommend the use of Bayesian statistical reasoning for evaluation of salient interaction effects.

Copyright © Informa Healthcare USA, Inc.

Address correspondence to Charles E. Green, Ph.D., 6431 Fannin Street, UT-Medical School at Houston, Houston TX 77030. Charles.Green@uth.tmc.edu.

Declaration of Interest

The authors report no conflicts of interest. The authors alone are responsible for the content and writing of the paper.

Preliminary results were presented at the annual meeting of the College on Problems of Drug Dependence in 2008.

Supplementary materials are available for this article. Go to the publisher's online edition of *The American Journal of Drug and Alcohol Abuse* for the following free supplemental resource: Appendix 1

Methods—A secondary data analysis of a previously published, double-blind, randomized controlled trial examines the interaction of decision-making, as measured by the Iowa Gambling Task, and citalopram in increasing longest sustained abstinence from cocaine use.

Results—Bayesian analysis indicated that there was a 99% chance that improved decision-making enhances response to citalopram. Given the strong positive nature of this finding, a formal, quantitative Bayesian approach to evaluate the result from the perspective of a skeptic was applied.

Conclusions—Bayesian statistical reasoning provides a formal means of weighing evidence for the presence of an interaction in scenarios where conventional, Frequentist analyses may be less informative.

Keywords

Bayesian; cocaine dependence; substance abuse

INTRODUCTION

The search for pharmacotherapies for cocaine dependence has not yet produced any medications approved by the FDA (1). This may be partially attributable to heterogeneity among cocaine dependent patients (2–6). Proposed sources of variability include gender, ethnicity, employment status/SES, severity of dependence, route of administration, primary drug dependence diagnosis, drug and alcohol comorbidities, psychiatric comorbidities, readiness to change, and differences in dopaminergic tone (3–5). Recommendations for characterizing this heterogeneity include the exploratory evaluation of possible subgroups (4, 6) which might inform the design and analysis of clinical trials (1, 3, 5, 6).

Subgroup Analyses: Methodological Issues

Subgroup analysis for exploring heterogeneity in clinical trials is controversial (7–10). Design and sample size determination for clinical trials characteristically focus on evaluating the main effects of treatment. Methodological concerns regarding subgroup analyses include difficulties in the definition of subgroups, as well as the inflation of Type I and Type II error rates (11–15).

Since study-wise Type I error increases with the number of statistical tests, multiple comparisons arising through subgroup analysis (i.e., Number of statistical tests = Number of subgroups × Number of outcomes) are problematic. Since Type II errors arise from small sample sizes and small effects, diminished sample sizes in subgroups increase the risk of these errors as well. Given an optimistic scenario in which an interaction effect is equal in magnitude to the treatment effect, a sample providing 80% power to detect the treatment effect yields 29% power to detect the interaction (15) with 20% and 71% Type II error rates, respectively. In this context, achieving 80% power for the interaction term requires approximately quadrupling the sample size (15). Since interaction effects are typically a fraction of the magnitude of the treatment effect, even more extreme Type II error rates may result (15). Conducting subgroup analyses by examining treatment effects within each subgroup exacerbates these problems: Increasing the number of statistical tests while

partitioning the sample into even smaller sub-samples results in larger Type I and Type II error rates. This approach also fails to test the hypothesis that the sub-groups differ from each other, instead evaluating only whether the simple effect of treatment, in each group, differs from zero. The current paper agrees with the general consensus that subgroup analyses should eschew analysis of within-groups analyses in favor of tests of interaction, with the caveat that follow-up examination of simple effects often helps characterize interactions.

Interest in exploring heterogeneity in treatment response and concomitant methodological problems have resulted in various guidelines for the appropriate conduct of subgroup analyses, (12, 14, 16, 17). While these recommendations constitute an attempt at mitigating the methodological difficulties of subgroup analysis, the problem of elevated Type I and II error rates persists. Refining methods for understanding heterogeneity in treatment response has resulted in the application of Bayesian approaches to subgroup analyses (18–22). Bayesian analysis permits conclusions regarding the probability that a subgroup effect of a specified magnitude exists compared to the forced dichotomy of statistical hypothesis testing. Bayesian statistical methods more effectively evaluate the weight of the existing evidence, a function that is not well-served by the conventional Fisher/Neyman-Pearson (Frequentist) approach to statistical testing (23, 24).

STATISTICAL APPROACH

Bayesian Statistical Reasoning

Two divergent approaches to modeling uncertainty currently dominate statistical practice (25). Frequentist reasoning frames the preponderance of statistical teaching and practice. Bayesian reasoning, though historically older, is less familiar to the data analytic community. Distinctions between these approaches result from different conceptions of probability. Frequentists interpret the probability of a repeatable event as the limit of the event's relative frequency occurring in an infinite sequence of similar events. Bayesians interpret the probability of an event, which need not be repeatable, as a judgment or degree of belief that the event will occur (26).

Both Frequentists and Bayesians represent a set of observations using an abstract model, a *random data-generating process*, indexed by an unobservable parameter θ . Both sample observations attempting to gain information regarding θ . Frequentists assume this parameter is *fixed and unknown*: uncertainty regarding estimates of θ arises from sampling error, resulting, for example, in the assertion that the 95% Confidence Interval has a 0.95 probability of capturing the true value of the parameter. Bayesian statistical reasoning, while acknowledging uncertainty due to sampling error, defines parameters as *random*: uncertainty regarding parameter estimates is characterized using a probability distribution. This results in the assertion, for instance, that there is a 0.95 probability that the true parameter value falls in the 95% Bayesian Credible Interval (BCI), and also permits statements regarding the likelihood that certain values of the parameter obtain (26). Considering both observations and parameters to be random with subjective probability distributions makes the full resources of probability theory available for Bayesian inference, in particular, Bayes' Theorem:

$$p(\theta|\text{data})=p(\text{data}|\theta)p(\theta)/p(\text{data}).$$

This non-controversial mathematical theorem describes the process for coherently combining a *prior* judgment regarding θ (i.e., $p(\theta)$) and new data or the *likelihood* (i.e., $p(\text{data}|\theta)$) to form a *posterior* judgment regarding θ (27).

Requiring a prior distribution to represent the investigator's information regarding the parameter of interest has troubled Frequentists: an investigator's prior could be idiosyncratic, open to inferential "chicanery" (28; p.470), or so "dogmatic" that no amount of data could alter it. The role of the prior is to summarize all relevant evidence regarding the parameter of interest so that the data from the current study supplements that information (29). Submission of priors to scientific scrutiny forces investigators to make beliefs explicit regarding the phenomenon under investigation (30). Several considerations affect the choice of a prior (25). With critical appraisal, results of previous studies can be incorporated into the prior. Previous Bayesian analyses are particularly suitable because their posterior distributions, adjusted for adequacy and relevance (27), can form priors for the current study. Finally, specification of priors should favor parameters that demonstrate causal or theoretical plausibility.

Prior distributions are normative *representations* of an investigator's degree of belief (28). An investigator may *temper* a prior to more reasonably represent degrees of belief in the scientific community (31). Prior distributions can be made *diffuse* or *vague* indicating minimal pre-existing information. Further, multiple priors, such as "skeptical," "indifference," and "enthusiastic" priors, can be separately constructed to reflect variation in scientific opinion (27). *Skeptical* priors would be located over a region indicating a treatment has minimal or even harmful effects. *Indifference* priors would be located over a region indicating equipoise; roughly equal probability of benefit or harm from the treatment. *Enthusiastic* priors would be located over a region representing the prior belief that the treatment will be beneficial. Assessing the data under each of these priors will generate separate posterior distributions thus indicating the degree to which empirical findings would alter respective prior judgments. Accrual of more information should result in the convergence of these separate posteriors towards a common posterior distribution. Thus prior distributions, far from contaminating the data, promote transparency and objectivity in the scientific process; creating a more realistic logic of scientific reasoning.

Following a Bayesian analysis, all the information from the likelihood and the prior are contained in the posterior distribution for which point estimates (e.g., the posterior expectation and variance) and interval estimates (i.e., *credible intervals*) provide summaries.

Summarizing, parameters controlling a data-generation process are conceptualized differently in Frequentist and Bayesian approaches. Frequentist approaches regard the parameter as fixed but unknown; uncertainty is associated with an estimator of a parameter over possible observable samples. Bayesian approaches regard the parameter as random, assessing the uncertainty of parameter over the possible parameter values. Frequentist estimators attempt to "capture" parameters with confidence intervals at some level of

frequency. Bayesian estimators directly assess the probability of the parameter reasoning that observations, once observed, are no longer random but fixed. This feature of the Bayesian statistical method makes it more informative in evaluating evidence for parameter values.

Bayesian Subgroup Analysis

Simon (22) provides a clear explanation of Bayesian subgroup analysis. Given two predictors and their interaction, the key parameter is the b-weight associated with the interaction term. Estimating the posterior distribution of this parameter relies critically on the specification of the prior distribution. Simon argues that the paucity of true interactions in most research, suggests using skeptical, informative, prior distributions: priors centered on the null hypothesis of no effect with a low chance of a clinically meaningful effect occurring. This constrains the variance of the prior and requires strong evidence from the data to revise the posterior in favor of concluding that an interaction is likely. The variance for this skeptical prior is calculated as follows (22; p. 2911, Equation (1)):

$$\pi = \Phi \left(\frac{\delta}{\sqrt{d_i}} \right),$$

Where δ is a clinically meaningful effect size and π the probability of such an effect occurring. Φ is the cumulative normal distribution function. Given investigator-specified values for δ and π , Φ permits calculation of d_i ; the variance of the skeptical prior distribution. Indifferent, diffuse priors are specified for the other model parameters: centering priors at the null hypothesis (i.e., no effect) with large variances indicating relative ignorance regarding possible values for these parameters. Specification of diffuse priors results in estimates very close to Frequentist analyses. The current authors, likely as a function of subject area, diverge from Simon's view that interactions are relatively rare. Since the current analyses are hypothesis generating and are neither being used to set policy, nor to argue for changes in clinical practice we will posit indifferent, diffuse priors for all parameters estimated in the statistical models. The authors acknowledge the potential criticism that an exploratory analysis, with diffuse priors may capitalize on chance variability in the data. As such, an additional analysis using skeptical priors derived via Simon's (22) approach will permit evaluation of the credibility of results based on diffuse priors.

ANALYTIC EXAMPLE

Hypothesis

In a secondary data analysis of a trial of citalopram for treating cocaine dependence (32), the current paper implements a Bayesian subgroup analysis to evaluate treatment response as a function of a baseline performance measure of decision-making. Specifically, how does decision-making, as measured by performance on the Iowa Gambling Task (IGT), moderate the effect of citalopram in reducing longest sustained abstinence measured by consecutive, cocaine-negative urines?

Rationale for the Subgroup Analysis

Impaired decision-making is a correlate of substance abuse in general and specifically of cocaine dependence (33, 34). Decision-making as measured by the IGT involves disregard for long-term consequences, which is related to the nonplanning aspects of the broader construct of impulsivity (35). Impulsivity is a predictor of treatment retention, an indicator of treatment effectiveness (36, 37). Impulsivity is associated with serotonergic (5HT) function. Moreover, 5HT polymorphisms are associated with differential vulnerability to cocaine contingent reward (38–45). Serotonergic 5-HT_{2A} antagonists and/or 5-HT_{2C} agonists are promising agents for addressing impulsivity with possible resulting effects on cocaine use among cocaine dependent participants (31). Given its role in impulsivity, which is implicated in substance use, and its potential responsiveness to serotonergic medications, baseline level of decision-making may moderate the effect of citalopram on cocaine use.

METHODS

Sample

Details of the trial design and sample composition may be found elsewhere (31). The trial (N = 76) was powered to detect a treatment effect due to citalopram on cognitive function and cocaine-free urines. While the original choice of citalopram was guided primarily by its relatively selective effects on receptor systems associated with impulsivity, the trial was not powered to detect an interaction between behavioral laboratory measures associated with impulsivity (e.g., the IGT) and citalopram. This is consequently a *post hoc* analysis of the interaction between impulsivity and citalopram.

Measures

Outcome—Longest sustained abstinence from cocaine-use was modeled as the longest consecutive number of cocaine-free urines (i.e., benzoylecgonine < 300 ng/ml), measured twice weekly. Following Kampman et al. (46) intermittent missing data were recoded as cocaine-positive, data missing due to dropout were left as missing.

Decision-Making

Iowa Gambling Task (47)—This computerized version of the original gambling task requires subjects to choose between four decks of 60 cards each with different theoretical monetary reward rates. Scoring is based on the total number of cards selected from the advantageous minus the disadvantageous decks across five blocks of 20 cards each. The net score of cards selected in each of the five blocks was used as a behavioral index of decision-making.

Statistical Analyses

All Bayesian statistical analyses utilized PROC BGENMOD (SAS v. 9.1.3) for the analysis of generalized linear models. Longest consecutive number of cocaine-free urines was modeled as a Poisson outcome. Violations of dispersion were addressed via a Pearson scaling coefficient. Based on desirable distributional properties all priors were specified on the log scale: priors specified for the log of the b-weights approximate a normal distribution

(27). Evidence for an interaction is examined from two perspectives: one in which the prior assumes very little regarding the likelihood of the interaction (diffuse prior), and one in which a meaningful interaction is assumed to have a low likelihood of occurring (skeptical prior).

Diffuse Priors

Log-scale, diffuse prior distributions were centered at zero, with a variance of 1×10^6 , indicating an initial assumption of no effect, with a 95% chance that the true log (parameter) falls within a broad range (i.e., $\pm 1.96 \times \sqrt{1 \times 10^6}$ or -1960 to 1960).

Skeptical Priors

Following Simon (22), all parameters except for the interaction term are specified as $\sim N(0, 1 \times 10^6)$. Calculation of the prior variance for the skeptical prior assumes that a clinically meaningful effect would be a 1% improvement in longest consecutive abstinence in the citalopram condition for every one-point increment in performance on the IGT (i.e., a risk-ratio = 1.01). Working in the log-scale, this effect takes the value $\delta = \log(1.01) = 0.00995$. Assuming that this effect is unlikely to occur we set $\pi = 0.025$. Centering the skeptical prior at the null hypothesis (i.e., $\log(\text{risk-ratio}) = 0$) and utilizing the cumulative normal distribution function yields a variance of $d_i = 2.577 \times 10^{-5}$. Thus the skeptical prior has the form $\sim N(0, 2.577 \times 10^{-5})$ indicating equipoise regarding the presence of an interaction, but that a clinically meaningful effect is unlikely.

RESULTS

For all analyses, estimation of the posterior distributions via Monte-Carlo Markov chain (MCMC) demonstrated adequate graphical (Trace Plot, Autocorrelation Plot) and quantitative (Geweke Diagnostics, Gelman-Rubin Diagnostics, and Heidelberger-Welsh Diagnostics) evidence of convergence. Explanations of these diagnostics may be found in the SAS documentation for the Bayesian procedures (48).

Figures 1(a) and 1(b) depict the prior and posterior distributions for the diffuse and skeptical analyses, respectively. Figure 1(c) shows three posterior distributions, indicating the likelihood of an interaction based on a diffuse prior, and two posterior distributions for the simple effects of IGT net score in the Citalopram and Placebo conditions, respectively. While subgroup analyses should rely on the evaluation of the interaction term, posterior distributions for the within group effects are provided to elucidate the nature of the subgroup differences.

Table 1 provides a quantitative summary of the posterior distributions, parameter estimates and 95% BCI's in exponentiated form to facilitate interpretation. Unless otherwise indicated all Bayesian parameter estimates derive from the model using diffuse priors. Controlling for the main effects of treatment and IGT performance, individuals in the Citalopram condition experienced an increase in longest sustained abstinence by a factor of 1.03 (or 3%) for every one-point increment in IGT net score (Figure 2). The 95% BCI does not include 1.00 suggesting that the effect is reliably different from this value. Table 1 shows that maximum

likelihood estimates from the Frequentist analysis are in approximate agreement with Bayesian values derived using diffuse priors, however, Bayesian estimates permit a more nuanced interpretation of the data.

Additional information, resides in the posterior distribution for the interaction term (Figure 1a), permitting evaluation of the strength of the evidence for an interaction effect of some specified size. A substantial proportion of the area under the curve (i.e., 0.998) lies to the right of a $\log(\text{risk-ratio}) = 0$ indicating a 99.8% chance that a beneficial subgroup effect obtains. Table 2 shows the probability of various, potentially “meaningful” effect sizes, provided in exponentiated form for interpretability. The interaction column shows probabilities of increasing risk ratios. For example, we find that there is an 83.7% probability that a risk ratio equal to or greater than 1.02 obtains. Evaluation of the simple effects elucidates the differences across subgroups. Again using a risk ratio of 1.02 or greater, there is a 98.7% chance it that this effect obtains within the Citalopram and only a 2.6% chance it obtains in the placebo condition. Examining the posterior distribution in more detail (Table 2), within the Citalopram condition the probability of IGT net score being positively associated with an incremental increase in sustained abstinence (risk-ratio = 1.0) is >99.9%, the likelihood of a risk-ratio = 1.03 is 63%. Within the Placebo condition the risk-ratio associated with IGT net score is 1.00 (95% BCI 0.982–1.020). The probability of higher IGT net scores being associated with a positive benefit in sustained abstinence (risk-ratio = 1.0) is 54.2% while the probability of a risk-ratio = 1.03 is only 0.1%.

Analyzing the same data with a skeptical prior specified for the interaction term still indicates that a large proportion (i.e., [0.89]) of the posterior distribution falls to the right of the $\log(\text{risk-ratio}) = 0$ suggesting that a skeptic would still conclude that there is an 89% chance that an interaction between IGT and citalopram obtains.

Using either prior distribution there appears to be strong evidence that decision-making interacts with citalopram in predicting longest sustained abstinence from cocaine use. Better decision-making as defined by the IGT net score is associated with stronger effects of citalopram on sustained abstinence.

DISCUSSION

Difficulties in developing effective medications for cocaine dependence have prompted calls for evaluation of heterogeneity in treatment response. Attendant methodological complications require consideration of quantitative strategies that will maximize the information gleaned from the observed data in order to most effectively weigh the evidence for dimensions characterizing this heterogeneity. Bayesian methods offer benefits over current Frequentist methods due to the information included in the posterior distribution. Leveraging this information requires acknowledgement that long-run frequencies are not the only way in which we might think of probability, but that probabilities may also represent the strength of belief we hold about phenomena based on the evidence at our disposal (49).

The current analysis uses this approach to address the issue of subgroup effects, broadly defined as moderators. Bayesian reasoning permits formal specification of prior beliefs

regarding the likelihood that an interaction exists. Incorporating the observed data yields a posterior distribution, permitting probabilistic estimates that interactions do occur. Specification of different priors permits consideration of the empirical evidence from various perspectives.

Information inherent in the posterior distribution may help in effectively focusing research efforts. Heterogeneity detected in Bayesian subgroup analysis may be incorporated into future trial designs. Further, informative, empirically defensible priors may lead to smaller more efficient trials. Finally, the same process of systematically revising prior distributions into posterior distributions based on the observed data provides not only a bridge across studies, but also provides a mechanism for data analysis required in adaptive designs. Use of such designs may offer further efficiency in the search for pharmacotherapies for cocaine dependence.

Acknowledgments

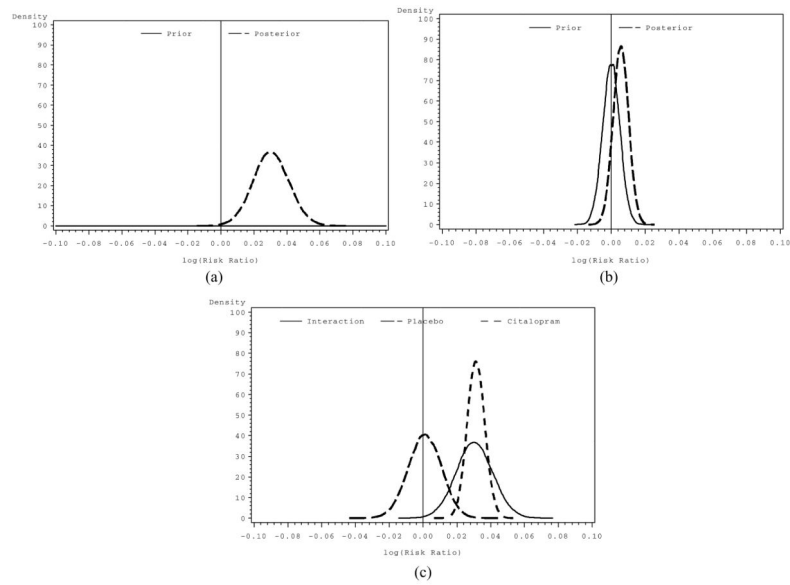
This study was funded by grants from the National Institute on Drug Abuse (5R01DA008425 and P50-DA-9262) and the University of Texas Medical School at Houston (Clinical Investigator Award).

References

1. de Lima MS, de Oliveira Soares BG, Pereira Reisser AA, Farell M. Pharmacological treatment of cocaine dependence: A systematic review. *Addiction*. 2002; 97:931–949. [PubMed: 12144591]
2. de Lima MS, de Oliveira Soares BG, Pereira Reisser AA, Farell M. Pharmacological treatment of cocaine dependence: a systematic review. *Addiction*. 2002; 97:931–949. [PubMed: 12144591]
3. Elkashev A, Vocci F. Biological markers of cocaine addiction: Implications for medications development. *Addict Biol*. 2003; 8:123–139. [PubMed: 12850771]
4. Carroll KM, Nich C, Rounsaville BJ. Variability in treatment-seeking cocaine abusers: Implications for clinical pharmacotherapy trials. *NIDA Res Monogr*. 1997; 175:137–157. [PubMed: 9467796]
5. Nunes EV. Methodologic recommendations for cocaine abuse clinical trials: A clinician-researcher's perspective. *NIDA Res Monogr*. 1997; 175:73–95. [PubMed: 9467793]
6. Elkashev A, Holmes TH, Bloch DA, Shoptaw S, Kampman K, Reid MS, Somoza E, Ciraulo D, Rotrosen J, Leiderman D, Montgomery A, Vocci F. Retrospective analyses of pooled data from CREST I and CREST II trials for treatment of cocaine dependence. *Addiction*. 2005; 100(Supplement. 1):91–101. [PubMed: 15730353]
7. Hernandez AV, Boersma E, Murray GD, Habbema JDF, Steyerberg EW. Subgroup analyses in therapeutic cardiovascular clinical trials: Are most of them misleading? *Am Heart J*. 2006; 151:257–264. [PubMed: 16442886]
8. Counsell CE, Clarke MJ, Slattery J, Sandercock PAG. The miracle of DICE therapy for acute stroke: Fact or fictional product of subgroup analysis? *British Med J*. 1994; 309:1677–1681.
9. ISIS-2 Collaborative Group. Randomized trial of IV streptokinase, oral aspirin, both, or neither among 17187 cases of suspected myocardial infarction. *Lancet*. 1988; 2:349–360. [PubMed: 2899772]
10. Feinstein AR. The problem of cogent subgroups: A clinicostatistical tragedy. *J Clin Epidemiol*. 1998; 51(4):297–299. [PubMed: 9539884]
11. Brookes ST, Whitely E, Egger M, Smith GD, Mulheran PA, Peters TJ. Subgroup analyses in randomized controlled trials: Quantifying the risks of false-positives and false negatives. *Executive Summary Health Technol Assess*. 2001; 5(33):1–4.
12. Assmann SF, Pocock SJ, Enos LE, Kasten LE. Subgroup analysis and other (mis)uses of baseline data in clinical trials. *Lancet*. 2000; 355:1064–1069. [PubMed: 10744093]

13. Pocock SJ, Assmann SE, Enos LE, Kasten LE. Subgroup analysis, covariate adjustment and baseline comparisons in clinical trial reporting: Current practice and problems. *Stat Med.* 2002; 21:2917–2930. [PubMed: 12325108]
14. Rothwell PM. Treating Individuals 2: Subgroup analysis in randomized controlled trials: Importance, indications, and interpretations. *Lancet.* 2005; 365:176–186. [PubMed: 15639301]
15. Brookes ST, Whitely E, Egger M, Smith GD, Mulheran PA, Peters TJ. Subgroup Analyses in randomized trials: Risks of subgroup-specific analyses; power and sample size for interaction tests. *J Clin Epidemiol.* 2004; 57:229–236. [PubMed: 15066682]
16. Adams KF. Post hoc subgroup analysis and the truth of a clinical trial. *Am Heart J.* 1998; 136(5): 753–758. [PubMed: 9812066]
17. Cook DI, GebSKI VJ, Keech AC. Subgroup analysis in clinical trials. *Med J Aust.* 2004; 180:289–291. [PubMed: 15012568]
18. Guillemain F. Primer: The fallacy of subgroup analysis. *Nat Clin Pract Rheumatology.* 2007; 3(7): 407–413.
19. Dixon DO, Simon R. Bayesian Subset Analysis. *Biometrics.* 1991; 47:871–881. [PubMed: 1742443]
20. Simon R, Freedman LS. Bayesian design and analysis of two X two factorial clinical trials. *Biometrics.* 1997; 53:456–464. [PubMed: 9192445]
21. Simon, R.; Dixon, DO.; Freidlin, B. Bayesian subset analysis of a clinical trial for the treatment of HIV infections. In: Berry, Donald; Stangl, Dalene, editors. *Bayesian Biostatistics (Statistics: a Series of Textbooks and Monographs)*. Marcel Dekker; New York: 1996.
22. Simon R. Bayesian subset analysis: Application to studying treatment-by-gender interactions. *Stat Med.* 2002; 21:2909–2916. [PubMed: 12325107]
23. Goodman SN. Toward Evidence-Based Medical Statistics. 1: The P Value Fallacy. *Ann Intern Med.* 1999; 130:995–1004. [PubMed: 10383371]
24. Johnstone DJ. Tests of significance in theory and practice. *Statistician.* 1986; 35:491–504.
25. Lucke JF. Fall prevention programs for the elderly: A Bayesian secondary meta-analysis. *Canadian J Nurs Res.* 2004; 36(3):48–64.
26. Howson, C.; Urbach, P. *Scientific reasoning: The Bayesian approach*. Open Court; Chicago: 2006.
27. Firebaugh G. Will Bayesian Inference Help? A Skeptical View *Sociological Methodol.* 1995; 25:469–472.
28. Spiegelhalter, DJ.; Abrams, KR.; Myles, JP. *Bayesian approaches to clinical trials and health-care evaluation*. John Wiley & Sons Ltd; West Sussex, England: 2004.
29. Kadane JB. Prime time for Bayes. *Control Clin Trials.* 1995; 16:313–318. [PubMed: 8582149]
30. Hacking, I. *An introduction to probability and inductive logic*. Cambridge University Press; Cambridge: 2001.
31. Shimony, A. *Search for a naturalistic world view: Scientific method and epistemology*. Cambridge University Press; Cambridge, UK: 1993. *Scientific inference*.
32. Moeller FG, Schmitz JM, Steinberg JL, Green CM, Reist C, Lai LY, Swann AC, Grabowski J. Citalopram combined with behavioral therapy reduces cocaine use: A double-blind placebo controlled trial. *Am J Drug Alcohol Abuse.* 2007; 33(3):367–378. [PubMed: 17613964]
33. Bechara A, Damasio H. Decision-making and addiction (part I): Impaired activation of somatic states in substance dependent individuals when pondering decisions with negative future consequences. *Neuropsychologia.* 2002; 40:1675–1689. [PubMed: 11992656]
34. Stout JC, Busemeyer JR, Lin A, Grant SJ, Bonson KR. Cognitive modeling analysis of decision-making processes in cocaine abusers. *Psychon Bull Rev.* 2004; 11(4):742–747. [PubMed: 15581127]
35. Moeller FG, Barratt ES, Dougherty DM, Schmitz JM, Swann AC. Psychiatric aspects of impulsivity. *Am J Psychiatry.* 2001; 158:1783–1793. [PubMed: 11691682]
36. Coffey SF, Gudleski GD, Saladin ME, Brady KT. Impulsivity and rapid discounting of delayed hypothetical rewards in cocaine-dependent individuals. *Exp Clin Psychopharmacol.* 2003; 11:18–25. [PubMed: 12622340]

37. Patkar AA, Murray HW, Mannelli P, Gottheil E, Weinstein SP, Vergare MJ. Pre-treatment measures of impulsivity, aggression and sensation seeking are associated with treatment outcome for African-American cocaine dependent patients. *J Addict D/s*. 2004; 23:109–122.
38. Peroutka SJ. Serotonin receptor variants in disease: New therapeutic opportunities? *Ann NY Acad Sci*. 1998; 861:16–25. [PubMed: 9928234]
39. Gothert M, Propping P, Bonisch H, Bruss M, Nothen MM. Genetic variation in human 5-HT receptors: potential pathogenetic and pharmacological role. *Ann NY Acad Sci*. 1998; 861:26–30. [PubMed: 9928235]
40. Cravchik A, Goldman D. Neurochemical individuality: Genetic diversity among human dopamine and serotonin receptors and transporters. *Arch Gen Psychiatry*. 2000; 57:1105–1114. [PubMed: 11115324]
41. Bjork JM, Moeller FG, Dougherty DM, Swann AC, Machado MA, Hanis CL. Serotonin 2a receptor T102C polymorphism and impaired impulse control. *Am J Med Genet*. 2002; 14:336–339. [PubMed: 11920859]
42. Glatt CE, Tampilic M, Christie C, DeYoung J, Freimer NB. Re-screening serotonin receptors for genetic variants identifies population and molecular genetic complexity. *Am J Med Genet B Neuropsychiatr Genet*. 2004; 124:92–100. [PubMed: 14681923]
43. Patkar AA, Mannelli P, Peindl K, Hill KP, Gopalakrishnan R, Berrettini WH. Relationship of disinhibition and aggression to blunted prolactin response to metachlorophenylpiperazine in cocaine-dependent patients. *Psychopharmacology (Be/t)*. 2006:1–10.
44. Mannelli P, Patkar AA, Peindl K, Tharwani H, Gopalakrishnan R, Hill KP, Berrettini WH. Polymorphism in the serotonin transporter gene and moderators of prolactin response to metachlorophenylpiperazine in African-American cocaine abusers and controls. *Psychiatry Res*. 2006; 144:99–108. [PubMed: 1700009]
45. Kreek MJ, Bart G, Lilly C, LaForge KS, Nielsen DA. Pharmacogenetics and human molecular genetics of opiate and cocaine addictions and their treatments. *Pharmacol Rev*. 2005; 57:1–26. [PubMed: 15734726]
46. Kampman KM, Pettinati H, Lynch KG, Dackis C, Sparkman T, Weigley C, O'Brien CP. A pilot trial of topirimate for the treatment of cocaine dependence. *Drug Alcohol Depend*. 2004; 75:233–240. [PubMed: 15283944]
47. Bechara A, Damasio H, Tranel D, Damasio AR. Deciding advantageously before knowing the advantageous strategy. *Science*. 1997; 275:1293–1295. [PubMed: 9036851]
48. SAS Institute Inc. Preliminary Capabilities for Bayesian Analysis in SAS/STAT® Software. Cary, NC: SAS Institute Inc; 2006.
49. Ramsey, FP. Truth and Probability. In: Braithwaite, RB., editor. *The Foundations of Mathematics and other Logical Essays*. Vol. VII. London: Kegan, Paul, Trench, Trubner & Co; New York: Harcourt, Brace and Company; 1931. p. 156-198.(1926)

**FIG. 1.**

(a) Non-informative prior distribution and the resulting posterior distribution for the interaction of IGT net score and citalopram. (b) Skeptical prior distribution and the resulting posterior distribution for the interaction of IGT net score and citalopram. (c) Posterior distributions for the interaction of IGT net score and citalopram, and the simple effects of IGT net score within each treatment condition.

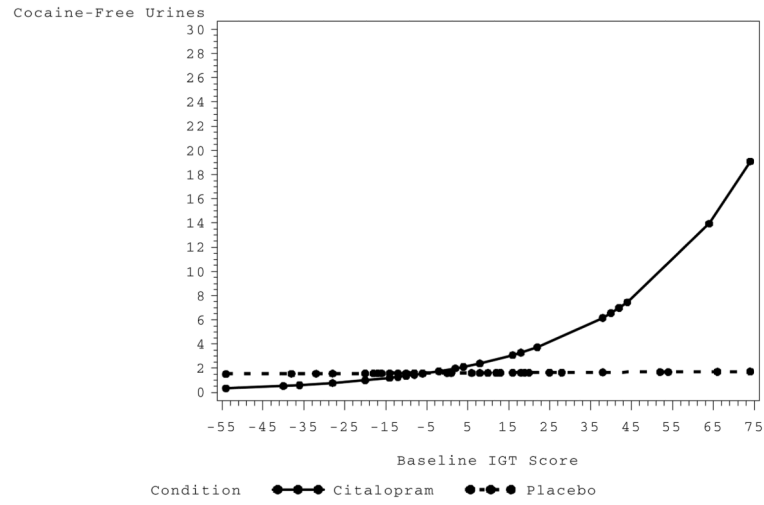


FIG. 2. Estimated values for longest number of consecutive days abstinent as a function of IGT score, treatment and their interaction.

TABLE 1

Descriptive statistics of the posterior samples and maximum likelihood estimates in exponentiated form

Parameter	Bayesian Estimates				Frequentist Estimates			
	Posterior Mean	95% Credible Interval		Maximum Likelihood Estimate	Wald 95% Confidence Limits		Upper Limit	
		Lower Limit	Upper Limit		Lower Limit	Upper Limit		
Intercept	1.60866	0.92821	2.61039	1.71910	1.04112	2.83885		
IGT	1.00098	0.98393	1.01796	1.00120	0.98452	1.01826		
Treatment	1.16416	0.52038	2.55640	1.14499	0.52967	2.47540		
IGT × Treatment	1.03087	1.00932	1.05317	1.03035	1.00894	1.05232		

TABLE 2

Posterior probabilities for selected effect sizes for the interaction term and simple effects (using non-informative priors)

Probability Estimates Derived from Bayesian Posterior Distributions			
Risk Ratio	Probability of the Risk Ratio		
	Interaction	IGT net score within the Citalopram Condition	IGT net score within the Placebo Condition
>1.0	0.998	>0.999	0.542
1.01	0.971	>0.999	0.180
1.02	0.837	0.987	0.026
1.03	0.528	0.630	0.001
1.04	0.207	0.065	<0.001
1.05	0.046	<0.001	<0.001

Author Manuscript

Author Manuscript

Author Manuscript

Author Manuscript