Clinical Therapeutics 46 (2024) 176–180

Contents lists available at ScienceDirect

Clinical Therapeutics

journal homepage: www.elsevier.com/locate/clinthera

Research Letter

Bayesian Analysis of the ICAT·COVID Randomized Clinical Trial ☆

Pierre Malchair, MD [1,*], Jesús Villoria, MD, BStat [2,3,*,#], Jordi Giol, MD, PhD [1], Javier Jacob, MD, PhD [1], Thiago Carnaval, MD [4,5], Sebastián Videla, MD, PhD [4,5]

[1] *Emergency Department, Bellvitge University Hospital, Barcelona, Spain*
[2] *Pharmacology Unit, Department of Pathology and Experimental Therapeutics, School of Medicine and Health Sciences, Institute of Neurosciences, University of Barcelona, L'Hospitalet de Llobregat, Spain*
[3] *Department of Design and Biometrics, Medicxact, Madrid, Spain*
[4] *Clinical Research Support Unit, Bellvitge Biomedical Research Institute (IDIBELL), Bellvitge University Hospital, Barcelona, Spain*
[5] *Pharmacology Unit, Department of Pathology and Experimental Therapeutics, School of Medicine and Health Sciences, Bellvitge Biomedical Research Institute (IDIBELL), University of Barcelona, Barcelona, Spain*

ARTICLE INFO

Key words:
Coronavirus infections
Odds ratio
Pandemics
SARS-CoV-2
Survival analysis

Introduction

This communication provides new effect measures in the multiplicative scale from the ICAT·COVID randomized clinical trial, obtained through Bayesian statistics. These could not be calculated using the traditional frequentist statistics included in the original publication because the benefits of icatibant (a competitive antagonist of the bradykinin B2 receptors) on top of standard care in patients with COVID-19 pneumonia were such that there were no events in the active group.[1] Additive effect measures (eg, risk differences) are the most appropriate measures for identifying the population groups that will benefit most from interventions in presence of interactions acting as effect modifiers.[2] However, an aspect that multiplicative measures provide where additive effect measures cannot, is an indication of how many times interventions or exposures increase or decrease disease risk (eg, risk ratio, hazard ratio). Furthermore, multiplicative measures are more commonly used in epidemiology, and are more appropriate for outcome measures with strictly positive values, such as counts and the numerators of incidence rates.[3]

Strictly speaking, when there are no events in one group, multiplicative effect measures are indeterminate because of the singularity involving multiplication or division by zero. However, even in these circumstances, the Bayesian approach provides a framework that enables their estimation.[4] This is because it relies on an entire distribution of potential values rather than on the singular (ie, collapsed) values of the outcomes observed in any given sample. Bayesian statistics, nevertheless, may suffer from imprecision and, more notably, biases derived from a biased selection of the prior distributions, particularly when the sample size is small,[5] as in our study. This has deterred some editors from publishing Bayesian statistics where they are not well prespecified in the study protocol.[6] This rules them out for post hoc analyses that, by definition, are carried out once the results are known.

For the aforesaid reasons, we did not include Bayesian post hoc analyses in the primary manuscript,[1] but we provide them here in this Research Letter to complement and support the conclusions on clinical response and mortality by affording unified effect measures in the multiplicative scale.

☆ Trial Registration: ClinicalTrials.gov identifier NCT04978051.

Address correspondence to: Jesús Villoria, MD, BStat, Department of Design and Biometrics, Medicxact, plaza Ermita 4 Alpedrete, 28430 Madrid, Spain.
E-mail address: villoriajesus@medicxact.es (J. Villoria).
* These authors contributed equally to this manuscript.

https://doi.org/10.1016/j.clinthera.2023.11.010
Accepted 13 November 2023

New Methods and Results

Participants and Methods

The ICAT·COVID was a Phase II proof-of-concept, randomized, open-label, controlled trial that assessed the safety and efficacy of icatibant to avoid severe COVID-19 progression into late and sometimes fatal stages. The Ethics Committee of the Bellvitge University Hospital and the Spanish Medicines Agency approved the trial protocol prior to start, and all patients provided written informed consent to participate. Included were 77 inpatients with COVID-19 pneumonia requiring supplemental but not high-flow oxygen or mechanical ventilation. These patients were randomly allocated three 30 mg subcutaneous doses of icatibant per day for 3 consecutive days on top of standard care (icatibant group, n = 39), or standard care alone (SoC group, n = 38). Outcomes were: clinical response at Day 10 and 28 days after initial discharge, time to cessation of supplemental oxygen and hospital discharge, COVID-19–related and all-cause mortality, and safety. Response was defined as achieving a score ≤2 in a clinical progression scale either with or without an additional safety criterion (stringent and lax response criteria, respectively). Further details on design, participants, procedures and outcomes can be found elsewhere.[1]

Statistical Analysis

Analyses involving dichotomous outcomes, including response rates, were performed using the Bayesian Binomial model with flat prior distributions for event rates, assuming that, in the absence of previous similar studies, the prior probability was the same in both study arms (ie, we were skeptical about the effect of the intervention; the Figure 1 shows

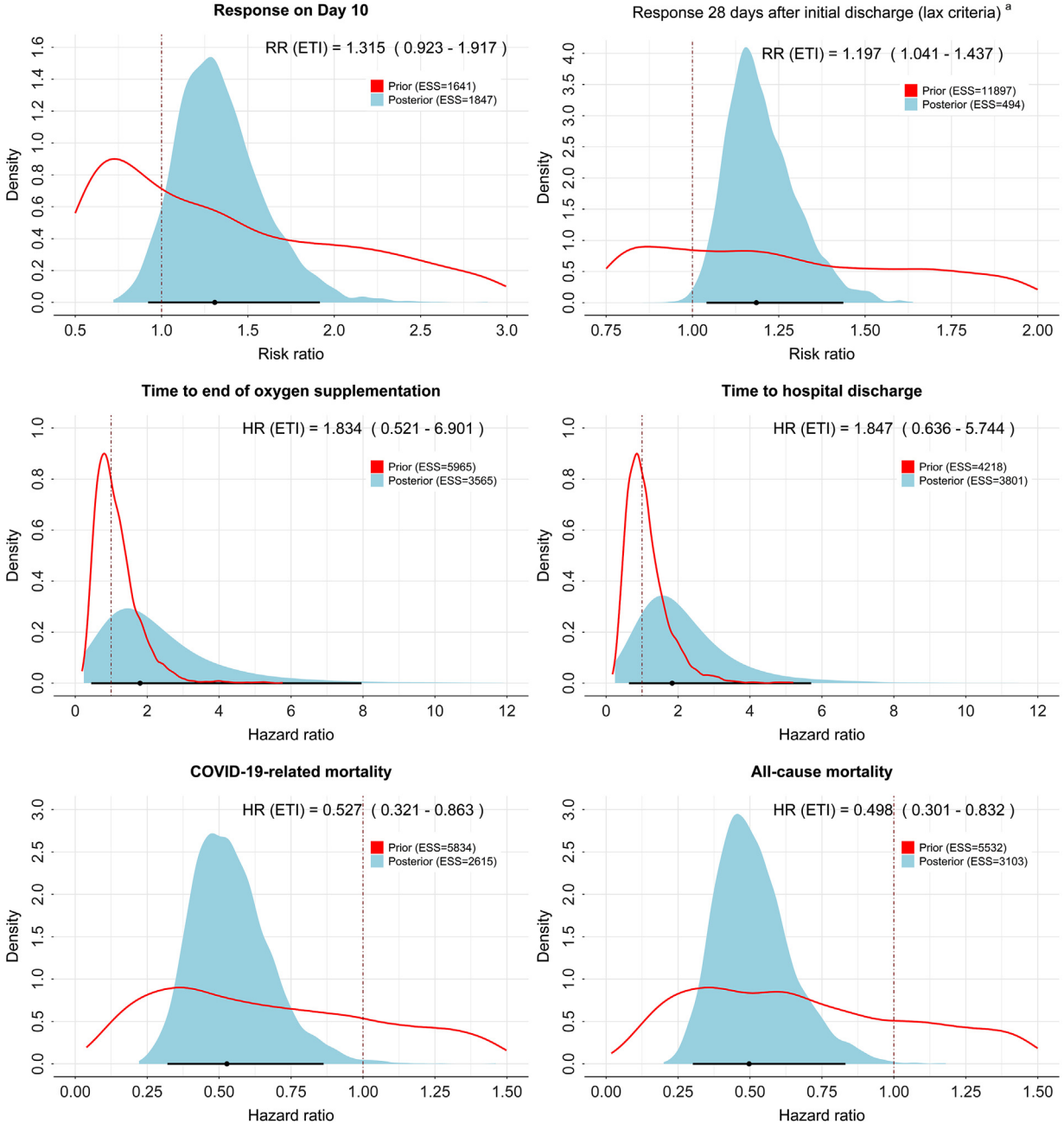

Figure 1. Bayesian prior and posterior distributions. ESS = effective sample size; ETI = equal-tailed 95% credible intervals; HR = hazard ratio; RR = risk ratio; SoC = standard of care. [a]The Bayesian risk ratio for the stringent criteria was marginally significant (RR [ETI]: 1.17 [1.00–1.44]).

how the empirical prior distributions outlined as red lines have peaks around the neutral value 1 or include it with wide margins). Multiplicative effect measures (odds ratios and risk ratios) were obtained using the canonical logit and noncanonical log links, respectively. In addition, additive effect measures (risk differences) were obtained using the noncanonical identity link.[7] Analyses involving time-to-event endpoints were done using the Bayesian Cox model with an uninformative prior for the hazard ratio. We assumed that the hazard rates would be the same in both arms (Figure 1, idem as previous). The performance of prior distributions was checked through empirical images obtained by dry-runs of the sampling algorithm without updating. This was particularly useful to avoid excessive widening of the parameter space, which could in turn have had anticonservative effects (see the discussion). The implementation was done with the *brms* R package,[8] used as interface to the Stan modeling language, using 4 independent Markov Chain Monte Carlo (MCMC) chains of sizes 2000 to 10,000 each. The size of the chains was adjusted on a model basis to achieve adequate effective sample sizes (ESS), and they were systematically checked for convergence. Inference was made through equal-tailed credible intervals (ETI) based on 2.5th, 50th, and 97.5th quantiles of the averaged posterior distributions. For convenience, the homologous results obtained during the original frequentist analyses are also provided, when available, to enable head-to-head comparisons.

Results

Seventy-three out of 77 patients had analyzable data. Baseline features were well matched between study arms (icatibant, n = 37 and SoC, n = 36). In the icatibant group, 73% and 100% (27 and 37 of 37) of patients met response criteria on Day 10 and 28 days after initial discharge, respectively, compared to 55.6% and 83.3% in the SoC group. The respective relative risks were 1.32 and 1.20 (Bayesian credible intervals: 0.92–1.92 and 1.04–1.44, Figures 1 and 2), and the probabilities that the response was higher in the icatibant group compared to SoC were 93.5% (Day 10) and 99.7% (28 days after discharge). Since the credibility interval for Day 10 includes the value 1, this result can be regarded as nonsignificant (but see the comments about statistical power in the discussion). The times on supplemental oxygen and hospitalization were shorter in the icatibant group than they were in the SoC group, although the differences did not reach statistical significance in the Bayesian case. The hazard ratios were 1.83 and 1.85, respectively (Bayesian credible intervals: 0.52–6.90 and 0.64–5.74, Figures 1 and 2). No patient in the icatibant group died compared to 6 out of 36 patients (17.6%) in the SoC group (5, 13.9%, conclusively due to COVID-19). The hazard ratios were 0.53 for COVID-19–related and 0.50 for all-cause mortality, both significantly below one (Bayesian credible intervals: 0.32–0.86 and 0.30–0.83, Figures 1 and 2).

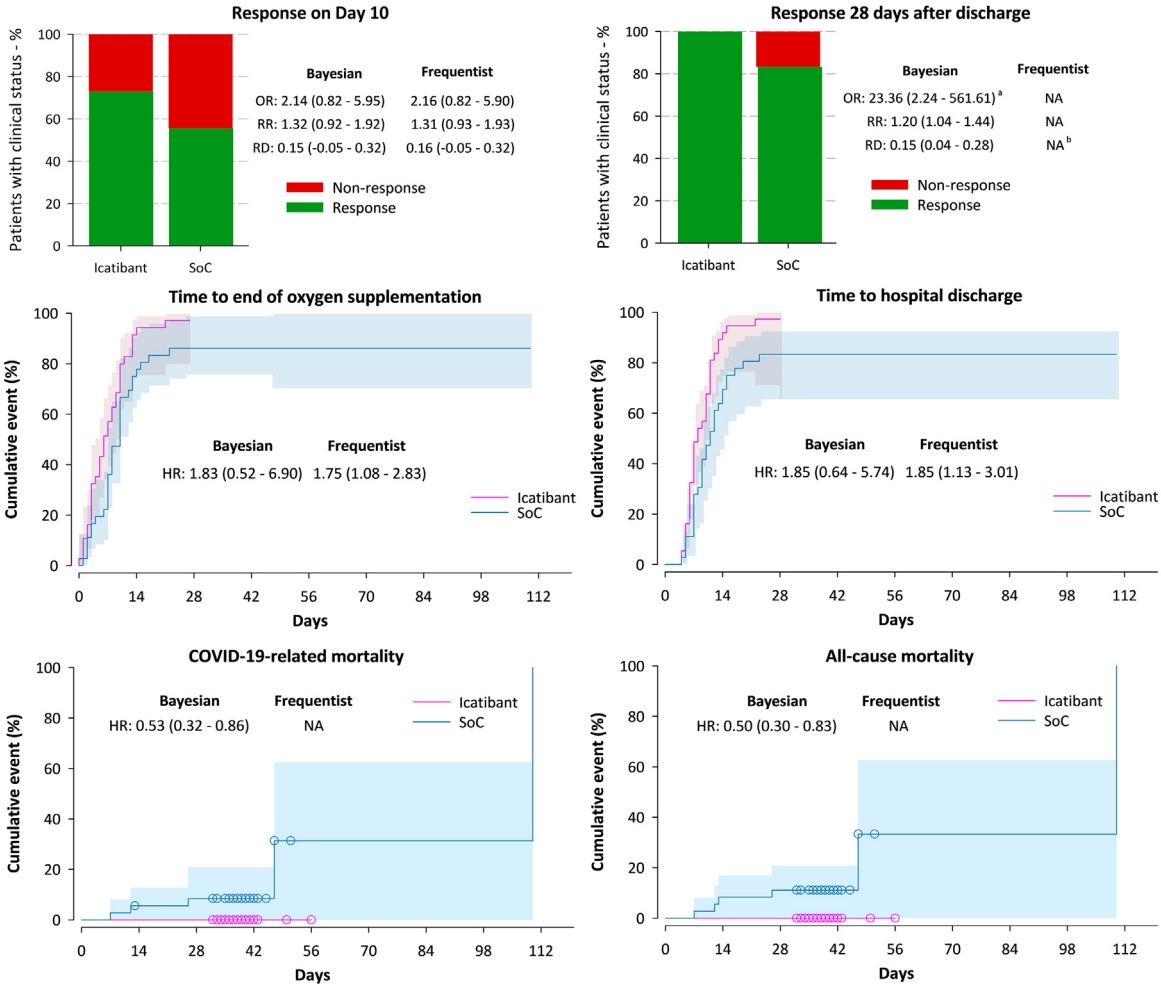

Figure 2. Results and effect measures for icatibant vs. SoC. CI = 95% equal-tailed credible interval (for Bayesian statistics) or 95% confidence interval (for frequentist statistics); HR = hazard ratio; NA = not applicable (division by zero); OR = odds ratio; RD = risk difference; RR = risk ratio; SoC = standard of care. [a]Odds ratios became impractical when there are no events in one of the groups even in the Bayesian case. [b]Although the risk difference can be calculated analytically, it cannot be estimated via maximum likelihood because the function is not defined when there are no events in one of the groups.

Discussion

Although all patients in the active group met the response criteria and none died, current effect measures in the multiplicative scale allow us to infer that the addition of icatibant could increase the response by 20% to 30% and cut mortality in half—and possibly up to 70% according to the lower limits of the credible intervals—in patients with COVID-19 pneumonia. In other aspects, the present Bayesian analyses confirm the results obtained in the primary analyses, which showed a hastening of clinical improvement and a reduction in mortality.

In this occasion, we did not use a competing risk model for the time to end of oxygen supplementation and hospital discharge. This was because the calculation of Bayesian cause-specific hazard ratios is quite complex and the efficiency inflation with regard to the frequentist approach was negligible. Notwithstanding, the Bayesian analysis was more conservative than the frequentist analysis. Significance was not reached, probably due to the frequentist approach relying heavily on only 3 patients from the control group with very long times.[1] This occurred even though the (partial) likelihood function of the Cox model intrinsically constrains the scale parameter. This is a good illustration of the versatility and strength of Bayesian statistics. In fact, if the prior distribution was forced to accommodate much variability (eg, placing a large-variance Gaussian prior for the treatment parameter) the results would reach significance at the expense of unreasonably large upper credible limits (as high as e,[4] data available on request). We therefore recommend checking the behavior of priors by drawing samples without updating the likelihood, as we did in these analyses (see Methods), although this option is not feasible when improper, nonparametric, generator priors are employed.

As in the frequentist analysis, statistical significance was not reached for response at Day 10 but did occur once 28 days elapsed after initial discharge. Statistical power considerations are more nuanced in the Bayesian case than in the frequentist case. This is because Bayesian inference is closer to the significance tests as originally proposed by Fisher, than to the now widely used hypothesis testing framework developed by Neyman and Pearson.[9] In consequence, the rather familiar false positive or negative (Type I and II) frequentist errors are not properly defined in the Bayesian case, and statistical power cannot be conceptualized in the same intuitive way as in the frequentist framework. This has obvious implications when performing sample size calculations, creating difficulties that are possibly the most dissuasive for the use of Bayesian statistics. However, there are theoretical developments to generalize frequentist sample size formulas, like the so-called “proto-prior” method,[10] in which the (collapsed) values of the expected differences are replaced by reasonable distributions with imputed variances (the interested reader can find more technical details on this topic in Ciarleglio et al[11,12]). Using this method, we found that for detecting a difference of about 17 percentage points in response rates, 126 and 72 patients would be required when the proportions in the control group were 55.6% and 83.3%, respectively; that is, those observed at Day 10 and 28 days after discharge. This suggests that, if we had been able to carry out the study with the initially planned sample of 120 patients (see the original publication[1]), we would have obtained significant results at Day 10 also in the Bayesian case; significance 28 days after discharge ensues from the fact that we had one more patient than required. To foster the use of Bayesian statistics, we would also like to point out that Monte-Carlo simulations can always be performed to obtain efficient estimations of the required sample size when planning somewhat complex clinical studies under the Bayesian framework that involve more than comparing 2 independent means or proportions.

As mentioned, the main limitation of these post hoc Bayesian analyses is that they were not pre-specified in the study protocol. However, we have ensured that the chosen prior distributions were neutral with regard to group allocation by inspecting them and providing their empirical layouts along with posterior distributions. The use of highest density intervals in place of traditional ETIs has been advocated for since this would avoid coverage of very low credibility values,[13] but we believe it is unlikely in our case given the symmetry of posterior distributions (Figure 1). Moreover, highest density intervals would have masked the excessive variability derived from large variance priors in the analysis of the time on supplemental oxygen, creating a false sense of stability and obscuring the sparseness of data. The current results provide an illustrative example of the robustness of Bayesian statistics against influential distant values, because the conformation of posterior distributions was modulated by the observed realizations of the random variables, narrowing in the case of clinical response and death, and widening in the case of clinical milestones (oxygen supplementation and hospital discharge), which were much sparser. Although we did not contemplate the use of Bayesian statistics from the beginning of this study (mainly because there were no previous similar studies to orientate prior selection) we encourage greater use of Bayesian statistics, because it can more naturally accommodate the inductive-deductive movement of scientific inquiry.

Conclusions

In conclusion, the present Bayesian analyses have provided multiplicative effect measures that are more familiar to clinicians and support that icatibant treatment for COVID-19 pneumonia patients seems to prompt clinical improvement and could substantially reduce mortality. Nonetheless, a larger Phase III confirmatory trial would be necessary to provide firm evidence.

Author’ contribution

Pierre Malchair and Jordi Giol participated in the concept and design of the study, developed the protocol, performed clinical evaluations of patients, participated in the interpretation of the data, supervised the study and drafted the manuscript.

Jesús Villoria and Thiago Carnaval designed and performed the statistical analyses, contributed to the interpretation of the data and drafted the manuscript.

Javier Jacob performed clinical evaluations and had substantial roles in the acquisition and interpretation of the data.

Sebastián Videla participated in the concept and design of the study, developed the protocol, contributed substantially to the analysis and interpretation of the data, supervised the study and drafted the manuscript.

All authors participated in the review and extraction of literature as well as in drafting or revising the manuscript. All authors had full access to the data and have provided their final approval of the version of the manuscript to be submitted for publication.

This collaborative research clinical trial received funding for expenses related to administrative procedures, study monitoring and the electronic data capture system from Takeda Farmacéutica España. Takeda also provided the study medication (icatibant), valued at over one million euros.

Declaration of Competing Interest

The authors do not have any competing interest to declare.

Acknowledgments

The authors give special thanks to Takeda Farmacéutica España for their support in obtaining study medication and funding for the study. We would also like to thank the BioClever 2005 S.L. (Full Service Clinical Research Organization) personnel involved in the study, and the Bellvitge University Hospital, IDIBELL and CERCA Program/Generalitat de Catalunya for institutional support.

References

1. Malchair P, Giol J, Garcia V. Three-day icatibant on top of standard care in patients with coronavirus disease 2019 pneumonia: a randomized, open-label, phase 2, proof-of-concept trial. *Clin Infect Dis*. 2023;76:1784–1792.
2. Saracci R. Interaction and synergism. *Am J Epidemiol*. 1980;112:465–466.
3. Hernan MA, Robins JM. *Why Model? Causal Inference: What If*. Boca Raton, FL: Chapman & Hall/CRC; 2020:139–148.
4. Moller S, Ahrenfeldt LJ. Estimating relative risk when observing zero events-frequentist inference and Bayesian credibility intervals. *Int J Environ Res Public Health*. 2021;18:5527.
5. Kruschke JK. *Bayes' Rule. Doing Bayesian Analysis: A Tutorial with R, JAGS and Stan* Elsevier Inc.; 2015:99–120.
6. Statistical Editors. The New England Journal of Medicine - Preparation instructions for authors—statistical reporting guidelines; Bayesian analyses. https://www.nejm.org/author-center/new-manuscripts. Accessed 7 January 2023.
7. Spiegelman D, Hertzmark E. Easy SAS calculations for risk or prevalence ratios and differences. *Am J Epidemiol*. 2005;162:199–200.
8. Bürkner P-C. brms: an R package for Bayesian multilevel models using Stan. *J Stat Softw*. 2017;80:1–28.
9. Goodman SN. p values, hypothesis tests, and likelihood: implications for epidemiology of a neglected historical debate. *Am J Epidemiol*. 1993;137:485–496 discussion 97-501.
10. Kruschke JK. *Goals, Power, and Sample Size. Doing Bayesian Analysis: A Tutorial with R, JAGS and Stan* Elsevier Inc.; 2015:359–398.
11. Ciarleglio MM, Arendt CD, Makuch RW, Peduzzi PN. Selection of the treatment effect for sample size determination in a superiority clinical trial using a hybrid classical and Bayesian procedure. *Contemp Clin Trials*. 2015;41:160–171.
12. Ciarleglio MM, Arendt CD, Peduzzi PN. Selection of the effect size for sample size determination for a continuous response in a superiority clinical trial using a hybrid classical and Bayesian procedure. *Clin Trials*. 2016;13:275–285.
13. Kruschke JK. *Bayesian Approaches to Testing a Point ("Null") Hypothesis. Doing Bayesian Analysis: A Tutorial with R, JAGS and Stan* Elsevier Inc.; 2015:335–358.