Title

A comparison of computational algorithms for the Bayesian analysis of clinical trials.

Journal

Clinical Trials, 21(6)

Authors

Chen, Ziming
Berger, Jeffrey
Castellucci, Lana
et al.

Publication Date

2024-12-01

DOI

10.1177/17407745241247334

Peer reviewed

HHS Public Access

Author manuscript

Clin Trials. Author manuscript; available in PMC 2024 December 01.

Published in final edited form as:

Clin Trials. 2024 December ; 21(6): 689–700. doi:10.1177/17407745241247334.

A comparison of computational algorithms for the Bayesian analysis of clinical trials

Ziming Chen[1], **Jeffrey S Berger**[2], **Lana A Castellucci**[3], **Michael Farkouh**[4], **Ewan C Goligher**[4], **Erinn M Hade**[2], **Beverley J Hunt**[5], **Lucy Z Kornblith**[6], **Patrick R Lawler**[4], **Eric S Leifer**[7], **Elizabeth Lorenzi**[8], **Matthew D Neal**[9], **Ryan Zarychanski**[10], **Anna Heath**[1,11,12]

[1]The Hospital for Sick Children, Toronto, ON, Canada

[2]New York University, New York, NY, USA

[3]Department of Medicine, Ottawa Hospital Research Institute, University of Ottawa, Ottawa, ON, Canada

[4]University Health Network, Toronto, ON, Canada

[5]King's College London, London, UK

[6]University of California, San Francisco, San Francisco, CA, USA

[7]National Heart, Lung, and Blood Institute, Bethesda, MD, USA

[8]Berry Consultants, Austin, TX, USA

[9]University of Pittsburgh Medical Center, Pittsburgh, PA, USA

[10]University of Manitoba, Winnipeg, MB, Canada

[11]Division of Biostatistics, Dalla Lana School of Public Health, University of Toronto, Toronto, ON, Canada

[12]Department of Statistical Science, University College London, London, UK

Abstract

Background: Clinical trials are increasingly using Bayesian methods for their design and analysis. Inference in Bayesian trials typically uses simulation-based approaches such as Markov

Corresponding author: Ziming Chen, The Hospital for Sick Children, Toronto, ON, Canada. jocelynczm@gmail.com.

Declaration of conflicting interests

The author(s) declared the following potential conflicts of interest with respect to the research, authorship, and/or publication of this article: JSB is on the advisory board Advisory for Amgen and Janssen outside the submitted work. LAC holds a Heart and Stroke Foundation of Canada National New Investigator Award, and a Tier 2 research Chair in Thrombosis and Anticoagulation Safety from the University of Ottawa. LZK is a consultant for Cerus, Gamma Diagnostics and the University of Maryland. MDN is served as the Co-Chair of the ACTIV4a trial.

Publication committee members

The co-authors of this manuscript are as follows: ZC, JSB, MD; LAC, MD; MEF, MD; ECG, MD, PhD; EMH, PhD; BJH, MD; LZK, MD; PRL, MD; ESL, PhD; EL, PhD; MDN, MD; RZ, MD; and AH, PhD. ZC analysed the data, drafted the manuscript and created the tables and figures. AH supervised the analysis and drafted the manuscript. JSB, LAC, MEF, ECG, EMH, BJH, LZK, PRL, ESL, EL, MDN, and RZ contributed to data collection and edited the manuscript. All authors approved the final manuscript for publication.

Supplemental material

Supplemental material for this article is available online.

Chain Monte Carlo methods. Markov Chain Monte Carlo has high computational cost and can be complex to implement. The Integrated Nested Laplace Approximations algorithm provides approximate Bayesian inference without the need for computationally complex simulations, making it more efficient than Markov Chain Monte Carlo. The practical properties of Integrated Nested Laplace Approximations compared to Markov Chain Monte Carlo have not been considered for clinical trials. Using data from a published clinical trial, we aim to investigate whether Integrated Nested Laplace Approximations is a feasible and accurate alternative to Markov Chain Monte Carlo and provide practical guidance for trialists interested in Bayesian trial design.

Methods: Data from an international Bayesian multi-platform adaptive trial that compared therapeutic-dose anticoagulation with heparin to usual care in non-critically ill patients hospitalized for COVID-19 were used to fit Bayesian hierarchical generalized mixed models. Integrated Nested Laplace Approximations was compared to two Markov Chain Monte Carlo algorithms, implemented in the software JAGS and stan, using packages available in the statistical software R. Seven outcomes were analysed: organ-support free days (an ordinal outcome), five binary outcomes related to survival and length of hospital stay, and a time-to-event outcome. The posterior distributions for the treatment and sex effects and the variances for the hierarchical effects of age, site and time period were obtained. We summarized these posteriors by calculating the mean, standard deviations and the 95% equitailed credible intervals and presenting the results graphically. The computation time for each algorithm was recorded.

Results: The average overlap of the 95% credible interval for the treatment and sex effects estimated using Integrated Nested Laplace Approximations was 96% and 97.6% compared with stan, respectively. The graphical posterior densities for these effects overlapped for all three algorithms. The posterior mean for the variance of the hierarchical effects of age, site and time estimated using Integrated Nested Laplace Approximations are within the 95% credible interval estimated using Markov Chain Monte Carlo but the average overlap of the credible interval is lower, 77%, 85.6% and 91.3%, respectively, for Integrated Nested Laplace Approximations compared to stan. Integrated Nested Laplace Approximations and stan were easily implemented in clear, well-established packages in R, while JAGS required the direct specification of the model. Integrated Nested Laplace Approximations was between 85 and 269 times faster than stan and 26 and 1852 times faster than JAGS.

Conclusion: Integrated Nested Laplace Approximations could reduce the computational complexity of Bayesian analysis in clinical trials as it is easy to implement in R, substantially faster than Markov Chain Monte Carlo methods implemented in JAGS and stan, and provides near identical approximations to the posterior distributions for the treatment effect. Integrated Nested Laplace Approximations was less accurate when estimating the posterior distribution for the variance of hierarchical effects, particularly for the proportional odds model, and future work should determine if the Integrated Nested Laplace Approximations algorithm can be adjusted to improve this estimation.

Keywords

Bayesian clinical trial analysis; proportional odds model; survival analysis; logistic regression; Integrated Nested Laplace Approximations; Markov chain Monte Carlo; JAGS; stan

Introduction

Bayesian methods offer a framework for statistical inference[1] and are increasingly employed in clinical trials.[2–6] Bayesian methods have been used for adaptive trials,[7] including for master protocol designs,[8,9] adaptive randomization,[10] sample size adjustment,[11] and interim monitoring and stopping.[12,13] Bayesian methods also allow historical and external information, that is, from previous or concurrent trials, to be incorporated into trial analysis.[14] They can also easily integrate different levels of hierarchy to account for similarities between patients, that is, if patients are recruited at different hospitals within different countries. Finally, Bayesian analyses calculate the probability of specific events,[15] which can provide an intuitive interpretation of results.[16–20]

However, Bayesian methods are often computationally complex since the quantity of interest, the posterior distribution for the treatment effect, cannot be directly computed.[21] This requires approximations, usually based on simulation approaches known as Monte Carlo Markov Chain (MCMC) methods.[22] To exploit the theory that MCMC methods are asymptotically exact, they require extensive simulations, especially when the proposed analyses are complicated.[23] Furthermore, they also often require knowledge of specialized software.[24–26]

A highly efficient alternative to MCMC, which does not use simulation, is the Integrated Nested Laplace Approximations (INLA) algorithm.[27] The accuracy of INLA has been explored using simulation studies, demonstrating that INLA is computationally efficient and accurate for generalized linear mixed models as well as b-spline nonparametric regression.[28,29] INLA has also been shown to be valid in complex spatial models.[30,31] However, there has been no comparison of INLA and MCMC methods using data from a clinical trial.

Thus, we aimed to determine whether INLA is an appropriate alternative to MCMC methods for clinical trials by analysing data from a Bayesian multi-platform adaptive trial that investigated if treatment with therapeutic-dose heparin improves a composite outcome of organ failure and death among hospitalized moderately ill COVID-19 patients.[32] We compared INLA and two MCMC algorithms, a Gibbs sampler[33] and Hamiltonian Monte Carlo,[34] based on their accuracy, computational efficiency and ease of implementation. We evaluated these aspects by reproducing the analysis from the trial using three types of Bayesian hierarchical generalized models; ordinal proportional odds models, binary logistic regression models, and Cox proportional hazard models. We adjusted these models for key covariates including site and age. This article introduces the data from the trial and the Bayesian methods used for analysis. It then presents the results of our comparison and concludes that INLA offers an efficient, accurate alternative to MCMC in clinical trials.

Methods

The ATTACC and ACTIV-4a trials

We analysed data from an international, adaptive, multi-platform, randomized, controlled trial that compared the effect of therapeutic-dose anticoagulation, either unfractionated

or low-molecular-weight heparin, to usual-care pharmacologic thromboprophylaxis in patients hospitalized for COVID-19.[32] Three clinical trial platforms contributed data to the original publication,[32] (Antithrombotic Therapy to Ameliorate Complications of COVID-19 (ATTACC), Accelerating COVID-19 Therapeutic Interventions and Vaccines 4 ACUTE (ACTIV-4A) and Randomized, Embedded, Multi-factorial, Adaptive Platform Trial for Community-Acquired Pneumonia (REMAP-CAP)). Our study restricts to data from the ATTACC and ACTIV-4A platforms as data from REMAP-CAP were not available.

Adult patients were enrolled in the trial if they were hospitalized with COVID-19 and not receiving intensive-care level organ support at the time of enrolment. Our initial data set included 2146 patients. We excluded participants if they withdrew from screening and randomization, or if their study record was missing values for age, site, date of randomization, sex, or treatment, leaving 1914 records for analysis. A detailed description of the trial design and data collection process including the inclusion/exclusion criteria is provided in the primary trial publication.[32]

Outcomes in the ATTACC/ACTIV-4A trials

We analysed seven outcomes using definitions and models from the trial protocol. The primary outcome was organ support-free days, a composite ordinal outcome that includes in-hospital death and the number of days free of cardiovascular or respiratory organ support. Ordinal values ranged from between −1 and 22. Patients who died during their index hospitalization or before day 21 were assigned the score of −1. Patients who survived and did not receive any organ support during hospitalization were assigned the outcome 22. Otherwise, organ support-free days is calculated as 21 minus the number of days on which the patient required organ support up to day 21 postrandomization. Figure 1 displays the organ support-free days distribution.

We also considered six secondary outcomes, five binary outcomes, and one time-to-event outcome. The binary outcomes, defined in the trial protocol, were survival to hospital discharge, survival without receipt of organ support, survival without receipt of invasive mechanical ventilation, survival without mechanical respiratory support, and whether the patient had a major thrombotic event or death.[32] The time-to-event outcome was the length of hospital stay, defined as the difference between the date the patients were discharged from the hospital and the date that the patients were randomized into the trial. Patients who died during hospitalization were censored at the date they died. Figure 2 shows the distribution of the length of hospital stay.

Bayesian modelling

We used generalized linear models with different error functions for each outcome type. For organ support-free days, we used a cumulative proportional odds model. For the binary outcomes, we used binary logistic regression and the time-to-event outcome was modelled using a Cox proportional hazards model. We adjusted all models for age, $\leqslant 39$, 40 – 49, 50 – 59, 60 – 69, 70 – 79 and $\geqslant 80$, sex, trial site, and enrolment period, in 2-week intervals.[32] Bayesian hierarchical models were used for the age, site, and enrolment period effects.[35]

Primary outcome model: organ support free days.—Organ support-free days was modelled using a cumulative proportional odds model.[36] For $y = -1, \ldots, 22$ the probability that organ support-free days for patients $i = 1, \ldots, 1914$ is less than or equal to y, $\pi_i^y = P(Y \leqslant y)$, is

$$log\left(\frac{\pi_i^y}{1-\pi_i^y}\right) = \alpha_y + \beta_1 sex_i + \beta_2 treatment_i + U_{site=site_i} + V_{age=age_i} + W_{time=time_i}, \quad (1)$$

where α_y relates to the cumulative probability of observing each category for organ support-free days, β_1 is the effect of sex and β_2 is the treatment effect of the therapeutic anticoagulation compared to usual care. We assume that $U_{site} \sim N(0, \sigma^2_{site})$ for $site = 1, \ldots, 83$, $V_{age} \sim N(0, \sigma^2_{age})$ for $age = 1, \ldots, 6$ and $W_{time} \sim N(0, \sigma^2_{time})$ for $time = 1, \ldots, 21$.

Organ-support or ventilator-free days are commonly used in critical care trials[37–39] as they combine mortality and the time survivors receive organ support or ventilation. This outcome provides higher statistical power to detect a treatment effect compared to mortality alone.[40] While many methods have been proposed to analyse this outcome, a standard method is not available.[41] Thus, used the proportional odds model to mimic the trial analysis.[32]

Secondary outcomes.—The binary secondary outcomes are analysed with logistic regression such that the probability of experiencing the event of interest for patient i, $i = 1, \ldots, 1914$, $\pi_i = P(Z_i = 1)$, is

$$log\left(\frac{\pi_i}{1-\pi_i}\right) = \delta + \gamma_1 sex_i + \gamma_2 treatment_i + U_{site=site_i} + V_{age=age_i} + W_{time=time_i} \quad (2)$$

where δ is related to the baseline probability of the event, γ_1 and γ_2 are the log-odds ratios of being female and receiving therapeutic anticoagulation, respectively. The distributions for U_{site}, V_{age} and W_{time} are normally distributed.

Finally, the length of hospital stay was modelled using a Cox proportional hazard model.[42] Let $\lambda_i(t)$ be the hazard of leaving hospital for patient $i = 1, \ldots, 1914$. The Cox proportional hazard model can be written as

$$\lambda_i(t) = \lambda_0(t) e^{\theta_1 sex_i + \theta_2 treatment_i + U_{site=site_i} + V_{age=age_i} + W_{time=time_i}} \quad (3)$$

where $\lambda_0(t)$ is the baseline hazard rate, θ_1 and θ_2 are the log hazard ratios for the sex and treatment effects, and U_{site}, V_{age} and W_{time} are normally distributed.

Defining the prior distributions.—Bayesian inference represents parameter uncertainty using probability distributions.[43] To achieve this, the uncertainty in the parameters before observing the data must be encoded in a *prior distribution*,[44] which is then updated by the data to determine the *posterior distribution*. The posterior distribution provides a complete summary of all the parameter information and is proportional to the product of the likelihood and the prior distribution. Once the posterior distribution is available, it can be summarized by, for example, calculating the posterior mean and 95% credible intervals (CI).

As in the original publication, we specified minimally informative prior distributions for each parameter to minimize their impact on the results.[44] We used normal priors with mean zero and precision 0.1 for the sex and treatment effects and the intercept terms for the proportional odds model and logistic regression models, except for the proportional odds model in INLA. In INLA, these intercept terms are reparameterized allowing a Dirichlet prior distribution with scale parameter 3 to be used. For the variance of the hierarchical effects, σ^2_{site}, σ^2_{age} and σ^2_{time}, we used half-t prior distributions with 3 degrees of freedom and a scale parameter of 2.5.[45]

Practical Bayesian inference

Except in a small number of cases, which are not applicable for the proposed analysis, it is not possible to obtain the posterior distribution analytically,[46,47] and it must be must be approximated. In this article, we compare two MCMC algorithms and the INLA algorithm for obtaining these approximations.

Monte Carlo Markov chain methods.—MCMC methods are a class of algorithms that simulate from probability distributions in settings where samples cannot be easily drawn.[48] For practical Bayesian analysis two MCMC methods are most commonly used,[49] the Metropolis-Hastings algorithm combined with Gibbs sampling (Metropolis-Hastings within Gibbs) and Hamiltonian Monte Carlo. To facilitate the use of MCMC methods, a range of different software has been developed.[24–26,50,51]

Gibbs sampling simulates from multivariate distributions by simulating each parameter from a univariate distribution conditional on the values of the other parameters.[33] In Bayesian analysis, the posterior distribution is not known so Gibbs sampling simulates from a distribution that approximates the posterior distribution, known as a proposal distribution. The ‘proposed’ parameter values are then either rejected or accepted based on the relative ‘likelihood’[†] of the current parameter values and the proposed parameter values. This is the Metropolis-Hastings step, which draws simulations from an unknown distribution.[53] The Metropolis-Hastings within Gibbs algorithm is fast when (a) the conditional distributions are easy to sample from and (b) the proposal distribution is close to the true posterior distribution. Thus, Metropolis-Hastings within Gibbs can encounter significant challenges in complex, hierarchical models.[54]

[†]The word likelihood is not used in the true statistical sense but the exact form the rejection step is outside the scope of this paper. For more information, please see.[52]

Hamiltonian Monte Carlo is more efficient than Metropolis-Hastings within Gibbs for complex models.[34] Hamiltonian Monte Carlo also uses the Metropolis-Hastings algorithm to reject proposed parameter simulations but it uses derivatives to generate 'better' parameter proposals for the posterior distribution.[55] Essentially, Hamiltonian Monte Carlo identifies parameter values that are 'far away' from the current parameter value, which reduces the correlation between the successive simulations.[56] This requires fewer simulations to approximate the posterior distribution but by requiring derivatives, Hamiltonian Monte Carlo can be slower for simple models where the proposal distribution for the Gibbs sampler provides a good approximation of the posterior distribution.[57]

Both these MCMC methods are asymptotically exact, that is, they simulate from the true posterior distribution with an infinite number of samples.[58] Thus, to approach an infinite number of simulations, MCMC methods require *a large* number of simulations to be accurate.[59]

Integrated nested Laplace approximation.

INLA is a fast alternative to MCMC that approximates the posterior distributions for a class of models known as *latent Gaussian models*.[47] All common outcome models in clinical trials, including generalized linear models and survival models, can be specified as latent Gaussian models, making INLA a suitable tool for analysis in clinical trials.[60] Latent Gaussian models require a hierarchical structure with three layers. The highest level of the hierarchy should include a small number of parameters (ideally less than 6) but has very few restrictions on the distribution. The middle layer must then follow a Gaussian distribution conditional on the top layer parameters. Crucially, the parameters in this layer must also have a specific correlation structure that allows for efficient computation. Finally, the data are modelled in the bottom layer conditional on the model parameters and can have a wide range of error distributions.[27] INLA uses a Laplace approximation to approximate the posterior distribution for the top layer parameters and then approximates the middle layer parameters at a grid of values from the posterior of the top layer parameters using another *nested* Laplace approximation. This process approximates the *marginal* distributions for all the parameters in the statistical model.[27] In contrast to MCMC, INLA does not simulate from the posterior distribution. INLA simply a numerical approximation for the posterior, which is compututationally efficient but will always approximate the true distribution.

Implementing Bayesian analysis in practice.—We have implemented these algorithms using three popular Bayesian software tools. We selected these tools as they integrate with R.[61] There are two widely used softwares for implementing Metropolis-Hastings within Gibbs, BUGS[24] and JAGS.[25] We chose JAGS, implemented through the package R2jags,[62] as it is actively maintained. R2jags is only an interface to run JAGS, meaning that the software JAGS was downloaded and installed separately. We implemented Hamiltonian Monte Carlo using the stan software[26] through the brms[63] package. stan was directly installed in R through the Rstan package. Finally, we implemented INLA using the INLA[64] package, which does not require additional installations. However, the INLA package is not available on the CRAN repository and must be installed following the instructions at https://www.r-inla.org/download-install.

To fit a Bayesian model in `R2jags`, the user must use the JAGS syntax to specify the likelihood for the data and the prior distributions. We also found that we frequently needed to specify a reasonable set of starting values for the Metropolis-Hastings within Gibbs algorithm. Survival models were also complex to fit in JAGS as the data had to be restructured to fit the model.[65] In contrast, fitting Bayesian models using `brms` and `INLA` packages uses similar syntax to fitting a generalized linear model in the widely used `glm` function in `R`. Thus, the user can fit different types of models by specifying the 'family' of the model. This makes the syntax for `brms` and `INLA` very similar across our seven examples as the same covariates are used to adjust each models. Thus, compared to `R2jags`, `brms` and`INLA` were easier to implement. To support implementation of these methods, we provide our code in the Supplementary Material.

One key issue with the `INLA` package is that a maximum of 10 ordinal categories are permitted for the proportional odds model. Therefore, we regrouped the 24 organ-support free days categories into 10 categories in this comparison. Theoretically, this should not affect the value of the treatment effect but can affect the precision.[66] To minimize the variation in the estimated treatment effect, we created separate categories for −1, in-hospital death, and 22, surviving to hospital discharge without receipt of organ support, and then combined the remaining 21 categories to create groups that were as even possible in terms of their number of patients.

Finally, to compare INLA to the MCMC algorithms, we simulated from the marginals estimated by INLA using the inla.posterior.sample function for the second layer parameters, for example, the treatment effect, and the inla.hyperpar.sample function for the top layer parameters, for example, the variances of the hierarchical effects.

Comparing the algorithms

To assess the alignment and feasibility of INLA compared to two MCMC methods, we obtained the posterior distributions for model parameters using our three algorithms for the seven outcomes of interest. For the MCMC methods, we undertook 12,000 simulations with four chains and a burn-in of 2000, thinned to a simulation size of 2000. We computed the posterior mean, standard deviation, and 95% equitailed CI for the treatment and sex effects and the hierarchical model variances. We also computed the deciles of the posterior distributions (see Supplementary Material). We plotted the posterior density of the treatment and sex effects and the hierarchical model variances and computed the posterior probability of a 'positive' effect of treatment, defined separately for each outcome. Finally, we recorded computation time. All analyses were performed in R version 4.1.3. As all methods are approximations, it is unknown which method is 'accurate' when differences occur. However, we consider stan to be the 'true' result due to its ability to model complicated distributions and its theoretical accuracy with many simulations.

Results

Data summary

The baseline characteristics are summarized in the Supplementary Material. Approximately 60% of patients were from the ATTACC platform and 40% from ACTIV-4A with 53.4% receiving therapeutic-dose anticoagulation, due to the response adaptive randomization. 9.5% of patients who received therapeutic-dose anticoagulation died compared to 9.3% who received usual care. The proportion of males was 60.1%, as a greater proportion of men were hospitalized for COVID-19 globally, and the mean age was around 61.

Model results

Primary outcome. Table 1 shows the posterior mean, standard deviation, and 95% equitailed CI for the key parameters in the model for organ support-free days ($\beta_1, \beta_2, U_{site}, V_{age}, W_{time}$). For the sex and treatment effects, the estimates of the posterior mean are similar for all three algorithms, with the JAGS estimate slightly higher than the stan and INLA estimates, with the same standard deviation. The 95% CI from stan and INLA are almost identical, with a wider interval for JAGS. Thus, INLA accurately estimated the posterior distribution for the fixed effects, while JAGS struggled with the proportional odds model. However, the estimate of the posterior mean for the variance of the effects for age, site and time do not match between INLA and the MCMC methods. This is confirmed in Figure 3 where the density curves of the posterior distributions overlap for the effects of sex and treatment but not for the variance of the age, site and time effects.

Secondary outcomes.—Table 1 shows the posterior mean, standard deviation, and 95% CI of the parameters of interest for the secondary outcomes. The posterior means and the 95% CI estimated using all three algorithms almost completely overlap for all model parameters. The average overlap of the 95% CI for the treatment and sex effects estimated using INLA was 96% and 97.6% compared with stan and JAGS, respectively. The posterior mean for the variance of the age, site, and time effects, estimated using INLA, are within the 95% CI estimated using MCMC. However, the average overlap of the 95% CIs estimated using stan is 77%, 85.6%, and 91.3%, respectively. Figure 3 displays the density plots for the five parameters for two secondary outcomes (other outcomes are in the Supplementary Material). These figures confirm the numerical results. Crucially, differences also exist between stan and JAGS as well as stan and INLA.

Decision-making using JAGS, stan, and INLA.—Table 2 reports the posterior probability that the treatment provides a positive effect, adjusted for whether each outcome is positive or negative. There are minimal differences between these probabilities, with a largest difference of 6% for the length of hospital stay. The original trial stopped for treatment superiority when the posterior probability of a positive treatment effect for the primary outcome exceeded 99%. Thus, with this decision criteria, all three methods would provide the same trial conclusion.

Runtime.—Table 3 shows the computation time of the three algorithms. INLA required between 4.2 and 31.28 s, which is between 85 and 269 times faster than stan and 26 and

1852 times faster than JAGS. The computational efficiency of INLA is more pronounced for more complex models (proportional odds and survival models) where JAGS and stan required minutes or hours and INLA required less than a minute. As expected, stan is more efficient than JAGS for complex models while JAGS is more efficient for the simpler binary logistic regression.

Discussion

We evaluated the efficiency, feasibility and accuracy of INLA and MCMC methods for Bayesian clinical trial analysis. First, INLA substantially reduces computational time compared to MCMC. Second, the R packages INLA and brms package facilitate the implementation of INLA and Hamiltonian Monte Carlo, while JAGS is more complex. Furthermore, the posterior distributions for the treatment and sex effects were similar across all three algorithms. Thus, INLA can accurately estimate the treatment effect, which is the main interest in clinical trials, while also being computationally efficient and simple to implement in R.

There are limitations to applying INLA in practice. First, INLA could only fit up to 10 categories for the ordinal outcome using the proportional odds model, which could lead to reduced precision in the posterior distribution.[66] However, we did compare the 10-category INLA model with the 24-category model fit using MCMC and found minimal differences. We found that including additional categories did lead to an increase in the computational time for all algorithms, more pronounced for MCMC. Thus, if INLA could be implemented with the 24-category outcome, it would likely still represent a significant computational saving.

Second, INLA was unable to accurately approximate the posterior distributions of the hierarchical variances for the proportional odds model. This is because the INLA algorithm is designed for less than six parameters in the top layer of the model. However, in the proportional model, INLA treats each of the intercepts as a top-layer parameter, alongside the hierarchical variances, leading to 13 parameters. We suspect that we could improve the estimation of the hierarchical variances by reducing the number of categories for the ordinal outcome. However, this would affect the precision of the sex and treatment effects, and it is uncertain how the organ support-free days outcome should be recategorized.

We also observed that the posterior distribution curves for the hierarchical variance effects using INLA did not completely agree with Hamiltonian Monte Carlo. However, we also observed differences between Hamiltonian Monte Carlo and Metropolis-Hastings within Gibbs. Thus, it is likely that INLA is as accurate as the MCMC methods with the number of simulations we used. A key difference is that the MCMC methods could be improved by increasing the simulations, while the INLA approximation is fixed. This would, however, lead to increases in computation time. Note that we used thinning in the MCMC sampling, which is commonly used but has also been criticized as it can increase computational time without improving accuracy.[67] Thus, the MCMC methods could potentially be improved without adding significantly to the computational cost.

A limitation of this study is that we only used data from one multi-platform trial to compare INLA to the MCMC methods, meaning that the conclusions may not be transferable. To mitigate this concern, we tested the performance of INLA across a range of outcomes, analysed using commonly used models in clinical trials. We also aimed to challenge INLA using data from a complex Bayesian adaptive trial, which adjusted for multiple covariates. Nevertheless, data from other trials could also have been considered and the range of outcome types extended. INLA could also be compared to MCMC methods using observational data to offer more confidence in the method, although we expect these results to hold for different origins of data. Nonetheless, comparisons using data from other trials and data sets would provide practitioners with further confidence in INLA.

Another limitation is that, to replicate exact analysis in the original clinical trial, we did not perform any sensitivity analyses. In practice, in addition to modelling the outcomes, the assumptions of the analysis should also be investigated and suitable sensitivity analyses should be run. This includes understanding whether the model assumptions are valid and the impact of the choice of priors. In our setting, the amount of data should overwhelm the priors meaning they have limited impact on the results but they should be investigated. Note, however, that running these sensitivity analyses requires that alternative posterior distributions are obtained. Thus, these additional analyses would add computational complexity, making the use of INLA potentially more efficient.

Finally, this comparison only tested three Bayesian software packages and restricted to software that integrates within R. Alternative Bayesian software, such as pyMC[50] in Python or LaplaceDemon[51] in R are available. However, they still rely on MCMC methods and are, thus, unlikely to outperform INLA in terms of computational time.

Conclusion

This study assessed the accuracy, computational time, and feasibility of INLA, compared to MCMC methods, in the analysis of data from an international, Bayesian adaptive platform clinical trial. We demonstrated that INLA provides accurate estimates of the treatment effect, identical decisions to MCMC methods, efficient computation, and feasible implementation. Thus, INLA is an important tool to facilitate the use of Bayesian methods in clinical trials.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: The ATTACC platform was supported by grants from the Canadian Institutes of Health Research, LifeArc, Thistledown Foundation, Research Manitoba, CancerCare Manitoba Foundation, Victoria General Hospital Foundation, Ontario Ministry of Health, and the Peter Munk Cardiac Centre. The ACTIV-4a platform was sponsored by the National Heart, Lung, and Blood Institute, National Institutes of Health, Bethesda and administered through OTA-20-011. The research was, in part, funded by the National Institutes of Health (NIH) Agreement 1OT2HL156812 through the National Heart, Lung, and Blood Institute (NHLBI) CONNECTS programme. The views and conclusions contained in this document are those of the authors and should not be interpreted as representing the official policies, either expressed or implied, of the NIH. JSB is supported by NIH

grant funding and AHA grant funding. LAC is a member of the Canadian Venous Thromboembolism Research Network (CanVECTOR); the Network received grant funding from the Canadian Institutes of Health Research (Funding Reference: CDT-142654). EMH is supported by NIH grant funding. RZ reports grants from Canadian Institutes of Health Research, LifeArc Foundation, Research Manitoba, Peter Munk Cardiac Centre, Thistledown Foundation, and National Institutes of Health. AH is supported by a Canada Research Chair in Statistical Trial Design and funded by the Discovery Grant Programme of the Natural Sciences and Engineering Research Council of Canada (RGPIN-2021-03366). The funding agreement ensured the authors' independence in designing the study, interpreting the data, writing, and publishing the report.

References

1. Van de Schoot R, Kaplan D, Denissen J, et al. A gentle introduction to Bayesian analysis: applications to developmental research. Child Dev 2014; 85(3): 842–860. [PubMed: 24116396]
2. Berry DA. Bayesian clinical trials. Nat Rev Drug Discov 2006; 5(1): 27–36. [PubMed: 16485344]
3. Crider KS, Devine O, Hao L, et al. Population red blood cell folate concentrations for prevention of neural tube defects: Bayesian model. BMJ 2014; 349: g4554. [PubMed: 25073783]
4. Muss HB, Berry DA, Cirrincione CT, et al. Adjuvant chemotherapy in older women with early-stage breast cancer. N Engl J Med 2009; 360(20): 2055–2065. [PubMed: 19439741]
5. Brophy JM. Bayesian interpretation of the EXCEL trial and other randomized clinical trials of left main coronary artery revascularization. JAMA Intern Med 2020; 180(7): 986–992. [PubMed: 32478838]
6. Montazerhodjat V, Chaudhuri SE, Sargent DJ, et al. Use of Bayesian decision analysis to minimize harm in patient-centered randomized clinical trials in oncology. JAMA Oncol 2017; 3(9): e170123. [PubMed: 28418507]
7. Berry SM, Carlin BP, Lee JJ, et al. Bayesian adaptive methods for clinical trials. Boca Raton, FL: CRC Press, 2010.
8. Woodcock J and LaVange LM. Master protocols to study multiple therapies, multiple diseases, or both. N Engl J Med 2017; 377(1): 62–70. [PubMed: 28679092]
9. Park JJ, Detry MA, Murthy S, et al. How to use and interpret the results of a platform trial: users' guide to the medical literature. JAMA 2022; 327(1): 67–74. [PubMed: 34982138]
10. Barry WT, Perou CM, Marcom PK, et al. The use of Bayesian hierarchical models for adaptive randomization in biomarker-driven phase II studies. J Biopharm Stat 2015; 25(1): 66–88. [PubMed: 24836519]
11. Vasishth S, Yadav H, Schad DJ, et al. Sample size determination for Bayesian hierarchical models commonly used in psycholinguistics. Comput Brain Behav 2022: 1–25.
12. Liu J, Wick J, Martin R, et al. Two-stage Bayesian hierarchical modeling for blinded and unblinded safety monitoring in randomized clinical trials. BMC Med Res Methodol 2020; 20(1): 1–17.
13. Broglio K, Meurer WJ, Durkalski V, et al. Comparison of Bayesian vs frequentist adaptive trial design in the Stroke Hyperglycemia Insulin Network Effort trial. JAMA Netw Open 2022; 5(5): e2211616. [PubMed: 35544137]
14. Van Rosmalen J, Dejardin D, Van Norden Y, et al. Including historical data in the analysis of clinical trials: is it worth the effort. Stat Methods Med Res 2018; 27(10): 3167–3182. [PubMed: 28322129]
15. Van de Schoot R, Depaoli S, King R, et al. Bayesian statistics and modelling. Nat Rev Methods Primers 2021; 1(1): 1–26.
16. Jack Lee J and Chu CT. Bayesian clinical trials in action. Stat Med 2012; 31(25): 2955–2972. [PubMed: 22711340]
17. Kadane JB. Bayesian methods and ethics in a clinical trial design. Hoboken, NJ: John Wiley & Sons, 2011.
18. Goligher EC, Tomlinson G, Hajage D, et al. Extracorporeal membrane oxygenation for severe acute respiratory distress syndrome and posterior probability of mortality benefit in a post hoc Bayesian analysis of a randomized clinical trial. JAMA 2018; 320(21): 2251–2259. [PubMed: 30347031]
19. Reefhuis J, Devine O, Friedman JM, et al. Specific SSRIs and birth defects: Bayesian analysis to interpret new data in the context of previous reports. BMJ 2015; 351: h3190. [PubMed: 26156519]

20. Albuquerque AM, Tramujas L, Sewanan LR, et al. Mortality rates among hospitalized patients with COVID-19 infection treated with tocilizumab and corticosteroids: a Bayesian reanalysis of a previous meta-analysis. JAMA Netw Open 2022; 5(2): e220548. [PubMed: 35226077]
21. Gó mez-Rubio V. Bayesian inference with INLA Boca Raton, FL: CRC Press, 2020.
22. Gelfand AE and Smith AF. Sampling-based approaches to calculating marginal densities. J Am Stat Assoc 1990; 85(410): 398–409.
23. Belloni A and Chernozhukov V. On the computational complexity of MCMC-based estimators in large samples. Ann Stat 2009; 37(4): 2011–2055.
24. Lunn D, Jackson C, Best N, et al. The BUGS book: a practical introduction to Bayesian analysis. London: Chapman Hall, 2013.
25. Plummer M JAGS: a program for analysis of Bayesian graphical models using Gibbs sampling. In: Proceedings of the 3rd international workshop on distributed statistical computing, Vienna, 20–22 March 2003, pp. 1–10.
26. Carpenter B, Gelman A, Hoffman MD, et al. Stan: a probabilistic programming language. J Stat Softw 2017; 76: 1. [PubMed: 36568334]
27. Rue H, Martino S and Chopin N. Approximate Bayesian inference for latent Gaussian models by using integrated nested Laplace approximations. J R Stat Soc Ser B Methodol 2009; 71(2): 319–392.
28. Held L, Schrödle B and Rue H. Posterior and cross-validatory predictive checks: a comparison of MCMC and INLA. In: Kneib T and Tutz G (eds) Statistical modelling and regression structures. New York: Springer, 2010, pp. 91–110.
29. Fong Y, Rue H and Wakefield J. Bayesian inference for generalized linear mixed models. Biostatistics 2010; 11(3): 397–412. [PubMed: 19966070]
30. Taylor BM and Diggle PJ. INLA or MCMC? A tutorial and comparative evaluation for spatial prediction in log-Gaussian Cox processes. J Stat Comput Simul 2014; 84(10): 2266–2284.
31. Carroll R, Lawson AB, Faes C, et al. Comparing INLA and OpenBUGS for hierarchical Poisson modeling in disease mapping. Spat Spatiotemporal Epidemiol 2015; 14–15: 45–54.
32. ATTACC Investigators, ACTIV-4a Investigators, REMAP-CAP Investigators, et al. Therapeutic anticoagulation with heparin in noncritically ill patients with Covid-19. N Engl J Med 2021; 385(9): 790–802. [PubMed: 34351721]
33. Gelfand AE. Gibbs sampling. J Am Statl Assoc 2000; 95(452): 1300–1304.
34. Girolami M and Calderhead B. Riemann manifold Langevin and Hamiltonian Monte Carlo methods. J R Stat Soc Ser B Methodol 2011; 73(2): 123–214.
35. Albert J and Hu J. Probability and Bayesian modeling. Boca Raton, FL: CRC Press, 2019.
36. Harrell FE. Ordinal logistic regression. In: Harrell FE(ed.) Regression modeling strategies. New York: Springer, 2015, pp. 311–325.
37. National Heart, Lung, and Blood Institute Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network, Matthay MA, Brower RG, et al. Randomized, placebo-controlled clinical trial of an aerosolized β2-agonist for treatment of acute lung injury. Am J Respir Crit Care Med 2011; 184(5): 561–568. [PubMed: 21562125]
38. Heart National, Lung, and Blood Institute Acute Respiratory Distress Syndrome (ARDS) Clinical Trials Network, Rice TW, Wheeler AP, et al. . Initial trophic vs full enteral feeding in patients with acute lung injury: the EDEN randomized trial. JAMA: the Journal of the American Medical Association 2012; 307(8): 795–803. [PubMed: 22307571]
39. Rice TW, Wheeler AP, Thompson BT, et al. Enteral omega-3 fatty acid, γ-linolenic acid, and antioxidant supplementation in acute lung injury. JAMA 2011; 306(14): 1574–1581. [PubMed: 21976613]
40. Schoenfeld DAand Bernard GRand ARDS Network. Statistical evaluation of ventilator-free days as an efficacy measure in clinical trials of treatments for acute respiratory distress syndrome. Crit Care Med 2002; 30(8): 1772–1777. [PubMed: 12163791]
41. Yehya N, Harhay MO, Curley MA, et al. Reappraisal of ventilator-free days in critical care research. Am J Respir Crit Care Med 2019; 200(7): 828–836. [PubMed: 31034248]
42. Cox DR. Regression models and life-tables. J R Stat Soc Ser B Methodol 1972; 34(2): 187–202.

43. Lee PM. Bayesian statistics London: Oxford University Press, 1989.
44. Quintana M, Viele K and Lewis RJ. Bayesian analysis: using prior information to interpret the results of clinical trials. JAMA 2017; 318(16): 1605–1606. [PubMed: 29067406]
45. Gelman A Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). Bayesian Anal 2006; 1(3): 515–534.
46. Qian SS, Stow CA and Borsuk ME. On Monte Carlo methods for Bayesian inference. Ecol Modell 2003; 159(2–3): 269–277.
47. Rue H, Riebler A, Sørbye SH, et al. Bayesian computing with INLA: a review. Ann Rev Stat Appl 2017; 4: 395–421.
48. Ekvall KO and Jones GL. Markov chain Monte Carlo, 2019.
49. Robert CP, Elvira V, Tawn N, et al. Accelerating MCMC algorithms. Wiley Interdiscip Rev Comput Stat 2018; 10(5): e1435. [PubMed: 30167072]
50. Fonnesbeck CJ, Patil A, Huard D, et al. PyMC documentation, 2017, https://readthedocs.org/projects/pymcmc/downloads/pdf/stable/
51. Hall B LaplacesDemon: an R package for Bayesian Inference, 2008, http://eio.usc.es/pub/mjginzo/descargas/leyenda/Documents/R/win-library/2.12/LaplacesDemon/doc/LaplacesDemon.pdf
52. Robert CP and Casella G. The metropolis – Hastings algorithm. In: Robert CPand Casella G(eds) Monte Carlo statistical methods. New York: Springer, 2004, pp. 267–320.
53. Chib S and Greenberg E. Understanding the metropolis-Hastings algorithm. Am Stat 1995; 49(4): 327–335.
54. Depaoli S, Clifton JP and Cobb PR. Just another Gibbs sampler (JAGS) flexible software for MCMC implementation. J Educ Behav Stat 2016; 41(6): 628–649.
55. Stan Development Team. Stan modeling language users guide and reference manual, 2016, https://pj.freefaculty.org/guides/crmda_workshops/sem/Archive/sem-4/literature/manuals/stan-reference-2.11.0.pdf
56. Betancourt M A conceptual introduction to Hamiltonian Monte Carlo. arXiv [preprint], 2017. DOI: 10.48550/arXiv.1701.02434.
57. Betancourt M and Girolami M. Hamiltonian Monte Carlo for hierarchical models. Curr Trends Bayesian Methodol Appl 2015; 79(30): 2–4.
58. Hill SD and Spall JC. Stationarity and convergence of the metropolis-hastings algorithm: Insights into theoretical aspects. IEEE Contr Syst Mag 2019; 39(1): 56–67.
59. Frigessi A, Martinelli F and Stander J. Computational complexity of Markov chain Monte Carlo methods for finite Markov random fields. Biometrika 1997; 84(1): 1–18.
60. Wang X, Yue Y and Faraway JJ. Bayesian regression modeling with INLA. Boca Raton, FL: Chapman & Hall/CRC, 2018.
61. R Core Team. R: a language and environment for statistical computing. Vienna: R Core Team, 2022, https://www.R-project.org/
62. Su YS, Yajima M, Su MYS, et al. Package ‘R2jags’. R package version 003–08, 2015, https://cran.r-project.org/web/packages/R2jags/index.html
63. Bürkner PC. brms: an R package for Bayesian multilevel models using Stan. J Stat Softw 2017; 80: 1–28.
64. Lindgren F and Rue H. Bayesian spatial modelling with R-INLA. J Stat Softw 2015; 63: 1–25.
65. Chen Q, Wu H, Ware LB, et al. A Bayesian approach for the cox proportional hazards model with covariates subject to detection limit. Int J Stat Med Res 2014; 3(1): 32. [PubMed: 24772198]
66. Strömberg U Collapsing ordered outcome categories: a note of concern. Am J Epidemiol 1996; 144(4): 421–424. [PubMed: 8712200]
67. Link WA and Eaton MJ. On thinning of chains in MCMC. Methods Ecol Evol 2012; 3(1): 112–115.

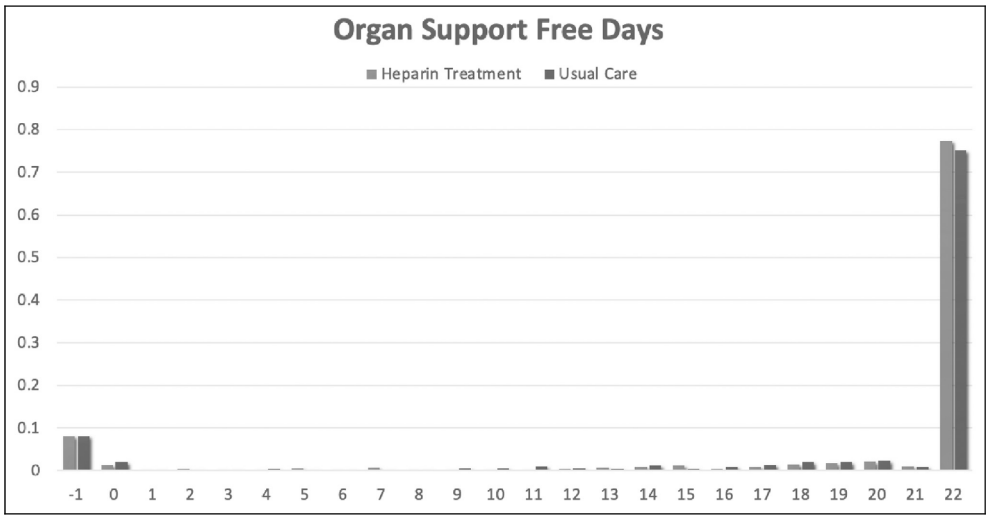

Figure 1.

The observed distribution of organ support free days, the primary outcome, from the combined data from the ATTACC/ACTIV-4A trials.

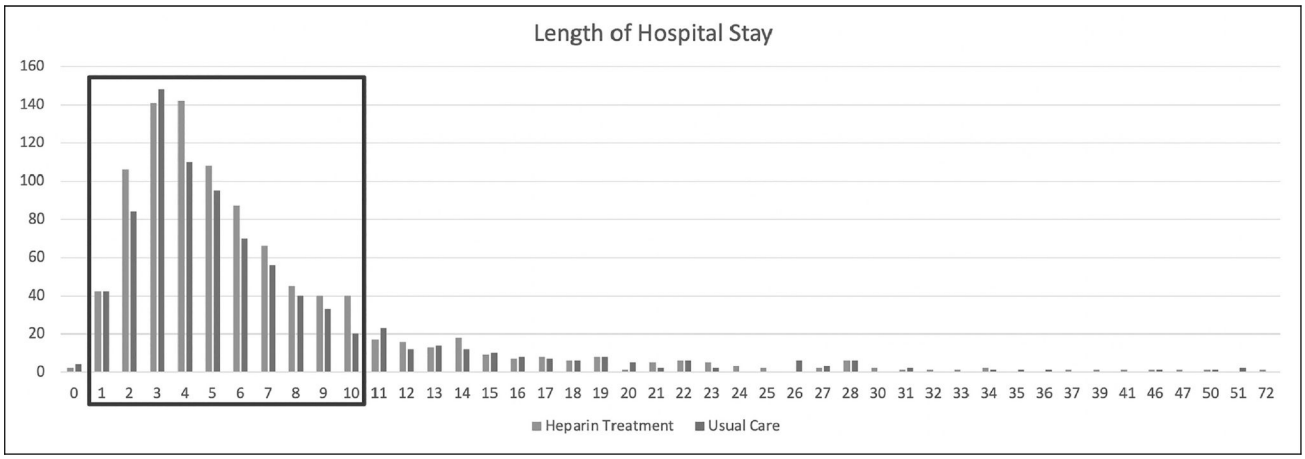

Figure 2.

The observed distribution of length of hospital stay, a key secondary outcome, from the combined data from the ATTACC/ACTIV-4A trials.

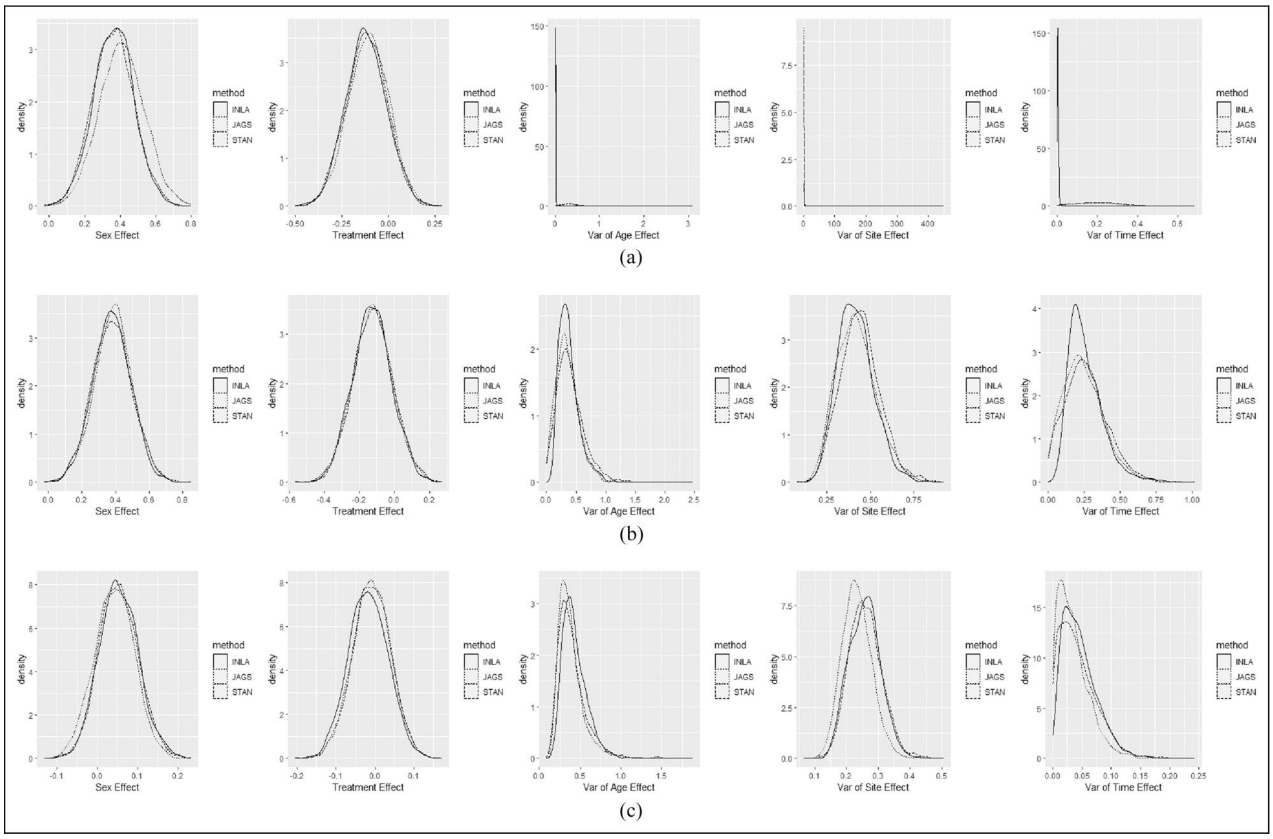

Figure 3.

Posterior density curves for five parameters of interest (from left to right), (1) the effect of sex and (2) treatment with therapeutic anti-coagulation on the outcome and the hierarchical variance for the effect of (3) age, (4) site and (5) time period. These are estimated for three outcomes, (a) Organ Support Free Days, (b) survival with no organ support, and (c) length of hospital stay. The posterior distributions are estimated with INLA, JAGS, and stan.

Table 1.

The estimated posterior mean (Mean), posterior standard deviation (SD), and 95% equitailed credible intervals (CI) for the seven outcomes considered in this analysis, (1) Organ Support Free Days, (2) survival with no organ support, (3) survival with no invasive mechanical ventilation (IMV), (4) survival with no mechanical respiratory support (MRS), (5) survival until hospital discharge, (6) major thrombotic event (TE) or death and (7) length of hospital stay. Results are reported for three algorithms, Integrated Nested Laplace Approximations (INLA), Hamiltonian Monte Carlo, implemented through stan (stan), and Gibbs sampling, implemented through JAGS, (JAGS).

	INLA			**stan**			**JAGS**		
Organ Support Free Days	**Mean**	**SD**	**CI**	**Mean**	**SD**	**CI**	**Mean**	**SD**	**CI**
Sex	0.366	0.112	(0.147, 0.587)	0.36	0.11	(0.14, 0.59)	0.383	0.112	(0.175, 0.612)
Treatment	−0.115	0.107	(−0.326, 0.095)	−0.11	0.11	(−0.31, 0.11)	−0.115	0.107	(−0.32, 0.099)
Age	0.01	0.005	(−0.002, 0.017)	0.39	0.26	(0.05, 1.01)	0.392	0.108	(0.197, 0.617)
Site	1.57	14.42	(−26.7, 29.84)	0.4	0.12	(0.2, 0.65)	0.348	0.224	(0.045, 0.937)
Time	0.004	0.003	(−0.002, 0.0l)	0.22	0.13	(0.02, 0.5)	0.213	0.12	(0.024, 0.480)
No Organ Support	Mean	SD	CI	Mean	SD	CI	Mean	SD	CI
Sex	0.381	0.115	(0.16, 0.607)	0.38	0.11	(0.16, 0.61)	0.381	0.112	(0.15, 0.618)
Treatment	−0.121	0.11	(−0.34, 0.095)	−0.12	0.11	(−0.33, 0.1)	−0.119	0.11	(−0.34, 0.091)
Age	0.38	0.17	(0.04, 0.71)	0.41	0.25	(0.06, 0.99)	0.395	0.214	(0.068, 0.914)
Site	0.42	0.1	(0.22, 0.62)	0.44	0.11	(0.23, 0.67)	0.432	0.119	(0.198, 0.674)
Time	0.26	0.12	(0.017, 0.51)	0.26	0.14	(0.02, 0.57)	0.247	0.139	(0.029, 0.549)
No IMV	Mean	SD	CI	Mean	SD	CI	Mean	SD	CI
Sex	0.336	0.15	(0.047, 0.631)	0.34	0.15	(0.05, 0.64)	0.347	0.149	(0.058, 0.636)
Treatment	0.004	0.14	(−0.274, 0.283)	0.01	0.14	(−0.27. 0.29)	0.008	0.146	(−0.281, 0.293)
Age	0.74	0.256	(0.24, 1.24)	0.9	0.4	(0.37, 1.93)	0.915	0.44	(0.385, 1.986)
Site	0.398	0.1	(0.2, 0.59)	0.4	0.12	(0.19, 0.65)	0.399	0.118	(0.184, 0.65)
Time	0.14	0.09	(−0.032, 0.31)	0.13	0.09	(−0.01, 0.34)	0.136	0.091	(0.009, 0.343)
No MRS	Mean	SD	CI	Mean	SD	CI	Mean	SD	CI
Sex	0.393	0.115	(0.17, 0.62)	0.4	0.12	(0.17, 0.61)	0.392	0.115	(0.165, 0.614)
Treatment	−0.114	0.111	(−0.331, 0.102)	−0.12	0.11	(−0.33, 0.1)	−0.115	0.11	(−0.332, 0.108)

	INLA			stan			JAGS		
Organ Support Free Days	**Mean**	**SD**	**CI**	**Mean**	**SD**	**CI**	**Mean**	**SD**	**CI**
Age	0.361	0.17	(0.02, 0.7)	0.39	0.24	(0.05, 1.03)	0.388	0.228	(0.039, 0.923)
Site	0.41	0.1	(0.21, 0.62)	0.44	0.12	(0.23, 0.68)	0.441	0.118	(0.232, 0.689)
Time	0.26	0.13	(0.02, 0.511)	0.25	0.15	(0.02, 0.56)	0.247	0.139	(0.022, 0.56l)
Survival at Hospital Discharge	Mean	SD	CI	Mean	SD	CI	Mean	SD	CI
Sex	0.253	0.185	(−0.105, 0.619)	0.26	0.19	(−0.12, 0.64)	0.265	0.188	(−0.089, 0.645)
Treatment	−0.037	0.178	(−0.385, 0.312)	−0.03	0.17	(−0.37, 0.3)	−0.025	0.176	(−0.37, 0.321)
Age	1.15	0.394	(0.28, 1.824)	1.53	0.73	(0.56, 3.37)	1.16	0.384	(0.31, 2.02)
Site	0.6	0.133	(0.33, 0.86)	0.63	0.16	(0.37, 0.97)	0.644	0.162	(0.365, l)
Time	0.15	0.1	(−0.02, 0.42)	0.15	0.12	(0.01, 0.43)	0.145	0.116	(0.004, 0.411)
Major TE or Death	Mean	SD	CI	Mean	SD	CI	Mean	SD	CI
Sex	−0.267	0.18	(−0.624, 0.082)	−0.28	0.18	(−0.63, 0.08)	−0.275	0.179	(−0.627, 0.069)
Treatment	0.069	0.172	(−0.27, 0.407)	0.06	0.17	(−0.28, 0.42)	0.059	0.175	(−0.288, 0.404)
Age	1.16	0.43	(0.244, 2.018)	1.6	0.73	(0.61, 3.33)	1.2	0.43	(0.22, 2.21)
Site	0.65	0.14	(0.364, 0.91)	0.67	0.16	(0.4, 1.02)	0.678	0.164	(0.395, 1.024)
Time	0.15	0.092	(−0.02, 0.333)	0.16	0.11	(0.01, 0.42)	0.165	0.116	(0.007, 0.43l)
Length of Hospital Stay	Mean	SD	CI	Mean	SD	CI	Mean	SD	CI
Sex	0.052	0.049	(−0.046, 0.148)	0.05	0.05	(−0.05, 0.15)	0.042	0.05	(−0.055, 0.137)
Treatment	−0.014	0.048	(−0.11, 0.08)	−0.01	0.05	(−0.1, 0.08)	−0.01	0.05	(−0.105, 0.085)
Age	0.4	0.12	(0.16, 0.63)	0.4	0.18	(0.19, 0.84)	0.38	0.15	(0.183, 0.773)
Site	0.29	0.05	(0.19, 0.393)	0.26	0.05	(0.17, 0.37)	0.233	0.05	(0.15, 0.331)
Time	0.06	0.033	(−0.0097, 0.12)	0.06	0.03	(0.001, 0.12)	0.035	0.027	(0.001, 0.101)

Table 2.

The posterior probabilities that there exists a positive treatment effect for the model analysing the seven outcomes considered in this analysis, (1) Organ Support Free Days, (2) survival with no organ support, (3) survival with no invasive mechanical ventilation (IMV), (4) survival with no mechanical respiratory support (MRS), (5) survival until hospital discharge, (6) major thrombotic event (TE) or death and (7) length of hospital stay for JAGS, stan and INLA.

	JAGS	stan	INLA
Organ support-free days	83%	86%	86%
Length of Hospital Stay	57%	57%	63%
Survival with No IMV	52%	53%	53%
Survival with No MRS	84%	84%	83%
Survival with No Organ Support	87%	86%	88%
Major Thrombotic Event or Death	63%	64%	67%
Survival Until Hospital Discharge	56%	55%	59%

Table 3.

The computation time required to estimate the posterior distribution for the model analysing the seven outcomes considered in this analysis, (1) Organ Support Free Days, (2) survival with no organ support, (3) survival with no invasive mechanical ventilation (IMV), (4) survival with no mechanical respiratory support (MRS), (5) survival until hospital discharge, (6) major thrombotic event (TE) or death and (7) length of hospital stay for INLA, stan and JAGS.

	JAGS	stan	INLA
Organ support-free days	1.5 h	30.5 min	31.28 s
Length of Hospital Stay	6.8 h	59.23 min	12.22 s
Survival with No IMV	7.35 min	7.73 min	4.2 s
Survival with No MRS	7.2 min	7.35 min	4.77 s
Survival with No Organ Support	2.05 min	6.91 min	4.74 s
Major Thrombotic Event or Death	7.43 min	14.5 min	4.44 s
Survival Until Hospital Discharge	7.21 min	8.4 min	4.78 s