Published in final edited form as:

Contemp Clin Trials. 2017 December ; 63: 73–83. doi:10.1016/j.cct.2017.02.004.

Bayesian clinical trial design using Markov models with applications to autoimmune disease

Barry S. Eggleston^a, Joseph G. Ibrahim^{b,*}, and Diane Catellier^a

^aRTI International, 3040 East Cornwallis Road, Research Triangle Park, NC 27709, USA

^bDepartment of Biostatistics, University of North Carolina at Chapel Hill, Chapel Hill, NC 27599, USA

Abstract

Immune Thrombocytopenia is an autoimmune disease associated with bleeding that is treated by increasing the platelet count to a level where the chance of uncontrollable bleeding is low. Failure occurs when platelet counts are not raised sufficiently (initial failure), or when high platelet counts are not maintained after initial success (relapse). In this paper, we propose a Bayesian clinical trial design that uses a Markov multistate model along with a power prior for the parameters which incorporates historical control data to estimate transition rates among two randomized groups as defined by the model. A detailed simulation is carried out to examine the operating characteristics of a trial to test whether a new treatment reduces the relapse rate by 40% relative to standard care when data from 60 historical controls treated with standard care is available. We also use simulated data to demonstrate effects of discordance between historical and randomized controls on the estimated hazard ratios. Finally, we use a simulated trial to demonstrate briefly what type of results the model can give and how those results can be used to address hypotheses regarding treatment effects. Using simulated data, we show that the model yields good operating characteristics when the historical and randomized controls are from the same population, and demonstrate how discordance between the control groups affects the operating characteristics.

Keywords

Markov models; Illness-death model; Bayesian methods; Clinical trial design

1. Introduction

Immune Thrombocytopenia (ITP) is an autoimmune disease associated with bleeding. The bleeding is the result of impaired platelet production that is often persistent and chronic and due to platelet destruction by the immune system [5]. ITP treatment consists of two objectives: quickly raising the platelet count to a therapeutic level to stop the ITP relating bleeding and maintaining the therapeutic platelet level to prevent future ITP-related bleeding. Current treatment uses established treatments to address the first objective, but clinicians need improved treatments to reduce the risk of relapse among initial treatment

*Corresponding author. jibrahim@email.unc.edu (J.G. Ibrahim).

responders. To identify these improved treatments, clinical researchers need efficient effect estimation models to properly account for failure to meet either objective.

This article proposes a three-state Markov Bayesian model that can incorporate historical control data if necessary to evaluate experimental ITP treatments while taking into account the two possible failure modes mentioned above. The multi-state nature of the model accounts for these two types of failure. The Bayesian model allows for inclusion of prior information and historical control data via the power prior of Ibrahim and Chen [7]. Using simulation, we demonstrate how the proposed Bayesian model can be used to estimate hazard rates and ratios from ITP trial data, its power and type 1 error characteristics, and how the model responds to differences between randomized and historical control relapse rates.

The article is organized as follows. In Section 2, we describe a clinical trial design that corresponds to the proposed three-state Markov model, we describe the Bayesian model used to estimate the parameters of the model, we describe the structure of the analysis dataset used in the model, and we illustrate the simulation process used to study and illustrate the Bayesian model. In Section 3, we describe the simulation results, and in Section 4 we conclude the article with a discussion. Financial funding for the research reported in this publication was given by the National Heart, Lung, And Blood Institute of the National Institutes of Health under Award Number U24HL114577. The sponsor had no other role in the research beyond funding.

2. Methods

2.1. Trial design

The ITP trial design, for which the Bayesian model is used to estimate treatment effects, consists of one active arm and one control arm, into which patients are randomized at the time of enrollment using a 1:1 ratio. After enrollment, both arms are treated with a selected therapy intended to raise the platelet counts to therapeutic levels. One possible outcome of this initial treatment is failure to respond. After this initial treatment period, initial responders in both arms have their treatment regimens modified to allow for a comparison of at least one experimental therapy with respect to ITP relapse rates. For example, responders in the control arm might receive standard relapse prevention care while the active arm might begin to take an experimental therapy intended to reduce relapse rates relative to standard care. Using data collected from the trial design, the Bayesian model is used to estimate initial and relapse failure rates and hazard ratios, as well as response rates to the initial treatment.

Another part of the design involves the platelet assessment schedule. Platelet levels are assessed periodically according to a pre-specified schedule. If the primary outcome is time-to-relapse and relapse is defined as time until platelet levels are less than some threshold, then the time-to-relapse will be interval censored. Therefore, the Bayesian model also takes interval censoring into account.

2.2. Three-state model

As noted earlier, ITP treatment failure occurs in two ways. A subject is an initial non-responder (an initial failure), if he or she does not experience therapeutically high platelet levels during initial treatments to raise platelet levels. On the other hand, an initial responder can experience relapse if the initial responder experiences critically low platelet levels at some later time. In order to accurately estimate the relapse failure rate, modelling needs to account for the possibility of initial failure, initial nonresponse. If these two paths to failure are combined with an initial state of critically low platelet counts, and a second state of initial treatment response, then a three-state model identical to an Illness-Death model can be used to accurately estimate the relapse failure rate without confounding from the initial failure rate, Fig. 1 [3,13].

State 1 of the three-state model represents a patient's initial condition of critically low platelet levels. State 2 represents initial treatment responders who have obtained a therapeutically high platelet count. A subject can reach State 2 only from State 1, a "1-2" transition. State 3 represents treatment failure, and subjects can reach State 3 from States 1 or 2. First, a subject with a critically low platelet count could fail to respond to initial treatment within a pre-specified period. This subject is an initial non-responder and reaches State 3 from State 1, a "1-3" transition. Second, a subject could fail by relapsing to a critically low platelet count after reaching State 2, a "2-3" transition. In an ITP trial that is evaluating a treatment for relapse, transition rates between States 1 and 2 should be high and transition rates between States 1 and 3 may be driven mostly by treatment protocol with respect to the definition of an initial non-responder. A subject must be observed to occupy State 2 before a "1-2" transition can be declared, and only subjects who are observed to occupy State 2 can contribute to the estimation of 2-3 transition model parameters. The proposed Bayesian model presented in this paper can be used to estimate all six transition rates, three for each arm, and all three hazard ratios.

2.3. Likelihood and priors for the randomized component of the Bayesian trial design

The MCMC models used in this article are based on exponential distributions with the mean transition time equal to a function of treatment. The likelihood function component for the 1-2 transition is given in Eq. (1), where $\delta_j = 1$ indicates whether or not the transition time is interval censored, (t_{12Lk}, t_{12Uk}) , $\delta_j = 2$ indicates a right censored at an unknown value, $[\hat{t}_{13Lk}, \infty]$, or $\delta_j = 3$ indicates a right censored at a known value $[t_{12Lk}, \infty]$. An unknown 1-2 right censored time occurs when an observed but interval censored 1-3 transition occurs. The values t_{12Lk} and t_{12Uk} are left and right interval time points for an interval censored 1-2 transition time similar to SAS's input structure for interval censoring in PROC LIFEREG. In case of a right censored time point, only one value is needed, so when $\delta_j = 3$ the value of t_{12Lk} represents the exact timing of a right censored 1-2 transition time. The value of t_{12Uk} can be any value allowable by the estimation algorithm. Similarly, when $\delta_j = 2$ the value of $\hat{t}_{13Lk}$ represents the predicted value of a 1-3 transition time, but due to competing risk between 1-2 and 1-3 transitions $\hat{t}_{13Lk}$ also is equal to the predicted time when the 1-2 transition was right censored. A predicted right censored 1-2 transition time is needed when a subject transitions out of State 1 into State 3, because the transition timing into State 2 is

right censored at an unknown value. The 1–3 transition time is interval censored, so the right censored transition time for the 1–2 transition is only known to occur within an interval. In such a case the present model predicts the unknown right censoring time for the 1–2 transition by making a random draw from a truncated version of 1–3 transition time distribution. The μ_{12k} is the exponential mean for the k th subject [11]. The parameter μ_{12k} is linked with the treatment variable by the log link, $\log(\mu_{12k}) = \beta_{120} + \beta_{121} Trt_k$, where Trt_k equals 1, if subject k is in the active arm and 0 otherwise.

$$L_{12} = \prod_{k=1}^n \left[\exp\left(-\frac{t_{12L}k}{\mu_{12k}}\right) - \exp\left(-\frac{t_{12U}k}{\mu_{12k}}\right) \right]^{\delta_k=1} \left[\exp\left(-\frac{\hat{t}_{13L}k}{\mu_{12k}}\right) \right]^{\delta_k=2} \left[\exp\left(-\frac{t_{12L}k}{\mu_{12k}}\right) \right]^{\delta_k=3} \quad (1)$$

In summary: using a convention similar to the SAS convention for representing interval censoring, if the subject experienced a 1–2 transition within the interval $(t_{12L}k, t_{12U}k)$, then δ_j equals 1. If the subject was right censored at an unknown value due to an observed but interval censored 1–3 transition, then the transition interval is represented by $(\hat{t}_{13L}k, C)$ and δ_j will equal 2, where $\hat{t}_{13L}k$ is a predicted value of when the 1–3 transition occurred and C is some arbitrary value larger than the possible maximum follow-up time. This C value was necessary given the JAGS 4.2 fitting algorithm used to estimate parameter posterior distributions. If the subject was right censored at a known value, then the transition interval is represented by $(t_{12L}k, C)$ and δ_j equals 3, where C is some arbitrary value larger than the possible maximum follow-up time.

Due to competing risks among 1–2 and 1–3 transitions, the likelihood function component for the 1–3 transition is similar to the likelihood function for the 1–2 transition and is given in Eq. (2), where δ_j indicates transition type, $\log(\mu_{13k}) = \beta_{130} + \beta_{131} Trt_k$, and $t_{13L}k, t_{13U}k$, and $\hat{t}_{12L}k$ are defined in a similar manner as $t_{12L}k, t_{12U}k$, and $\hat{t}_{13L}k$. In the case of interval censored 1–3 transition, the values $t_{13L}k$ and $t_{13U}k$, are left and right endpoints of a 1–3 transition. In the case of right censoring at a known value, $t_{13L}k$ is equal to the known timing of the right censoring. In the case of right censoring at an unknown value, $\hat{t}_{12L}k$ is equal to the predicted timing of a 1–2 transition which due to competing risks is equal to the predicted time of a right censored 1–3 transition time.

$$L_{13} = \prod_{k=1}^n \left[\exp\left(-\frac{t_{13L}k}{\mu_{13k}}\right) - \exp\left(-\frac{t_{13U}k}{\mu_{13k}}\right) \right]^{\delta_k=1} \left[\exp\left(-\frac{\hat{t}_{12L}k}{\mu_{13k}}\right) \right]^{\delta_k=2} \left[\exp\left(-\frac{t_{13L}k}{\mu_{13k}}\right) \right]^{\delta_k=3} \quad (2)$$

To estimate the 1–2 transition and 1–3 transition rates using the Bayesian model, priors need to be placed on β_{120} , β_{121} , β_{130} , and β_{131} . In this paper, non-informative $N(\mu=0, \sigma=100)$ priors are used.

Since $\hat{t}_{12L}k$ and $\hat{t}_{13L}k$ are predicted values, predictive distributions need to be identified for $\hat{t}_{12L}k$ and $\hat{t}_{13L}k$. The values of $\hat{t}_{12L}k$ will be drawn from a truncated version of the model's estimated exponential distribution for the 1-2 transition times, since the censoring mechanism for the 1-3 transition is equal to the 1-2 transition mechanism. For $\hat{t}_{12L}k$ the truncation will be driven by subject specific left and right values of the observed interval censored 1-2 transition. For $\hat{t}_{13L}k$ the truncation will be driven by subject specific left and right values of the observed interval censored 1-3 transition. Given the assumption of exponential transition times, the truncated distribution for $\hat{t}_{12L}k$ is given in Eq. (3), where β_{120} and β_{121} are the corresponding parameters for the 1-2 transition while $t_{12L}k$ and $t_{12U}k$ are left and right endpoints of the interval censored 1–2 transition time.

$$\pi(\hat{t}_{12L}k|\beta) \propto e^{-(\beta_{120}+\beta_{121}Trt_k)} \exp \left[-e^{-(\beta_{120}+\beta_{121}Trt_k)} \hat{t}_{12L}k \right] \cdot I(t_{12L}k < \hat{t}_{12L}k \leq t_{12U}k) \quad (3)$$

A similar truncated distribution is used to predict the values of $\hat{t}_{13L}k$ for the 1-3 transition and it is given in Eq. (4).

$$\pi(\hat{t}_{13L}k|\beta) \propto e^{-(\beta_{130}+\beta_{131}Trt_k)} \exp \left[-e^{-(\beta_{130}+\beta_{131}Trt_k)} \hat{t}_{13L}k \right] \cdot I(t_{13L}k < \hat{t}_{13L}k \leq t_{13U}k) \quad (4)$$

The likelihood for the 2–3 transition is somewhat different, because the exact time in State 2 is unknown and all right censoring observations are observed. The likelihood for the 2–3 transition and the k th subject is given in Eq. (5), where δ_j indicates whether or not the transition is interval censored (=1) or right censored at a known value (=3), $t_{23L}k$ and $t_{23U}k$ are left and right interval time points for an interval censored 2–3 transition time, and $\hat{t}_{12L}k$ is the predicted value of the 1–2 transition time. Similar to the 1–2 and 1–3 transitions, the parameter μ_{23k} is linked with the treatment variable by the log link, $\log(\mu_{23k}) = \beta_{230} + \beta_{231}Trt_k$, where Trt_k equals 1, if subject k is in the active arm and 0 otherwise.

$$L_{23} = \prod_{k=1}^{n_{23}} \left[\exp \left(-\frac{t_{23L}k - \hat{t}_{12L}k}{\mu_{23k}} \right) - \exp \left(-\frac{t_{23U}k - \hat{t}_{12L}k}{\mu_{23k}} \right) \right]^{\delta_k=1} \left[\exp \left(-\frac{\hat{t}_{23L}k - \hat{t}_{12L}k}{\mu_{23k}} \right) \right]^{\delta_k=3} \quad (5)$$

Priors for β_{230} and β_{231} are the same as those assigned to the corresponding parameters in the 1–2 and 1–3 transition models, $N(\mu=0, \sigma=100)$.

Relative to the 1–2 and 1–3 transitions, the 2–3 transition likelihood component is different in that the predicted timing of the 1–2 transition is subtracted from both interval time points for the interval censored 2–3 transition time, and there is no need to predict a right censoring time due to a lack of any competing risks. Subtracting the predicted 1–2 transition time from the 2–3 transition interval endpoints is necessary to ensure the estimated hazard ratio is for time-to-relapse using the transition time into State 2 as time zero. The total likelihood for the randomized component of the Bayesian model is $L = L_{12} \cdot L_{13} \cdot L_{23}$, and the Posterior for $\beta = [\beta_{120} \beta_{121} \beta_{130} \beta_{131} \beta_{230} \beta_{231}]$ is given in Eq. 6. [4,6].

$$\pi(\beta|data) \propto L \cdot \pi(\beta_{120}) \cdot \pi(\beta_{121}) \cdot \pi(\beta_{130}) \cdot \pi(\beta_{131}) \cdot \pi(\beta_{230}) \cdot \pi(\beta_{231}) \quad (6)$$

2.4. Power prior for incorporating historical control data

If historical interval and right censored control data are included in the analysis, then the power prior for β_{230} , and β_{231} needs to be added to the model given in Eq. (6) and the power prior is given in Eq. (7), where $0 < a_0 < 1$, δ_j indicates whether or not the transition is interval censored (=1) or right censored at a known value (=3), t_{12HLk} and t_{12HUK} are left and right interval time points for an interval censored 2–3 transition time among historical controls, and $\log(\mu_{23k}) = \beta_{230} + \beta_{231} Trt_k$ [7,8,9]. If $a_0 = 0$, then the historical data is ignored. If $a_0 = 1$, then all of the information in historical data is used to estimate the parameters. If $0 < a_0 < 1$ then the information in the historical data is discounted [2,12].

$$\pi(\beta_{230}, \beta_{231}) = \left(\prod_{k=1}^{n_H} \left[\exp\left(-\frac{t_{23HLk}}{\mu_{23k}}\right) - \exp\left(-\frac{t_{23HUK}}{\mu_{23k}}\right) \right]^{\delta_k=1} \left[\exp\left(-\frac{t_{23HLk}}{\mu_{23k}}\right) \right]^{\delta_k=3} \right)^{a_0} \quad (7)$$

After incorporating the historical data through the power prior for β_{230} , and β_{231} the posterior for $\beta = [\beta_{120} \beta_{121} \beta_{130} \beta_{131} \beta_{230} \beta_{231}]$ is given in Eq. 8.

$$\pi(\beta|data) \propto L \cdot \pi(\beta_{120}) \cdot \pi(\beta_{121}) \cdot \pi(\beta_{130}) \cdot \pi(\beta_{131}) \cdot \pi(\beta_{230}, \beta_{231}) \cdot \pi(\beta_{230}) \cdot \pi(\beta_{231}) \quad (8)$$

While incorporating historical data, the 1–2 transition rates and hazard ratio are estimated using the part of the total likelihood given in Eq. (1) and the $N(\mu=0, \sigma=100)$ priors placed on β_{120} , and β_{121} . Similarly, the 1–3 transition rates and hazard ratio are estimated using the part of the total likelihood given in Eq. (2) and the $N(\mu=0, \sigma=100)$ priors placed on β_{130} , and β_{131} . The 2–3 transition rates and hazard ratio on the other hand are estimated using the

likelihood/power prior combination given in Eqs. (5) and (7) as well as the $N(\mu=0, \sigma=100)$ priors placed on β_{230} and β_{231} .

2.5. Assessable hypotheses via Bayesian model

This Bayesian model can be used to assess three main hypotheses. The first hypothesis considers the hazard ratio between the two arms with respect to the 1–2 transition, via the log hazard ratio for the 1–2 transition, β_{121} [1]. The second hypothesis considers the hazard ratio between the two arms with respect to the 1–3 transition, via the log hazard ratio for the 1–3 transition, β_{131} [1]. The third and main hypothesis considers the hazard ratio between the two arms with respect to the 2–3 transition, via the log hazard for the 2–3 transition, β_{231} [1]. Table 1 lists all 3 of these hypotheses in terms of null and alternative hypotheses, one pair per column.

2.6. Expected analysis data structure

For each patient in an ITP trial, the recorded trajectories of platelet count measurements can be used to identify intervals within which transitions between the states have occurred. These intervals can have endpoints equal to time since randomization. For use in the proposed model and JAGS algorithm (Appendix B), each patient needs to have either two or three observations in the analysis dataset. Each observation needs to have a variable that identifies the type of transition represented, ‘1–2’, ‘1–3’, or ‘2–3’. If the patient experienced no transitions, then the patient will have two right censored observations, one for the 1–2 transition and one for the 1–3 transition. The left endpoint of right censored observations needs to equal the time of right censoring while the right endpoint will need to equal some arbitrary value that is allowable by the estimation algorithm. In this paper we used a value 0.1 larger than the maximum possible follow-up time, which was chosen solely due to the JAGS 4.2 program used for posterior estimation (Appendix B). For both observations, a Status variable will equal 3 indicating the timing of the right censoring is exactly known. If the patient experiences a 1–3 transition, then the patient will have two observations. One for the observed 1–3 transition, and one for the unobserved 1–2 transition. For the 1–3 transition observation, the Status variable will equal 1 and the left and right endpoints will equal the interval within which the 1–3 transition is known to have occurred. On the other hand, the Status variable for the 1–2 transition will equal 2, while the left and right endpoints will equal the same values as the 1–3 transition observation. For the 1–2 transition records the interval values indicate the interval within which the 1–2 right censoring time is known to have occurred. If the patient experiences a 1–2 transition, then the patient will have three observations. One observation for each possible transition, 1–2, 1–3, and 2–3. The 1–2 and 1–3 transition observations will have interval endpoints equal to the times spanning the known 1–2 transition time, but the Status variable will equal 1 for the 1–2 transition observation and the Status variable will equal 2 for the 1–3 transition. Finally, the 2–3 transition observation will be either an interval or right censored observation and the Status variable will equal 1 or 3 respectively. With respect to the Status variable, it equals 1 for an interval censored 2–3 transition, and three for a right censored 2–3 transition at a known time. The data management techniques necessary to build up this dataset are discussed in chapter 8 of [13]. Example data for 25 patients are given in Appendix A. A JAGS program

to analysis data with this structure is given in Appendix B. The JAGS program in Appendix B was run using JAGS 4.2.

2.7. Simulated analysis data

Hypothetical datasets were constructed from the proposed Illness-Death Markov model to illustrate the use of the model. Each hypothetical dataset consisted of data from 220 patients randomized using a 1:1 ratio into two arms, active treatment and standard care, for the treatment of ITP and followed for 3 years. Also, each hypothetical dataset consisted of relapse data, 2–3 transition data, from 60 historical patients who had hypothetically received standard care in a previous trial and had an estimated hazard rate of 0.5 ITP relapses per year. The relapse times from these 60 historical patients are either interval or right censored. Both randomized arms of the hypothetical trial received 1–2 and 1–3 transition times representative of receiving an initial treatment regimen which would produce a 1–2 transition rate of 3.29 transitions per person year and would produce a 1–3 transition rate of 0.84 transitions per person year, thus producing hazard ratios for the 1–2 and 1–3 transition that are equal to 1. The active arm of the hypothetical trial received 2–3 transition times representative of receiving an experimental treatment regimen that reduced the relapse rate among initial responders by 40% relative to standard care, while the control arm received 2–3 transition times representative of receiving standard relapse preventive care.

2.8. Operating characteristics of proposed Bayesian design

Using data generated with the characteristics described in Section 2.7, four sets of simulations are carried out to describe the output and to study the operating characteristics of the proposed Bayesian design. First, the Bayesian model given in Eq. (8) is applied to illustrate power and Type 1 error when including historical control data from the 2–3 transition using a power prior with a_0 ranging from 0 to 1. In this first set of simulations, the randomized and historical control subjects have 1–2 and 1–3 hazard rates equal to those defined in Section 2.7, and the 2–3 hazard rate in both control groups is equal to 0.5 ITP relapses per year. At different sample sizes, a set of 100 trial simulations are completed using a_0 values of 0.0, 0.33, 0.67, and 1.0, when the 2–3 hazard ratio is equal to 0.60 (power assessment) and 1.0 (Type 1 error assessment). The second simulation set is a single simulated trial producing data with the setup described in Section 2.7. This simulated trial dataset is analyzed using the model in Eq. (8) to illustrate how the model can be used to approximate the posterior distributions for the 1–2, 1–3, and 2–3 transition hazard rates and hazard ratios. This simulated trial data has 1–2 and 1–3 transition hazard ratios of 1.0, with rates equal to those defined in Section 2.7, while the 2–3 transition ratio is 0.6, with rates equal to 0.5 in both historical and randomized controls. The third set of simulations uses the Bayesian model defined in Eq. (8) to analyze three datasets of simulated trial data using the setup for the 1–2 and 1–3 transitions defined in Section 2.7, but the randomized and historical control subjects differ by a set amount with respect to the 2–3 transition rate. One dataset has both randomized and historical control subjects that have the same 2–3 transition hazard rate of 0.50. Another dataset has the randomized control subjects with a hazard rate of 0.3, while the historical subjects have a hazard rate of 0.5 (representing a 40% decrease in 2–3 transition hazard rate relative to historical controls). The final dataset has the randomized control subjects with a hazard rate of 0.7, while the historical subjects have a

hazard rate of 0.5 (representing a 40% increase in 2–3 transition hazard rate relative to historical controls). In this third set of simulations, the data is repeatedly analyzed using a_0 values of 0, 0.2, 0.4, 0.6, 0.8, and 1.0. Since only the 2–3 transition has different hazard rates among the groups, only analysis results from the 2–3 transition hazard ratio estimation is presented. This restricted focus on the 2–3 transition is justified by the fact that most current ITP trials focus on the goal of relapse prevention. Finally, the fourth set of simulations uses the Bayesian model defined in Eq. (8) to illustrate the effects on power and type 1 error when historical and randomized controls differ in the 2–3 transition hazard rate [14]. This fourth simulation uses the same three datasets of simulated trial data as the third simulation set, but repeatedly simulates and analyzes the data using a_0 values ranging from 0, 0.33, 0.67, and 1.0. At different sample sizes and a_0 values, a set of 100 trial simulations are completed when the 2–3 hazard ratio is equal to 0.60 (power assessment) and 1.0 (Type 1 error assessment).

3. ITP trial simulation results

3.1. Power and type 1 error

First, we consider the Power and Type 1 characteristics of the design when the randomized and historical control subjects do not differ with respect to the 2–3 transition hazard rate. Figs. 2 and 3 illustrate the power and type 1 error characteristics of the Bayesian model under the scenario of a relapse trial that focused only on the 2–3 transition hazard ratio. The trial randomizes an equal number of subjects to active and control arms while incorporating 60 historical control subjects. All subjects start the trial in State 1 (low platelet counts), and those who progress into State 2 are included in the analysis used to test the hypotheses of $H_{03}:\beta_{231} = 0$ vs. $H_{a3}:\beta_{231} > 0$. Fig. 2 shows that when the historical data are not used ($a_0 = 0$), the model needs 116 subjects per arm for 80% power. With a_0 equal to 0.33, the number reduces to 100 subjects per arm, and with a_0 equal to 0.67 or 1.0 the number of subjects needed reduces down to 90 and 76 subjects per arm respectively. The value of a_0 is affecting the variance of the posterior for the 2–3 transition hazard ratio. As a_0 increases, the variance of the posterior decreases thus increasing the power. Fig. 3 shows that as a_0 increases to 1 the type one error reduces down to a value close to 0.038, so the model properly controls for Type 1 error.

3.2. Example trial analysis

In this example analysis, a simulated dataset generated with characteristics described in Sections 2.6 and 2.7 is used as the analysis dataset. The simulated trial involves 110 subjects randomized to active treatment and 110 subjects randomized to standard care with 2–3 transition data on 60 historical control subjects included using a power prior with $a_0 = 1.0$. Table 2 shows numerical summaries for the hazard rates in each group and transition type as well as numerical summaries and posterior plots for the hazard ratio of each transition. Table 2 also reports the posterior probability of the transition hazard ratio being <1.0 . Table 2 shows that the model accurately estimated the hazard rates and ratios for each transition. If this was an actual trial, then the results would support the conclusion that the active treatment reduced the relapse hazard rate, 2–3 transition rate, by about 40%, hazard ratio of 0.61 and a 95% credible interval of 0.43 to 0.84. The posterior probability of the 2–3 hazard

ratio being <1 is 0.999. Similarly the posterior plot for the 2–3 transition hazard ratio shows very little area above 1.0, while the posteriors for the 1–2 and 1–3 transition hazard ratios clearly are centered on a value near 1.0 indicating no treatment effect.

3.3. Effects of discordance between historical and randomized controls

Although not presented in Section 3.1 above, the effects of a_0 on power and type 1 depends on the correspondence between randomized and historical control subjects. If the hazard rate among randomized control subjects differs from historical control subjects, then the location of the posterior distributions change as a_0 changes. Table 3 illustrates the bias caused by discordant randomized and historical control subjects in the case when the 2–3 transition hazard ratio between the two randomized groups is always 0.6 regardless of the hazard rate among randomized controls. The first column represents the case where the hazard rate among randomized controls is 0.3, the hazard rate among randomized and actively treated subjects is 0.18, and the hazard rate among historical control subjects is 0.5. This first column of Table 3 shows that increasing a_0 moves the posterior mean hazard ratio down when randomized controls have a lower hazard rate than historical controls. This is caused by the fact that including more information from historical control subjects moves the posterior estimate of the hazard rate among controls upward towards 0.5. Nevertheless, the posterior mean of the hazard rate among actively treated subjects is still going to be about 0.18. As a result the estimated hazard ratio for the treatment effect relative to standard care decreases from 0.60 down to 0.48. On the other hand, the last column of Table 3 shows that the effect of increasing a_0 reverses when the randomized control subjects have a higher hazard rate than the historical control subjects. For the case where the randomized controls has a hazard rate of 0.7 and the historical controls have a hazard rate of 0.5, the estimated hazard ratio for the treatment effect relative to standard care increases from 0.60 up to 0.70. Notice that in both of these columns representing discordant cases, the variance still decreases as a_0 increases. The middle column shows that when no discordance is present between randomized and historical controls, increasing a_0 simply improves the posterior precision. The posterior plots in Fig. 4 graphically illustrate this shifting of the posterior as a consequence of including data from discordant historical control subjects and illustrate the possible precision improvements that result from including data from concordant historical control subjects. The bias imposed on the model estimate of the 2–3 transition hazard ratio has an effect on power and type 1 error. Figs. 2, 5, and 6 show this effect. Fig. 5 contains power curves when the randomized controls have a hazard rate of 0.3 compared to a hazard rate of 0.5 among the historical controls (a 40% decrease in hazard rate). As more of the historical control information is included in the estimation of the 2–3 transition hazard ratio, the estimated hazard ratio becomes smaller since the overall estimate of the hazard rate among controls increases towards 0.5 while the estimate of the hazard rate among active subjects remains around 0.18. The reduction of the estimated hazard ratio increases power. Fig. 2 contains power curves when the historical and randomized controls have the same hazard rate of 0.5. As more of the historical information is included in the estimation, more events are available for estimating the common hazard rate thus reducing the posterior variance. The consequence is an increase in power. Fig. 6 contains power curves when the randomized controls have a hazard rate of 0.7 compared to a hazard rate of 0.5 among the historical controls (representing a 40% increase in hazard rate). As more of the historical

control information is included in the estimation of the 2–3 transition hazard ratio, the estimated hazard ratio becomes larger since the overall estimate of the hazard rate among controls decreases towards 0.5 while the estimate of the hazard rate among active subjects remains around 0.42. The increase of the estimated hazard ratio towards 0 decreases power. Finally, Figs. 3 and 7 contains curves for type 1 error as a function of a_0 . Fig. 7 shows that when historical and randomized controls differ in the 2–3 transition hazard rate the type 1 error increases as more of the historical control information is included in estimation of the 2–3 transition hazard ratio. Fig. 3 and the black line in Fig. 7 show that when historical and randomized controls have the same 2–3 transition hazard rate, Type 1 error is controlled by a_0 .

4. Discussion

In this paper, we proposed a multistate Bayesian Markov model to compare treatments in diseases such as ITP. Also, we demonstrated that incorporating historical control data through the power prior can greatly increase the power of an ITP trial that focuses on relapse while estimating treatment effects on initial failure and success, while also controlling Type I error to an acceptable level. Under the scenario of identical hazard rates between randomized and historical controls, the required sample size was reduced by 34%. This reduction, however, comes at a price, in that a good understanding of possible hazard rates among randomized control subjects relative to the known hazard rate among the historical controls is needed. Incorporating discordant data from historical controls will bias the estimates and consequently adversely affect operating characteristics of the trial design.

One limitation of the design is that it does not allow a study team to a priori engineer adaptive techniques into the design to protect against the use of inappropriate historical data when all evidence indicates that the trial control data are not compatible with the historical control data. Consequently, some thought needs to be given to the possible range of discordance between randomized and historical controls, with respect to some effect parameter such as a hazard ratio. Once this range is quantified, simulations can be used to assess discordance effects on operating characteristics with this range. These simulation results can be used to a priori select an a_0 value that reduces the risk of adverse effects on trial operating characteristics while increasing the feasibility of the trial design due to reducing overall randomized sample size. This chosen value of a_0 will be a function of the maximum acceptable difference between the two control groups with respect to the chosen effect parameter. For example, one may find an a_0 value that bounds Type 1 error and power at some acceptable level given the expected range of possible hazard ratios between historical and randomized controls. If no non-zero value of a_0 can be agreed upon using these simulations, then the trial may need to be designed using randomized data alone. If $a_0 = 0$ is not necessary, but no clear evidence exist for choosing a value for a_0 , then used an a_0 value of 0.5 [10]. If a value of a_0 can be agreed upon, then the trial should be designed using this agreed upon value, call it $a_{0\text{plan}}$. If the trial is designed to include historical control data, the actual analysis should involve a study of the two control datasets to ensure acceptable concordance. Also, a model selection criterion such as DIC should be used to select the actual best value of a_0 , where a_0 ranges from 0 to 1, call this value $a_{0\text{est}}$ [8]. If $a_{0\text{est}}$ is greater than $a_{0\text{plan}}$ then the study team will need to make sure the control groups are not more

different with respect to the effect parameter used during the planning stage. If the effect difference between the two control groups is more than the agreed upon limit during trial planning, the actual analysis should use $a_{0\text{plan}}$, otherwise actual analysis could use $a_{0\text{est}}$. Regardless, if historical data is incorporated into the final analysis, sensitivity analyses should be completed using a_0 values centered on the a_0 value used in final analysis and plots similar to Fig. 4 constructed.

Overall, the Bayesian model proposed in this paper, which incorporates historical data using a power prior, correctly separated initial treatment failure from treatment failure due to relapses using a multistate Markov model. Also, it enables historical control data to be incorporated into the analysis while controlling type I error inflation and increasing power.

Acknowledgments

The authors would like to thank James B. Bussell for providing study data that motivated this research and served as a basis for simulations.

Appendix A. Necessary data structure, data prep in R, and JAGS call in R to fit model

In order to use the JAGS model in Appendix B, the raw dataset from randomized subjects and historical controls needs to have data for at least six variables, one treatment variable that must be coded 0/1. One variable needs to be a subject ID. One variable that records the transition time from start of study into a given transition state, “Start.val”, and another variable records the transition time from start of study out of a given transition state, “End.val”. A third variable needs to indicate the type of transition the record represents (a 1–2 transition, 1–3 transition, or a 2–3 transition), “Trans”. Another variable, “Status”, needs to indicate transition status for the record where:

- 1 = transition observed within an interval,
- 2 = transition time is right censored but censoring time only known within an interval (for example: when a 1–2 transition is interval censored then the 1–3 transition is right censored but the right censoring time is only known within an interval),
- 3 = right censored observation with exact censoring time known.

ID	Start.val	End.val	Trans	Status	trt
1	0.0000	0.0833	12	2	1
1	0.0000	0.0833	13	1	1
2	0.3333	0.4167	12	1	0
2	0.3333	0.4167	13	2	0
2	1.1667	1.2500	23	1	0
3	0.0000	0.0833	12	1	1
3	0.0000	0.0833	13	2	1
3	0.5833	0.6667	23	1	1
4	0.0000	0.0833	12	2	0

4	0.0000	0.0833	13	1	0
5	0.8333	0.9167	12	1	1
5	0.8333	0.9167	13	2	1
5	1.1667	1.2500	23	1	1
6	0.0833	0.1667	12	1	1
6	0.0833	0.1667	13	2	1
6	3.0000	NA	23	3	1

The last value needs to be a treatment indicator, “trt”. The raw dataset needs to have one record for the 1–2 transition and one record for the 1–3 transition. In these two records the value of `Start.val` and `End.val` will be the same. Only the `Status` value may differ, depending on whether a 1–2 or 1–3 transition occurred. The data from historical controls used in this analysis had the same structure as the randomized subjects described above, yet only records for the 2–3 transition were present. I call the historical control dataset `Hcontrol.data23`. The raw dataset for the randomized subjects is divided into two datasets. One contains only the records for 1–3 transitions contained in the raw dataset. I call this dataset `Trial.data13`. The other dataset is a transformed dataset from 1–2 and 2–3 transition records. I call this dataset `Trial.data1223`. This second dataset has information about the 2–3 transition concatenated column-wise onto the data for the 1–2 transition, and it has the following structure:

ID	Start.val	End.val	Trans	Status	treatment	start.val23	End.val23	Status23
1	0.0000	0.0833	12	2	1	0.0000	0.0000	0
2	0.3333	0.4167	12	1	0	1.1667	1.2500	1
3	0.0000	0.0833	12	1	1	0.5833	0.6667	1
4	0.0000	0.0833	12	2	0	0.0000	0.0000	0
5	0.8333	0.9167	12	1	1	1.1667	1.2500	1
6	0.0833	0.1667	12	1	1	3.0000	NA	3

Once these two derived datasets are created from the raw data in R, then I also create the following vectors in R:

```
Nval1223 <- dim(Trial.data1223)[1]
tL_vec12 <- Trial.data1223$Start.val
tR_vec12 <- Trial.data1223$End.val
tL_vec12[is.na(tL_vec12)] <- (CensorValue + 0.1) #Where CensorValue is the
largest follow-up time possible.
tR_vec12[is.na(tR_vec12)] <- (CensorValue + 0.1)trt_vec1223 <-
Trial.data1223$treatment #1=active trt, 0 = control
delta12 <- Trial.data1223$Status
tL_vec23 <- Trial.data1223$Start.val23
tR_vec23 <- Trial.data1223$End.val23
tL_vec23[is.na(tL_vec23)] <- (CensorValue + 0.1)tR_vec23[is.na(tR_vec23)] <-
(CensorValue + 0.1)delta23 <- Trial.data1223$Status23
Nval13 <- dim(Trial.data13)[1]
```

```

tL_vec13 <- Trial.data13$Start.val
tR_vec13 <- Trial.data13$End.val
tL_vec13[is.na(tL_vec13)] <- (CensorValue + 0.1)tR_vec13[is.na(tR_vec13)] <-
(CensorValue + 0.1)
trt_vec13 <- Trial.data13$treatment #1=active trt, 0 = control
delta13 <- Trial.data13$Status
Nval23H <- dim(Hcontrol.data23) [1]
tL_vec23H <- Hcontrol.data23$Start.val
tR_vec23H <- Hcontrol.data23$End.val
tL_vec23H[is.na(tL_vec23H)] <- (CensorValue + 0.1)
tR_vec23H[is.na(tR_vec23H)] <- (CensorValue + 0.1)trt_vec23H <-
Hcontrol.data23$treatment #always equal to 0
delta23H <- Hcontrol.data23$Status

```

Finally, a run of the following code will fit the Bayesian model given in Appendix B and produce a summary of the hazard rate and ratio posteriors.

```

#Build up JAGS code for Power Prior model
mod.data <- list(tL_12 = tL_vec12, tR_12 = tR_vec12, trtl223 = trt_vec1223,
delta12 = delta12,
  tL_23=tL_vec23, tR_23=tR_vec23, delta23=delta23, N1223 = Nval1223,
  tL_13 = tL_vec13, tR_13 = tR_vec13, trtl3 = trt_vec13, delta13 = delta13,
N13 =
  Nval13, tL_H = tL_vec23H, tR_H = tR_vec23H, trtH = trt_vec23H, deltaH =
  delta23H, NH = Nval23H, a0 = 1.0)
#Build a list of initial values
initslist <- list(intercept12 = 1, trtbeta12 = 0, intercept13 = 1, trtbeta13
= 0,
  intercept23 = 1, trtbeta23 = 0)
#Call JAGS to build the model
library(rjags)
jags.m <- jags.model (file = [location of Bugs model), data = mod.data,
inits = initslist,
  n.chains = 3, n.adapt = 1000)
#Prepares to creates three chains of the following parameters and statistics:
#Define the parameters for which chains will be created:
params <- c("trt012", "trtl12", "trtHR12",
  "trt013", "trtl13", "trtHR13",
  "trt023", "trtl23", "trtHR23")
#Generate MCMC samples and summarize the Posteriors (11000 samples in each
chain, use first 1000
#as burn-in
samps <- coda.samples(jags.m, params, n.iter = 11000)
burn.in <- 1000

```

```
samps.sum <- summary(window(samps, start = burn.in), quantities=c(0.025, 0.5,
0.975))
samps.sum
```

Appendix B. JAGS model code

```
data {
  for (i in 1:N1223){ #Set up zero vector for group that had 12
    #transitions observed
    zeros12[i] <- 0
    zeros23[i] <- 0
  }
  for (j in 1:N13){#Set up zero vector for group that had 13
    #transitions observed
    zeros13[j] <- 0
  }
  for (j in 1:NH){#Set up zero vector for historical group
    #that had 23 transitions observed
    zerosH[j] <- 0
  }
}
{
model {
  C <- 10000 # this just has to be large enough to ensure all
    # phi[ij] 's > 0
  #Set up Likelihood component for the 12 transition
  for (i in 1:N1223) {
    logLambdal2[i] <- intercept12 + trtbetal2*trt1223[i]
    logLpart1_12[i] <- (deltal2[i]-1)*
      (log (exp {-1*exp (logLambdal2 [i]) *tL_12 [i]})-
        exp(-1*exp(logLambdal2[i])*tR_12[i])))
    #logLpart1_12 is used when the 12 transition time is
    #interval censored.
    LogLpart2_12 [i] <- (deltal2[i]==2)*{-1*exp(LogLambdal2[i])*
      TransPl2_13[i])
    #logLpart2_12 is used when the right censor time for
    #the 12 transition is predicted due to an observed 13
    #transition. Here you need to use the predicted time
    #for the 13 transition as the right censoring time
    #point for the 12 transition.
    logLpart3_12[i] <- (deltal2[i]==3)*(-1*exp(logLambdal2[i])*
      tL_12[i])
    #logLpart3_12 is used when the right censor time for
    #the 12 transition is observed and does not need
    #predicting.
```



```

# Build up the Loglikelihood for the 12 transition
logL12[i] <- logLpart1_12[i]+logLpart2_12[i]+
  logLpart3_12[i]
phil2[i] <- -logL12[i] + C
zeros12[i] ~ dpois(phil2[i])
} #end of likelihood for the 12 transition.
#Set up Likelihood component for the 13 transition
for (j in 1:N13) {
  logLambdal3 [j] <- intercept13 + trtbetal3*trt13 [j]
  logLpart1_13[j] <- (deltal3[j]==1)*
    (log(exp(-1*exp(logLambdal3[j])*tL_13[j])-
      exp(-1*exp (logLambdal3 [j]) *tR_13 [j]))))
  #logLpart1_13 is used when the 13 transition time is
  #interval censored.
  logLpart2_13[j] <- (deltal3[j]==2)*(-1*exp(logLambdal3[j])*
    TransP13_12[j])
  #logLpart2_13 is used when the right censor time for the 13
  #transition is predicted due to an observed 12 transition.
  #Here you need to use the predicted time for the 12
  #transition as the right censoring time point for the 13
  #transition.
  logLpart3_13[j] <- (deltal3[j]==3)*(-1*exp{logLambdal3[j])*
    tL_13[j])
  #logLpart3_13 is used when the right censored time for the
  #12 transition is observed and does not need predicting.
  # Build up the loglikelihood for the 13 transition
  logL13[j] <- logLpart1_13[j]+logLpart2_13[j]+
    logLpart3_13[j]
  phil3[j] <- -logL13[j] + C
  zeros13[j]~ dpois (phil3 [j])
} #end of likelihood for the 12 transition.
#Set up Likelihood component for the 23 transition
for (k in 1:N1223) {
  logLambda23[k] <- intercept23 + trtbeta23*trt1223[k]
  logLpart1_23[k] <- (deltal2[k]==1)*(delta23[k]==3)*
    (-1*exp(logLambda23[k])*(tL_23[k]-TransP12[k]))
1#logLpart1_23 represents a right censored observation.
  #Note that if deltal2[k] is zero the record contributes
  #nothing to the log likelihood. In such a case the
  #predicted 12 transition time does not influence the log
  #likelihood and in fact it shouldn't since such a record
  #represents a subject that did not transition into state 2.
logLpart2_23[k] <-
  deltal2[k]==1) *(delta23[k]==1)*(log(exp{-

```

```

1*exp(logLambda23[k]*(tL_23[k]-TransP12_12[k]))-exp(-
1*exp(LogLambda23[k]*(tR_23[k]-TransP12_12[k]))))
#logLpart2_23 represents an interval censored observation.
#Note that if delta12[k] is zero the record contributes
#nothing to the log likelihood. In such a case the
#predicted 12 transition time does not influence the log
3likelihood and in fact it shouldn't since such a record
#represents a subject that did not transition into state 2.
# Build up the loglikelihood for the 23 transition
logL23[k] <- logLpart1_23[k]+logLpart2_23[k]
phi23[k] <- -logL23[k] + C
zeros23[k] ~ dpois(phi23[k])
} #end of likelihood for the 23 transition.
#Set up Likelihood component for the 23 transition in the
#historical controls
for (j in 1:NH) {
logLambdaH[j] <- intercept23 + trtbeta23*trtH[j]
logLpart1_H[j] <- (deltaH[j]==1)*
(log(exp(-1*exp(logLambdaH[j])*tL_H[j])-
exp(-1*exp(logLambdaH[j])*tR_H[j]))))
#logLpart1_H is used when the 23 transition time for the
historical record is interval censored.
logLpart2_H[j] <- (deltaH[j]==2)*(-1*exp(logLambdaH[j]*1 *
tL_H[j]))
#logLpart2_H is used when the 23 transition time for the
#historical record is right censored.
#Set up Likelihood component for the 23 transition among
#historical controls
logLH[j] <- a0*(logLpart1_H[j]+logLpart2_H[j])
phiH[j] <- -logLH[j] + C
zerosH[j] ~ dpois(phiH[j])
} #end of likelihood for the 23 transition among historical
#controls.
#Set up the priors for the predicted transition times in the
#12 group
for (i in 1:N1223){
TransP12_12[i] ~ dexp(exp(intercept12 +
trtbetal2*trtl223[i])) T(tL_12[i],tR_12[i])
#TransP12_12 is a prediction of the 1-2 transition time.
#These predictions are only used in the 2-3 likelihood for
#those subjects who actually transition into State 2.
TransP12_13 [i] ~ dexp (exp (intercept13 +
trtbetal3*trtl223[i])) T(tL_12[i],tR_12[i])
#TransP12_13 is a prediction of the 1-3 transition time.

```

```

#Note: tL_12[i] and tR_12[i] are not just the interval
#values for an observed 1-2 transition time, but they
#represent the interval values for observed 1-2 transition
#times, actually known right censored 1-2 transition times,
#as well as interval values for the observed 1-3
#transition time which bound the right censored time of a
#1-2 transition when a 1-3 transition occurs. Every record
#in the 1-2/2-3 dataset gets a TransP12_13 value, but only
#values of TransP12_13 for those records representing an
#actual transition into State 3 instead of State 2
#contribute to the 1-2 transition likelihood.
}

#Set up the priors for the predicted transition times in the
#13 group
for (j in 1:N13){
  TransP13_12[j] ~ dexp(exp(intercept12 +
    trtbetal2*trtl3(j))) T(tL_13[j],tR_13[j]) #Predict the 12
    transition time.
  TransP13_12 is a prediction of the 1-2 transition time.
  #Note: tL_13[i] and tR_13[i] are not just the interval
  #values for an observed 1-3 transition time, but they
  #represent the interval values for observed 1-3 transition
  #times, actually known right censored 1-3 transition times,
  #as well as interval values for the observed 1-2
  #transition time which bound the right censored time of a
  #1-3 transition when a 1-2 transition occurs. Every record
  #in the 1-3 dataset gets a TransP13_12 value, but only
  #values of TransP13_12 for those records representing an
  #actual transition into State 2 instead of State 3
  #contribute to the 1-3 transition likelihood.
}

#Note: TransP12_12 and TransP13_12 are both predictions of a
#1-2 transition time, but they are used in separate parts of
#the likelihood which is segmented by those who do and do not
#transition into State 2. Perhaps a different data structure
#would allow the removal of this redundancy, but given the data
#structure used, double drawing from the 1-2 transition
#distribution was necessary. TransP13_12 represents  $\hat{t}_{12L}k$  in
#Equation 2, while TransP12_12 represents  $\hat{t}_{12L}k$  equation 5.
Also, note that TransP12_13 represents  $\hat{t}_{13L}k$  #Equation 1.
#Set up the priors for the 12 transitions
intercept12 ~ dnorm(0,0.001)
trtbetal2 ~ dnorm(0,0.001)
trt012 <- exp(intercept12)

```

```

trtl12 <- exp(intercept12+trtbetal2)
trtHR12 <- exp(trtbetal2)
#Set up the priors for the 13 transitions
intercept13 ~ dnorm (0,0.001)
trtbetal3 ~ dnorm(0,0.001)
trt013 <- exp(intercept13)
trtl13 <- exp(intercept13+trtbetal3)
trtHR13 <- exp(trtbetal3)
#Set up the priors for the 23 transitions
intercept23 ~ dnorm(0,0.001)
trtbeta23 ~ dnorm(0,0.001)
trt023 <- exp(intercept23)
trtl23 <- exp(intercept23+trtbeta23)
trtHR23 <- exp(trtbeta23)
} #end JAGS model code

```

References

1. Allison, PD. Survival Analysis Using the SAS System, A Practical Guide. SAS Institute, Inc.; Cary NC: 1995.
2. Berry, SM., Carlin, BP., Lee, JJ., Müller, P. Bayesian Adaptive Methods for Clinical Trials. CRC Press; Florida: 2011.
3. Beyersmann, J., Allignol, A., Schumacher, M. Competing Risks and Multistate Models. R. Springer; New York: 2012.
4. Bolstad, WM. Introduction to Bayesian Statistics. Wiley; New Jersey: 2007.
5. Cines DB, Blanchette VS. Immune thrombocytopenic purpura. N Engl J Med. 2002; 346(13):995–1008. [PubMed: 11919310]
6. Hogg, RV., Craig, AT. Introduction to Mathematical Statistics. fifth. Prentice Hall; New Jersey: 1995.
7. Ibrahim JG, Chen MH. Power prior distributions for regression models. Stat Sci. 2000; 15:46–60.
8. Ibrahim JG, Chen MH, Chu H. Bayesian methods in clinical trials: a Bayesian analysis of ECOG trials E1684 and E1690. BMC Med Res Methodol. 2012; 12:183. [PubMed: 23194570]
9. Ibrahim, JG., Chen, MH., Sinha, D. Bayesian Survival Analysis. Springer; New York: 2001.
10. Ibrahim JG, Chen MH, Sinha D. On optimality properties of the power prior. J Am Stat Assoc. 2003; 98(461):204–213.
11. Lawless, J. Statistical Models and Methods for Lifetime Data. second. Wiley; New Jersey: 2003.
12. Spiegelhalter, DJ., Abrams, KR., Myles, JP. Bayesian Approaches to Clinical Trials and Healthcare Evaluation. Wiley; Chichester, UK: 2004.
13. Therneau, TM., Grambsch, PM. Modeling Survival Data, Extending the Cox Model. Springer; New York: 2000.
14. Viele K, Berry S, Neuenschwander B, Amzal B, Chen F, Enas N, Hobbs B, Ibrahim J, Kinneersley N, Lindborg S, Micallef S, Roychoudhury S, Thompson L. Use of historical control data for assessing treatment effects in clinical trials. Pharm Stat. 2014; 13:41–54. [PubMed: 23913901]

Fig. 1.
Illness-Death model for ITP.

Fig. 2. Power curves for Bayesian estimation model, assuming same hazard rate among historical and randomized controls.

Fig. 3.
Type one error by a_0 .

Fig. 4.

Posterior plots for 2–3 hazard ratio when (a) randomized controls have a hazard rate equal to 0.3 and historical controls have a hazard rate of 0.5, (b) randomized and historical controls both have a hazard rate of 0.5, (c) randomized controls have a hazard rate equal to 0.7 and historical controls have a hazard rate of 0.5.

Fig. 5.

Power with 2–3 transition rate among historical controls is equal to 0.5, among randomized controls is equal to 0.3, and in randomized active subjects is equal to 0.18 (hazard ratio of 0.6).

Fig. 6. Power with 2–3 transition rate among historical controls is equal to 0.5, among randomized controls is equal to 0.7, and in randomized active subjects is equal to 0.42 (hazard ratio of 0.6).

Fig. 7.

Type 1 error by a_0 . (+40%) implies a 2–3 transition rate among historical controls equal to 0.5 and a 2–3 transition rate among randomized controls equal to 0.7. (–40%) implies a 2–3 transition rate among historical controls equal to 0.5 and a 2–3 transition rate among randomized controls equal to 0.3. (0%) implies a 2–3 transition rate among historical and randomized controls equal to 0.5. In all cases the 2–3 transition ratio between randomized active subjects and randomized controls is equal to 1.0.

Table 1

Hypotheses which can be addressed using the basic Bayesian model.

$H_{01}:\beta_{121} = 0$	$H_{02}:\beta_{131} = 0$	$H_{03}:\beta_{231} = 0$
$H_{a1}:\beta_{121} \neq 0$	$H_{a2}:\beta_{131} \neq 0$	$H_{a3}:\beta_{231} \neq 0$

Table 2

Example analysis results.

	1–2 transition	1–3 transition	2–3 transition
Median hazard rate, controls	3.28 (2.63, 4.02)	0.86 (0.55, 1.26)	0.50 (0.41, 0.60)
Median hazard rate, active	3.28 (2.64, 4.02)	0.86 (0.55, 1.26)	0.30 (0.23, 0.40)
Median hazard ratio (HR)	1.00 (0.74, 1.35)	1.00 (0.56, 1.80)	0.61 (0.43, 0.84)
Posterior	0.50	0.50	0.999
Pr (HR < 1.0)			
Plots of Hazard Ratio Posteriors

Table 3

Effects of historical/randomized control discordance on 2–3 transition hazard ratio. Hazard rate among controls is always 0.5.

	Randomized control hazard rate = 0.3		Randomized control hazard rate = 0.5		Randomized control hazard rate = 0.7	
	Median	95% CI	Median	95% CI	Median	95% CI
Trial data only	0.60	(0.39, 0.93)	0.60	(0.41, 0.87)	0.60	(0.42, 0.84)
Trial & historical data, $a_0 = 0.2$	0.56	(0.36, 0.85)	0.60	(0.42, 0.86)	0.63	(0.45, 0.88)
Trial & historical data, $a_0 = 0.4$	0.53	(0.35, 0.80)	0.60	(0.42, 0.85)	0.65	(0.47, 0.90)
Trial & historical data, $a_0 = 0.6$	0.51	(0.34, 0.76)	0.61	(0.43, 0.85)	0.67	(0.48, 0.92)
Trial & historical data, $a_0 = 0.8$	0.49	(0.33, 0.73)	0.61	(0.43, 0.84)	0.69	(0.50, 0.93)
Trial & historical data, $a_0 = 1.0$	0.48	(0.32, 0.70)	0.61	(0.43, 0.84)	0.70	(0.51, 0.95)