

Bayesian approach to the assessment of the population-specific risk of inhibitors in hemophilia A patients: a case study

Ji Cheng^{1,2}
Alfonso Iorio^{2,3}
Maura Marcucci⁴
Vadim Romanov⁵
Eleanor M Pullenayegum^{6,7}
John K Marshall^{3,8}
Lehana Thabane^{1,2}

¹Biostatistics Unit, St Joseph's Healthcare Hamilton, ²Department of Clinical Epidemiology and Biostatistics, ³Department of Medicine, McMaster University, Hamilton, ON, Canada; ⁴Geriatrics, Fondazione Ca' Granda Ospedale Maggiore Policlinico, Università degli Studi di Milano, Milan, Italy; ⁵Baxter HealthCare, Global Medical Affairs, Westlake Village, CA, USA; ⁶Child Health Evaluation Sciences, Hospital for Sick Children, ⁷Dalla Lana School of Public Health, University of Toronto, Toronto, ⁸Division of Gastroenterology, Hamilton Health Science, Hamilton, ON, Canada

Background: Developing inhibitors is a rare event during the treatment of hemophilia A. The multifacets and uncertainty surrounding the development of inhibitors further complicate the process of estimating inhibitor rate from the limited data. Bayesian statistical modeling provides a useful tool in generating, enhancing, and exploring the evidence through incorporating all the available information.

Methods: We built our Bayesian analysis using three study cases to estimate the inhibitor rates of patients with hemophilia A in three different scenarios: Case 1, a single cohort of previously treated patients (PTPs) or previously untreated patients; Case 2, a meta-analysis of PTP cohorts; and Case 3, a previously unexplored patient population – patients with baseline low-titer inhibitor or history of inhibitor development. The data used in this study were extracted from three published ADVATE (antihemophilic factor [recombinant] is a product of Baxter for treating hemophilia A) post-authorization surveillance studies. Noninformative and informative priors were applied to Bayesian standard (Case 1) or random-effects (Case 2 and Case 3) logistic models. Bayesian probabilities of satisfying three meaningful thresholds of the risk of developing a clinical significant inhibitor (10/100, 5/100 [high rates], and 1/86 [the Food and Drug Administration mandated cutoff rate in PTPs]) were calculated. The effect of discounting prior information or scaling up the study data was evaluated.

Results: Results based on noninformative priors were similar to the classical approach. Using priors from PTPs lowered the point estimate and narrowed the 95% credible intervals (Case 1: from 1.3 [0.5, 2.7] to 0.8 [0.5, 1.1]; Case 2: from 1.9 [0.6, 6.0] to 0.8 [0.5, 1.1]; Case 3: 2.3 [0.5, 6.8] to 0.7 [0.5, 1.1]). All probabilities of satisfying a threshold of 1/86 were above 0.65. Increasing the number of patients by two and ten times substantially narrowed the credible intervals for the single cohort study (1.4 [0.7, 2.3] and 1.4 [1.1, 1.8], respectively). Increasing the number of studies by two and ten times for the multiple study scenarios (Case 2: 1.9 [0.6, 4.0] and 1.9 [1.5, 2.6]; Case 3: 2.4 [0.9, 5.0] and 2.6 [1.9, 3.5], respectively) had a similar effect.

Conclusion: Bayesian approach as a robust, transparent, and reproducible analytic method can be efficiently used to estimate the inhibitor rate of hemophilia A in complex clinical settings.

Keywords: inhibitor rate; meta-analysis; multicentric study; Bayesian; hemophilia A

Background

Challenges in estimating the development of inhibitors

The development of inhibitory antibodies, usually IgG 4, is a natural response of the host immune system to exogenous factor VIII in patients with hemophilia, who are devoid of endogenous factor VIII and therefore not naturally tolerant to it. Some forms of immunological reaction are observed in more than half of the patients at the first exposure to factor VIII. In many cases, the antibodies are able to block factor VIII

Correspondence: Lehana Thabane
Biostatistics Unit, St Joseph's Healthcare
Hamilton, Third Floor Martha Wing,
Room H325, 50 Charlton Avenue East,
Hamilton, ON L8N 4A6, Canada
Tel +1 905 522 1155 Ext. 33720/34905
Fax +1 905 528 7386
Email thabanl@mcmaster.ca

submit your manuscript | www.dovepress.com

Dovepress    

<http://dx.doi.org/10.2147/JBM.S103087>

Journal of Blood Medicine 2016:7 239–253

239



© 2016 Cheng et al. This work is published and licensed by Dove Medical Press Limited. The full terms of this license are available at <https://www.dovepress.com/terms.php> and incorporate the Creative Commons Attribution – Non Commercial (unported, v3.0) License (<http://creativecommons.org/licenses/by-nc/3.0/>). By accessing the work you hereby accept the Terms. Non-commercial uses of the work are permitted without any further permission from Dove Medical Press Limited, provided the work is properly attributed. For permission for commercial use of this work, please see paragraphs 4.2 and 5 of our Terms (<https://www.dovepress.com/terms.php>).

clotting activity, and their presence can be detected and quantitated using a specific test, the Bethesda assay. Inhibitory responses of more than five Bethesda units are called high responding antibodies. Approximately one-third of inhibitors disappear spontaneously, while others are persistent, making exogenous factor VIII ineffective and requiring the use of bypassing agents to stop bleeding or prolonged administration of high-dose factor VIII to eradicate the inhibitor (immune tolerance induction treatment).

Developing inhibitors against factor VIII concentrates is the most severe and costly complication of the treatment of hemophilia A.¹ Patients who develop inhibitors tend to have more severe bleeds, either because they are often treated on demand more than prophylactically, or because their bleeds are more difficult to stop. As to the utilization of factor concentrates, this might be lower, when patients are preferentially treated with bypassing agents, or higher, when low-titer or subclinical inhibitors are treated with factor VIII, usually requiring a neutralizing dose (formally calculated or empirically determined) in addition to the usual dose of concentrates required to prevent bleeding.

There are several reasons that complicate studying the determinants of inhibitor development.² The first is that the development of inhibitory antibodies is a combination of different events, more than a single one, with nothing as simple as black and white.³⁻⁶ The second is the uncertainty of explaining both known and unknown risk factors associated with the development of inhibitors.^{7,8} The third is the rarity of the disease, which hampers the opportunity to obtain substantial comparative data.²

In this challenging scenario, it is important to determine the risk associated with specific brands or classes of factor concentrates, because the type of product is one of the few actionable risk factors in the field.^{9,10} Other characteristics of the treatment regimen such as dose, frequency, indication, and concomitant treatments or exposures also contribute to the risk of inhibitor development.¹¹⁻¹³ Progress in this field requires a close collaboration with complementary expertise. Knowledge of immunology and basic science can help gain a broader and deeper understanding of the molecular and cellular mechanisms driving the development or breach of tolerance.¹⁴⁻¹⁷ Clinical investigators can work to dissect the common characteristics among the heterogeneous clinical manifestations of inhibitory responses. Epidemiologists and biostatisticians can develop more powerful and efficient ways of looking at the available data and generating new ones.

There are several unmet needs in the statistical models used to analyze observational data about inhibitor development,

which relate to the rarity of adverse events in an already rare disease.¹⁸⁻²⁴ The first critical issue is the scarcity of evidence, which emphasizes the need for incorporating external evidence to increase the power and the informative value of small and otherwise weak cohorts.¹⁹ A second issue is the need for an efficient way to analyze the intricate relationship between treatment, time, and the varying risk of events over time.²⁵ A third is the need to adjust for covariates (known risk factors) when performing multivariable exploration of, for example, inhibitor rates in previously untreated patients (PUPs). The fourth and last is the proper assessment and comparison of event rates generated by nonparallel cohorts.²⁶⁻²⁸ In the current article, we address the first and fourth issues.

A powerful approach to the abovementioned problems might be a Bayesian framework. The Bayesian approach to interpreting experimental data from a clinical study consists of modeling the logical process leading to a change in opinion from before to after the availability of new information (the evidence provided by a new observation).

The benefit of the Bayesian approach derives from the opportunity of making use of existing knowledge in the assessment of data that extends to either incorporating that knowledge in the final results or using it as a standard to quantify the comparison between the new evidence and the previous knowledge. That knowledge could be a similar measure in a similar unrelated trial, or a threshold of clinical importance.

The concepts behind Bayesian inference

Statistical inference is the process of fitting a probability model to a set of observed samples from a population to summarize the results by a probability distribution on the parameters of interests to make a general statement about the population and predictions for new observations. In the classical (frequentist) approach, the statistical modeling involves only fitting a probability distribution to the observed experimental data to model the likelihood of the observed experimental data for a given estimate of interest such as treatment effect and incidence rate. Unlike the classical approach, the Bayesian approach combines experimental and prior or external information via the Bayes theorem, to produce the posterior distribution that is used to make all inferences about the estimate of interest.

$$p(\delta | \text{data}) \sim p(\text{data} | \delta) \times p(\delta) \quad (1)$$

where δ is the parameter of interest.

$$\begin{aligned} \text{Posterior distribution} &\sim \text{data likelihood} \\ &\times \text{prior distribution} \end{aligned} \quad (2)$$

As shown in Equations 1 and 2, $p(\delta)$ represents prior distribution of the parameter of interest (henceforth to be referred to as “the parameter”), ie, the prior or external information about the estimate of treatment effect, incidence rate, etc; $p(\text{data}|\delta)$, the likelihood function, specifies the statistical model of the observed experimental data given the parameter; and $p(\delta|\text{data})$ is the posterior distribution of the parameter, which is essentially a combination of the evidence provided by the observed experimental data and prior relevant data from clinical experience or past research evidence.^{29,30} In many cases, the posterior distribution $p(\delta|\text{data})$ is intractable, and therefore to make inferences about the parameter the Bayesian approach uses Monte Carlo Markov Chain to obtain samples from the posterior $p(\delta|\text{data})$.³¹ Monte Carlo Markov Chain is an iterative process, with each iteration yielding a realization or observation from the posterior distribution $p(\delta|\text{data})$. Typically, investigators will conduct a large number of iterations or simulations: 1,000–10,000, or even more. These are used to inform posterior inferences about the parameter. For example, the posterior mean or median is used to estimate the parameter,

while the 2.5th and 97.5th observations are used as the 95% credible interval (CrI) for the parameter. The probability of the point estimate less than a certain threshold is abstained by calculating the proportion of the cases lower than the threshold over the total number of the interactions. Table 1 provides a brief summary of the comparison of main features of the frequentist and Bayesian approaches in clinical trials. The difference between Bayesian and frequentist statistical inference is presented in Table 1, and the definitions and descriptions of Bayesian terminologies used in this article are given in Table 2.

Objectives of this article

By using real clinically observed data from the recently published articles,^{32–34} we aimed to demonstrate the benefits or advantages of Bayesian approach in estimating the risk of developing inhibitors among patients with hemophilia A from the following aspects: increasing the credibility of the results and the understanding of the underlying mechanisms by incorporating external evidence; generating probabilities to be used in a physician-to-patient interaction.

Table 1 Brief comparison of the frequentist and Bayesian approaches in clinical trials

| Feature | | Frequentist approach | Bayesian approach |
|--|--|--|--|
| Interpretation of probability | | The proportion of times an event will occur in an infinitely long series of repeated identical situations | The “degree of belief” an event (or a number of repeatable events) will occur |
| Main question | | What is the probability of data (trial result), given the hypothesis (treatment effect)? | What is the probability of the hypothesis (treatment effect), given the data (trial result)? |
| Design features | | Hypotheses, type I and II errors | Hypotheses, prior or external information |
| Reasoning paradigm | | Deductive reasoning | Inductive reasoning |
| Trial monitoring | | Prespecified with adjustments for type I error for interim analyses | Adaptive by design based on accumulating evidence |
| Condition of drawing statistical inference | | Inference based on observed experimental data | Inference based on observed experimental data and prior information |
| Information for analysis | Use of external information/pre-belief | Informally considered only at study design stage, eg, sample size calculation | Formally incorporated in the design, analysis, and interpretation as a prior |
| | Experimental data | Summarized via the likelihood function, which captures all information provided by the observed data regarding any unknown population parameters | Summarized via the likelihood function, which captures all information provided by the observed data regarding any unknown population parameters |
| Results summaries | Point estimate | The “best estimate” obtained from observed experimental data | A weighted point estimate from the posterior distribution derived by combining all relevant sources of information including the external information and observed experimental data |
| | Interval estimates | 95% CI: an interval that we are 95% confident that the true value of the unknown parameter would be as low as its lower bound and as high as its upper bound | 95% CrI: an interval has a 0.95 probability that the unknown parameter would lie within, given the observed experimental data |
| | Probabilities | P-value, the chance of observing a result as extreme as what is seen in the experiment when the null hypothesis of no effect is true | Posterior probabilities |
| Decision making | Framework | Not straightforward and hard to apply in clinical practice | Intuitive and based on minimizing expected losses; easy to apply in clinical practice |

Abbreviations: CI, confidence interval; CrI, credible interval.

Table 2 Definitions of Bayesian statistical inference

| Terms | Definition | Examples of interpretation |
|--------------------------------|--|--|
| Bayesian statistical inference | $p(\delta \text{data}) \sim p(\text{data} \delta) \times p(\delta)$, posterior distribution is the joint distribution of prior distribution and the distribution of data likelihood | Obtaining the most updated evidence by combining the information from the study data with other sources |
| Prior | $p(\delta)$, the probability distribution of the unknown parameter of interest | Preexisting knowledge or evidence, or external information of treatment effect or incidence rate |
| Noninformative prior | A prior distribution is called noninformative if the prior has little information to impact the posterior distribution | A noninformative prior is also called vague or flat prior |
| Informative prior | A prior is called informative if the prior contains some information which impacts the posterior distribution | The information presented in the prior can be the external evidence, pre-believe, or expert opinion |
| Likelihood function | $p(\text{data} \delta)$, the probability distribution of the data observed over the parameter of interest, which embeds the statistical model | Data obtained for a study designed to investigate certain treatment effects or diseases |
| Posterior | $p(\delta \text{data})$, the updated probability distribution of the parameter of the interest given data, which is obtained by combining data likelihood and prior distribution | Updated evidence after combining the evidence from the study data and other information |
| 95% CrI | 95% CrI, the estimated interval has a 95% probability (credibility) that the parameter of interest lies within it | 95% probability is a direct quantity attached to the obtained interval of the posterior estimates of the parameter of interest |

Abbreviations: CrI, credible Interval; δ , parameter of interest.

Methods

Overall study design

This article is built around three case studies and uses a standardized multistep approach to show 1) how the Bayesian results compare with those based on the commonly used classical approach; 2) the impact of different sources of external information used to construct a Bayesian prior; 3) the use of different sources of external information as thresholds against which to benchmark Bayesian posterior estimates of risk; and 4) the impact of the size of information on the Bayesian posterior estimates. In this section, we describe the three case studies, the statistical details, and the data source we used for the simulation.

Case study scenarios and patient populations

Case 1: analyzing a rare adverse event in a single cohort

The first case was set to represent the analysis of a single study, where all patients were treated with the same FVIII product, aiming to assess the rate of inhibitor development in the cohort of patients with hemophilia A including previously treated patients (PTPs) and PUPs. For this case, we reanalyzed the same cohort already published by Oldenburg et al.³² We used this case to explore and discuss the basics of the Bayesian approach and the pros and cons of choosing different priors.

Case 2: analyzing a rare adverse event by pooling a set of studies in a meta-analysis

The second case was conceived to represent a meta-analysis of studies assessing the rate of inhibitors in a set of

independent but similar studies in comparable populations of patients with hemophilia A; for this case, we used an earlier article we published.³³ The main goal of this case was to show how the evidence synthesized from meta-analysis can be further enhanced using Bayesian approach by incorporating external information through priors.

Case 3: analyzing the inhibitor rate in a previously unexplored setting

The third case illustrates how Bayesian method can be used to explore a new clinical setting for which no obvious priors are available in the literature. Although the study design and data collection were similar to the second case, the patient population was defined as the patients with low-titer inhibitors at baseline or positive personal history of inhibitors, which was definitely different and not directly comparable to any existing ones.³⁴ The challenge presented in this case was how to properly choose priors to explore an unstudied population. In addition to that, the data used in this case were extremely sparse in a multicenter/meta-analytical setting, where no outcomes were observed in some centers/studies (the so-called zero event).

Source data and outcome

The individual data sets used to build our cases for illustration purposes^{32–34} were from the ADVATE post-authorization safety studies (PASS) program. The study population in PASS studies were hemophilia A patients undergoing treatment (prophylaxis or on-demand) with ADVATE in routine clinical use. The primary safety outcome in these

studies is defined as inhibitor levels during the study period including de novo, recurrent, and persistent inhibitors. We adopted the cutoffs of inhibitor classification specified in the original PASS protocols: 1.0 Bethesda unit (BU) for US, EU, and Australian PASS; and 0.6 BU for Japan, Italian, Korean, and Taiwan PASS (studies adopting the Nijmegen modification).^{32,33} More detailed description of the data such as hemophilia severity and FVIII exposure history and the case-specific definitions of inhibitors are presented in Table 3.

Setting basic models with noninformative priors

For all the three cases, Bayesian statistical models with non-informative priors were introduced first as the basic starting model. For Case 1, classical logistic and Bayesian logistic models were used. For Case 2 and Case 3, classical random-effects logistic and Bayesian hierarchical random-effects logistic models were adopted, through which the patients from the same study were treated as within the same cluster. The random-effects model was adopted as the most common choice for individual patient data (IPD) meta-analyses. As for noninformative prior for log odds, the normal distribution with mean equals to zero (equal odds between developing or not developing inhibitor) and precision equals to 0.00001 were specified. The uniform distribution was used to specify the prior for the between-study standard deviation in the random-effects logistic model. The details of prior specifications are presented in Table S1.

Choosing informative priors

As a second step, we replaced noninformative priors with information-rich priors to incorporate the preexisting external information or knowledge from earlier studies into the analysis of the current study data. Unlike noninformative priors, informative ones contribute information to the posterior estimates, which can be looked at as a “combination” of the preexisting evidence with evidence generated by the current experiment. To this scope, we sought relevant comparable priors and tested two different sets of informative priors. The first set 1) comprised the data obtained during the treatment with a certain molecule (eg, rAHF-PFM) in different studies; specifically i) estimates of inhibitor rates from the manufacturer’s pivotal studies³⁵; ii) estimates of inhibitor rates from a meta-analysis³⁶; and iii) estimates of inhibitor rates from an independent prospective multicentric cohort.³⁷ The second set of informative priors 2) comprised pooled inhibitor rates for any FVIII concentrate, including: i) a meta-analysis³⁶ and ii) an independent prospective multicenter cohort.³⁷ The details of generating informative priors, published data, are presented in Table S1.

In this study, Case 3 was specifically chosen not to have a study on the same patient population already available; thus, no obvious informative priors can be located in the literature. Notwithstanding, we wanted to show the value of the Bayesian approach in exploring how the rate of inhibitor development in this population would change when the known rate in PTPs and that in PUPs is added in. Consequently, in addition

Table 3 Information of study data and outcome

| Study information | Case 1 (n=428) | Case 2 (n=1,188) | Case 3 (n=219) |
|--|---------------------------|--|---|
| Patient population | PTP and PUP, all severity | PTP and PUP, moderate-to-severe patients | Patients with baseline low-titer inhibitor or history of inhibitor |
| Source of PASS data | Europe | Australia Europe Italy Japan The USA | Australia Europe Italy Japan The USA Taiwan South Korea |
| Hemophilia severity | 308 (72.0) | 883 (74.3) | 198 (90) |
| Severe, n (%) | | | |
| FVIII exposure history | 393 (91.8) | 1089 (91.6) | 214 (97.7) |
| >50 EDs, n (%) | | | |
| Defined outcome: developing inhibitor ^a , n (%) | 6 (1.4) | 21 (1.8) | 6 (2.7) |

Notes: ^aInhibitor defined in Case 1: inhibitor >1.0 BU; inhibitor defined in Case 2: de novo, recurrent, or persistent inhibitor; inhibitor defined in Case 3: titer increase or inhibitor recurrence. De novo inhibitor was defined as inhibitor occurring in patient with negative history of inhibitors and with negative titer at baseline; recurrent inhibitor was defined as the inhibitor observed for the patient with a history of inhibitor; persistent inhibitor was defined as the inhibitor that was present at the beginning of the study and carried through during the study.

Abbreviations: BU, Bethesda unit; ED, exposure day; PASS, post-authorization safety studies; PTP, previously treated patient; PUP, previously untreated patient.

to the informative priors used in the first two cases, we also added the inhibitor rates for PUPs reported in the European Hemophilia Safety Surveillance (EUHASS) study³⁷ for 1) the specific molecule and for 2) all products. The key rationale was to assess the robustness or sensitivity of the posterior inhibitor rates seen when a prior based on a truly high-risk population is used, ie, the worst-case scenarios. The details of generating priors for Case 3 are presented in Table S1.

Calculating probabilities

To make this more evident and show another peculiar property of the Bayesian framework, we further calculated, for the third case, the Bayesian probabilities of the posterior inhibitor rates being lower than three specific clinically meaningful thresholds, two high rates (10/100 and 5/100)³⁸ recommended by International Society on Thrombosis and Haemostasis (ISTH), and the US Food and Drug Administration (FDA)-mandated cutoff rate in PTPs (1/86).^{39,40} The probability was calculated as the cumulated proportion of the posterior inhibitor rate less than the threshold over the total number of iterations, ie, the number of resamplings used in each Bayesian model.

Weighting prior information

We then moved to show the effect of scaling down or discounting the value of the prior information in Case 3, for which due to the inability to use a full consistent informative prior, one might want to assign less weight to the information carried by the selected priors. The weight of the prior was reduced in the following ways: 1) by decreasing the precision, ie, enlarging the variance of the priors depending on how relevant the particular piece of information is to the study we are assessing.^{30,41} In our third case, we discount the precision of the rates of inhibitor in PUPs in EUHASS for the specific molecule and for any factor VIII concentrate by 75% and 95% each, respectively. This is equal to the human process of any perceived information: “you told me that the rate of event is this, but I only 25% trust your information”; and 2) a second approach to obtain the same objective, ie, to undervalue the contribution of the priors, is scaling up the weight of the study data by increasing the precision assigned to the experimental data. One easy and understandable way to do this is to simulate the impact on posterior estimates of increasing the study sample size. Thus, we showed the effect of increasing the study data sample size by two and ten times for all three study cases. The increment of sample was performed in two ways: 1)

for all the three cases, increasing the number of events and number of patients in each center proportionally, ie, multiplying the numbers of events and patients for each center by two or ten times; and 2) for Cases 2 and 3, increasing the numbers of centers by two or ten times while keeping the numbers of events and patients in each center as the same as in the original data. We reran the Bayesian model with noninformative priors using the purposely inflated data for all the three cases to test the idea and then further reanalyzed the effect of data inflation for Case 3 with all informative priors previously used.

Analysis and reporting

Throughout this study, posterior inhibitor rates (our results) were reported as percentage rates with 95% associated confidence interval (CI) in the case of classic statistics, or 95% CrI in the case of Bayesian statistics. Graphic, descriptive statistics and classical meta-analyses were performed using STATA 13.1 (StataCorp LP, College Station, TX, USA). Bayesian analyses were performed using WinBUGS software 1.4.3 (<http://www.mrc-bsu.cam.ac.uk/bugs/>). In every Bayesian analysis, two chains were run simultaneously, and the convergence of the Bayesian models was assessed based on the history trace, posterior density, and autocorrelation plots for parameters of interest. The codes of Bayesian models that detailed setups on the Bayesian simulations are presented in the Table S2.

Results

Raw inhibitor rates in three cases

For Case 1, six inhibitors (1.4%) were reported among 428 patients. For Case 2, five studies were included in the IPD meta-analysis, and 21 inhibitors (1.8%) were reported in 1,188 patients. For Case 3, IPD were extracted from seven PASS, and six inhibitors (2.7) were reported in 219 patients (Table 4).

Comparing results from Bayesian approach to classical approach

As expected, the results obtained from classical analytical approach and Bayesian statistical model using noninformative priors were similar for all the three cases. For the single cohort study (Case 1), the estimates were the same to one decimal place (percent rate [95% CI]: 1.3% [0.5, 2.7]; percent rate [95% CrI]: 1.3% [0.5, 2.7]). For the pooled analysis (Case 2), the Bayesian posterior estimates gave a slightly wider 95% CrI (1.9% [0.8, 4.5]) toward the lower end as compared to the

Table 4 Inhibitor rate (on percentage scale) for three different cases

| Approaches and priors | Case 1 | Case 2 | Case 3 |
|--|---------------------------------|------------------------------------|--|
| Method | Single study | Meta-analysis | Multicenter cohort – no appropriate priors |
| Test data (number of inhibitors/number of patients) | PASS data ³⁰ (6/428) | PASS data ³¹ (21/1,188) | PASS data ³² (6/219) |
| Classical statistical analysis: percent rate (95% CI) | 1.3 (0.5, 2.7) | 1.9 (0.8, 4.5) | 2.6 (1.0, 6.8) |
| Bayesian statistical analysis: percent rate (95% CrI) | | | |
| Non-informative prior | 1.3 (0.5, 2.7) | 1.9 (0.6, 6.0) | 2.3 (0.5, 6.8) |
| Informative prior: Baxter pivot study (1/102) | 1.3 (0.5, 2.5) | 1.6 (0.6, 4.1) | 1.8 (0.5, 4.8) |
| Informative prior: meta-analysis of OS (seven ADVATE studies; 3/569) | 0.9 (0.4, 1.9) | 1.0 (0.4, 2.2) | 0.9 (0.3, 2.3) |
| Informative prior: meta-analysis of OS (38/3,866) | 1.0 (0.8, 1.4) | 1.0 (0.8, 1.4) | 1.0 (0.8, 1.4) |
| Informative prior: EUHASS study of de novo inhibitor in PUPs, ADVATE (37/141) | N/A | N/A | 23.4 (17.5, 30.7) |
| Informative prior: EUHASS study of de novo inhibitor in PUPs (108/417) | N/A | N/A | 24.9 (21.1, 29.2) |
| Informative prior: EUHASS study of inhibitors in PTPs, ADVATE (5/707) | 1.0 (0.5, 1.8) | 1.1 (0.5, 2.1) | 1.0 (0.4, 2.1) |
| Informative prior: EUHASS study of inhibitors in PTPs (all FVIII) (26/3,736) | 0.8 (0.5, 1.1) | 0.8 (0.5, 1.1) | 0.7 (0.5, 1.1) |
| Discounted prior: discounting EUHASS in PUPs, ADVATE by 75% | N/A | N/A | 16.9 (9.0, 29.4) |
| Discounted prior: discounting EUHASS in PUPs, ADVATE by 95% | N/A | N/A | 5.3 (2.2, 16.0) |
| Discounted prior: discounting EUHASS in PUPs, all by 75% | N/A | N/A | 22.2 (15.7, 30.4) |
| Discounted prior: discounting EUHASS in PUPs, all by 95% | N/A | N/A | 12.3 (5.4, 25.8) |
| Enhanced data: enhancing study data by two times – increasing number of patients (with noninformative prior) | 1.4 (0.7, 2.3) | 2.0 (0.6, 6.4) | 2.2 (0.5, 6.6) |
| Enhanced data: enhancing study data by two times – increasing number of studies (with noninformative prior) | N/A | 1.9 (0.9, 4.0) | 2.4 (0.9, 5.0) |
| Enhanced data: enhancing study data by ten times – increasing number of patients (with noninformative prior) | 1.4 (1.1, 1.8) | 2.1 (0.6, 6.6) | 1.6 (0.4, 5.4) |
| Enhanced data: enhancing study data by ten times – increasing number of studies (with noninformative prior) | N/A | 1.9 (1.5, 2.6) | 2.6 (1.9, 3.5) |

Abbreviations: CI, confidence interval; CrI, credible interval; EUHASS, European Hemophilia Safety Surveillance; OS, observational study; PASS, post-authorization safety studies; PTP, previously treated patient; PUP, previously untreated patient; N/A, not applicable.

estimate from the classical approach (1.9% [0.8, 4.5]). For the cohort of patients with previous/current inhibitor (Case 3), the posterior estimates from the Bayesian model showed a slightly lower percent rate and wider 95% CrI (2.3% [0.5, 6.8]) as compared to the estimate from classical approach (2.6% [1.0, 6.8]; Table 4, Figure 1A–C).

Impact of using informative priors in Bayesian analysis

For Case 1 and 2, using external information as priors consistently narrowed the CrIs and lowered the central estimate of percentage rates. The range of the inhibitor percentage rates for the single cohort (Case 1) was 0.8%–1.3% and for the pooled analysis was 0.8%–1.9%. For the cohort of patients with previous/current inhibitor (Case 3), the posterior estimates of inhibitor percentage rates changed depending on the external information brought in by priors. The lowest percentage rates with narrowest 95% CrI was obtained using the informative prior generated from the EUHASS study of PTPs for all FVIII products: 0.7% (0.5, 1.1). The highest percentage rate was gained using the

informative priors generated from the EUHASS study of PUPs for all FVIII products: 24.9% (21.1, 29.2; Table 4, Figure 1A–C).

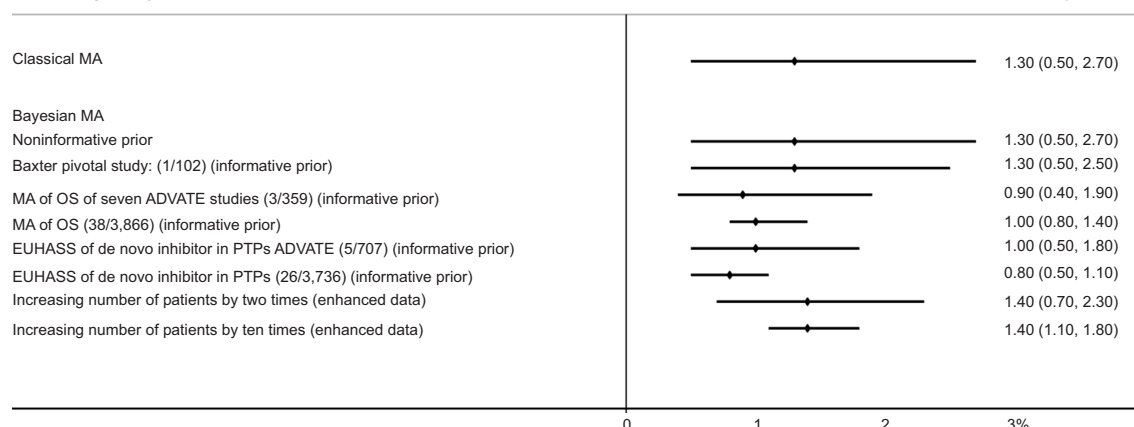
Comparing posterior inhibitor rates to thresholds

For the cohort of patients with previous/current inhibitor (Case 3), the posterior estimates of inhibitor rates were compared to the selected thresholds, and the probabilities of posterior inhibitor rates lower than the thresholds were then calculated. A probability of 1 means that the calculated rate of inhibitors is certainly below the threshold, and a probability of 0.5 means a 50% likelihood that the rate is below the threshold. Testing a threshold for the rate of inhibitors of 10%, six out of eight comparisons (when noninformative prior and informative priors were generated from the studies of PTPs) showed a probability >0.99. In contrast, when the EUHASS PUP study results were used as priors, the probabilities of a rate lower than 10% dropped dramatically to <0.001. Similar findings were obtained when the threshold was dropped to lower than 5%. When

A

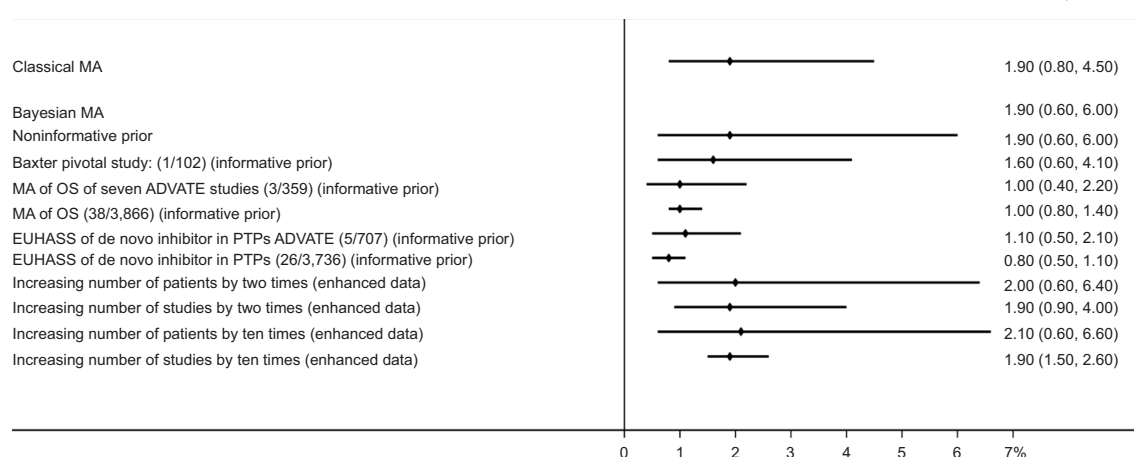
Case 1: Single study

Percent rate (95% CI/CrI)

**B**

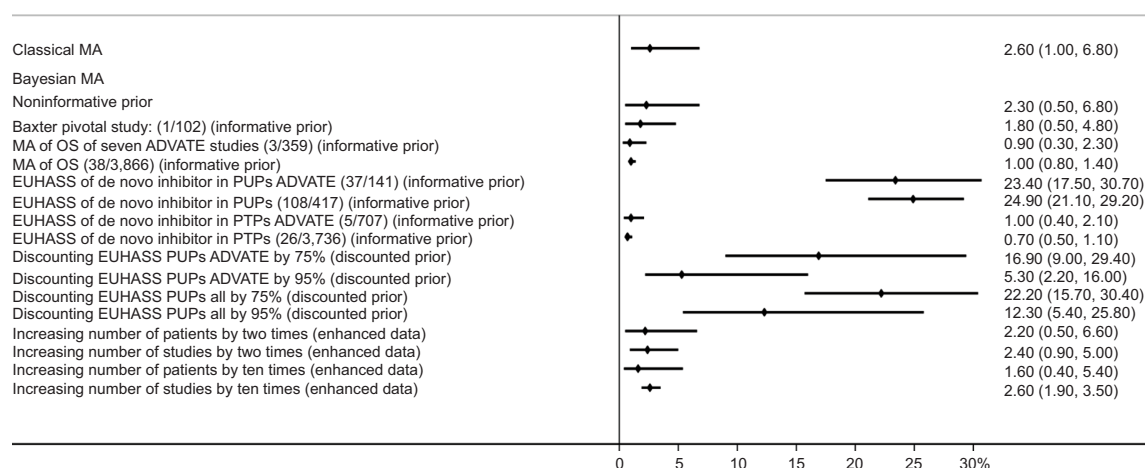
Case 2: MA

Percent rate (95% CI/CrI)

**C**

Case 3: Multicenter cohort – no appropriate priors

Percent rate (95% CI/CrI)

**Figure 1** Estimates of inhibitor rate obtained using classical or Bayesian approaches for three study cases **A**, **B**, **C** respectively.

Abbreviations: CI, confidence interval; CrI, credible interval; EUHASS, European Hemophilia Safety Surveillance; MA, meta-analysis; OS, observational study; PTP, previously treated patient; PUP, previously untreated patient.

the threshold was dropped further to the FDA-approved rate for PTPs of 1/86, only the probability using priors for the EUHASS study in PTPs for all products was >0.9, but all estimations using PTPs as priors were still above 0.65 (Table 5).

Impact of sample size of data

We found that increasing the number of patients narrowed the CrI for Case 1 (ie, the mimic of single-center studies), but had little impact on the posterior estimates of Cases 2 and 3, which represented multicenter study settings. However, when more centers were added to get the same sample size, the CrIs noticeably narrowed. Another interesting observation is that when the number of patients was increased in each center while the number of centers remained the same, the posterior inhibitor rates decreased for Case 3, in which three of seven centers reported no inhibitor event in the original data (Table 3; more exploratory results are found in Table S2).

Discussion

In this study, we used three cases from the real-world data to examine the abilities of the Bayesian approach in analyzing and interpreting rare events observed in a rare population. In comparison of the frequentist (or classical) approach, the

advantages of Bayesian framework majorly lay on the following three points.

First, unlike the classical frequentist statistics, Bayesian approach provides a way to dynamically update the evidence around a certain event using all the available information. In fact, posterior estimates are weighted averages based on prior and current experimental evidence/data, with the weights determined by the precision of the corresponding evidence. As in our study cases, we updated the estimate of the inhibitor rate that is represented by the posterior distribution by combining the information on the inhibitor rate contained in our data and the knowledge on the inhibitor rate found in the external study. This nature leads to the continuous updates of the evidence upon the availability of new data. The most informative example of using the Bayesian framework is where the inhibitor rate among patients with low-titer inhibitor at baseline or personal history of the inhibitor had been given little consideration to date. Therefore, comparing the posterior estimates of the inhibitor rate obtained from the study data commonly used as clinical thresholds will provide clinicians with meaningful ways to interpret the results.

Second, the Bayesian posterior probability is fundamentally different from the *P*-value used for frequentist hypothesis test. For instance, when testing “the inhibitor rate (in our population) is <10%” in our case, the *P* obtained

Table 5 Probabilities calculated for the inhibitor rates lower than specific thresholds for Case 3

| Prior | Case 3: PASS | Threshold 1 | Threshold 2 | Threshold 3 |
|--|---|-------------|-------------|-------------|
| Bayesian statistical analysis: percent rate (95% CrI) | Multicenter study – no appropriate priors | <10/100 | <5/100 | <1/86 |
| Noninformative prior | 2.3 (0.5, 6.8) | 0.994 | 0.921 | 0.165 |
| Informative prior: Baxter pivot study (1/102) | 1.8 (0.5, 4.8) | >0.999 | 0.979 | 0.225 |
| Informative prior: meta-analysis of OS (seven ADVATE studies) (3/569) | 0.9 (0.3, 2.3) | >0.999 | >0.999 | 0.677 |
| Informative prior: meta-analysis of OS (38/3,866) | 1.0 (0.8, 1.4) | >0.999 | >0.999 | 0.782 |
| Informative prior: EUHASS study of de novo inhibitor in PUPs, ADVATE (37/141) | 23.4 (17.5, 30.7) | <0.001 | <0.001 | <0.001 |
| Informative prior: EUHASS study of de novo inhibitor in PUPs (108/417) | 24.9 (21.1, 29.2) | <0.001 | <0.001 | <0.001 |
| Informative prior: EUHASS study of inhibitors in PTPs, ADVATE (5/707) | 1.0 (0.4, 2.1) | >0.999 | >0.999 | 0.658 |
| Informative prior: EUHASS study of inhibitors in PTPs (all FVIII) (26/3,736) | 0.7 (0.5, 1.1) | >0.999 | >0.999 | 0.988 |
| Discounted prior: discounting EUHASS in PUPs, ADVATE by 75% | 16.9 (9.0, 29.4) | 0.051 | <0.001 | <0.001 |
| Discounted prior: discounting EUHASS in PUPs, ADVATE by 95% | 5.3 (2.2, 16.0) | 0.876 | 0.449 | 0.001 |
| Discounted prior: discounting EUHASS in PUPs, all by 75% | 22.2 (15.7, 30.4) | <0.001 | <0.001 | <0.001 |
| Discounted prior: discounting EUHASS in PUPs, all by 95% | 12.3 (5.4, 25.8) | 0.306 | 0.016 | <0.001 |
| Enhanced data: enhancing study data by two times – increasing number of patients (with noninformative prior) | 2.2 (0.5, 6.6) | 0.995 | 0.932 | 0.161 |
| Enhanced data: enhancing study data by two times – increasing number of studies (with noninformative prior) | 2.4 (0.9, 5.0) | 0.998 | 0.967 | 0.305 |
| Enhanced data: enhancing study data by ten times – increasing number of patients (with noninformative prior) | 1.6 (0.4, 5.4) | 0.998 | 0.976 | 0.067 |
| Enhanced data: enhancing study data by ten times – increasing number of studies (with noninformative prior) | 2.6 (1.9, 3.5) | >0.999 | >0.999 | <0.001 |

Abbreviations: CI, confidence interval; CrI, credible interval; EUHASS, European Hemophilia Safety Surveillance; OS, observational study; PASS, post-authorization safety studies; PTP, previously treated patient; PUP, previously untreated patient.

from the frequentist test tells us the probability of wrongly rejecting the null hypothesis – “the inhibitor rate equals to 10%” in our case. This P -value reflects the type I error. In addition, thus the statistical significance from frequentist approach is built on an arbitrary cutoff for tolerating this type I error, say 0.05.⁴² In our example, if the “probability” of “the inhibitor rate is $\geq 10\%$ ” is < 0.05 , we conclude that this hypothesis can be rejected. However, this probability is not in fact a probability directly related to the acceptance of the testing hypothesis, but a level of confidence that the chance of mistakenly rejecting the null hypothesis is low. In fact, when $P < 0.05$, we can reject the null hypothesis, but we are never able to say that the probability of “the inhibitor rate being $< 10\%$ ” is truly 0.95. On the other hand, the Bayesian probability is a quantity of the testing hypothesis, ie, the “degree of truth” of the study hypothesis. The Bayesian can really test the probability that the rate of inhibitor in our sample is $< 10\%$. If $P = 0.95$, we are confident that the probability of “the inhibitor rate being $< 10\%$ ” is actually 0.95.

Third point is the different interpretations between the CI in classical approach and CrI in Bayesian approach. Back to our example, the 95% CI is interpreted as “the estimates of the inhibitor rate will fall in between these two boundaries 95% of the time if the data can be repeated infinitely”. It cannot be used to make an assertion about the current test based on a single sample set without the assumption of the infinite repetition. In comparison, the 95% CrI tells us a straightforward story, “given the data and the model, the chance of the true inhibitor rate fall in this interval is 95%”.

Some further considerations are hopefully of value. For Cases 1 and 2, the Bayesian models with noninformative priors yielded results comparable to the classical approach. For third study case, the point estimate of inhibitor rate obtained from the Bayesian random-effects logistic model was lower than that obtained from the classical random-effects logistic model. The reason is that the data used for this example are extremely sparse. In three out of seven pooled studies, there were no inhibitors observed. The classical logistic model directly takes event as outcome and thus fails to generate the estimates when no event is in the data. Therefore, when classical random-effect logistic was used to pool the data from seven individual studies, the three studies without outcomes were ignored, and the inhibitor rate was estimated from the four studies with observed inhibitor. Unlike the classical model, the Bayesian model resamples the data for certain times (eg, 100,000) based on the information provided by the current data and then generates the estimates in accordance before reporting the

posterior estimates, which usually are the median of the entire estimates. In our example, when data reported no event, the Bayesian model resampled the data using the probability of event sampled around zero. By doing so, the Bayesian model was able to incorporate those studies reporting no inhibitor into the posterior estimates and thus gave a lower inhibitor rate. On the other hand, for the same example, the 95% CrIs were wider than the 95% CIs. This is because the Bayesian model introduced more random uncertainty through noninformative priors that had very large variance. When the study data are not large enough, random uncertainty will be added in the posterior estimates. In our example, we had seven studies that were not even able to provide saturated information for estimating the between-study variance. Therefore, the model borrowed information from noninformative priors that only added uncertainty to estimate the between-study variance.

The Bayesian approach provides a very useful tool for exploring our Case 3 data, which presents the unexplored patient population with no existing evidence. We modeled the effect of observing six inhibitors in ~200 patients from the unlikely expectation that the inhibitor rate would have been as in Research Of Determinants of INhibitor (RODIN) study¹¹, to the optimistic expectation that the rate would have not been different from that in PTPs. We also showed how we could model the “strength” of this belief, by “discounting” the previous information. Critics of the Bayesian approach would certainly say that by adding “discounts” you may play with data until you show what you want. We would object that this would be the case if you were only using one set of priors (maybe even discounted). If you instead show the results produced by using a whole range of priors, you explore the relevance of your previous beliefs and assumptions. Along the same lines, we found that Bayesian modeling can be used to simulate the effect by increasing the number of patients in the same centers or by increasing the number of centers. The information found through these attempts can be easily used on estimating the sample size for the future studies. All of this richness of information is completely unavailable when using the frequentist approach. The reader needs to be aware, at this point, that most of the modeling of the impact of health care interventions on economics of health care systems or quality of life of patient population is generally obtained via Monte Carlo chain simulations which are, in essence, Bayesian probability applications.^{30,43}

Bayesian estimates can be used to inform clinical decisions in patient management in complex clinical settings. The success of integrating all relevant evidence through a

Bayesian approach depends on two aspects: 1) how to properly choose the clinically relevant priors, and 2) how to statistically formulate the clinical knowledge. These tasks need the joint force of clinicians and statisticians. Properly implementing the Bayesian results in the clinical decision making depends on the comprehensive understanding of the evidence, in particular, for the findings obtained from the first-time-ever exploration regarding new study settings or populations.

In regard to the purpose of serving as case studies for conducting Bayesian analysis in hematology area to analyze and generate evidence for rare events among rare study populations, we chose to use Bayesian random-effect logistic regression throughout the entire project for three scenarios for simplicity. We are aware that other statistical models such as random-effect Poisson regress may be a better choice for rare event data with zero outcomes. It is worth noting that properly setting up priors can be challenging because it depends on the types of outcomes and Bayesian models.

To come back to the clinical ground, we showed how the Bayesian posterior distribution can be interrogated to get, for example, the posterior probability that the rate of inhibitors in a population like the one we studied (eg, patients with previous history of inhibitors) was above or below a given (clinically meaningful) threshold. This is what, in our opinion, is needed for clinical decisions, and, indirectly, for policy-making decisions like taking into account the 30% of patients with a previous history of developing inhibitors in the proportion of population to be suitable to switch concentrate as a result of a tender process.

Conclusion

The Bayesian estimates of the inhibitor rate of patients undergoing treatment with ADVATE provide a broader understanding for the clinicians, which can be utilized to inform clinical decisions in the management of patients with hemophilia A. Bayesian approach as a robust, transparent, and reproducible analytic method can be efficiently used to answer the complex clinical questions through updating, enhancing, and exploring the evidence by incorporating all the available information. Furthermore, the Bayesian probability can be directly used by clinicians to quantify individual patient's risk on developing certain conditions.

Acknowledgments

The authors would like to acknowledge Jessica Fowler (Baxter Healthcare, Deerfield, Illinois, USA) for editorial services in reviewing the manuscript. This work was supported by a research contract agreement between Baxter Healthcare and St Joseph's HealthCare Hamilton.

Author contributions

JC, AI, MM, VR, and LT designed the study and drafted the manuscript. JC conducted the statistical analysis. AI, MM, VR, and JKM provided the clinical expertise. JC, AI, MM, EMP, and LT provided input on statistical concept. All authors contributed toward data analysis, drafting and critically revising the paper and agree to be accountable for all aspects of the work. All authors revised the manuscript for important clinical and statistical contents and approved the final manuscript.

Disclosure

JC, MM, EMP, and JKM report no conflicts of interest in this work. AI has worked as a consultant for and held research contracts sponsored by Bayer AG, Biogen Idec, Novo Nordisk A/S, and Pfizer, Inc. VR is an employee of Baxter Healthcare. LT has worked as a consultant for and held research contracts sponsored by many pharmaceutical companies including GlaxoSmithKline plc, AstraZeneca plc, Serono Canada Inc, Hoffman-La Roche Ltd., Pfizer, Inc., Theralase Inc, CanReg Inc, and Merck Frosst – Schering Pharmaceuticals and Proctor and Gamble Pharmaceuticals Canada Inc.

References

1. W Collins P, Chalmers E, Hart D, et al. Diagnosis and management of acquired coagulation inhibitors: a guideline from UKHCDO. *Br J Haematol*. 2013;162(6):758–773.
2. Iorio A. Epidemiology of inhibitors in hemophilia. In: Lee CA, Berntorp EE, Hoots WK, editors. *Textbook of Hemophilia*. Oxford: John Wiley & Sons, Ltd; 2014:53–58.
3. Caram C, de Souza RG, de Sousa JC, et al. The long-term course of factor VIII inhibitors in patients with congenital haemophilia A without immune tolerance induction. *Thromb Haemost*. 2011;105(1):59–65.
4. Hay CRM, Palmer B, Chalmers E, et al; United Kingdom Haemophilia Centre Doctors' Organisation (UKHCDO). Incidence of factor VIII inhibitors throughout life in severe hemophilia A in the United Kingdom. *Blood*. 2011;117(23):6367–6370.
5. Stieltjes N, Torchet MF, Misrahi L, et al. Epidemiological survey of haemophiliacs with inhibitors in France: orthopaedic status, quality of life and cost – the “Statut Orthopédique des Patients Hémophiles” avec Inhibiteur study. *Blood Coagul Fibrinolysis*. 2009;20(1):4–11.
6. Gringeri A, Monzini M, Tagariello G, et al. Occurrence of inhibitors in previously untreated or minimally treated patients with haemophilia A after exposure to a plasma-derived solvent-detergent factor VIII concentrate. *Haemophilia*. 2006;12(2):128–132.
7. Astermark J, Altisent C, Batorova A, et al; European Haemophilia Therapy Standardisation Board. Non-genetic risk factors and the development of inhibitors in haemophilia: a comprehensive review and consensus report. *Haemophilia*. 2010;16(5):747–766.
8. Coppola A, Santoro C, Tagliaferri A, Franchini M, Di Minno G. Understanding inhibitor development in haemophilia A: towards clinical prediction and prevention strategies. *Haemophilia*. 2010;16(suppl 1):13–19.
9. Iorio A, Puccetti P, Makris M. Clotting factor concentrate switching and inhibitor development in hemophilia A. *Blood*. 2012;120(4):720–727.
10. Lee CA, Lillicrap D, Astermark J. Inhibitor development in hemophiliacs: the roles of genetic versus environmental factors. *Semin Thromb Hemost*. 2006;32(suppl 2):10–14.
11. Gouw SC, van den Berg HM, Fischer K, et al; PedNet and Research of Determinants of INhibitor development (RODIN) Study Group. Intensity of factor VIII treatment and inhibitor development in children with severe hemophilia A: the RODIN study. *Blood*. 2013;121(20):4046–4055.

12. Gouw SC, van den Berg HM. The multifactorial etiology of inhibitor development in hemophilia: genetics and environment. *Semin Thromb Hemost.* 2009;35(8):723–734.
13. Gouw SC, van der Bom JG, Marijke van den Berg H. Treatment-related risk factors of inhibitor development in previously untreated patients with hemophilia A: the CANAL cohort study. *Blood.* 2007; 109(11):4648–4654.
14. Pavlova A, Delev D, Lacroix-Desmazes S, et al. Impact of polymorphisms of the major histocompatibility complex class II, interleukin-10, tumor necrosis factor-alpha and cytotoxic T-lymphocyte antigen-4 genes on inhibitor development in severe hemophilia A. *J Thromb Haemost.* 2009; 7(12):2006–2015.
15. Astermark J, Lacroix-Desmazes S, Reding MT. Inhibitor development. *Haemophilia.* 2008;14(suppl 3):36–42.
16. Saint-Remy J-M, Reipert BM, Monroe DM. Models for assessing immunogenicity and efficacy of new therapeutics for the treatment of haemophilia. *Haemophilia.* 2012;18(suppl 4):43–47.
17. Martino D, Lillicrap D, Astermark J, et al. Switching clotting factor concentrates: considerations in estimating the risk of immunogenicity. *Haemophilia.* 2014;20(2):200–206.
18. Kesselheim AS, Myers JA, Avorn J. Characteristics of clinical trials to support approval of orphan vs nonorphan drugs for cancer. *JAMA.* 2011;305(22):2320–2326.
19. Iorio A, Marcucci M. Clinical trials and haemophilia : does the Bayesian approach make the ideal and desirable good friends ? *Haemophilia.* 2009;15(4):900–903.
20. Chow S-C, Chang M. Adaptive design methods in clinical trials – a review. *Orphanet J Rare Dis.* 2008;3:11.
21. Honkanen VE, Siegel AF, Szalai JP, Berger V, Feldman BM, Siegel JN. A three-stage clinical trial design for rare disorders. *Stat Med.* 2001; 20(20):3009–3021.
22. Lilford RJ, Thornton JG, Braunholtz D. Clinical trials and rare diseases: a way out of a conundrum. *BMJ.* 1995;311(7020):1621–1625.
23. Behera M, Kumar A, Soares HP, Sokol L, Djulbegovic B. Evidence-based medicine for rare diseases: implications for data interpretation and clinical trial design. *Cancer Control.* 2007;14(2):160–166.
24. Committee on Strategies for Small-Number-Participant Clinical Research Trials, Board on Health Sciences Policy Committee on Strategies for Small-Number-Participant Clinical Research Trials B on HSP, editor. *Small Clinical Trials: Issues and Challenges.* Washington, DC: National Academy Press; 2001.
25. Harrell FE, Lee KL, Mark DB. Multivariable prognostic models: issues in developing models, evaluating assumptions and adequacy, and measuring and reducing errors. *Stat Med.* 1996;15(4):361–387.
26. Vandenbroucke JP. When are observational studies as credible as randomised trials? *Lancet.* 2004;363(9422):1728–1731.
27. Hornberger J, Wrona E. When to base clinical policies on observational versus randomized trial data. *Ann Intern Med.* 1997;127(8 pt 2):697–703.
28. Von Elm E, Altman DG, Egger M, et al; STROBE Initiative. The strengthening the reporting of observational studies in epidemiology (STROBE) statement: guidelines for reporting observational studies. *Epidemiology.* 2007;18(6):800–804.
29. Kalil AC, Sun J. Bayesian methodology for the design and interpretation of clinical trials in critical care medicine. *Crit Care Med.* 2014; 42:2267–2277.
30. Spiegelhalter DJ, Abrams KR, Myles JP. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation.* 1st ed. Chichester, West Sussex, England: John Wiley & Sons Ltd; 2004.
31. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Stat Comput.* 2000;10:325–337.
32. Oldenburg J, Goudemand J, Valentino L, et al. Postauthorization safety surveillance of ADVATE [antihaemophilic factor (recombinant), plasma/albumin-free method] demonstrates efficacy, safety and low-risk for immunogenicity in routine clinical practice. *Haemophilia.* 2010;16(6):866–877.
33. Iorio A, Marcucci M, Cheng J, et al. Patient data meta-analysis of post-authorization safety surveillance (PASS) studies of haemophilia A patients treated with rAHF-PFM. *Haemophilia.* 2014;20(6):777–783.
34. Romanov V, Marcucci M, Cheng J, Thabane L, Iorio A. Evaluation of safety and effectiveness of factor VIII treatment in hemophilia A patients with low titer inhibitors or a personal history of inhibitor. *Thromb Haemost.* 2015;113(3):56–64.
35. Tarantino MD, Collins PW, Hay CRM, et al; RAHF-PFM Clinical Study Group. Clinical evaluation of an advanced category antihaemophilic factor prepared using a plasma/albumin-free method: pharmacokinetics, efficacy, and safety in previously treated patients with haemophilia A. *Haemophilia.* 2004;10(5):428–437.
36. Xi M, Makris M, Marcucci M, Santagostino E, Mannucci PM, Iorio A. Inhibitor development in previously treated hemophilia A patients: a systematic review, meta-analysis, and meta-regression. *J Thromb Haemost.* 2013;11(9):1655–1662.
37. Fischer K, Lassila R, Peyvandi F, et al; EUHASS Participants. Inhibitor development in haemophilia according to concentrate four-year results from the European haemophilia safety surveillance (EUHASS) project. *Thromb Haemost.* 2015;113(5):968–975.
38. Dimichele DM, Lacroix-Desmazes S, Peyvandi F, Srivastava A, Rosendaal FR. Design of clinical trials for new products in hemophilia: communication from the SSC of the ISTH. *J Thromb Haemost.* 2015;13(5):876–879.
39. Aledort LM. Harmonization of clinical trial guidelines for assessing the risk of inhibitor development in hemophilia A treatment. *J Thromb Haemost.* 2011;9(3):423–427.
40. Department of Health and Human Services Food And Drug Administration, Center For Biologics Evaluation And Research, Lister Hill Center National Institutes of Health. Workshop on Factor VIII Inhibitors. November 21, 2003, Miller Reporting Co., Inc, Washington D.C.
41. Salpeter SR, Cheng J, Thabane L, Buckley NS, Salpeter EE. Bayesian meta-analysis of hormone therapy and mortality in younger postmenopausal women. *Am J Med.* 2009;122(11):1016.e–1022.e.
42. Greenland S, Poole C. Living with P values. *Epidemiology.* 2013;24(1): 62–68.
43. Lindley DV. An introduction to Bayesian inference and decision. *J Oper Res Soc.* 1974;25(2):336–337.

Supplementary materials

Figure S1 Bayesian codes

Number of updates, 300,000; number of chains, 2; number of thin, 5; burn in, 10,000; seed, 314,159.

Model 1: Simple logistic regression

```
model {
  r ~ dbin(p, n)
  logit(p) <- mu #log odds
  mu ~ dnorm(0, 1.0E-5) # non-informative
  #mu ~ dnorm(-4.615, 0.990) # PIVOT 1/102, log_odds=log(1/101), var=1/1+1/101
  #mu ~ dnorm(-5.24, 2.984) # Meta-OS ADVATE, 3/569
  #mu ~ dnorm(-4.613, 37.626) # Meta-OS ADVATE, 38/3866
  #mu ~ dnorm(-4.944, 4.965) # EUHASS PTP ADVATE, 5/707
  #mu ~ dnorm(-4.961, 25.819) # EUHASS PTP ALL, 26/3736
  odds<-exp(mu)
  prop<-odds/(1+odds)
  perc<-prop*100
}
```

Note: “number of thin” is a term of Bayesian analysis, which means every 1 sample taking into account over certain numbers of samples generated.

Model 2: Random-effects logistic model

```
model {
  for( i in 1 : Num ) {
    r[i] ~ dbin(p[i], n[i])
    logit(p[i]) <- mu[i] #log odds
    mu[i] ~ dnorm(d, tau)
  }
  d ~ dnorm(0,1.0E-5) # Non-informative prior
  # Prior1: Baxter Pivot Trial, 1/102
  #d ~ dnorm(-4.62,0.99) # log_odds, variance=1/1+1/(102-1), precision=1/var
  #d ~ dnorm(-5.24, 2.984) # Meta-OS ADVATE, 3/569
  #d ~ dnorm(-4.613, 37.626) # Meta-OS ADVATE, 38/3866
  #d ~ dnorm(-1.033, 27.291) # EUHASS PUP ADVATE, 37/141
  #d ~ dnorm(-1.051,80.029) # EUHASS PUP ALL, 108/417
  #d ~ dnorm(-4.944, 4.965) # EUHASS PTP ADVATE, 5/707
  #d ~ dnorm(-4.961, 25.819) # EUHASS PTP ALL, 26/3736
  #d ~ dnorm(-1.033, 6.822) # EUHASS PUP ADVATE(-75%), 9.25/35.25
  #d ~ dnorm(-1.033, 1.365) # EUHASS PUP ADVATE(-95%), 1.85/7.05
  #d ~ dnorm(-1.051, 20.007) # EUHASS PUP ALL(-75%), 27/104.25
  #d ~ dnorm(-1.052, 4.001) # EUHASS PUP ALL(-95%), 5.4/20.85
  tau<-1/(sigma*sigma)
  sigma~dunif(0,2) # between study variance is estimated from PASS1
  odds<-exp(d)
  prop <- exp(d)/(1+exp(d))
  perc<-prop*100
  ppos1<-step(10/100-prop)
  ppos2<-step(5/100-prop)
  ppos3<-step(1/86-prop)
}
```

Table S1 Analysis methods and the choice of priors

| Approaches and priors | Case 1 | Case 2 | Case 3 |
|--|--------------------------------|--|--|
| Method | Single study | Meta-analysis | Multicentric cohort – no appropriate priors |
| Test data | PASS data ¹ (6/428) | PASS data ² (21/1,188) | PASS data ³ (6/219) |
| Classical statistical analysis | Logistic model | Random-effects logistic model | Random-effects logistic model |
| Bayesian statistical analysis | Logistic model | Hierarchical (random-effects) logistic model | Hierarchical (random-effects) logistic model |
| Noninformative prior | Ok | Ok | Ok |
| Informative prior: Baxter pivotal study (1/102) | Ok | Ok | Ok |
| Informative prior: meta-analysis of OS (seven ADVATE studies) (3/569) | Ok | Ok | Ok |
| Informative prior: meta-analysis of OS (38/3,866): do you need data per study? | Ok | Ok | Ok |
| Informative prior: EUHASS study of de novo inhibitor in PUPs, ADVATE (37/141) | No | No | Ok |
| Informative prior: EUHASS study of de novo inhibitor in PUPs (108/417) | No | No | Ok |
| Informative prior: EUHASS study of inhibitors in PTPs, ADVATE (5/707) | Ok | Ok | Ok |
| Informative prior: EUHASS study of inhibitors in PTPs (all FVIII) 22/3,736 | Ok | Ok | Ok |
| Discounted prior: discounting EUHASS in PUPs, ADVATE by 75% | No | No | Ok |
| Discounted prior: discounting EUHASS in PUPs, ADVATE by 95% | No | No | Ok |
| Discounted prior: discounting EUHASS in PUPs, all by 75% | No | No | Ok |
| Discounted prior: discounting EUHASS in PUPs, all by 95% | No | No | Ok |
| Enhanced data: enhancing study data by two times – increasing number of patients | Ok | Ok | Ok |
| Enhanced data: enhancing study data by two times – increasing number of studies | No | Ok | Ok |
| Enhanced data: enhancing study data by ten times – increasing number of studies | Ok | Ok | Ok |
| Enhanced data: enhancing study data by ten times – increasing number of studies | No | Ok | Ok |

Note: “OK” means it is valid choice thus will be used in the analysis. “No” means it is no a valid choice and won’t be used in the analysis.

Abbreviations: CI, confidence interval; CrI, credible interval; EUHASS, European Hemophilia Safety Surveillance; OS, observational study; PASS, post-authorization safety studies; PTP, previously treated patient; PUP, previously untreated patient.

Table S2 Assessing the impact of sample size change and choice of priors on the Bayesian posterior estimates

| | Original data | Increasing sample size by two times | | Increasing sample size by ten times | |
|---|---------------------------------------|---|--|---|--|
| Test data (number of inhibitors/number of patients); number of centers | Case 3: PASS (6/219) ³ ; 7 | Increasing number of patients in each center: (12/438); 7 | Increasing number of centers: (12/438); 14 | Increasing number of patients in each center: (60/2,190); 7 | Increasing number of centers: (60/2,190); 70 |
| Noninformative prior | 2.3 (0.5, 6.8) | 2.2 (0.5, 6.6) | 2.4 (0.9, 5.0) | 1.6 (0.4, 5.4) | 2.6 (1.9, 3.5) |
| Informative prior: Baxter pivotal study (1/102) | 1.8 (0.5, 4.8) | 1.8 (0.5, 4.6) | 2.1 (0.8, 4.2) | 1.4 (0.4, 3.9) | 2.6 (1.8, 3.4) |
| Informative prior: meta-analysis of OS (seven ADVATE studies) (3/569) | 0.9 (0.3, 2.3) | 0.9 (0.3, 2.4) | 1.3 (0.5, 2.7) | 0.8 (0.3, 2.0) | 2.3 (1.5, 3.1) |
| Informative prior: meta-analysis of OS (38/3,866) | 1.0 (0.8, 1.4) | 1.0 (0.8, 1.4) | 1.1 (0.8, 1.5) | 1.0 (0.7, 1.4) | 1.4 (1.0, 1.8) |
| Informative prior: EUHASS study of de novo inhibitor in PUPs, ADVATE (37/141) | 23.4 (17.5, 30.7) | 23.2 (17.3, 30.4) | 21.4 (15.6, 28.1) | 22.8 (17.0, 29.8) | 11.2 (7.7, 15.4) |
| Informative prior: EUHASS study of de novo inhibitor in PUPs (108/417) | 24.9 (21.1, 29.2) | 24.8 (20.1, 29.1) | 24.1 (20.4, 28.3) | 24.7 (20.8, 28.9) | 19.2 (16.1, 22.5) |
| Informative prior: EUHASS study of inhibitors in PTPs, ADVATE (5/707) | 1.0 (0.4, 2.1) | 1.0 (0.4, 2.1) | 1.2 (0.6, 2.4) | 0.9 (0.4, 1.9) | 2.3 (1.5, 3.0) |
| Informative prior: EUHASS study of inhibitors in PTPs (all FVIII) (26/3,736) | 0.7 (0.5, 1.1) | 0.8 (0.5, 1.1) | 0.8 (0.5, 1.2) | 0.7 (0.5, 1.1) | 1.2 (0.8, 1.7) |

Note: Data presented as percent rate (95% CrI).

Abbreviations: CI, confidence interval; CrI, credible interval; EUHASS, European Haemophilia Safety Surveillance; OS, observational study; PASS, post-authorization safety studies; PTP, previously treated patient; PUP, previously untreated patient.

References

1. Spiegelhalter DJ, Abrams KR, Myles JP. *Bayesian Approaches to Clinical Trials and Health-Care Evaluation*. 1st ed. Chichester, West Sussex, England: John Wiley & Sons Ltd; 2004.
2. Lunn DJ, Thomas A, Best N, Spiegelhalter D. WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Stat Comput*. 2000;10:325–337.
3. Romanov V, Marcucci M, Cheng J, Thabane L, Iorio A. Evaluation of Safety and Effectiveness of factor VIII treatment in Hemophilia A patients with low titer inhibitors or a personal history of inhibitor. *Thromb Haemost*. 2015;114(1):56–64.