The Use of Bayesian Analysis Techniques in Pediatric Research

In a recent article appearing in *The Journal*, Faustino et al reported the findings of a multicenter intensive care unit (ICU) study investigating the frequency of deep venous thrombosis (DVT) among critically ill adolescents (13-17 years of age).[1] The authors indicate that DVT has been well studied in the adult population of critically ill patients, but such studies in the pediatric population are sparse. As such, the frequency of DVT among critically ill adolescents is not well known, and pediatric protocols to use pharmacologic prophylaxis to reduce risk of DVT vary widely from center to center. The prospective study by Faustino et al was aimed at estimating the frequency of DVT and providing clear recommendations for pharmacologic prophylaxis among this adolescent population.[1] Though many clinicians may be unfamiliar with Bayesian analyses, the authors carefully chose this approach to include knowledge about the frequency of DVTs from the adult population, account for multiple ICU heterogeneity, and estimate the probability that the frequency of DVT was above the treatment threshold.

See related article, J Pediatr 2018:201:176-83

What is a "Bayesian" Analysis?

Although Bayesian analysis is not new, it has recently become practical as an emerging analytic approach because of continuously improving computational capabilities.[2,3] The use of Bayesian analysis methods in a study indicate that the investigators wish to follow a course of discovery that begins with prior information about a parameter of interest, in this case, the frequency of DVT, obtained from a related adult population. This prior knowledge is mathematically combined with observed data from the current study to produce an evidence-based updated understanding of the parameter known in Bayesian analysis as a "posterior." This type of analysis offers 2 crucial advantages over more traditional statistical methods when applied to practical clinical questions, especially in pediatrics.

(1) The ability to mathematically incorporate prior, related clinical information to sharpen the conclusions of a current inquiry. This is important in pediatrics because we often have information from adult populations that is at least somewhat informative to the study question.
(2) The ability to estimate probabilities for parameters of interest (eg, frequency of DVT), allowing calculation of specific probabilities regarding relationships to established treatment thresholds—in the Faustino et al article, "what is the probability that the DVT frequency in the ICU population lies above 10%?".[1]

It has been noted that Bayesian techniques are a natural fit for clinicians because of the similarity to the process of clinical reasoning. Their inclusion of prior information coupled with current data to arrive at a probable conclusion is similar to how a physician incorporates evidence, clinical experience, and specific data from an individual patient to create a prioritized differential diagnosis and treatment plan.[4] Pediatric medicine especially lends itself to Bayesian inference, as there is often a relative paucity of direct clinical evidence for interventions in pediatric patients compared with the adult population.

DVT	Deep venous thrombosis
ICU	Intensive care unit

Prior Information

In a Bayesian analysis, the posterior, or final result, reflects a weighted averaging of the prior information and the evidence in the data. In data rich settings, the effect of the prior information on the posterior can be minimal, and results will be similar to those obtained using traditional statistical methods. In settings where data are sparse, however, using prior information can have a profound effect on the posterior and infuse the analysis with statistical power. The quality of prior information in such settings is critical and should be carefully validated from literature and expert opinion when possible.

When Should Prior Information be Used?

A common concern among those more familiar with traditional methods of analysis is that using prior information may introduce bias into the final estimates. However, there are at least 2 situations in medicine where the introduction of prior knowledge makes intuitive sense.

(1) Biological plausibility—For example, if a study aims to estimate a mean systolic blood pressure for a population of ambulatory adolescents, investigators could reasonably set upper and lower bounds on a prior that excluded biologically impossible values (negative numbers, numbers greater than 300, etc). In this sense, estimation is focused on biologically plausible values, reserving statistical power for values that make the most biological sense. This can improve the certainty of our final estimate.
(2) Prior related information is available—the Faustino et al DVT study falls into this category. In this case, we have information on a similar group of individuals (adults) that share commonalities (coagulation system, etc) with the pediatric population of interest. This information can be mathematically formalized and added to the statistical model to augment information from the small pediatric study. Although not perfect, as there can be differences between the population informing the prior and the

The authors declare no conflicts of interest.

0022-3476/$ - see front matter.
https://doi.org10.1016/j.jpeds.2018.11.059

population being studied, it is a step that can be used to sharpen our estimates as we gather more data on our target pediatric population. It is important to note that as clinicians, this calculation is often done intuitively as we extrapolate adult experience to the pediatric world. Bayesian inference offers a mathematically standardized method for a practice which we already employ.

How Should Prior Information be Used?

For use in a Bayesian analysis, the prior information must be quantified into a mathematical expression of probability that reflects the prior information and its uncertainty. A strong "prior" indicates high confidence in the prior information and a low level of uncertainty. A weak "prior," on the other hand, indicates less confidence in the prior information and is accompanied by a higher level of uncertainty. Faustino et al refer to these as informative and minimally informative priors, respectively.[1] In their article, the informative priors were used for estimating the frequency of DVT and hail from information synthesized from 5 adult studies giving them high confidence and low uncertainty in their prior information. However, for the secondary outcomes of the study such as potential risk factors and the duration of support, the authors note "an absence of pertinent prior information" and so for analysis, use minimally informative priors reflecting low confidence and high uncertainty in the prior information. The minimally informative priors will essentially let the data provide all the information for the posterior, yielding results that should be similar to those of a traditional analysis. A point of interest in this study is that the authors repeated the Bayesian analysis to estimate the frequency of DVT using minimally informative priors (**Figure**) to gauge the effect of the informative prior information on the posterior. This is commonly referred to as a sensitivity analysis.

Clustered (Hierarchical) Data

One advantage to Bayesian analyses is the natural way "clustered" data can be incorporated.[5] Faustino et al encounter such data as a result of the study being conducted at 6 different ICUs.[1] Although the frequency of DVT may not vary widely from center to center, it is probably not reasonable to assume that rates of occurrence will be identical because of the differences in staff, care protocols, and catchment areas. The goal of the study was not to determine estimates of the frequency of DVT for those 6 specific ICUs, but rather to develop an overall estimate of the frequency of DVT in the pediatric ICU setting. Such an estimate should acknowledge the variation seen among the ICUs in their study is representative of the variation of the larger population of all pediatric ICUs. As such, the authors use a hierarchical modeling approach to include ICU variation in the model (hence, the phrase "ICU of admission as random effect"). This process allows the model to not only include the assumed similarities between ICUs, but also the variability between them when final estimates are computed.

Posterior Results/Parameter Estimation

All results and conclusions in a Bayesian analysis come from the "posterior." This is viewed as the most complete picture

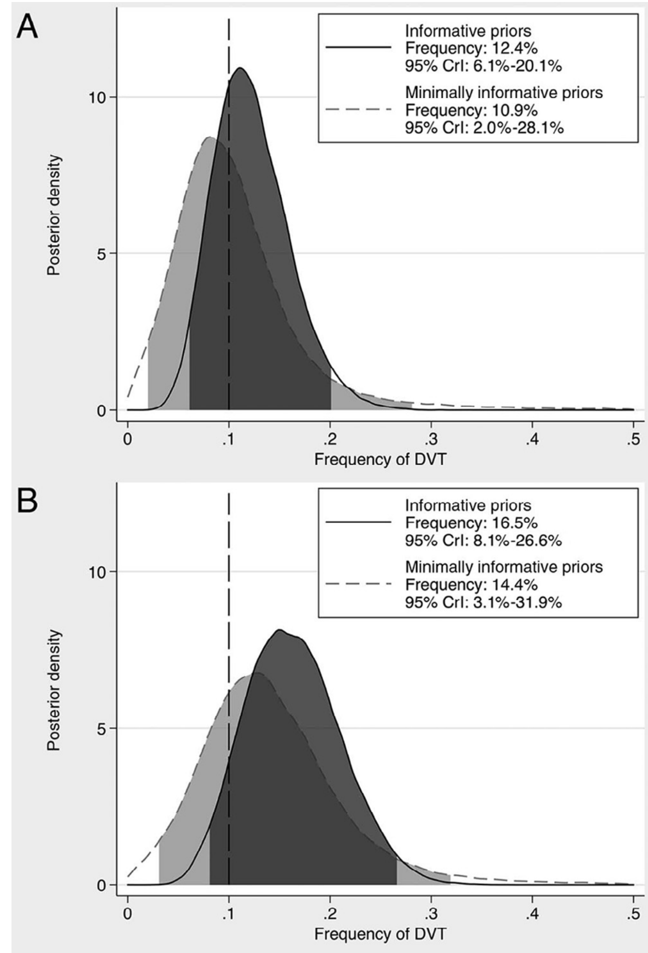

Figure. Posterior distributions of the frequency of lower extremity DVT among **A,** all adolescents and **B,** those invasively mechanically ventilated. The threshold frequency of lower extremity DVT to recommend pharmacologic prophylaxis in adults is 10%. *Shaded areas* represent the 95% equal-tailed credible intervals (CrI; 2.5th and 97.5th percentiles of the posterior distribution) around the posterior mean. Originally published in *J Pediatr* 2018:201:176-83.

of the parameter as it incorporates both knowledge about it prior to the study and information from the observed data. The posterior is a mathematical expression of probability that yields point estimates in the form of the posterior mean/median (which are analogous to the point estimate in traditional statistics) as well as a measure of the uncertainty of the those estimates. A 95% "credible interval" can be obtained from the posterior and indicates the probability the parameter lies in the interval is 0.95. For example Faustino et al plot the posteriors of the frequency of DVT among critically ill adolescents in **Figure**, A with (*solid line*) and without (*dashed*) informative priors.[1] When informative priors were used, the posterior estimate of the mean frequency of DVT is 12.4% with 95% credible interval of 6.1%-20.1%. This means there is a 95% probability that the true frequency of DVT is in that interval. This is in contrast to the 95% confidence interval in traditional statistics, which has the somewhat more complicated interpretation of "If 100 samples of the same size were drawn from the same population, each yielding a confidence interval constructed in the same manner, 95 of them would contain the true parameter of interest." The difference between these 2 intervals (credible and confidence) can be subtle but leads to very different interpretations. The credible interval is a direct probability statement about the parameter of interest, whereas the confidence interval is a probability statement about the interval only and how likely it is to be one that contains the true parameter.

A key advantage of a Bayesian analysis that makes it especially useful for decision-making is that the posterior allows direct statements about the probabilities of parameter values. Recalling that pharmacologic prophylaxis is recommended when the frequency of DVT is greater than 10%,[6] one could consult the posterior for insight on this probability. Traditional statistical methods would typically lead to a test to see if the frequency of DVT differed significantly from 10%, and then to a decision to reject or fail to reject that claim. With a Bayesian analysis on the other hand, the posterior allows Faustino et al to make the more useful statement that "The probability of frequency of DVT >10% was 0.70."[1] This probabilistic statement comes from noting the area of the posterior under the solid curve in **Figure**, A is 70%. Individual readers can weigh this information knowing there is a 30% chance that the frequency of DVT is 10% or less and reach their own decision about how effective pharmacologic prophylaxis use will be in this population.

Interpretation of the Study Findings

Faustino et al used prior information on the frequency of DVT in adult populations to augment the data of their pediatric study.[1] The use of informative and minimally informative priors allows clinicians with varying prior beliefs to see how the posterior would change over a range of possible prior beliefs. The authors also appear to have considered the issue of multiple ICUs in the study producing clustered data and appropriately included this in the modeling process. The authors report comparisons of baseline characteristics, interventions, and durations of support between DVT groups with 95% Bayesian credible intervals. For the main research question regarding the model adjusted frequency of DVT in critically ill adolescents, the authors report findings as probabilistic statements such as "the probability of frequency of DVT >10% was 0.70" and when speaking of a subset of the study subjects that were invasively mechanically ventilated that had "a probability of frequency of DVT >10% of 0.92." These statements carry intuitive meaning and can be powerful for clinicians deciding whether the use of pharmacologic prophylaxis is appropriate in this population. ■

Michael P. Anderson, PhD
Department of Biostatistics and Epidemiology, College of Public Health, The University of Oklahoma Health Sciences Center
Oklahoma City, OK

M. Townsend Cooper, Jr, MD, MS
Department of Pediatrics, College of Medicine
The University of Oklahoma Health Sciences Center
Oklahoma City, OK

References

1. Faustino E, Shabanova V, Pinto M, Li S, Trakas E, Miksa M, et al. Epidemiology of lower extremity deep venous thrombosis in critically ill adolescents. J Pediatr 2018;201:176.e2-183.e2.
2. Cowles K, Kass R, O'Hagan T. "What is Bayesian analysis?" [about 2 screens]. The International Society for Bayesian Analysis [homepage on the internet]. Duke University, Durham NC; 2018 [updated 2018]. https://bayesian.org/what-is-bayesian-analysis/. Accessed October 5, 2018.
3. Van de Schoot R, Kaplan D, Denissen J, Asendorpf JB, Neyer FJ, van Aken MA. A gentle introduction to Bayesian analysis: applications to developmental research. Child Dev 2014;85:842-60.
4. Gill CJ, Sabin L, Schmid C. Why clinicians are natural Bayesians. BMJ 2005;330:1080-3.
5. Lunn D, Jackson C, Best N, Thomas A, Spiegelhalter D. The BUGS book: a practical introduction to Bayesian analysis. Boca Raton: CRC Press; 2013.
6. Alhazzani W, Lim W, Jaeschke RZ, Murad MH, Cade J, Cook DJ. Heparin thromboprophylaxis in medical-surgical critically ill patients: a systematic review and meta-analysis of randomized trials. Crit Care Med 2013;41:2088-98.