

# Bayesian Inference: Statistical Gimmick or Added Value?



## Bayesian Inference

Use of the Bayesian inferential paradigm, although uncommon, is increasing in the rheumatology literature. The Bayesian school of statistical inference is based on the work of Reverend Thomas Bayes<sup>1</sup>. Using Bayes's theorem, observations are used to calculate the probability that a hypothesis is true. Pre-existing data or knowledge are expressed as a prior probability distribution or "prior." New observations are made and expressed as a likelihood. Using the new data, the degree of confidence in the hypothesis is recalculated. This process thus informs the reader on how this new information should change the way we think<sup>2</sup>. This contrasts with the "frequentist" statistical paradigm. Using frequentist inference, a null hypothesis of "no effect" and an alternative hypothesis that a "treatment effect exists" are specified before the conduct of a study. An allowable false-positive rate (i.e., level of significance, *p* value) is often arbitrarily set at 5%. A treatment is said to be effective if the *p* value is  $< 0.05$  or the 95% confidence interval excludes a null effect.

Bayesian statistics was the dominant statistical paradigm — prior to introduction of the now mainstream frequentist statistics. Its wide-scale application was limited by its computational complexity. In recent decades, its application became feasible with the introduction of faster computers and user-friendly statistical software. The more common day-to-day usage of the Bayesian inferential paradigm can be seen in meteorology (e.g., 20% probability of rain tomorrow) or the computer sciences (90% probability that this e-mail is spam and thus should go to the junk mail folder). In clinical medicine, we regularly apply Bayes's theorem when interpreting a diagnostic test. The post-test probability of disease (e.g., inflammatory back disease) is determined by our pre-test probability of disease (usually based on history and physical examination) and the results of a diagnostic test (e.g., the likelihood based on the HLA-B27 test result). We are now seeing increased application of this statistical paradigm in medical research, and in particular, rheumatology research.

## Applications in Rheumatology

To illustrate some applications of Bayesian statistics, this editorial highlights 4 recent examples in the rheumatology literature.

*Example 1.* In the setting of amitriptyline use for pain reduction in children with polyarticular juvenile inflammatory arthritis, Bayesian metaanalysis was used to analyze multiple N-of-1 trials<sup>3</sup>. By using this method, small numbers of patients ( $n = 6$ ) provided sufficient information to generate estimates of the population effect and the likelihood of treatment benefit. These data are of great value prior to the initiation of a large, potentially expensive, multicenter trial. In this case, the investigators demonstrated a very low probability of a beneficial treatment effect, thereby preventing the undertaking of a futile trial.

*Example 2.* Using the example of the efficacy of methotrexate in systemic sclerosis (SSc), Bayesian statistics was used to make inferences about treatment effects in an uncommon condition where the sample size is small and the study has insufficient power to detect a treatment effect using the frequentist inferential paradigm<sup>4</sup>. Raw data from a clinical trial in SSc suggested a beneficial treatment effect of methotrexate on skin score<sup>5</sup>. However, due to the small numbers of patients recruited into the trial and to limited power, the investigators had to correctly conclude that there was insufficient evidence to reject the null hypothesis of no treatment effect. This perpetuated the belief that methotrexate is ineffective in SSc. A Bayesian analysis demonstrated that methotrexate has a high probability of a beneficial treatment effect on skin score in SSc (consistent with the raw data). The Bayesian paradigm allowed clinically useful inferences to be made with the data at hand.

*Example 3.* In the setting of warfarin use for improving survival in SSc-associated pulmonary arterial hypertension,

---

See Bayesian analysis of CZP versus other anticytokines for RA, page 835

---

the Bayesian paradigm has been used to quantify and illustrate international experts' beliefs about treatment effect<sup>6</sup>. This has at least 2 important applications. First, investigators were able to scientifically demonstrate the presence of community equipoise — a necessary prerequisite prior to the conduct of a clinical trial. Second, in a setting where good quality data are scarce (or absent), clinicians readily rely on experts in the field to guide clinical practice. This quantification of expert knowledge can be used in combination with observational data from longitudinal cohorts to evaluate evidence of clinical benefit, and evaluate if investigators should proceed with a clinical trial to answer this question.

*Example 4.* Bernatsky, *et al* estimated the prevalence of polymyositis and dermatomyositis at the population level using administrative data (provincial physician billing and hospitalization data)<sup>7</sup>. Case ascertainment was dependent on accurate International Classification of Diseases, 9th edition, coding of these diseases and different diagnostic algorithms. As a result, these sources of data are susceptible to measurement error, leading to misclassification. The investigators applied a Bayesian latent class regression model to take into account variability in the sensitivity and specificity of the diagnostic algorithms, and variations in prevalence due to patient demographics. These methods allowed the investigators to make population-level estimates of disease prevalence despite imperfect data sources and different case ascertainment approaches.

Indeed, the Bayesian inferential paradigm confers a number of advantages in the analysis of clinical research data. In this issue of *The Journal*, Launois, *et al* apply Bayesian methods of metaanalysis to evaluate the non-inferiority of certolizumab with other anticytokine agents for the treatment of rheumatoid arthritis<sup>8</sup>. Traditionally, metaanalysis is used to synthesize data from multiple studies. Compared with the individual studies, it provides improved power and more precise estimates of a treatment effect<sup>9</sup>. Bayesian methods of metaanalysis confer some additional advantages. Bayesian indirect comparison analyses can be performed in a situation where several studies have been conducted evaluating several treatments that have been compared in different combinations, and the investigator wishes to make inferences about specific treatment contrasts<sup>2</sup>. Bayesian metaanalysis (i.e., indirect comparison analyses) facilitates estimation of treatment effects that have never been directly measured<sup>2</sup>. Launois, *et al* utilized 19 studies that evaluated 7 treatments (infliximab, etanercept, adalimumab, golimumab, certolizumab, anakinra, tocilizumab) versus placebo to evaluate the non-inferiority of certolizumab in the treatment of rheumatoid arthritis. This analysis provided a valuable signal of treatment effect in a setting where no head-to-head trial evidence exists.

## Reading a Bayesian Study with a Critical Eye

Consumers of the medical literature should critically appraise Bayesian studies to the same standard that one would expect from any other design and/or analytic strategy. For example, with any metaanalysis, the final estimate of treatment effect is only as good as the quality of the data going into the analysis. Thus the quality of the trials included in the analysis is typically evaluated with a measure of quality, such as the commonly used Jadad scale<sup>10</sup>. The Bayesian portion of a study demands an added level of scrutiny. A number of guidelines for the components of a Bayesian analysis that should be reported have been published. These include ROBUST (Reporting of Bayes Used in Clinical Studies)<sup>11</sup>, BaSiS (Bayesian Standards in Science)<sup>12</sup>, and BayesWatch<sup>13</sup>. Reporting of the components recommended by these guidelines will help the reader interpret the analysis, and allow the reader to have confidence that the approach was appropriate<sup>11</sup>. At the minimum, Bayesian studies should include some description of the prior distribution (which may include justification or rationale for this choice), the analysis conducted (analytic model, analytic technique), and the results (measures of central tendency and dispersion). It has also been recommended that Bayesian studies report an “interpretation” component that may supplement the results section or be included in the discussion section of an article<sup>14,15</sup>. This section may summarize the posterior distribution or discuss the sensitivity of the analysis to alternative priors<sup>12,13</sup>. In doing so, it would help the reader position the new data into the context of existing knowledge<sup>14</sup>.

## Conclusion

Critics of the Bayesian paradigm have challenged investigators to improve the methodologic rigor of Bayesian studies, and to clearly demonstrate that this approach actually offers some advantage over the existing paradigm<sup>16</sup>. Bayesian clinical investigators are rising to these challenges. At the minimum, the Bayesian paradigm complements existing analytic strategies. Given some of the issues that challenge rheumatology research (uncommon connective tissue disease, small sample sizes resulting in limited power, imperfect data sources, clinical trials competing for patient recruitment), the Bayesian paradigm clearly confers added value. In the coming years, we can likely anticipate more innovative applications of Bayesian statistics in the rheumatology literature.

**SINDHU R. JOHNSON, MD,**

Division of Rheumatology,  
Department of Medicine,  
University Health Network,  
Mount Sinai Hospital and  
University of Toronto,  
Toronto, Ontario, Canada

Address correspondence to Dr. S. Johnson, Division of Rheumatology, Ground Floor, East Wing, Toronto Western Hospital, 399 Bathurst Street, Toronto, Ontario, Canada, M5T 2S8. E-mail: [Sindhu.Johnson@uhn.on.ca](mailto:Sindhu.Johnson@uhn.on.ca)  
Dr. Johnson is supported by a Canadian Institutes of Health Research Clinician Scientist Award and the Norton-Evans Fund for Scleroderma Research.

## REFERENCES

1. Barnard G, Bayes T. Studies in the history of probability and statistics: IX. Thomas Bayes's essay towards solving a problem in the doctrine of chances. *Biometrika* 1958;45:293-315.
2. Spiegelhalter D, Abrams K, Myles J. Bayesian approaches to clinical trials and health-care evaluation. Chichester: John Wiley & Sons, Ltd.; 2004.
3. Huber AM, Tomlinson GA, Koren G, Feldman BM. Amitriptyline to relieve pain in juvenile idiopathic arthritis: a pilot study using Bayesian metaanalysis of multiple N-of-1 clinical trials. *J Rheumatol* 2007;34:1125-32.
4. Johnson SR, Feldman BM, Pope JE, Tomlinson GA. Shifting our thinking about uncommon disease trials: the case of methotrexate in scleroderma. *J Rheumatol* 2009;36:323-9.
5. Pope JE, Bellamy N, Seibold JR, Baron M, Ellman M, Carette S, et al. A randomized, controlled trial of methotrexate versus placebo in early diffuse scleroderma. *Arthritis Rheum* 2001;44:1351-8.
6. Johnson SR, Granton JT, Tomlinson GA, Grosbein HA, Hawker GA, Feldman BM. Effect of warfarin on survival in scleroderma-associated pulmonary arterial hypertension (SSc-PAH) and idiopathic PAH. Belief elicitation for bayesian priors. *J Rheumatol* 2011;38:462-9; Epub 2010 Dec 15.
7. Bernatsky S, Joseph L, Pineau CA, Belisle P, Boivin JF, Banerjee D, et al. Estimating the prevalence of polymyositis and dermatomyositis from administrative data: age, sex and regional differences. *Ann Rheum Dis* 2009;68:1192-6.
8. Launois R, Avouac B, Berenbaum F, Blin O, Bru I, Fautrel B, et al. Comparison of certolizumab pegol with other anticytokine agents for treatment of rheumatoid arthritis: a multiple-treatment Bayesian metaanalysis. *J Rheumatol* 2011;38:835-45.
9. Sung L, Beyene J, Hayden J, Nathan PC, Lange B, Tomlinson GA. A Bayesian meta-analysis of prophylactic granulocyte colony-stimulating factor and granulocyte-macrophage colony-stimulating factor in children with cancer. *Am J Epidemiol* 2006;163:811-7.
10. Jadad AR, Moore RA, Carroll D, Jenkinson C, Reynolds DJ, Gavaghan DJ, et al. Assessing the quality of reports of randomized clinical trials: is blinding necessary? *Control Clin Trials* 1996;17:1-12.
11. Sung L, Hayden J, Greenberg ML, Koren G, Feldman BM, Tomlinson GA. Seven items were identified for inclusion when reporting a Bayesian analysis of a clinical study. *J Clin Epidemiol* 2005;58:261-8.
12. The BaSiS Group. Bayesian Standards in Science. Standards for reporting of Bayesian analysis in the scientific literature 2001. [Internet. Accessed February 8, 2011.] Available from: <http://lib.stat.cmu.edu/bayesworkshop/2001/BaSiS.html>
13. Spiegelhalter D, Myles J, Jones D, Abrams K. Bayesian methods in health technology assessment: a review. *Health Technol Assess* 2000;4:1-130.
14. Spiegelhalter DJ, Freedman LS, Parmar MK. Applying Bayesian ideas in drug development and clinical trials. *Stat Med* 1993;12:1501-11; discussion 13-7.
15. Wijesundera DN, Austin PC, Hux JE, Beattie WS, Laupacis A. Bayesian statistical inference enhances the interpretation of contemporary randomized controlled trials. *J Clin Epidemiol* 2009;62:13-21;e5.
16. Moye LA. Bayesians in clinical trials: asleep at the switch. *Stat Med* 2008;27:469-82; discussion 83-9.

*J Rheumatol* 2011;38:794-6; doi:10.3899/jrheum.110013