Publishing a Bayesian study – a sometimes overlooked hurdle

Ros Walley, Foteini Strimenopoulou and Andy Grieve

PSI Conference 2018, Amsterdam

Inspired by patients.
Driven by science.

Outline

- Context
- Principles for Bayesian reporting
- Potentially contentious topics
- Examples of published studies
- Conclusions

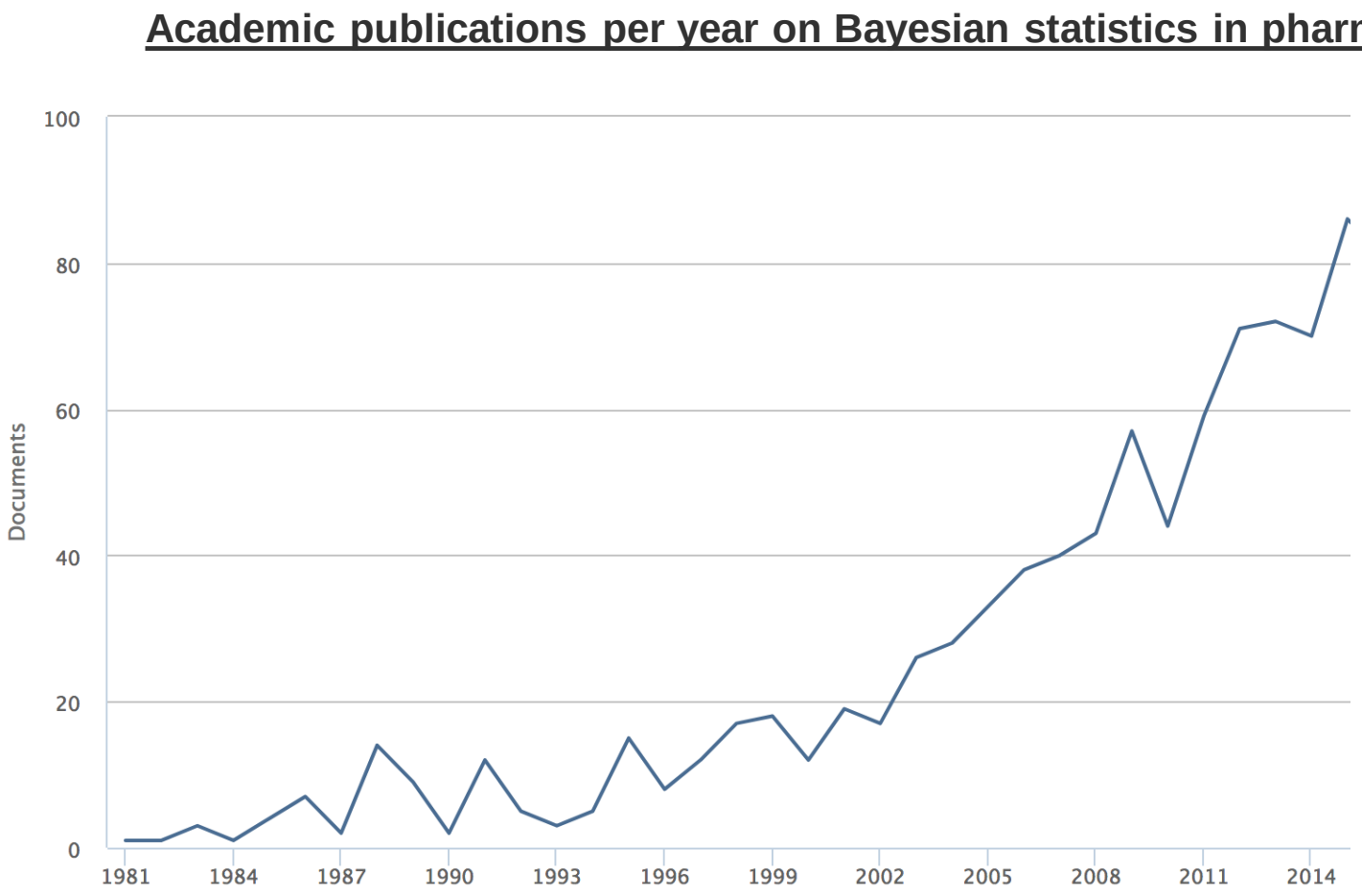
Context

"A drug is a substance which, if injected into a rabbit, produces a paper."

Otto Loewi

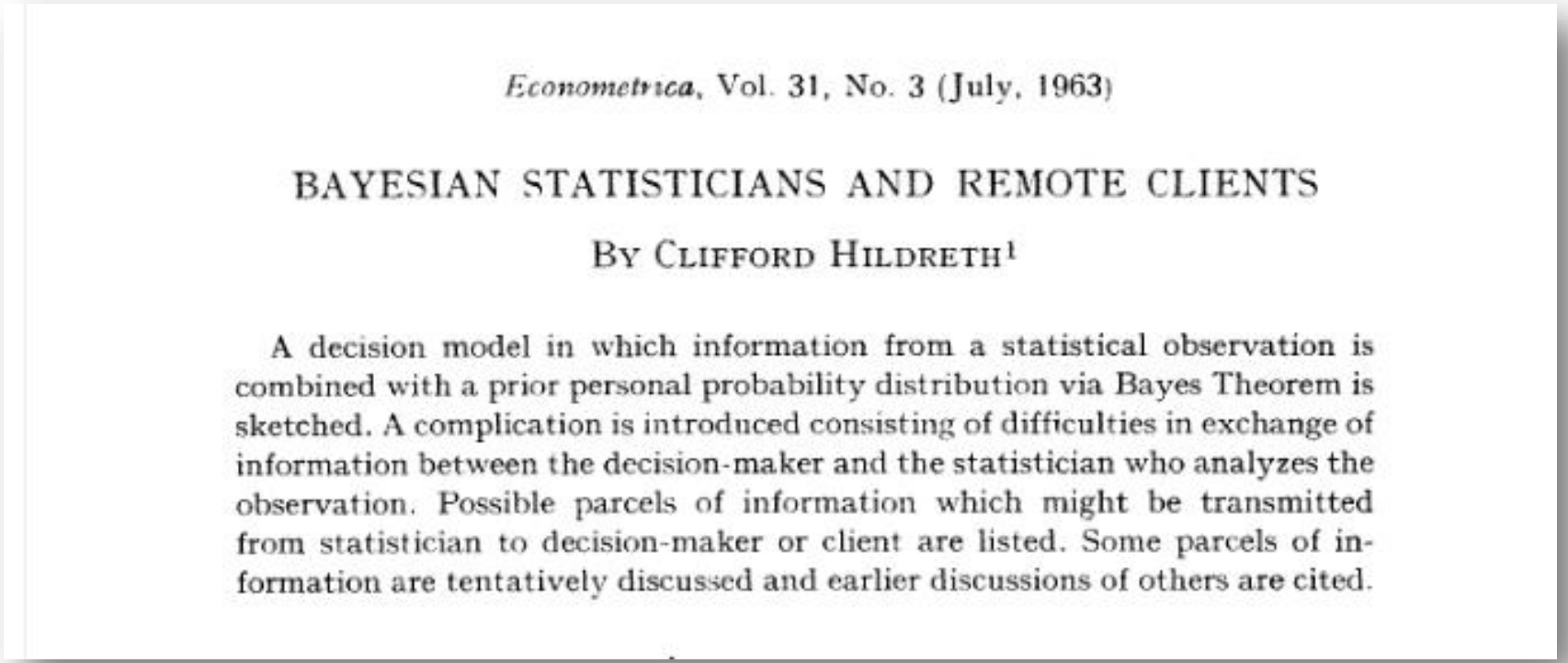
Source of image: https://www.pets4homes.co.uk/pet-advice/vaccinations-in-rabbits.html

Increasing number of Bayesian publications in pharma



Source: SCOPUS
Search: bayesian AND statistic*

The problem is not new…..



Econometrica, Vol. 31, No. 3 (July, 1963)

BAYESIAN STATISTICIANS AND REMOTE CLIENTS

By Clifford Hildreth[1]

A decision model in which information from a statistical observation is combined with a prior personal probability distribution via Bayes Theorem is sketched. A complication is introduced consisting of difficulties in exchange of information between the decision-maker and the statistician who analyzes the observation. Possible parcels of information which might be transmitted from statistician to decision-maker or client are listed. Some parcels of information are tentatively discussed and earlier discussions of others are cited.

"The results are summarised in a bulletin, article, or book, which is essentially a more or less extended memorandum "to who it may concern".

…

<The readers'> purposes, utility functions, prior ideas and other sources of information may differ widely. Uses of the analysis may extend over a considerable period of time after the statisticians work has been completed."

Additional stakeholders at publication stage

New study team members

Internal publications team

Medical writer

Internal governance

Journal editor

Manuscript referees

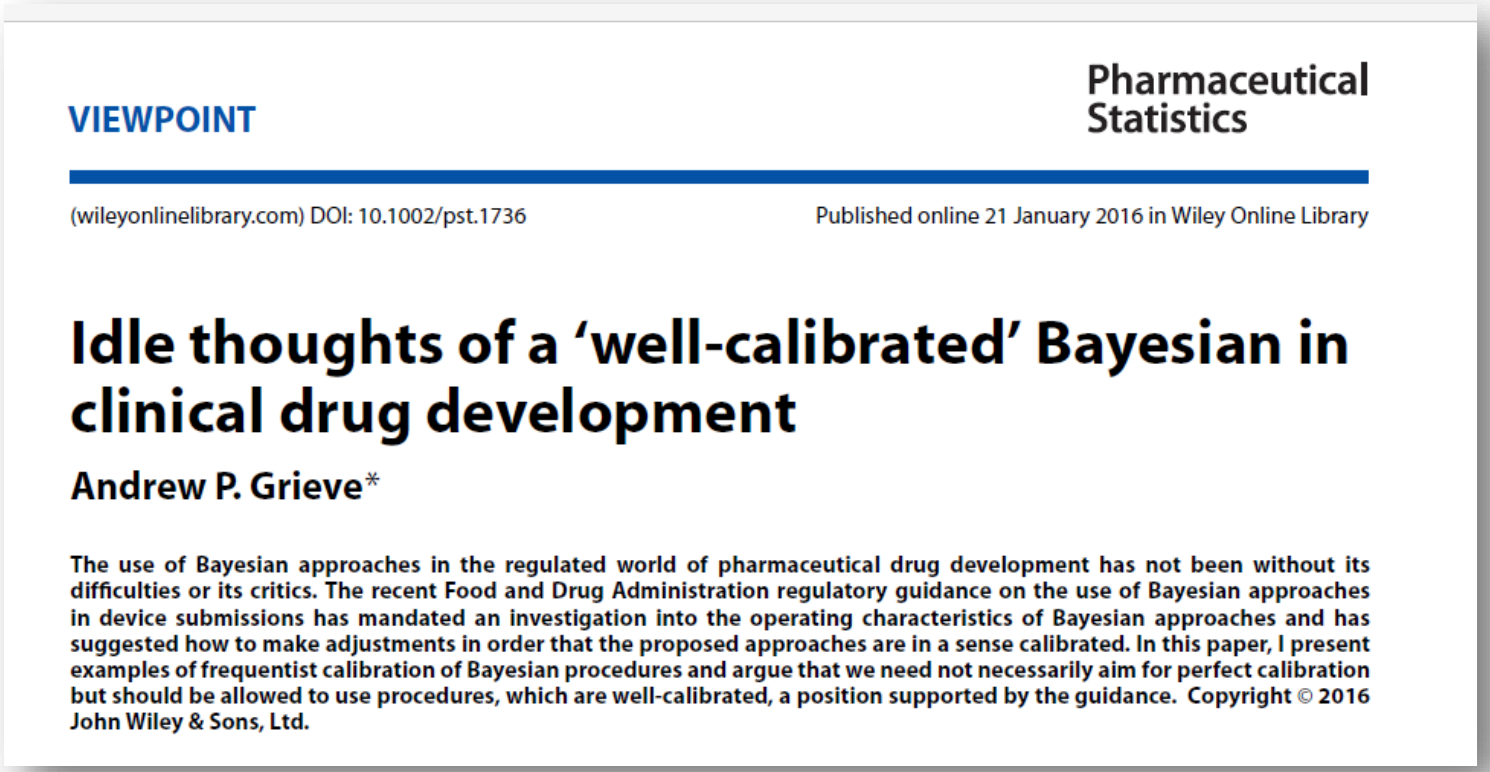
The ultimate audience….

Journal readership

Images: https://www.shutterstock.com/video/clip-25302104-stock-footage-gold-colored-metal-stanchions-with-red-velvet-rope-for-crowd-control.html

Principles for Bayesian reporting

Bayesian reporting guidelines



VIEWPOINT

Pharmaceutical Statistics

(wileyonlinelibrary.com) DOI: 10.1002/pst.1736

Published online 21 January 2016 in Wiley Online Library

Idle thoughts of a 'well-calibrated' Bayesian in clinical drug development

Andrew P. Grieve*

The use of Bayesian approaches in the regulated world of pharmaceutical drug development has not been without its difficulties or its critics. The recent Food and Drug Administration regulatory guidance on the use of Bayesian approaches in device submissions has mandated an investigation into the operating characteristics of Bayesian approaches and has suggested how to make adjustments in order that the proposed approaches are in a sense calibrated. In this paper, I present examples of frequentist calibration of Bayesian procedures and argue that we need not necessarily aim for perfect calibration but should be allowed to use procedures, which are well-calibrated, a position supported by the guidance.

Has a section on "Reporting of Bayesian Analyses of Clinical Trials"

Describes both academic guidance and the regulatory view

Academic guidelines for reporting Bayesian analysis

Health Technology Assessment 2000; Vol. 4: No. 38

Methodology

Bayesian methods in health technology assessment: a review

DJ Spiegelhalter
JP Myles
DR Jones
KR Abrams

ELSEVIER

Journal of Clinical Epidemiology 58 (2005) 261–268

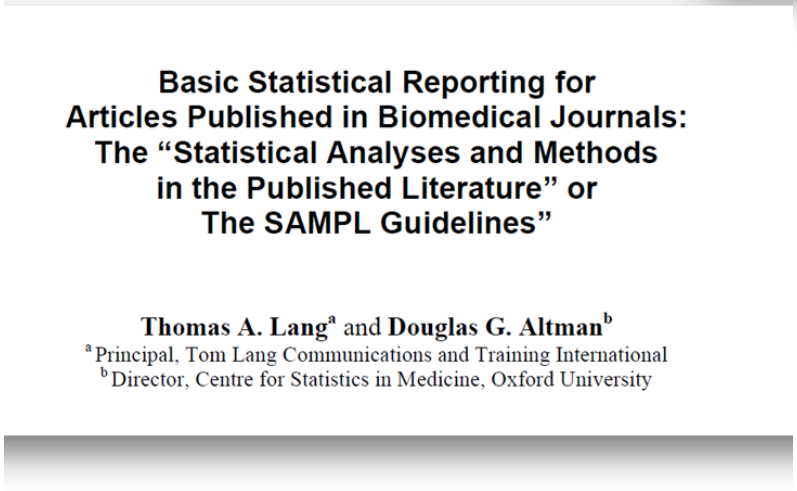
Journal of Clinical Epidemiology

Seven items were identified for inclusion when reporting a Bayesian analysis of a clinical study

Lillian Sung[a,b,c,d,*], Jill Hayden[b,e], Mark L. Greenberg[a,c], Gideon Koren[a,b,c,d], Brian M. Feldman[a,b,c,d,f], George A. Tomlinson[f,g,h]

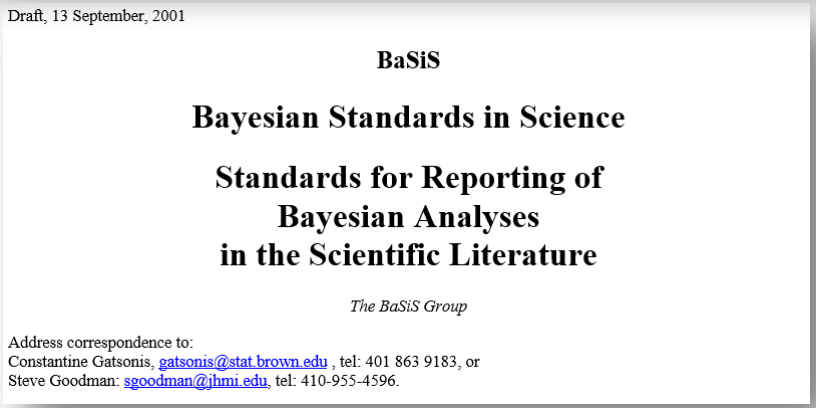
[a]Division of Hematology/Oncology, Department of Pediatrics, Hospital for Sick Children, Toronto, Ontario, Canada
[b]Department of Health Policy Management and Evaluation, University of Toronto, Toronto, Ontario, Canada
[c]Department of Pediatrics, Hospital for Sick Children, Toronto, Ontario, Canada
[d]Program in Population Health Sciences, Hospital for Sick Children, Toronto, Ontario, Canada
[e]Institute for Work and Health, Toronto, Ontario, Canada
[f]Department of Public Health Sciences, University of Toronto, Toronto, Ontario, Canada
[g]Department of Medicine, University of Toronto, Toronto, Ontario, Canada
[h]Department of Medicine, University Health Network, Toronto, Ontario, Canada

Accepted 30 August 2004



Basic Statistical Reporting for Articles Published in Biomedical Journals: The "Statistical Analyses and Methods in the Published Literature" or The SAMPL Guidelines"

Thomas A. Lang[a] and **Douglas G. Altman**[b]
[a] Principal, Tom Lang Communications and Training International
[b] Director, Centre for Statistics in Medicine, Oxford University



Draft, 13 September, 2001

BaSiS

Bayesian Standards in Science

Standards for Reporting of Bayesian Analyses in the Scientific Literature

The BaSiS Group

Address correspondence to:
Constantine Gatsonis, gatsonis@stat.brown.edu , tel: 401 863 9183, or
Steve Goodman: sgoodman@jhmi.edu, tel: 410-955-4596.

Academic Guidelines for Reporting Bayesian Analyses

ROBUST

- Prior Distribution
  - Specified
  - Justified
  - Sensitivity analysis
- Analysis
  - Statistical model
  - Analytical technique
- Results
  - Central tendency
  - SD or Credible Interval

BAYESWATCH

- Introduction
  - Intervention described
  - Objectives of study
- Methods
  - Design of Study
  - Statistical model
  - Prior / Loss function?
    - When constructed
    - Prior / Loss descriptions
  - Use of Software
    - MCMC , starting values, run-in, length of runs, convergence, diagnostics
- Results
- Interpretation
  - Posterior distribution summarized
  - Sensitivity analysis if alternative priors used

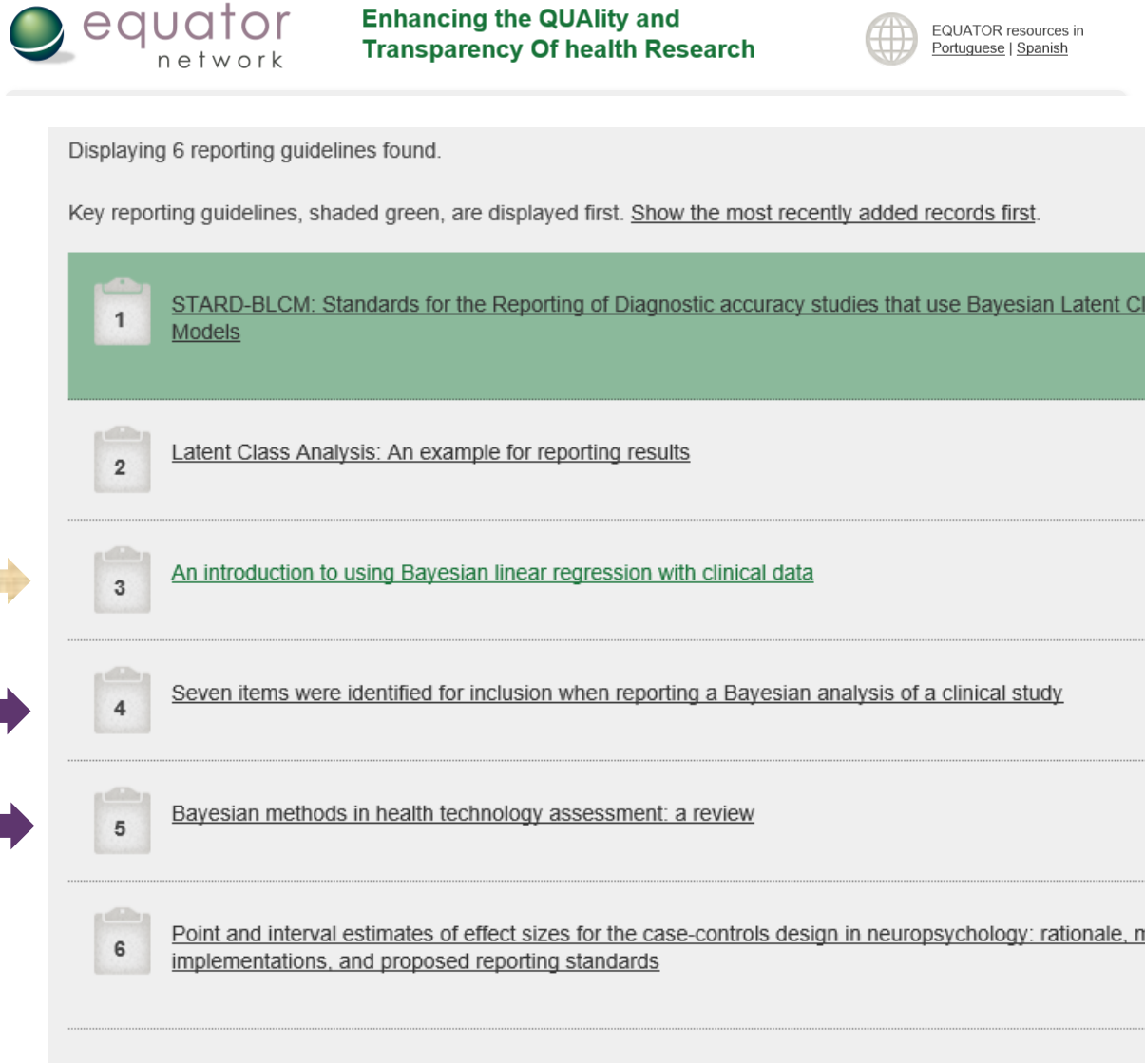
BASIS

- Research Question
- Statistical model
  - Likelihood, structure, prior & rationale
- Computation
  - Software - convergence if MCMC, validation, methods for generating posterior summaries
- Model checks, sensitivity analysis
- Posterior Distribution
  - Summaries used: i). Mean, std, quintiles ii) posterior shape, (iii) joint posterior for mult comp, (iv) Bayes factors
- Results of model checks and sensitivity analyses
- Interpretation of Results
- Limitation of Analysis

SAMPL

- Prior Distribution
  - Specified
  - Justified
  - Sensitivity analysis
- Analysis
  - Statistical model
  - Analytical technique
  - Software
- Results
  - Central tendency
  - SD or Credible Interval

EQUATOR list of guidelines



Kostoulas et al

Schreiber

Baldwin and Larson

ROBUST

BAYESWATCH

Crawford et al

More recent publications on guidelines for publishing

Behaviour Research and Therapy 98 (2017) 58–75

Contents lists available at ScienceDirect

Behaviour Research and Therapy

journal homepage: www.elsevier.com/locate/brat

An introduction to using Bayesian linear regression with clinical data

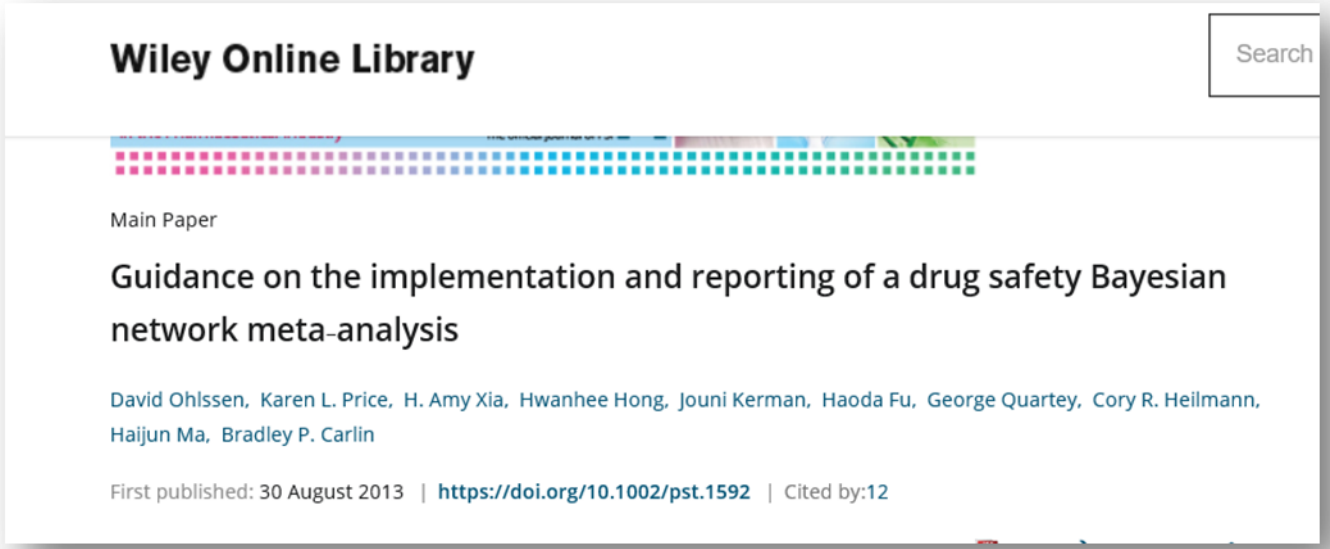
Scott A. Baldwin [a, *], Michael J. Larson [b]

[a] Department of Psychology, Brigham Young University, USA
[b] Department of Psychology and Neuroscience Center, Brigham Young University, USA

Minimum reporting guidelines:

- **Complete description of likelihood and all priors**
- **Details of convergence**
- **Software details**
- **Provide data and scripts where possible**
- **Be explicit how posterior is summarised**
- **Report no of draws from the posterior as well as the effective sample size**

More recent publications on guidelines for publishing



Specific context of network meta-analysis for safety but some good general points, including:

- **Rationale for choice of prior; justification of exchangeability assumption; when prior was specified**
- **Stats model: rate of missingness; imputation of missing covariates; structure of levels of model….**
- **Impact of prior on posterior**
- **Possible limitations of the analysis**

Decision/success criteria

- **Clinical studies typically have decision criteria associated with them**
  - May not be explicitly stated. E.g. safety objectives, criteria related to $p<0.05$
- **Bayesian guidelines do not focus on decision criteria**
- **Decision criteria for internal benchmarking may not be "set in stone"**
  - Leads to a reluctance to document and publicise

However,

- **Studies need a sample size justification**
  - So typically at least one decision criterion is at least implied by the protocol
- **If decision criteria are included in the protocol or SAP, then**
  - Easy to demonstrate they were "prespecified"
  - Thus easier to convince sceptical reviewers

Potentially contentious areas

Comments from stakeholders at publication stage

The Bayesian method hasn't worked (the estimated treatment diff doesn't match the summary stats)

Can we include p-values?

Statistical analysis should be performed paired and unpaired fashion for all data

NB. Not all these issues are specifically Bayesian

Comments from stakeholders at publication stage (2)

NB. Not all these issues are specifically Bayesian

Comments from stakeholders at publication stage (3)

Not all decision-criteria were described in protocol

of the prior in the Supplemental data should also show the treatment group and how the prior impacted its shift.

Can we include p-values?

e Bayesian method hasn't worked (the estimated treatment diff doesn't match the summary stats)

Standard of care threshold is now out of date and we don't want to mention it

We are extremely limited by the number of additional words that we can add to the main body of the manuscript

Statistical analysis should be performed paired and unpaired fashion for all data

There is nothing wrong with Bayesian approaches to analyses, but the choice of the parameters for the priors is unsupported

Need to justify a non-zero threshold based on standard of care (in-house assessment of a competitor)

NB. Not all these issues are specifically Bayesian

More complex analysis can be viewed with suspicion

Dynamically downweighting priors e.g. mixture priors

- “Are you changing the prior after the study has started?”

Extending the likelihood to downweight outliers e.g using a t-distribution or mixture likelihood

- “Which outliers have been excluded?”

Assessment of the design is more complex than in the frequentist setting. E.g.

- type 1 error/bias when there are informative priors
- performance of dynamically downweighting priors
- Assessment of prior data conflict
- “too much detail” or “what are you hiding?”

Examples: excerpts from published papers

Secukinumab POC in ankylosing spondylitis

Articles

Anti-interleukin-17A monoclonal antibody secukinumab in treatment of ankylosing spondylitis: a randomised, double-blind, placebo-controlled trial

Dominique Baeten, Xenofon Baraliakos, Jürgen Braun, Joachim Sieper, Paul Emery, Désirée van der Heijde, Iain McInnes, Jacob M van Laar, Robert Landewé, Paul Wordsworth, Jürgen Wollenhaupt, Herbert Kellner, Jacqueline Paramarta, Jiawei Wei, Arndt Brachat, Stephan Bek, Didier Laurent, Yali Li, Ying A Wang, Arthur P Bertolino, Sandro Gsteiger, Andrew M Wright, Wolfgang Hueber

Summary

Methods: The primary efficacy endpoint was the percentage of patients with a 20% response according to the Assessment of SpondyloArthritis international Society criteria for improvement (ASAS20) at week 6 (Bayesian analysis).

Findings: At week 6, ASAS20 response estimates were 59% on secukinumab versus 24% on placebo (99.8% probability that secukinumab is superior to placebo)

PDE5 POC in diabetic nephropathy

CLINICAL RESEARCH www.jasn.org

Phosphodiesterase Type 5 Inhibition Reduces Albuminuria in Subjects with Overt Diabetic Nephropathy

Wim Scheele,* Susan Diamond,† Jeremy Gale,* Valerie Clerin,* Nihad Tamimi,‡ Vu Le,* Rosalind Walley,‡ Fernando Grover-Páez,§ Christelle Perros-Huguet,* Timothy Rolph,* and Meguid El Nahas‖

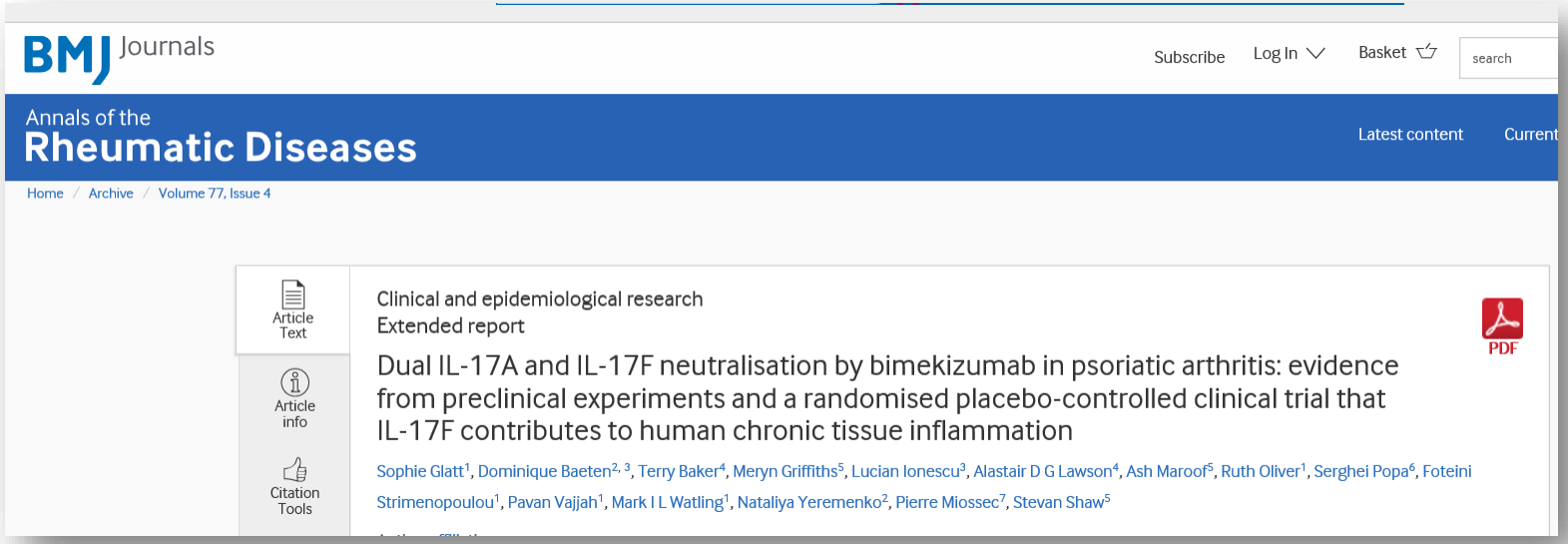
*Pfizer Inc., Cambridge, Massachusetts; †San Antonio Kidney Disease Center, San Antonio, Texas; ‡Pfizer Ltd.,

Abstract: Using the predefined primary assessment of efficacy (Bayesian analysis with informative prior), we observed ….

Statistical analysis:

- **A two-part decision criteria for efficacy and futility was used at the end of this study which is described in the Supplemental Material**
- **More details of the statistical design and analysis including the outlier robust approach can be found in ….**

Bimekizumab POC in psoriatic arthritis (1)



Statistical methods:

The analyses of ACRn and ACR20 at week 8 followed the Bayesian paradigm to improve the operating characteristics of the study design. The Bayesian methodology was used for the analysis of ACRn and ACR20 at week 8 assuming informative priors only on the placebo response (vague prior distributions were assumed for all other parameters). For both endpoints, the prior placebo response, in terms of ACR20 response rate, was approximately 25% and the effective sample size was 32. Normal likelihood model and logistic model were assumed for ACRn and ACR20, respectively.

Bimekizumab POC in psoriatic arthritis (2)

Statistical methods:

Two study efficacy criteria were predefined based on ACRn at week 8. For declaring PoC, we required a high (≥97.5%) posterior probability that bimekizumab is superior to placebo. To give further confidence about the effect size, the posterior probability that bimekizumab improvement over placebo exceeds a clinically relevant effect (ie, approximately 25% difference from placebo in ACR20 terms) was required to be ≥70%.

Conclusions

Conclusions

There are additional stakeholders at publication stage

- They don’t have the benefit of understanding/inputting to the design in advance of the study
- May be expecting a frequentist discussion
- They may be suspicious/inexperienced with Bayesian methods
- Include a respected stats expert for the contentious points

Consider the potential publication(s) at the design stage

- Prespecify and document as much as possible
- Consider alternative priors/demonstrate robustness of conclusions to these
- Can’t easily demonstrate pre-specification if its not in the protocol or SAP
- Clear wording in protocol and SAP
- Consider publishing design in stats journal (more sympathetic to technical details)

In the publication

- Follow available guidelines for publication
- Be clear what's been pre-specified
- Be clear what's the primary analysis and what's not
- Warn other authors that glossing over details may lead to suspicion
- If authors wont agree, at least put the wording in the supplemental documents

References (1)

Baeten D, Baraliakos X, Braun J, Sieper J, Emery P, Van Der Heijde D, McInnes I, Van Laar JM, Landewé R, Wordsworth P, Wollenhaupt J. Anti-interleukin-17A monoclonal antibody secukinumab in treatment of ankylosing spondylitis: a randomised, double-blind, placebo-controlled trial. The Lancet. 2013 Nov 23;382(9906):1705-13.

Baldwin SA, Larson MJ. An introduction to using Bayesian linear regression with clinical data. Behaviour research and therapy. 2017 Nov 30;98:58-75.

The BaSiS Group. Bayesian standards in science (BaSiS) [Internet]. Draft: Sept. 13, 2001. Available at: http://lib.stat.cmu.edu/bayesworkshop/2001/BaSis.html.

Crawford JR, Garthwaite PH, Porter S. Point and interval estimates of effect sizes for the case-controls design in neuropsychology: rationale, methods, implementations, and proposed reporting standards. Cognitive neuropsychology. 2010 May 1;27(3):245-60.

Glatt S, Baeten D, Baker T, et al. Dual IL-17A and IL-17F neutralisation by bimekizumab in psoriatic arthritis: evidence from preclinical experiments and a randomised placebo-controlled clinical trial that IL-17F contributes to human chronic tissue inflammation. Annals of the Rheumatic Diseases 2018;77:523-532

Grieve AP. Idle thoughts of a 'well-calibrated'Bayesian in clinical drug development. Pharmaceutical statistics. 2016 Mar 1;15(2):96-108.

Hildreth C. Bayesian statisticians and remote clients. Econometrica: Journal of the Econometric Society. 1963 Jul 1:422-38.

References (2)

Kostoulas P, Nielsen SS, Branscum AJ, Johnson WO, Dendukuri N, Dhand NK, Toft N, Gardner IA. STARD-BLCM: Standards for the Reporting of Diagnostic accuracy studies that use Bayesian Latent Class Models. Prev Vet Med. 2017; 138:37-47

Lang TA, Altman DG. Basic statistical reporting for articles published in biomedical journals: the “Statistical Analyses and Methods in the Published Literature” or the SAMPL Guidelines”. Handbook, European Association of Science Editors. 2013;256:256.

Ohlssen D, Price KL, Amy Xia H, Hong H, Kerman J, Fu H, Quartey G, Heilmann CR, Ma H, Carlin BP. Guidance on the implementation and reporting of a drug safety Bayesian network meta-analysis. Pharmaceutical statistics. 2014 Jan 1;13(1):55-70.

Pennello G, Thompson L. Experience with reviewing Bayesian medical device trials. Journal of biopharmaceutical statistics. 2007 Dec 10;18(1):81-115.

Schreiber JB. Latent Class Analysis: An example for reporting results. Res Social Adm Pharm. 2016.

Spiegelhalter DJ, Myles JP, Jones DR, Abrams KR. Bayesian methods in health technology assessment: a review. Health Technol Assess 2000;4(38):1–130

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. Journal of clinical epidemiology. 2005 Mar 1;58(3):261-8.

Backup

Some differences between reporting frequentist and Bayesian analysis

	Frequentist	Bayesian
Readership	Familiarity with methods	May be less familiar and even suspicious
Success criterion	Assumed to be 5% statistical significance for primary comparison (based on a 2-sided test)	No such assumption
Prior belief	Described in intro and discussion	Specified mathematically
Approach	Typically use analytical results e.g. Transform and remove outliers to assume normally distributed data	Can easily use different distributions for the model. No particular driver for simpler models.
Estimand	Summary measures relates to analytical approach e.g. odds ratio in logistic regression	No restriction

Bayesian reporting guidelines

Journal of Biopharmaceutical Statistics, 18: 81–115, 2008

ISSN: 1054-3406 print/1520-5711 online
DOI: 10.1080/10543400701668274

Taylor & Francis
Taylor & Francis Group

EXPERIENCE WITH REVIEWING BAYESIAN MEDICAL DEVICE TRIALS

Gene Pennello and Laura Thompson
Division of Biostatistics, Rockville, Maryland, USA

The purpose of this paper is to present a statistical reviewer's perspective on some technical aspects of reviewing Bayesian medical device trials submitted to the Food and Drug Administration. The discussion reflects the experiences of the authors and should not be misconstrued as official guidance by the FDA. A variety of applications are described, reflecting our experience with therapeutic and diagnostic devices. In addition to Bayesian analysis of trials, Bayesian trial design and Bayesian monitoring are discussed. Analyses were implemented in WinBUGS (http://www.mrc-bsu.cam.ac.uk/bugs/winbugs/contents.shtml), with the code provided.

Describes FDA reviewers perspective of reviewing Bayesian medical device trials

N.B. Refers to a regulatory review rather than a peer review for a journal