Using Bayesian Statistics in Health Psychology: A comment on Depaoli et al. (2017)

Emma Beard & Robert West

Accepted author version posted online: 10 Jul 2017.

Submit your article to this journal ⬈

View related articles ⬈

View Crossmark data ⬈

Using Bayesian Statistics in Health Psychology: A comment on Depaoli et al. (2017)

Health Psychology Review

Emma Beard[1,2] & Robert West[2]

[1] Research Department of Clinical, Educational and Health Psychology, University College London, London

[2] Department of Behavioural Science and Health, University College London, London

Word count: 1220 (max 1500)

Using Bayesian Statistics in Health Psychology: A comment on Depaoli et al (2017)

Health Psychology Review

Emma Beard[1,2] & Robert West[2]

[1] Research Department of Clinical, Educational and Health Psychology, University College London, London

[2] Department of Behavioural Science and Health, University College London, London

Word count: 1220 (max 1500)

Health psychology has advanced our understanding of the influence of biopsychosocial factors on disease development and has been instrumental in improving health-care services and the treatment of chronic conditions, but like other areas of psychology has been criticised for publishing underpowered studies. This may be in part be due to many researchers relying on intuition and rules of thumb when determining sample size (Maxwell, 2004). There are also strong incentives to engage in research practices that make findings publishable more quickly (Bakker et al., 2016) and that produce novel statistically significant results (typically $p < .05$; Fanelli , 2011; Nosek, Spies, & Motl, 2012).

One proposed solution has been the publication of pre-specified analysis plans which have for some time been a gold standard for randomised controlled trials (Head et al., 2015). A suitable platform for this is the Open Science Framework, which allows documents to be published online for free in a timely manner, and is now a recommended repository for the Nature Publishing Group's data journal (Munafò et al., 2017). Another proposed solution is the use of Bayesian analysis.

Historically, psychologists have tended to focus on frequentist statistics, following a hybrid of the Fisherian and Neyman-Pearson approach. In brief, this involves specifying a null and alternative hypothesis, deciding on a significance level, collecting the data and calculating the tests statistic and p-value, and then comparing this p-value to the significance level. As this approach focuses on the probability of the data given the null hypothesis, it cannot tell us the converse i.e. the probability of the null hypothesis given the data, although p-values are often misinterpreted in this way (Open Science Collaboration, 2012).

In the Bayesian framework, one can evaluate the extent of evidence for and against both the null and alternative hypothesis. It is also entirely appropriate to collect data until a hypothesis has been confirmed or disconfirmed with a specified level of confidence and to conduct analyses on small sample sizes without violating underlying assumptions (Wagenmakers et al., 2011). Researchers should of course be warned against selecting this approach in order to justify an underpowered study. In fact, Bayesian estimates in the presence of a miss-specified prior distribution can actually perform worse than a frequentist approach with small sample corrections (Edwards, Lindman, &

Savage, 1963) and still require a study to have a greater number of observations relative to parameters (McNeish, 2016).

Bayesian analysis works by combining in a natural manner our past experience (our prior knowledge or prior probability) with the results of our current experiment (data) to result in a revised probability (posterior probability):

$$P(hypothesis|data) \propto P(hypothesis) \times P(data|hypothesis)$$

The fact that we use the word 'our' indicates the subjective nature of this method of statistical inference. It is this concern and the belief that its concepts are difficult to understand and are hard to implement that has probably contributed to its under-use in Health Psychology (Lee & Song, 2016). Since 2010 there has been growth in the absolute number of published articles using or referencing Bayesian analysis across subject areas (see Figure 1) (partly accountable no doubt to a growth in the total number of articles published), but only two of these articles were published in Health Psychology affiliated journals: *Health Psychology*, *Journal of Health Psychology*, *Journal of Occupational Psychology Review*, *Health Psychology Research*, *Health Psychology and Behavioural Medicine*, *Health Psychology Open*, *International Journal of Clinical and Health Psychology*, *Psychology and Health*, and Behavioural Medicine (Lesaffre & Lawson, 2012; Warner et al., 2016).

This is despite the fact that psychology students are adept at acquiring an intuitive understanding of a number of Bayesian concepts and that advances in software means that implementation is relatively easy (Change et al., 2012). Several psychology journals are also now actively encouraging individuals to consider a Bayesian approach, including the journal Basic and Applied Social Psychology, which has completely banned the use of null hypothesis significant tests. One can however, as we discuss below, use a Bayesian and Frequentist approach in unison in the form of a Bayes Factor.
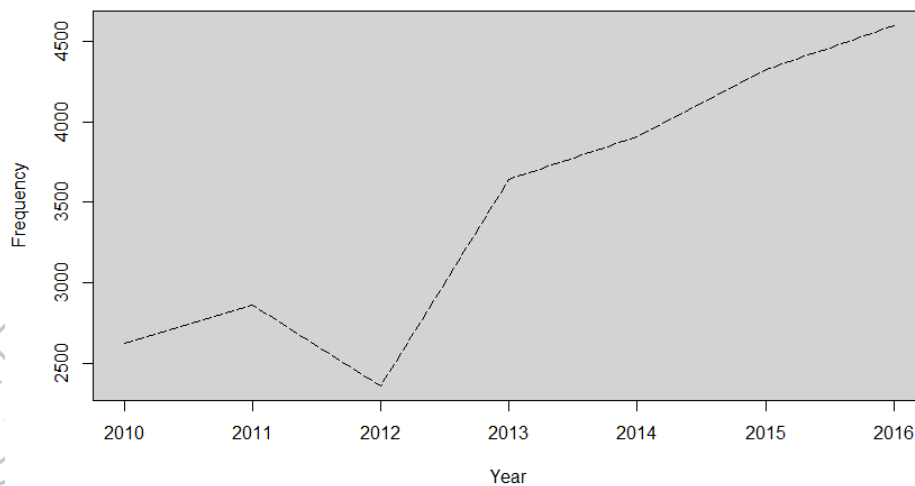

Figure 1: Number of articles published which reference Bayes or Bayesian analysis in the title/abstract (identified by PubMed)

As such, the article by Depaoli, Rus, Clifton, van de Schoot and Tiemensma (2017) in this issue of Health Psychology Review is very welcome. The authors provide a detailed introduction to Bayesian analysis for normally distributed data and a worked example of implementing Bayesian estimation in

the context of blood pressure changes after experience of an acute stressor. Although Bayesian analysis is not available in SPSS, the authors offer syntax for a regression analysis in many different statistical programs including Mplus and R (the latter being free). Importantly, there is a clear detailed checklist for the reporting of Bayesian analysis which includes the need to specify how prior probabilities were arrived at and assessing the sensitivity of the model to these. The authors acknowledge that it is unlikely that this article alone will be sufficient to learn how to conduct or implement a Bayesian analysis, and readers are referred to a number of online materials and advanced readings. For example, the authors do not discuss non-normally distributed data which may require an understanding of the binomial, gamma, beta and/or uniform distributions.

There is also a possibility of taking a semi-Bayesian rather than a full-Bayesian approach (for a detailed discussion of the differences see Díaz, 2010). This is done through the calculation of Bayes Factors from output from Frequentist analysis procedures. Bayes Factors are simply the ratio of the (average) likelihood of two hypotheses being correct given a set of data and can be seen as equivalent to a likelihood ratio:

$$Bayes\ Factor = \frac{P(data|alternative\ hypothesis)}{P(data|null\ hypothesis)}$$

A Bayes Factor equivalent of the Frequentist p<0.05 threshold has been proposed by Jeffreys (Rouder & Morey, 2010): a Bayes Factor greater than 3 was suggested as a reasonable threshold for considering that the experimental hypothesis is correct while a Bayes Factor less than 1/3rd would lead to a judgement that the experimental hypothesis was incorrect. Several software packages are available for the calculation of Bayes Factors including an online calculator developed by Zoltan Dienes (http://www.lifesci. sussex.ac.uk/home/Zoltan_Dienes/inference/Bayes.htm) and a modified version by John Christie using R code (Kruschke, 2011). Both approaches require the specification of an expected effect size (i.e. a plausible range of predicted values based on previous studies, judgement or clinical significance), the published effect size (e.g. mean difference or log odds ratio) and standard error of this parameter (for further details see Jeffreys, 1961).

In summary, there is need for Health Psychologists to consider alternatives to the traditional Frequentist statistical analysis approach. Bayesian statistics, be it a full analysis as suggested by Depaoli and colleagues or a Bayes Factor approach, has several advantages including the ability to make inferences on small sample sizes and to draw inferences for and against the experimental hypothesis. There are limitations to its implementation including the complexity of the analysis, but Depaoli et al. (2017)'s article provides an excellent introduction to the area. Researchers should of course be warned against selecting this approach in order to justify an underpowered study and pre-specified analysis plans with expected effect sizes should be made available. This is of course difficult when publication record and career success seem to be maximized through the conduction of small studies with 10-40% statistical power (Christie, 2011). Bayesian analysis is also not the sole solution to the reproducibility problem inherent in psychological research. This will also require a consideration of how to protect studies from cognitive biases, improving methodological training, increased collaborations, encouraging transparency and open data, diversifying peer review and rewarding correct statistical practices (Head et al., 2015).

References

Bakker M, Hartgerink CHJ, Wicherts JM, et al. Researchers' Intuitions About Power in Psychological Research. Psychological Science 2016;**27**(8):1069-77 doi: 10.1177/0956797616647519[published Online First: Epub Date]|.

Chang DS, Kang OS, Kim HH, et al. Pre-existing beliefs and expectations influence judgments of novel health information. Journal of health psychology 2012;**17**(5):753-63 doi: 10.1177/1359105311421044[published Online First: Epub Date]|.

Depaoli, S., Rus, H., Clifton, J., van de Schoot, R., & Tiemensma, J. (2017). An introduction to Bayesian statistics in health psychology. *Health Psychology Review*. doi: 10.1080/17437199.2017.1343676

Díaz C. Psychology Students' Understanding of Elementary Bayesian Inference. Secondary Psychology Students' Understanding of Elementary Bayesian Inference 2010.

Edwards W, Lindman H, Savage LJ. Bayesian statistical inference for psychological research. Psychological review 1963;**70**(3):193

Fanelli D. Negative results are disappearing from most disciplines and countries. Scientometrics 2011;**90**(3):891-904

Head ML, Holman L, Lanfear R, et al. The extent and consequences of p-hacking in science. PLoS Biol 2015;**13**(3):e1002106

Jeffreys H. Theory ofprobability: Oxford: Clarendon Press, 1961.

Kruschke JK. Bayesian assessment of null values via parameter estimation and model comparison. Perspectives on Psychological Science 2011;**6**(3):299-312

Lee S-Y, Song X-Y. Evaluation of the Bayesian and Maximum Likelihood Approaches in Analyzing Structural Equation Models with Small Sample Sizes. Multivariate Behavioral Research 2004;**39**(4):653-86 doi: 10.1207/s15327906mbr3904_4[published Online First: Epub Date]|.

Lesaffre E, Lawson AB. *Bayesian biostatistics*: John Wiley & Sons, 2012.

Maxwell SE. The persistence of underpowered studies in psychological research: causes, consequences, and remedies. Psychological methods 2004;**9**(2):147-63 doi: 10.1037/1082-989x.9.2.147[published Online First: Epub Date]|.

Munafò MR, Nosek BA, Bishop DV, et al. A manifesto for reproducible science. Nature Human Behaviour 2017;**1**:0021

Nosek BA, Spies JR, Motyl M. Scientific utopia: II. Restructuring incentives and practices to promote truth over publishability. Perspectives on Psychological Science 2012;**7**(6):615-31

Open Science Collaboration. An open, large-scale, collaborative effort to estimate the reproducibility of psychological science. Perspectives on Psychological Science 2012;**7**(6):657-60

Rouder JN, Morey RD. A Bayes factor meta-analysis of Bem's ESP claim. Psychonomic Bulletin & Review 2011;**18**(4):682-89

Wagenmakers EJ, Wetzels R, Borsboom D, et al. Why psychologists must change the way they analyze their data: the case of psi: comment on Bem (2011). 2011

McNeish D. On Using Bayesian Methods to Address Small Sample Problems. Structural Equation Modeling: A Multidisciplinary Journal 2016;**23**(5):750-73 doi: 10.1080/10705511.2016.1186549[published Online First: Epub Date]|.

Warner LM, Wolff JK, Ziegelmann JP, et al. Revisiting self-regulatory techniques to promote physical activity in older adults: null-findings from a randomised controlled trial. Psychology & health 2016;**31**(10):1145-65 doi: 10.1080/08870446.2016.1185523[published Online First: Epub Date]|.