STATISTICS IN MEDICINE
Statist. Med. 2008; **27**:619–624

LETTER TO THE EDITOR

Comments on 'Trying to be precise about vagueness'

by Stephen Senn, *Statistics in Medicine* 2007; **26**:1417–1430

From: *Paul C. Lambert*, *Alex J. Sutton*, *Paul R. Burton*, *Keith R. Abrams and David R. Jones*
Centre for Biostatistics & Genetic Epidemiology, *Department of Health Sciences*,
University of Leicester, *Leicester*, *U.K.*

In his paper 'Trying to be precise about vagueness' [1], Stephen Senn critiques a recent paper of ours where we use simulation methods to assess the impact of using vague prior distributions in hierarchical models [2]. We are pleased that Professor Senn agrees with us that 'vague priors be treated with a degree of caution'. However, there are a number of issues raised by Senn that we feel that we should comment on.

Much of Senn's paper is concerned with various meta-analysis issues that, although of some interest, are not directly relevant to the main focus of our paper, i.e. the use of vague prior distribution in hierarchical models with particular emphasis on scale parameters. We used meta-analysis as an example of a hierarchical model, but our simulation study is relevant to many types of hierarchical models. However, we will briefly comment on some of these meta-analysis issues before returning to the main issue of the use of vague prior distributions. The Otitis Media example is useful in illustrating some of the problems of the use of vague prior distributions whether or not particular studies should have been included in the original Cochrane meta-analysis. We apologize for incorrectly referencing the original meta-analysis. We agree that in the Otitis Media example and in the various simulation scenarios the difference in the *point estimate* of the study effect between fixed and random effect models will be small. It is of course for this reason why the main focus of our paper was on the specification of prior distributions for the between-study standard deviation and their impact on the uncertainty of study effect estimate and the estimate of the between-study standard deviation itself.

Senn questions the use of simulation methods to assess the use of different prior distributions. He states that '. . .the simulation is unfair and irrelevant to any Bayesian who truly believed what the prior distributions represented'. In addition, he states that 'Bayesians are bound to disagree since there is no such thing as necessary agreement'. Taken to extreme, these statements imply that a 'True' Bayesian would regard any analysis irrelevant unless the prior distributions used coincided with his/her own. However, many applied statisticians use Bayesian methods as a means to fit complicated models and perhaps regard the use of prior distributions as a necessary evil, hoping that the analysis will have good frequentist properties in terms of bias and coverage. Many of the papers that use Bayesian methods in this journal use only vague prior distributions. With the arrival of software such as WinBUGS, like it or not, vague prior distributions are being used in real analyses with important decisions being made in medicine and other areas. We believe that

Table I. Applying Senns's prior distribution to simulated meta-analysis of five studies with between-study standard deviations of 0.001, 0.3 and 0.8.

	Five studies		
Professor Senn's prior distribution	SD = 0.001	SD = 0.3	SD = 0.8
Treatment effect (θ)*			
$\alpha = 0.05, \beta = 0.1$	**0.324** 0.117 *96.6* per cent	**0.341** 0.144 *85.8* per cent	**0.462** 0.413 *61.6* per cent
$\alpha = 0.05, \beta = 1$	**0.307** 0.149 *99.1* per cent	**0.327** 0.208 *93.3* per cent	**0.395** 0.533 *89.5* per cent
$\alpha = 0.2, \beta = 0.1$	**0.322** 0.139 *99.0* per cent	**0.328** 0.178 92.2 per cent	**0.371** 0.380 *86.4* per cent
$\alpha = 0.2, \beta = 1$	**0.306** 0.159 *99.7* per cent	**0.315** 0.216 94.4 per cent	**0.364** 0.512 *92.4* per cent
Between-study standard deviation (τ)†			
$\alpha = 0.05, \beta = 0.1$	**0.056** 0.071 *6.4* per cent	**0.112** 0.101 *60.7* per cent	**0.489** 0.235 *67.2* per cent
$\alpha = 0.05, \beta = 1$	**0.126** 0.149 *0.7* per cent	**0.237** 0.214 *98.9* per cent	**0.744** 0.472 *91.3* per cent
$\alpha = 0.2, \beta = 0.1$	**0.113** 0.122 *0.5* per cent	**0.210** 0.156 *99.4* per cent	**0.596** 0.247 *85.9* per cent
$\alpha = 0.2, \beta = 1$	**0.147** 0.166 *0.1* per cent	**0.265** 0.225 *98.5* per cent	**0.753** 0.464 *92.9* per cent

*Mean values of pooled median effect size (bold font)—true effect is 0.323, mean of the standard deviation of the pooled effect size (normal font) and coverage for the 95 per cent credible interval for the pooled effect size (italic font).

†Mean values of median between-study standard deviation (bold font)—true effect is given in the column headings, mean of the standard deviation of the between-study standard deviation (normal font) and coverage for the 95 per cent credible interval for the between-study standard deviation (italic font).

these users may not be fully aware of the impact the choice of vague prior distribution can have and thus hope our paper raises awareness of this issue.

Senn argues that if a random effects analysis is desired it will be valuable to perform a frequentist analysis. We agree that the two methods quoted may be preferable in some situations to a Bayesian analysis. However, for more complex models, Bayesian methods will often be used, and awareness of the potential impact of vague prior distributions is crucial. Bayesian methods are used for

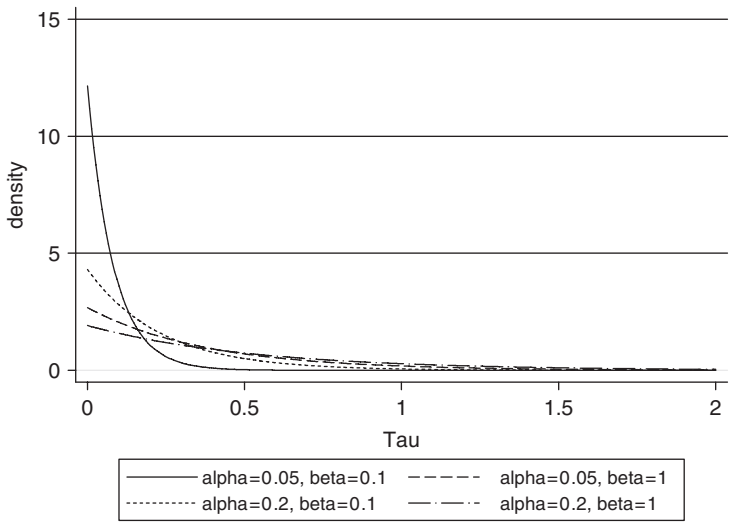

Figure 1. Senn's prior distributions.

a whole range of hierarchical models and, if there are problems in the simple models used in our paper, the problems will be at least as great for more complex models with more random effects.

Senn offers 'without any great enthusiasm' a conditional prior that allows dependence of the between-study standard deviation (τ) and the treatment effect (θ). We share Senn's lack of enthusiasm about this prior distribution. Firstly, we remain to be convinced about the need to link the two parameters in this way; is there any evidence that doing so will be beneficial? Secondly, through the choice of α and β, the analyst needs to assume some knowledge of what the relationship between τ and θ is. Of course we could perform multiple sensitivity analyses for a 'suitable' range of α and β, but we find this unappealing. We have applied Professor Senn's prior distributions to our simulated data sets. The results of meta-analysis with five studies are shown in Table I. It can be seen that when the between-study standard deviation is 0.3 or 0.8 the estimated between-study standard deviation is downwardly biased. This is because Senn's prior distributions give low prior probability to values of τ greater than 0.3 (Figure 1). In addition, these simulations demonstrate that Senn's prior can lead to some bias in the estimate of the treatment effect, θ. This comes through linking the between-study standard deviation (τ) and the treatment effect (θ) in the prior distribution in a situation when they are not in fact related, but are imposing a known relationship between them through the prior distribution. This degree of bias was not evident in the any of the priors in our original paper. Thus, the usefulness of a simulation approach is highlighted, as many users using vague prior distributions would not be happy with such poor frequentist properties.

We have no argument that 'philosophical' considerations should also be used when using vague prior distributions, but, if a statistician desires to have a model with good bias and coverage properties, but needs/wants to use Bayesian methods, then we believe simulation is a very good way of establishing this.

REFERENCES

1. Lambert PC, Sutton AJ, Burton PR, Abrams KR, Jones DR. How vague is vague? A simulation study of the impact of the use of vague prior distributions in MCMC using WinBUGS. *Statistics in Medicine* 2005; **24**:2401–2428.
2. Senn S. Trying to be precise about vagueness. *Statistics in Medicine* 2007; **26**:1417–1430.

Published online 11 September 2007 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/sim.3043

AUTHORS' REPLY

In their original article in this journal, Lambert *et al*. [1] referred to the 'philosophical advantages of the Bayesian approach'. These advantages include coherence and the ability to take into account all information. To include information that is not present in the data requires its incorporation into the prior distribution, which is a convenient representation of the prior information. To be *genuinely* coherent requires careful thought about what that prior information is, and to adopt a Bayesian framework commits one to certain attitudes about prior information and data: for example, that they are, to the degree defined by the model, exchangeable. (This latter point is sometimes misunderstood. For a recent example, see Peters *et al*. [2].) Any Bayesian who is committed to the philosophical advantages of the approach might as well see how it behaves when some of the data are changed as when the prior distribution is changed.

However, in their original article, Lambert *et al*. adopted an approach, now reinforced by their letter, which abandoned these philosophical advantages. There is no question of worrying about coherence and no attempt to include all sources of information. Indeed, they question the advantage of this writing 'Is there any evidence that doing so will be beneficial?' They use a simulation to judge the performance of a particular prior distribution for a meta-analysis when the universe represented by the simulation is quite different from that represented by the prior distribution.

In closely related contexts, inappropriate simulation has caused researchers to come to erroneous conclusions. Consider, for example, the claim in this journal by Chambless and Roeback [3–5] that in randomized clinical trials treatment estimates that have been adjusted for baseline differences between groups using analysis of covariance are not conditionally unbiased when the covariates are measured with an error. In a sense, the claim that they make is correct. It is correct, for example, that in a randomized trial in hypertension in which the difference in mean *true* diastolic blood pressure in mmHg is Δ adjusting using an observed difference $d = \Delta$ will not produce a conditionally unbiased estimate and, indeed, Chambless and Roeback demonstrated this fact.

The point is, however, that this fact, and hence the simulation, is irrelevant. No trialist is ever faced with the situation in which the difference is known to be Δ. On the other hand, researchers can be faced with situations where the observed difference is d. Now, when considering the variances and covariances of d and Δ we have $\sigma_{d\Delta} = \sigma^2_\Delta < \sigma^2_d$. Therefore, whereas the regression of d on Δ is $\sigma_{d\Delta}/\sigma^2_\Delta = 1$, the regression of Δ on d is $\sigma_{d\Delta}/\sigma^2_d < 1$. Any rational statistician having observed that the blood pressure difference was d ought to bet that for these patients if measured again it would (in the absence of treatment) be less than d. This fact is not represented by any simulation from a true difference of Δ and hence such a simulation is irrelevant [6, 7].

