Comments on Report of Devlin et al (2018) by
Mónica Hernández-Alava, Stephen Pudney & Allan Wailoo
ScHARR, University of Sheffield

Whilst I have read the whole of the report by Hernández-Alava et al I will only comment on their critique of the Bayesian modelling aspects and not the data quality, though I note some worrying concerns regarding the latter. I should also state that I may have had further comments had I had access to the WinBUGS models and data, which given that the work of Devlin et al was publicly funded, I find concerning.

The concerns raised by Hernández-Alava et al regarding the implementation of the Markov Chain Monte Carlo (MCMC) methods and the specification of the Bayesian model appear valid. In particular, it is extremely disappointing to see that the work of Devlin et al did not contain a series of robust sensitivity analyses as it standard practice in any Bayesian analysis (Spiegelhalter et al, 2003). These sensitivity analyses should consider the use of prior distributions (especially for variance parameters – see Lambert et al 2005), length of burn-in and samples used, starting values and auto-correlation in the MCMC sampler. It is worth stating that in any Bayesian analysis using MCMC methods it is not possible to *prove* convergence of the sampler, but rather to claim not to have found evidence of non-convergence.

As mentioned the use of prior distributions is a crucial component of any Bayesian analysis, and a thorough sensitivity analysis is vital to assess the impact that these have on any results. Whilst in models with many parameters use of overly vague prior distributions can be problematic, the use of plausibly vague prior distributions requires justification as Hernández-Alava et al state – as well as assessing the sensitivity to different prior distributions. For variance parameters specification of prior distributions is particularly problematic and again the lack of sensitivity analyses which Hernández-Alava et al highlight is extremely concerning. Devlin et al adopt Gamma prior distributions for precision (reciprocal of the variance) parameters – this is theoretically justified as it approximates a Jeffreys' prior distribution, but is not necessarily a vague prior distribution, placing greater weight on smaller values, as Hernández-Alava et al state. A better approach is to put prior distributions on the standard deviation scale (either Uniform or half-Normal), which being on the same scale as the observed data are easier to interpret and specify (Lambert et al, 2005).

Hernández-Alava et al are correct to highlight the fact that high auto-correlation is a concern. In fact Figure 3.4 appears to show that the sampler is either stuck in a part of the posterior distribution or the model is mis-specified. Without the multiple starting values being assessed as part of a sensitivity analysis (again highlighted by Hernández-Alava et al) it is impossible to know whether it is the former or the latter – in either case inferences based on the posterior distribution should be made with extreme caution as if it is the former then they could be *completely* erroneous.

Whilst I have concentrated on Hernández-Alava et al's review of the Bayesian analysis, I agree with them that ignoring the fact that there are multiple observations per patient, i.e. within patient clustering, will lead to erroneous levels of uncertainty, I nevertheless feel that this is probably minor in comparison to some of the other concerns raised, both in terms of potential model mis-specification and the wholly inadequate, and potentially mis-leading, implementation of the MCMC methods used.

References

Devlin N.J., Shah K.K., Feng Y., Mulhern B., van Hout B. (2018). Valuing health-related quality of life: An EQ-5D-5L value set for England. Health Economics, 27:7–22. https://doi.org/10.1002/hec.3564

Feng Y, Devlin NJ, Shah KK, Mulhern B, van Hout B. (2018). New methods for modelling EQ-5D-5L value sets: An application to English data. Health Economics, 27:23–38. https://doi.org/10.1002/hec.3560.

Lambert PC, Sutton AJ, Burton P, Abrams KR, Jones DR. How vague is vague? Assessment in the use of vague prior distributions for variance components. *Statistics in Medicine*. 2005;**24**(14):2401-2428.

Spiegelhalter DJ, Abrams KR, Myles JP. *Bayesian Approaches to Clinical Trials and Health-care Evaluation*. Chichester: Wiley and Sons. December 2003 ISBN 0-471-49975-7.