# Comments from authors

## Comments from Ralph

|  |  |  |  |
| --- | --- | --- | --- |
| **#** | **p** | **Comment** | **Response and action** |
| 1 | 3 | REFERENCE to the Dirichlet description ? | I don’t think this is needed, as it’s a common statistical distribution.  **Action**: ***None, unless RC provides a specific reference.*** |
| 2 | 4 | IF the source paper reports only the sample utility mean and standard deviation (in each category) of the utilities (as I except) you have to calculate the SEM from those 2 statistics first.  See <http://en.wikipedia.org/wiki/Sampling_distribution> Do not confuse the observed sample standard deviation (SD) of the individual patient-level utilities and the standard error of the mean (SEM)  While the statistic “Mean of the utilities” is normally distributed as a consequence of the Central Limit Theorem, the observed utilities are AS A RULE ARE NOT normally distributed. Empirically observed utilities present plenty of weird problems (see my 2010 and 2012 papers for example with references therein | This is a very important point that I’ve been unclear about in the manuscript, partly because I misremembered the process between writing the code and the manuscript.  The values in parentheses of table 3 of Rivero-Arias are \*standard deviations\* of the values within each mRS state. Uncertainty about the Ns in each state is also simulated in the model, using the Dirichlet distribution, so if SEs were to be calculated at this stage, it would have to be for each of simulated N combinations.  My view is that using SDs rather than SEs at this stage is appropriate, as what the PSA needs are distributions representing uncertainty in expected values of the combined distributions (dep stroke, indep stroke) rather than the expected values of the component distributions. This final stage (getting expectations of indep and dep) is done by bootstrapping the means of the weighted mixtures of the component distributions.  Using SDs to simulate the component distributions (mRS 1, 2 etc) means some of the predicted values at this stage could be extreme, and produce weird problems if they were used as a final stage. However as they’re only used as intermediate stage assumptions, and then bootstrapped at the end, means this shouldn’t be highly problematic.  **Action**: ***JM to revise text to be clear about use of SDs rather than SEs at this stage. [Done]***  ***RC to decide whether these intermediate stage outcomes should be simulated as expected rather than predicted values then let me know.*** |
| 3 | 4 | What do you mean exactly by that ????, be more precise, which parameter(s)?, the mixing proportions ?  In response to:  Within the first example, illustrated in Figure 1, the independent state category (Node 4a) is comprised of a mix of the three component states mRS 0, mRS 1 and mRS 2, and the dependent state category is comprised of a mix of the three component states mRS 3, mRS 4 and mRS 5. However, neither the independent state category nor the dependent state categories are composed of equal amounts of each component state, and so an equal weighting should not be assumed. It would also be wrong to disregard parameter uncertainty due to the finite sample size on which these estimates are based. This is why the dirichlet distribution was used. | I mean this is why I used the Dirichlet distribution.  I’ve added:  For example, there might be three mutually exclusive states A B C, and two samples of patients. In the first sample A=5, B=10, C=5, and in the second sample A=50, B=100, C=50. In both cases the proportion of the sample in each category was the same – 25%, 50%, 25% - but in the first case the smaller sample size means there is more uncertainty about the true proportion of patients in each category.  **Action**: ***RC to decide whether this explains the rationale clearly enough and let me know.*** |
| 4 | 5 | So this gives you what? An estimate of a utility multiplier and its standard deviation for each of the aggregate new states ie Independent and dependent ? | I’ve added this description:  The results of this process are distribution of predicted values for the utility associated with independent and dependent stroke states which takes into account both uncertainty in the predicted utilities for each of the component states, and uncertainty in the true proportion of each component state within each collapsed state.  **Action**: ***RC to review and say whether this describes what’s going on clearly enough.*** |
| 5 | 5 | What is the difference with just reweighting ?  (To use of the term ‘dynamically reweighted’) | This paragraph is now rewritten.  **Action**: ***RC/MS to say look at new paragraph and suggest further amendments if necessary.*** |
| 6 | 6 | In FIG 1 how can you have arrows from 4a to 7b and from 4b to 7a ???? | This is an error.  **Action:** ***JM to correct this figure. [DONE]*** |
| 7 | 9 | Why do you talk about Credible intervals , Why do you use Bayesian vocabulary while you are not working in a Bayesian framework ? These are just 95% confidence intervals . The fact that they result from simulation does not make them Bayesian | I’ve changed the term from credible intervals (CrIs) to predictive intervals (PrIs).  I think the use of simulation makes them closer to credible intervals than confidence intervals, but I’m not completely sure. I think they fit Gelman & Hill (2007)’s definition of ‘informal Bayesian inference’ (p. 143).  “Bayesian inference refers to statistical procedures that model unknown parameters ... as random variables”, which is exactly what I’m doing here.  Informal in the sense that the sampling from the random variables is done using a noninformative prior.  I think there are debates about whether cost-effectiveness models are ‘inherently bayesian’ or not for similar reasons, when the inputs are random variable rather than constants by using PSA. However I’m not going to try and resolve that debate in this comment!  **Action:** ***RC to say whether he agrees with the change from credible interval to predictive interval.*** |
| 8 | 9 | How do you get to the actual utiltiies from the Multiplier, what number do you apply the multiplier to ?  Because what I am interested in is the mean and 95% CI of the estimated utility in the new categories.  (same comment as above) | The utility multipliers were converted into utility values using age-gender specific mean utility values from  <http://www.shef.ac.uk/polopoly_fs/1.43390!/file/HEDS-DP-09-11.pdf>  However, that’s not the focus of this paper.  If someone doesn’t want to do this additional stage then the utility multipliers can be used as utility values, which is equivalent to assuming baseline utility scores of 1 (i.e. perfect health if no stroke/ICH).  **Action:** ***RC to decide whether the clarification above addresses this issue, and whether any further description of this is required in the manuscript itself.*** |
| 9 | 10 | The question is why would I like to do that ie to reduce the number of Health states (ie collapse them) when I have information at a more detailed level ie the non-collapsed states ? | I’ve added the following description.  This is useful where one type of information of interest to modellers, such as cost or transition probabilities to other states, is presented at a coarser level of aggregation (i.e. fewer states) than another type of information of interest to modellers, such as utility scores.  Basically, it’s easier to go from more to fewer categories than from fewer to more categories, so where utilities, costs and trans probs are reported by different numbers of categories, the model is likely to use however many categories the most aggregated measure was reported at.  This method is about how to go about ‘coarsening’ the data in a relatively smart way.  **Action:** ***RC to suggest rewording which would make this message/argument clearer.*** |

## Comments from Jon

|  |  |  |  |
| --- | --- | --- | --- |
| **#** | **p** | **Comment** | **Response and action** |
| 1 | 8 | Figure 1: 4a should go to 7a not 7b. To correct in VUE | **Action**: ***JM to correct this figure in VUE [DONE]*** |