Reviewer's report

Title: Even more suicide attempts in clinical trials with paroxetine randomised against placebo.

Version: Date: 27 September 2006

Reviewer: David Spiegelhalter

Reviewer's report:

General

Major Compulsory Revisions (that the author must respond to before a decision on publication can be reached)

The authors’ previous paper apparently led to GSK’s briefing document of April 2006, which explicitly states that “The results provide evidence of an increase in suicide attempts in adults with MDD treated with paroxetine compared to placebo”. Apart from providing publicity for this interesting admission, and replying to GSK’s correspondence following their article, the current paper’s purpose appears to be to conclude that the P-value of 0.058 reported by GSK understates the evidence. To this end the authors conduct a Bayesian analysis which provides high probabilities (0.98 to 0.99) of an increase in suicide risk.

As a statistician who promotes Bayesian methods, I strongly feel that any Bayesian analysis must be accompanied by an indication of the parameter values supported by the likelihood, rather than simply reporting prior and posterior: see, for example, the recommendations in Spiegelhalter, Abrams and Myles (2004) [Bayesian Methods in Clinical Trials and Healthcare Evaluation, Wiley]. The data analysed by the authors comprise 11 events in 601 patient–years vs 1 event in 333 patient–years. The appropriate classical analysis is based a conditional likelihood, equivalent to observing 11/12 events from a Binomial distribution with true proportion 601r/(333+601r), where r is the unknown rate ratio. The resulting likelihood for r is easily plotted. Crucially, the one-sided mid P-value from this Binomial distribution assuming r=1 is 0.022 (essentially the same tail–area as the authors report for their ‘optimistic prior’). If instead of their ‘informative priors’ they used a Jeffreys prior for each of their Poisson rates (proportional to 1/sqrt(rate)), they would have obtained a 1-sided tail area 0.016 (based on 1000000 iterations).

All these values are very similar, and so the crucial difference between these classical and ‘reference’ Bayesian analyses, and the GSK P-value of 0.058, is that their P-value is 2-sided and not a 1-sided mid P-value. Should GSK report a one-sided or two-sided P-value, and what exactly is the appropriate P-value for such comparisons? These issues have been widely debated and no definitive answer is available. The evidence seems quite strong anyway, and simply showing that the one-sided tail area can be made slightly smaller or larger by using an optimistic or pessimistic prior does not seem to add anything, and instead diverts attention away from the evidence in the data alone. Of course, this is not to say that a serious Bayesian analysis would not be interesting, but this would involve using data from different trials, different populations, making appropriate exchangeability assumptions and so on.

I therefore feel that the Bayesian analysis does not really add anything, and in any case has not been very well carried out and reported. The crucial methodological issue is what P-value to calculate when comparing rare event rates.

If this article were greatly shortened and the Bayesian emphasis severely reduced, it could make a reasonable reply to the correspondence from GSK.

The author

Minor Essential Revisions (such as missing labels on figures, or the wrong use of a term, which the author can be trusted to correct)

Discretionary Revisions (which the author can choose to ignore)
Which journal?: Too insignificant to warrant publication in BMC Psychiatry

What next?: Reject because too insignificant for publication in any BMC journal

Quality of written English: Acceptable

Statistical review: No

Declaration of competing interests:

I have acted as a paid consultant and trainer to GSK regarding the application of Bayesian methods in the pharmaceutical industry. However I have played no part in the GSK response to the previous article by the authors, and have not communicated with any GSK personnel concerning this referee’s report.