Reviewer's report

Title: Prevalence and acquisition of MRSA amongst patients admitted to a tertiary-care hospital in Brazil

Version: 2 Date: 6 July 2010

Reviewer: Ben Cooper

Reviewer's report:

Major Compulsory REvisions

None

Minor Essential Revisions

1. Full results of the multivariate analysis should be reported, not just those which are significant.

2. The intervention in reference 10 is described as cohort isolation. This should be changed to single room and cohort isolation.

3. The paper should report how random selection of patients was carried out (as described in the response to my earlier question). Reporting methods of randomisation is recommended for RCTs (in the CONSORT statement), and there are similar reasons for doing it in other studies too.

Discretionary revisions

1. My main concern in the previous version was the validity of the main statistical analysis. Notwithstanding the cited article by Barros & Hirakata, the analyses using a Poisson model (as in the previous version) or log-binomial model (as in the current version) both have disadvantages and can sometimes fail badly (as Barros & Hirakata point out). For example, the log-binomial model allows probabilities greater than 1 which sometimes causes convergence problems and gives rise to confidence intervals which are too narrow. Also, as the article reports (and shows) Cox regression will in general give rise to confidence intervals that are too wide. As Zou points out, using a standard (rather than modified) Poisson regression will also result in confidence intervals that are too wide (Am J Epidemiol 2004; 159(7):702-6).

However, in light of a useful simulation study by McNutt et al. (Am J Epidemiol 2003; 157:940-943) I am now reasonably convinced that these distortions will be modest, and not a major cause for concern. I have therefore somewhat softened my stance on the original analysis, but the new analysis (the log-binomial model) is certainly an improvement and simulation studies published by McNutt et al suggest that performance is adequate for publication. Any further revisions to this analysis should therefore be considered discretionary. I do also think the the
authors should cite McNutt 2003 in order to justify their approach - this is a much better reference than Barros & Hirakata which doesn't include any simulation studies.

Note that the justification for using such approaches instead of the more standard logistic regression (given in Barros & Hirakata) is that the odds ratio commonly obtained from logistic regression (by exponentiating model coefficients) can be a very poor estimator of the prevalence ratio, which is of primary interest. This is certainly true. However, the authors don't point out that the prevalence ratio (or risk ratio) can easily be estimated from a logistic regression. If the model is

$$\logit(p/(1-p)) = a + bX$$

then

$$P_0[X=0] = \frac{\exp(a)}{1+\exp(a)}$$
$$P_1[X=1] = \frac{\exp(a+bX)}{1+\exp(a+bX)}$$

and $P_1/P_0$ gives the prevalence ratio (though of course this will depend on values of other covariates).

The tricky bit is estimating the variance of $P_1/P_0$, and this is not so easy in standard software (in contrast estimating the variance of an odds ratio is easy). I can think of at least three ways this could be done should the authors wish to do so:

i) approximately using a Taylor Series expansion (i.e. the delta method)

ii) using a bootstrap approach

iii) exactly using an MCMC algorithm (this could easily be implemented in free software, WinBUGS, for example, or in R using Gary King's zelig package: Goodrich & Lu 2007.)

2. 1. The approach of including only variables with $P<0.20$ in the multivariate model, though common, is misguided and will tend to lead to overfitting. See, for example, Babyak 2004 for a good clear discussion of the issue. Better not to drop anything unless collinearity is high. Again, unless the journal has a policy on such issues I would consider revision discretionary in light of the fact that this is such common practice.

References


Level of interest: An article whose findings are important to those with closely related research interests

Quality of written English: Needs some language corrections before being published

Statistical review: Yes, and I have assessed the statistics in my report.

Declaration of competing interests:
I declare that I have no competing interests