Reviewer's report

Title: Antibiotic rotation strategies to reduce antimicrobial resistance in Gram-negative bacteria in European Intensive Care Units: A cluster-randomized cross-over trial

Version: 1 Date: 12 February 2014

Reviewer: Graeme MacLennan

Reviewer's report:

This protocol describes a cluster randomised cross-over trial designed to assess the effectiveness of interventions to reduce antimicrobial resistance patterns in Gram-negative bacteria in European ICUs. The case for this study is clear, but I have a number of concerns about the clarity of the protocol as currently written. I think that the protocol may benefit from a local professional statistician to help re-draft the sample size and analysis sections.

Major points:

1. The primary hypothesis is not stated anywhere in the protocol. ClinicalTrials.gov lists the two interventions as active arms, which is fine, but is not clear what the primary comparison of interest is from the protocol. It should be a test of superiority comparing Mixing to Cycling?

2. Outcomes: In the outcome section there is a description of how the primary outcome will be collected, but not an explicit statement on the primary outcome. At http://clinicaltrials.gov/show/NCT01293071 where the trial is registered the primary outcome is described as "Mean prevalence of ICU patients colonised with antimicrobial resistant Gram-negative pathogens " . I would therefore expect to see the opening sentence in the Outcomes section along the lines of "The primary outcome is mean prevalence of ICU patients colonised with antimicrobial resistant Gram-negative pathogens. This is defined as ....". We then need a definition of mean prevalence; I am guessing the mean of the nine months of serial point prevalence measurements during each intervention period? Similarly for the secondary outcome are listed explicitly at clinicaltrials.gov, but not so in the manuscript.

3. Sample size: The sample size section is very confusing and currently as written the calculation cannot be replicated. One issue is that the sample size is described as that of a cluster randomised trial where clusters are randomised but data are collected at the level of the individual. That is true here, but the analysis section (more on that below) and the primary outcome appear to be aggregated up to a cluster-level analysis. There is nothing wrong with this, but readers will notice sample size and analysis plan do not match. It is perfectly reasonable to use a more conservative sample size than required by the primary outcome, but that needs explained. However, even so, the details provided as not adequate to
replicate. Firstly, there is no detail provided on the sample size calculation unadjusted for clustering. It is stated that 200 patients per intervention period are required to detect a relative decrease of 10% with 95% certainty. Because of uncertainty around the definition of the primary outcome it is difficult to understand what this means. Is this a sample size for a difference in means (as the analysis section would suggest that primary outcome is being treated as continuous); we need to know what the baseline value is in this calculation; where does 10% relative reduction come from; what does 95% certainty mean (do you mean power here?); does that mean that 800 patients are required in total ignoring clustering; essentially all the ingredients and assumptions underpinning the basic sample size calculation need included but are not currently. As the authors have correctly noted, once a plain sample size is decided upon it needs adjusted for clustering, this requires an ICC (reported) and then usually one or other of the number of clusters or average cluster size is fixed (due to logistic or financial restraints). There is no description of what was fixed and why. The final sample size is stated 1476 cultures per intervention arm. Here there is a change of terminology, now it is not cultured patients, but cultures, does that mean that the same patient could provide more than one culture if they were in the ICU long enough? There is a change from intervention period to intervention arms, this is confusing. The last sentence in this section then states that study should be adequately powered to detect a 6% difference prevalence with 95% certainty. Is this now a 6% absolute difference? If not, why the change from 10% being important to 6%? What does adequate power mean? Again, I am not sure what 95% certainty refers to in this context? One last point, the discussion talks about approximately 10,000 patients being recruited, which would seem to contradict the sample size section.

4. Analysis: There are two analyses to be carried out: cluster-level and individual-level. It is not clear how the primary outcome will be analysed. I would expect here to see an explicit statement on the method(s) to be used. For example, there are 8 ICUs, each getting both intervention, therefore a simple analysis would be a paired t-test on the 8 paired observation periods aggregated up over each 9 month period? Is that what is planned? Or are the serial measurements going to used here, it is just not clear. For the individual-level analyses not enough detail is provided. Merely stating "advanced regression methods" is not enough, which methods?

Minor points:

1. There are no page numbers.

2. Reference 29 is the same as reference 25

3. The tense that the manuscript is written is inconsistent, for example Data collection section, (my emphasis) "Dedicated staff WILL collect data...." (correct) compared to "Weekly point-prevalence measurements of antibiotic consumption ARE collected..." (incorrect). There ARE in this last sentence should read "WILL BE". This is a protocol which is pre-specifying what will take place in the study, written before the study commences. (Although from ClinicalTrials.gov
recruitment is finished? Start date and end date of recruitment would be useful to know in the manuscript).

4. Design effect is 3.69 uses European convention of decimal comma, please change to decimal point.