Reviewer's report

Title: Effect of intermediate care on mortality following emergency abdominal surgery. The InCare-trial: rationale, design and feasibility of a randomised multi-centre trial.

Version: 2 Date: 28 December 2012

Reviewer: John Norrie

Reviewer's report:

Discretionary revisions

The authors present a clear protocol. There are some issues that could benefit from clarification, as follows:

1. The randomisation takes place when the patient is ready for transfer to the surgical ward within 24 hours of surgery - subject to the availability of an IC bed. The authors indicate that there are 7 ICU beds per 100,000 Danes, but only 1.3 intermediate care beds. The protocol states that if no bed is available, then the patient is not randomised - couple of points:
   a. how often is this happening? The authors indicate that with less than half the 400 target completed, the inclusion criteria of an APACHE score was reduced to 10, suggesting recruitment was slow?
   b. But as well as slow recruitment, there are the issues of generalisability and implementability. Generalisability because if at busy times more patients are turned away, there may be a selection bias for busy vs. quieter times? The authors indicate that they will collect mortality data on those eligible but not randomised, which is useful - but they should perhaps consider splitting this between those who refused to take part and those that were willing but no bed was available?
   c. And implementability because if there are many fewer beds than are needed to even maximise the recruitment into the trial, how useful will the intervention be in practice if proved to be safe and effective?
   d. It would be useful to describe how many intermediate care beds each of the 7 participating hospitals had, and what there availability was during the trial to date, and whether ICU beds could be used, and how the provision of this IC beds in these 7 hospitals compares with (a) the rest of Denmark (they mention n=23 units total) and (b) other healthcare settings?

2. The tables listing what comprises IC and standard surgical ward care are very useful - they show that some components are the same, others quite different in terms of frequency and method of measurement. But what of the staff involved? To what extent would any difference seen between the two randomised groups be down to differently experienced and skilled staff between IC and surgical ward sites? Is this something the authors will try to tease out or do they see this as an
integral part of the package of care?

3. Will the analysis adjust for staff within centre and centre overall? The authors mention that the lack of blinding and the fact that since the 'same surgeons' treat both groups might 'insinuate a learning bias' but it isn't clear whether either an overall staff and/or centre effect will be included, nor whether a sensitivity analysis exploring this possible contamination / learning effects will be undertaken?

4. The text states that the 'compliance with the trial protocol' measures are listed in Table 3 - and that the comparison is at 48 hours - but the column in the Table seems to indicate these measures are made at 2 and 14 days? It would be unusual to measure an important outcome at different times between the two randomised groups? How would you deal with the differential death rate, for example?

5. On the primary outcome of 30 day mortality the protocol states 'Mortality data will be retrieved from the Danish Civil Registration System (CRS) ... by CTU at interim analysis and at day 30 after the last patient has been enrolled' - couple of points:
   a. what is CTU? Clinical Trials Unit?
   b. How complete and accurate is the DRS?
   c. Wouldn't there be a lag in getting any death registered on the system?
   d. Why all cause mortality and not an appropriate cause specific death specification - would the intervention be expected to influence all deaths regardless of cause?
   e. Why 30 days, and not say in-hospital (shorter) or 6-12 months (longer)?

6. Not sure I understood the rationale behind the DMC stopping rules - specifically it seems to treat benefit and harm symmetrically - but if there is evidence of harm at P<0.001 at 200 patients, is it really acceptable to go on to recruit another 100 patients to have another look at n=300?

7. Not clear why the authors are conducting both logistic and Cox models on the primary outcome?

8. What is the current status of recruitment?