Reviewer’s report

Title: The Danish Neonatal Clinical Database is valuable for epidemiologic research in respiratory disease in preterm infants

Version: 2 Date: 17 September 2013

Reviewer: Brendan P Murphy

Reviewer’s report:

The clinical question posed here is to ascertain the quality of the data - the information on the use of surfactant and the use of and duration of nasal continuous positive airway pressure (nCPAP), oxygen supplementation, and mechanical ventilation in the Danish Neonatal Clinical Database (NeoBase). This question posed is very clear. However, in the methods, it becomes apparent that what is being compared is not simply the “gold standard” nurses’ daily bedside documentation (the prospective real time data acquisition on infants condition and clinical course) compared to the NeoBase, but in truth the accuracy of the selected dichotomous and continuous variables as extracted from this “gold standard” by a single person (SA) retrospectively as a dedicated research study compared to the same dichotomous and continuous variables extracted retrospectively from the medical and nurses’ charts (the same “gold standard”) by multiple clinicians (medical and/or nursing of unclear seniority or possibly designated administrative staff at or around (but again not clear exactly when this occurs) the time of infants discharge from hospital. Both the “gold standard of nurses’ bedside documentation, NeoBase and indeed assessment by reviewer SA all appear to be paper based with both NeoBase and reviewers data acquisition retrospective at some distance from clinical care of infant. Not unsurprisingly, a single reviewers data extraction is more accurate, and recordings of dichotomous variables have greater concordance then continuous variables and greater concordance in continuous variables if a greater degree of latitude is given to outcome collected (100% concordance for ± 7 days!)

In the results section, it is unclear whether 11/164 infants admitted to the 2 NNU’s had data entered into NeoBase although nurses charts i.e the “gold standard” was missing some data for retrospective comparison afterwards and thus these 11 records were excluded from comparison. In addition, how are the ICD diagnostic codes referred to in the results assigned to infants, – by whom and when in relation to time of data collection of NeoBase structured form. Is the ICD diagnosis record paper based or an electronic database either of clinical or non clinical origin.

Table 1 reveals much as to why results noted were seen. Mechanical ventilation is for a very brief period, median 0 days with range of 0-35 days for 44 infants. Usually mechanical ventilation is very clearly documented in both medical and nursing charts whereas longer duration of cpap and oxygen therapy is less clearly documented with a resultant lesser concordance between NeoBase and
the review by SA.

No information is given as to data definitions recorded by either NeoBase or by SA relating to treatment variables – whereas dichotomous variables are likely to be relatively clear, for continuous variables relating to duration of treatment, does this include part of day, or full 24 hour periods.

The degree of concordance between SA review and NeoBase is clearly documented with sensitivity, specificity, PPV and NPV correctly assigned, although this might be more easily interpreted by the reader by uses of a simple 2x2 table to better illustrate these results.

The discussion with respect to high concordance between NeoBase and SA’s retrospective review with greater concordance for dichotomous then for continuous variables and greater concordance for continuous if greater discrepancy in days permitted is as one would expect. Some greater discussion with respect to NeoBase, its composition and data definitions and standardization compared to international neonatal databases such as Vermont Oxford Network or EuroNeonet and how this database might be used not just for internal but also national and international benchmarking would add to the overall value of the paper.

The paper addresses a simple question clearly with almost self-evident results, but could add some useful information about value of national neonatal databases with some greater descriptions for the non Danish reader as outlined above.

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

**Quality of written English:** Acceptable

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**

I declare that I have no competing interests.