Author’s response to reviews

Title: A systematic review identifying common data items in neonatal trials and assessing their completeness in routinely recorded United Kingdom national neonatal data

Authors:

Sena Jawad (s.jawad@imperial.ac.uk)

Neena Modi (n.modi@imperial.ac.uk)

A Prevost (a.prevost@imperial.ac.uk)

Chris Gale (christopher.gale@imperial.ac.uk)

Version: 2 Date: 16 Oct 2019

Author's response to reviews:

Response to reviewers: A systematic review identifying common data items in neonatal trials and assessing their completeness in routinely recorded United Kingdom national neonatal data

Editors comments:

Comment 1: Please consider that the costs of trials in the Introduction ("the median cost of phase III trials of therapeutic agents in 2015-2016 was $19 million") is (a) based on an estimate and (b) refers to drug approval studies (which is a very specific situation and typically not possible using routinely collected data like EHR). It might be an estimate that can't be well generalized to many other RCTs.
Response: we have removed this citation and included a different estimate that is derived from a systematic review of the costs of a broader range of randomised controlled trials. “the median cost of randomised controlled trials was estimated between $43 and $103,254 per participant…”

Reviewer 1 comments:
The authors did a thoughtful job addressing reviewer and editor comments and the overall manuscript is substantially improved.

Comment 2: Page 6 line 33 - did you mean randomization?
Response: we acknowledge the reviewers comment however we wanted to clarify how the data items were used in randomisation. We have included randomisation in brackets to clarify this. “items used in stratification or minimisation (randomisation)…”

Comment 3: Page 12 line 29 ---impact 'than' in trials (left out word)
Response: the statement is not a comparison therefore this has not been changed.
Reviewer 2 comments:
Comment 4: In the Discussion section, 1st paragraph, line 5: In the following sentence "...We find that these 14 data items can be obtained from the NNRD with high completeness for most items."
: Please list here these 14 items, as it is important for the readers to clearly understand what these 14 common items that can be also found in NNRD are exactly.
Response: We have added reference to table 3 here, where the data items are listed.

Comment 5: In Table 3 (Data completeness in the NNRD for the data items reported in 20% of studies or more): Why the completeness for only 12 (and not 14) common items is listed?
Response: The reviewer is correct that completeness was listed for only 12 items. The remaining two items are infection and age. The discussion explains how we were unable to calculate the completeness for this:
“A further limitation of our study was that some items identified were dichotomous, for example presence or absence of infection prior to trial enrolment and it was not possible to calculate completeness for such items as absence of the condition is not always actively recorded.”
Further explanation has been added for age in the discussion:
“Age was found to be a common data item, however it is calculated using gestational age which is highly complete in the NNRD and therefore completeness for age was not calculated”
We have also altered the main findings in the discussion to clarify this:
“We find that 12 of these 14 data items can be obtained from the NNRD with high completeness for most items”

Comment 6: In the Results section, page 7, line 59 "Fourteen data items were reported by at least 20% of studies;....": Please ref Table 2 here, as this is the main study finding ( "Fourteen data items were reported by at least 20% of studies (Table 2);....)
Response: we have included this reference to table 2 in the results section:
“Fourteen data items were reported by more than 20% of studies (table 2).”

Comment 7: In the Results section, page 8, line 40 "The completeness of common data items in the NNRD are summarised by age groups in table 4.": Please kindly correct. There is no Table 4 in the revised mscpt. Should be Table 3.
Response: The reference has been changed to table 3.
“The completeness of common data items in the NNRD are summarised by age groups in table 3.”

Comment 8: In the Results section, page 8, line 42 : "Data completeness in the NNRD is 99.9% for gestational age at birth, 99.9% for sex, 100% for birth weight, 99.7% for multiple birth and 100% for respiratory support on day 1 (table 2).": Please kindly correct. Data completeness in the NNRD is listed in current Table 3 (not Table 2).
Response: This reference has been changed:
“Data completeness in the NNRD is 99.9% for gestational age at birth, 99.9% for sex, 100% for birth weight, 99.7% for multiple birth and 100% for respiratory support on day 1 (table 3)”

Comment 9: Re reply to comment #49, in the Discussion section, page 11, line 40, the following sentence should be re-edited "An additional limitation stems from the fact that some data items
collected in clinical trials did not directly align with data items in the NNRD, therefore there may be a loss of information from aggregating several data items into a common data item held by the NNRD to assess data quality" : This phrase should be re-edited to read: ".... therefore there may be a loss of information from aggregating several data items into a common data item held by the NNRD, that can compromise the quality and granularity of the NNRD recorded clinical information.

Response: We disagree that the lack of alignment between clinical trial data items and those in the NNRD compromises the quality of the NNRD.