Accounting for Multiple Births in Randomised Trials: A Systematic Review

Lisa Nicole Yelland\textsuperscript{a,b}, Thomas Richard Sullivan\textsuperscript{b}, Maria Makrides\textsuperscript{a,c}

\textsuperscript{a}Women’s and Children’s Health Research Institute, The University of Adelaide, Australia
\textsuperscript{b}School of Population Health, The University of Adelaide, Australia
\textsuperscript{c}South Australian Health and Medical Research Institute, Adelaide, Australia

Author for Correspondence: Dr Lisa Nicole Yelland, School of Population Health, Mail Drop DX650 511, The University of Adelaide, SA 5005, Australia, Email: lisa.yelland@adelaide.edu.au, Phone: +61 8 8313 3215, Fax: +61 8 8223 4075.

Keywords: Multiple births, twins, evidence based medicine, statistics

Word Count: 2500
ABSTRACT

Objectives: Multiple births are an important subgroup to consider in trials aimed at reducing preterm birth or its consequences. Including multiples results in a unique mixture of independent and clustered data, which has implications for the design, analysis and reporting of the trial. We aimed to determine how multiple births were taken into account in the design and analysis of recent trials involving preterm infants, and whether key information relevant to multiple births was reported.

Design: We conducted a systematic review of multicentre randomised trials involving preterm infants published between 2008 and 2013. Information relevant to multiple births was extracted.

Results: Of the 56 trials included in the review, 6 (11%) excluded multiples and 24 (43%) failed to indicate whether multiples were included. Among the 26 trials that reported multiples were included, only 1 (4%) accounted for clustering in the sample size calculations and 8 (31%) took the clustering into account in the analysis of the primary outcome. Of the 20 trials that randomised infants, 12 (60%) failed to report how infants from the same birth were randomised.

Conclusions: Information on multiple births is often poorly reported in trials involving preterm infants, and clustering due to multiple births is rarely taken into account. Since ignoring clustering could result in inappropriate recommendations for clinical practice, clustering should be taken into account in the design and analysis of future neonatal and perinatal trials including infants from a multiple birth.
INTRODUCTION

Preterm birth is a leading cause of neonatal mortality and morbidity, and numerous randomised trials have assessed interventions aimed at reducing preterm birth or its negative health consequences. Multiple births are an important subgroup to consider in such trials, since multiples account for around a quarter of all preterm births and half of all twins are born preterm.[1] Including multiple births in a trial can be challenging due to the correlation between outcomes of infants from the same birth that results from shared environmental and genetic factors.[2 3] Correlated or clustered data are common in health research and methods for analysing this type of data are widely discussed.[4-8] Since trials involving preterm infants often include a mixture of singletons and multiples, or independent and clustered data, the implications of clustering in this unique setting require special attention.

Analysis methods that either ignore or account for clustering due to multiple births have been examined over the last decade.[2 3 9-14] Failure to account for clustering in this setting can result in underestimated variances[9] and inflated type I error rates,[12] potentially leading to ineffective treatments being identified as effective. It is therefore recommended that clustering be taken into account in the analysis when infants from a multiple birth are included,[2 9 11-13] especially when the multiple birth rate is not low.[3 10] Despite these recommendations, clustering is rarely taken into account in trials involving preterm infants. A systematic review of multicentre randomised trials involving preterm infants revealed that only 4 (21%) of the 19 trials published between August 2003 and August 2008 that included multiples took the
clustering into account in the analysis.[11] It is unclear whether the practice of ignoring clustering in the analysis has continued in recent years.

Clustering not only has implications for the analysis but also for the design and reporting of trials. Failure to account for clustering at the design stage could result in an underpowered trial that fails to identify a beneficial treatment, since clustering often reduces the effective sample size.[15] Whether clustering due to multiple births is taken into account in sample size calculations, and how this is achieved, remains to be investigated. At the reporting stage, details such as the prevalence of multiple births are important for assessing the appropriateness of the statistical methods and the generalisability of the trial findings. Inadequate reporting of neonatal and perinatal trials has previously been highlighted[11 16] but many aspects of reporting relevant to multiple births are yet to be considered. The aims of this systematic review are to determine how multiple births have been taken into account in the design and analysis of recently published multicentre randomised trials involving preterm infants, and whether key information relevant to multiple births has been reported.

METHODS

This systematic review was conducted according to a pre-specified protocol. Consistent with the systematic review by Hibbs et al.,[11] the search focused on preterm infants, however defined in each individual trial, since multiple births are an important subgroup in this vulnerable population, and on multicentre randomised trials for feasibility. The search was conducted in PubMed on June 24, 2013 using the search terms “(preterm or premature) and (multicent* or
multi-cent*)” and the filters “Randomized Controlled Trial; published in the last 5 years; Humans; English; Newborn: birth-1 month”.

Articles were eligible for inclusion if they reported the results of the primary outcome of a multicentre randomised trial, where the primary outcome was either measured on the infant or could be attributed to the infant. Where multiple articles reported on the primary outcome for the same trial, only the first published article was included to avoid duplication of information included in the review. The primary outcome was defined as the outcome identified by the authors as primary. If multiple primary outcomes were identified by the authors, it was defined as the outcome used to determine the sample size, provided only one outcome was used to determine the sample size, otherwise it was the first outcome identified by the authors as primary that met the eligibility criteria. If no primary outcome was identified by the authors, it was defined as the first outcome used to determine the sample size that met the eligibility criteria. Articles were excluded if none of the potential primary outcomes met the eligibility criteria, they described the methods of a trial only, or they reported the results of a pilot, phase I or phase II trial, a follow-up study of a trial, secondary outcomes or analyses of a trial, or multiple trials, including meta-analyses. However, articles primarily reporting the results of a single eligible trial and then adding these results to an existing meta-analysis were included.

Titles and abstracts of all articles identified by the search were independently examined by two reviewers (including LNY) and classified as ineligible or eligible. The full text of all eligible titles and abstracts were then independently examined by two reviewers (LNY and TRS) to confirm eligibility and extract information using a purpose-specific data extraction form.
When assessing how multiple births were taken into account in the sample size and analysis, only the primary outcome was considered. Any differences between reviewers were resolved by discussion.

RESULTS

The search identified 164 articles, of which 56 (34.1%) met the inclusion criteria (Web Appendix) and were included in the review (Figure 1). For these 56 trials, the median (interquartile range (IQR)) sample size was 11 (4-22) centres and 214 (150-538) infants. There were 26 (46.4%) trials that indicated multiples were included and 24 (42.9%) trials where it was unclear whether multiples were included, either because the inclusion/exclusion criteria were not clearly stated (n=2 trials), or multiple births were eligible according to the inclusion/exclusion criteria but not otherwise mentioned (n=22 trials). Multiple births were excluded from the remaining 6 (10.7%) trials by either excluding women with a multiple pregnancy (n=5 trials) or excluding all infants from a multiple birth (n=1 trial). Only one trial justified the exclusion on the basis of the intervention previously being shown to be ineffective in women with multiple pregnancies.

Characteristics of the 26 trials reporting that multiples were included are given in Table 1 and their interventions and primary outcomes are provided in the Web Appendix. The median (IQR) percentage of infants from a multiple birth was 29.9% (23.9-33.3; n=24 trials where information clear). In 8 trials where infants were randomised and the method of randomising multiples was described, infants from the same birth were either randomised to the same
treatment group[17-22] or independently.[23 24] Two trials defined the primary outcome on the cluster level by assessing any occurrence of the outcome across infants from the same birth, while the primary outcome was defined on the infant level for the 24 remaining trials.

Overall, 12 (46.2%) of the 26 trials including infants from a multiple birth took this into account in the design and/or analysis in some way (Table 2). Accounting for clustering was unnecessary for the two trials with a cluster level primary outcome. Clustering was otherwise neglected in the sample size calculations for all but one trial,[18] where the sample size was calculated assuming independence and then multiplied by a design effect of 1.12, with no justification provided for this choice of design effect. Clustering was addressed in the analysis of the primary outcome for 8 (30.8%) trials in either the primary analysis[18 21 25 26] or a sensitivity analysis.[19 20 22 23] In the latter case, treatment effects were clearly significant or non-significant and accounting for clustering did not alter the conclusions. The most common analysis approach used to account for clustering was generalised estimating equations.[18 19 21 23 25] Other approaches included fitting a mixed effects model,[26] performing an adjusted chi-square test[19] and analysing one infant per birth as a sensitivity analysis.[20 22] Where clustering was ignored in the sample size calculations or analysis, no justification was provided. Other approaches used to account for multiple births that do not address clustering included: stratifying on multiple birth in the randomisation;[19 27] adjusting for multiple birth in the analysis as a fixed effect in a sensitivity analysis;[24 28] testing for a treatment by multiple birth interaction;[26 27] and performing subgroup analyses for singletons to assess the influence of including twins,[20] or both singletons and multiples to assess consistency of treatment effects.[27 29]
For the 26 trials including infants from a multiple birth, the percentage of missing data for the primary outcome ranged from 0% to 7.2% and was unclear in 6 (23.1%) trials. Only one trial performed multiple imputation to account for missing data and no information was provided on how multiple births were taken into account in the imputation process.[21]

Key details relevant to multiple births were not reported in some trial publications (Table 1). The percentage of infants from a multiple birth could be determined in all but 2 (7.7%) trials, while the percentage of women with a multiple pregnancy could only be determined in 7 (26.9%) trials. Among the 20 trials including multiples and randomising infants, 12 (60.0%) did not specify how infants from the same birth were randomised. No trial reported the intracluster correlation coefficient (ICC) as a measure of the magnitude of clustering present due to multiples.

**DISCUSSION**

Our systematic review demonstrates that clustering due to multiple births is rarely considered in the design and remains ignored in the analysis of many multicentre randomised trials involving preterm infants. It also highlights that information relevant to multiple births is poorly reported. It was unclear whether infants from a multiple birth were even included in almost half of the trials in our review, making it impossible to assess whether appropriate methods were used to calculate the sample size and analyse the data. When infants from a multiple birth were included, details such as how infants from the same birth were randomised were often not provided. This information is important, since the impact of clustering on the
power of the trial depends on the method used to randomise multiples.[30] Overall, our review indicates that the problems identified in the previous systematic review[11] are yet to be resolved, and identifies new areas where multiple births are inadequately handled. Clearly, substantial room for improvement remains in the way multiple births are dealt with in trials involving preterm infants.

To avoid conducting an underpowered trial, it is important to account for clustering due to multiple births in the sample size calculations, which typically requires an estimate of the ICC.[5 15] Unfortunately, ICCs relevant to multiple births are lacking in the literature and none of the trials in our review reported ICCs. This may help to explain why clustering was only addressed in the sample size calculations for one trial in our review. ICCs should be reported to comply with reporting guidelines[31] and assist with sample size planning for future trials.[15 32]

Ignoring clustering in the analysis can lead to substantially biased standard errors and potentially result in false conclusions being drawn from the data.[33] In practice, this could mean an effective intervention is missed or an ineffective intervention is recommended. While conclusions about the effectiveness of treatment were unaltered in the few trials in our review that reported analysing the data both with and without adjustment for clustering, we have previously demonstrated that adjustment for clustering due to multiple births can change conclusions.[12] The most popular analysis method used to account for clustering in our review was generalised estimating equations. This method is relatively simple to implement, is available in many statistical software packages, and has been shown to perform well for analysing trials
including both singletons and multiples,[12] making it a good choice in practice. Another popular analysis approach identified in our review was removing the clustering and analysing only one infant per mother. This method is not recommended, since it reduces the sample size and hence the power of the trial, and has the potential to produce different results depending on the infants selected.[2]. For a more detailed discussion of methods for analysing trials including infants from a multiple birth, see [2 3 9-14].

While use of methods that account for clustering was limited among the trials in our review, multiple births were sometimes considered in the design and analysis in other ways, which suggests an awareness of the potential for multiples to have an impact on the trial results. At the design stage, some trials used multiple birth as a stratification variable in the randomisation, which ensures approximate balance between treatment groups in the multiple birth rate, independent of how infants from the same birth are randomised. This approach may be useful in trials where multiple birth status is strongly related to the outcome, since a chance imbalance could influence the estimated treatment effect and reduce trial credibility. Stratification can also reduce type I error rates and increase power.[34] At the analysis stage, some trials adjusted for multiple birth, tested for a treatment by multiple birth interaction, or performed subgroup analyses based on multiple birth status. Adjustment controls for chance imbalance in the multiple birth rate between treatment groups and can increase power.[35] Interaction tests and subgroup analyses assess whether the magnitude of the treatment effect differs between singletons and multiples, and may help to determine which infants benefit from treatment, but should be interpreted with caution as they are likely to be underpowered.[36] While these approaches may provide useful information about imbalance and effect
modification, they do not overcome the problems associated with ignoring clustering in the analysis and should not be considered essential.

Imputation is a useful method for dealing with missing data in randomised trials, although only one trial included in our review applied this method. When clustering is present in the data, this should be taken into account in the imputation model but current strategies are infeasible in neonatal and perinatal trials due to the small cluster sizes.[37-38] Developing imputation methods suitable for handling clustering due to multiple births is an important area for future research.

A limitation of this systematic review is that information was extracted from the published articles only. Some details that were missing or unclear may have been obtainable from published trial protocols or by contacting the authors. As a result, we may have over or underestimated the prevalence of published trials that ignored clustering due to multiple births. Improved reporting around multiple births is important for future trials so that the appropriateness of the chosen statistical methods can adequately be assessed. This can be achieved through publication of additional information in supplementary materials online, stricter adherence to general reporting guidelines,[31-39] and the development of reporting guidelines specifically for trials including infants from multiple births.

In conclusion, clustering due to multiple births continues to be ignored in the design, analysis and reporting of many trials involving preterm infants. Given the potential impact of ignoring clustering, including underpowered trials and incorrect conclusions that could lead to
inappropriate recommendations for clinical practice, clustering should be taken into account in the design and analysis of future neonatal and perinatal trials including infants from multiple births.
ACKNOWLEDGMENTS

The authors would like to thank Ms Dominique Condo and Ms Lenka Malek for assessing titles and abstracts for eligibility and obtaining the full text articles.

COMPETING INTERESTS

None to declare.

FUNDING

Lisa Yelland and Maria Makrides were supported by Australian National Health and Medical Research Council (NHMRC) Fellowships (Early Career Fellowship ID 1052388 for LNY, Senior Research Fellowship ID 1061704 for MM). The sponsor had no involvement in the study design; collection, analysis and interpretation of data; writing the report; or the decision to submit the paper for publication.

CONTRIBUTORSHIP STATEMENT

All authors contributed to the design of the study and the interpretation of the results. LNY and TRS conducted the systematic review and MM advised on the review methodology. LNY drafted the initial manuscript and all authors critically revised the manuscript and approved the submitted manuscript.
What is already known on this topic

• Including multiple births in randomised trials results in clustering in the data.
• Clustering has implications for the design, analysis and reporting of trials.

What this study adds

• Clustering due to multiple births is often ignored in the trial design and analysis but should be taken into account.
• Information relevant to multiple births is poorly described and reporting guidelines are needed.
REFERENCES


**LICENCE STATEMENT**

The Corresponding Author has the right to grant on behalf of all authors and does grant on behalf of all authors, an exclusive licence (or non-exclusive for government employees) on a worldwide basis to the BMJ and co-owners or contracting owning societies (where published by the BMJ on their behalf), and its Licensees to permit this article (if accepted) to be published in Archives of Disease in Childhood and any other BMJ products and to exploit all subsidiary rights, as set out in our licence.
**Table 1: Characteristics of trials including infants from multiple births**

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>Trials (N=26)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Percentage of women with a multiple pregnancy</td>
<td></td>
</tr>
<tr>
<td>- median (IQR)*</td>
<td>20.0 (1.5-22.2)</td>
</tr>
<tr>
<td>- unclear: n (%)</td>
<td>19 (73.1)</td>
</tr>
<tr>
<td>Percentage of infants from a multiple birth</td>
<td></td>
</tr>
<tr>
<td>- median (IQR)*</td>
<td>29.9 (23.9-33.3)</td>
</tr>
<tr>
<td>- unclear: n (%)</td>
<td>2 (7.7)</td>
</tr>
<tr>
<td>Highest order multiples included in the trial: n (%)</td>
<td></td>
</tr>
<tr>
<td>- twins</td>
<td>7 (26.9)</td>
</tr>
<tr>
<td>- higher order multiples</td>
<td>9 (34.6)</td>
</tr>
<tr>
<td>- unclear</td>
<td>10 (38.5)</td>
</tr>
<tr>
<td>Who was randomised: n (%)</td>
<td></td>
</tr>
<tr>
<td>- mother</td>
<td>6 (23.1)</td>
</tr>
<tr>
<td>- infant</td>
<td>20 (76.9)</td>
</tr>
<tr>
<td>Timing of the intervention: n (%)</td>
<td></td>
</tr>
<tr>
<td>- prenatal</td>
<td>6 (23.1)</td>
</tr>
<tr>
<td>- postnatal</td>
<td>20 (76.9)</td>
</tr>
<tr>
<td>Who received the intervention: n (%)</td>
<td></td>
</tr>
<tr>
<td>- mother</td>
<td>6 (23.1)</td>
</tr>
<tr>
<td>- infant</td>
<td>18 (69.2)</td>
</tr>
<tr>
<td>- mother/parents and infant</td>
<td>2 (7.7)</td>
</tr>
<tr>
<td>Level of the primary outcome: n (%)</td>
<td></td>
</tr>
<tr>
<td>- cluster level</td>
<td>2 (7.7)</td>
</tr>
<tr>
<td>- infant level</td>
<td>24 (92.3)</td>
</tr>
</tbody>
</table>

* Based on the trials where the percentage was not unclear
Table 2: Approaches used to account for multiple births in the design and analysis of trials including infants from multiple births

<table>
<thead>
<tr>
<th>Approach</th>
<th>n (%)</th>
<th>Trials (N=26)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>Accounted for clustering in the sample size calculation for the primary outcome</td>
<td>1 (3.9)</td>
<td>[18]</td>
<td></td>
</tr>
<tr>
<td>Accounted for clustering in the analysis of the primary outcome</td>
<td>8 (30.8)</td>
<td>[18-23 25 26]</td>
<td></td>
</tr>
<tr>
<td>Stratified by multiple birth in the randomisation</td>
<td>2 (7.7)</td>
<td>[19 27]</td>
<td></td>
</tr>
<tr>
<td>Adjusted for multiple birth as a fixed effect for the primary outcome</td>
<td>2 (7.7)</td>
<td>[24 28]</td>
<td></td>
</tr>
<tr>
<td>Tested for treatment x multiple birth interaction for the primary outcome</td>
<td>2 (7.7)</td>
<td>[26 27]</td>
<td></td>
</tr>
<tr>
<td>Performed analysis in subgroup(s) based on multiple birth for the primary outcome</td>
<td>3 (11.5)</td>
<td>[20 27 29]</td>
<td></td>
</tr>
</tbody>
</table>
Web Appendix

1. References for 56 Trials Included in the Systematic Review

Articles 1-26 indicated that multiples were included, it was unclear whether multiples were included in articles 27-50, and multiples were excluded in articles 51-56


2. Interventions and Primary Outcomes for the 26 Trials Including Multiple Births

<table>
<thead>
<tr>
<th>Trial</th>
<th>Interventions</th>
<th>Primary Outcome*</th>
</tr>
</thead>
<tbody>
<tr>
<td>Blanken et al.[1]</td>
<td>Monthly palivizumab injections vs placebo</td>
<td>Total number of parent reported wheezing days in the first year of life</td>
</tr>
<tr>
<td>Schmidt et al.[2]</td>
<td>Pulse oximters displaying oxygen saturations 3% higher vs lower than the true value</td>
<td>Composite of death, gross motor disability, cognitive or language delay, severe hearing loss or bilateral blindness at 18 months corrected age</td>
</tr>
<tr>
<td>Roos et al.[3]</td>
<td>Maintenance tocolysis with nifedipine orally vs placebo</td>
<td>Composite of perinatal death, chronic lung disease, neonatal sepsis, intraventricular haemorrhage &gt;grade 2, periventricular leukomalacia &gt;grade 1 or necrotizing enterocolitis up to 6 months of age</td>
</tr>
<tr>
<td>Fergusson et al.[4]</td>
<td>Transfusion of red blood cells stored 7 days or less vs standard issue red blood cells</td>
<td>Composite of necrotizing enterocolitis, retinopathy of prematurity, bronchopulmonary dysplasia, intraventricular haemorrhage or death while in the neonatal intensive care unit up to</td>
</tr>
<tr>
<td>Study</td>
<td>Intervention</td>
<td>Outcome</td>
</tr>
<tr>
<td>------------------------------</td>
<td>-------------------------------------------------------------------------------</td>
<td>-------------------------------------------------------------------------</td>
</tr>
<tr>
<td>van der Ham et al.[5]</td>
<td>Expectant management vs induction of labour</td>
<td>Neonatal sepsis</td>
</tr>
<tr>
<td>van der Ham et al.[6]</td>
<td>Expectant management vs induction of labour</td>
<td>Neonatal sepsis</td>
</tr>
<tr>
<td>Leaf et al.[7]</td>
<td>Early vs late commencement of enteral feeds</td>
<td>Time to achieve full enteral feeding sustained for 72 hours</td>
</tr>
<tr>
<td>Ramanathan et al.[8]</td>
<td>Nasal intermittent positive pressure ventilation vs nasal continuous positive airway pressure</td>
<td>Need for mechanical ventilation via endotracheal tube at 7 days of age</td>
</tr>
<tr>
<td>Cignacco et al.[9]</td>
<td>Oral sucrose vs facilitated tucking vs combined oral sucrose and facilitated tucking</td>
<td>Pain response measured on the Bernese Pain Scale for Neonates</td>
</tr>
<tr>
<td>Dunn et al.[10]</td>
<td>Prophylactic surfactant vs intubate-surfactant-extubate vs nasal continuous positive airway pressure</td>
<td>Death or bronchopulmonary dysplasia at 36 weeks postmenstrual age</td>
</tr>
<tr>
<td>Gopel et al.[11]</td>
<td>Surfactant without ventilation vs standard care</td>
<td>Need for mechanical ventilation or not being ventilated but having $\text{pCO}_2&gt;65\text{mmHg}$ or $\text{FiO}_2&gt;0.60$ or both for more than 2 hours between 25 and 72 hours of age</td>
</tr>
<tr>
<td>New et al.[12]</td>
<td>Transfer to an open cot at 1600 vs 1800g</td>
<td>Average daily weight gain over the first 14 days following transfer to an open cot</td>
</tr>
<tr>
<td>Rubin et al.[13]</td>
<td>Carotenoid supplemented formula vs control formula</td>
<td>Plasma lutein concentration</td>
</tr>
<tr>
<td>Zachariassen et al.[14]</td>
<td>Unfortified vs fortified mother's milk</td>
<td>Weight</td>
</tr>
<tr>
<td>Mintz-Hittner et al.[15]</td>
<td>Conventional laser therapy vs intravitreal bevacizumab monotherapy</td>
<td>Recurrence of neovascularization in one or both eyes arising from the retinal vessels and requiring retreatment by 54 weeks postmenstrual age</td>
</tr>
<tr>
<td>Mercier et al.[16]</td>
<td>Inhaled nitric oxide vs placebo gas</td>
<td>Survival without development of bronchopulmonary dysplasia at 36 weeks postmenstrual age</td>
</tr>
<tr>
<td>Carlo et al.[17]</td>
<td>Target range of oxygen saturation of 85-89% vs 81-95%</td>
<td>Composite of severe retinopathy of prematurity or death before hospital discharge</td>
</tr>
<tr>
<td>Sandri et al.[18]</td>
<td>Prophylactic surfactant followed by nasal continuous positive airway pressure vs early nasal</td>
<td>Need for mechanical ventilation in the first 5 days of life</td>
</tr>
<tr>
<td>Study</td>
<td>Intervention</td>
<td>Comparator</td>
</tr>
<tr>
<td>-------</td>
<td>--------------</td>
<td>------------</td>
</tr>
<tr>
<td>Ortenstrand et al.[19]</td>
<td>Family care ward vs standard care ward</td>
<td>Total length of hospital stay</td>
</tr>
<tr>
<td>Carbonell-Estrany et al.[20]</td>
<td>Monthly motavizumab vs palivizumab injections</td>
<td>Positive respiratory syncytial virus (RSV) test and hospitalised for respiratory symptoms or new onset of RSV-positive lower respiratory illness with worsening respiratory status while already in hospital or death caused by RSV</td>
</tr>
<tr>
<td>Garite et al.[21]</td>
<td>Single rescue course of betamethasone vs placebo</td>
<td>Composite of respiratory distress syndrome, bronchopulmonary dysplasia, severe intraventricular haemorrhage, periventricular leukomalacia, blood culture proven sepsis, necrotizing enterocolitis or perinatal death</td>
</tr>
<tr>
<td>Carr et al.[22]</td>
<td>Granulocyte-macrophage colony stimulating factor vs standard management</td>
<td>Survival without an episode of culture positive systemic sepsis to 14 days from trial entry</td>
</tr>
<tr>
<td>Makrides et al.[23]</td>
<td>High vs standard docosahexaenoic acid enteral feeds</td>
<td>Bayley Mental Development Index at 18 months corrected age</td>
</tr>
<tr>
<td>Murphy et al.[24]</td>
<td>Multiple courses of antenatal corticosteroids vs placebo</td>
<td>Composite of perinatal or neonatal mortality, severe respiratory distress syndrome, intraventricular haemorrhage &gt;grade 2, periventricular leukomalacia, bronchopulmonary dysplasia or necrotizing enterocolitis</td>
</tr>
<tr>
<td>Koldewijn et al.[25]</td>
<td>Infant Behavioural Assessment and Intervention Program vs standard care</td>
<td>Bayley Mental Development Index at 6 months corrected age</td>
</tr>
<tr>
<td>Rouse et al.[26]</td>
<td>Magnesium sulfate vs placebo</td>
<td>Composite of stillbirth or infant death by 1 year corrected age or moderate or severe cerebral palsy at or beyond 2 years corrected age</td>
</tr>
</tbody>
</table>

* According to the definition described in the Methods. All outcomes relate to the infant.
Figure 1: Flow diagram

Articles identified in PubMed search (n=164)

Full text articles assessed for eligibility (n=68)

Articles included in the review (n=56)

Articles excluded during title and abstract review (n=96)
- trial protocol (n=15)
- pilot/phase I/phase II trial (n=8)
- multiple trials (n=4)
- follow-up study (n=17)
- secondary outcomes/analyses (n=37)
- non-randomised (n=11)
- single centre (n=2)
- primary outcome relates to mother (n=2)

Full text articles excluded (n=12)
- pilot/phase I/phase II trial (n=1)
- follow-up study (n=2)
- secondary outcomes/analyses (n=3)
- single centre (n=4)
- second report of primary outcome (n=1)
- eligibility could not be confirmed (n=1)