Lisa Nicole Yelland, Thomas Richard Sullivan, Maria Makrides
**Accounting for multiple births in randomised trials: a systematic review**
Archives of disease in childhood. Fetal and neonatal edition, 2015; 100(2):116-120

**Accounting for Multiple Births in Randomised Trials: A Systematic Review**

Lisa Nicole Yelland[a,b], Thomas Richard Sullivan[b], Maria Makrides[a,c]

[a]Women's and Children's Health Research Institute, The University of Adelaide, Australia

[b]School of Population Health, The University of Adelaide, Australia

[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**


1. Tucker J, McGuire W. ABC of preterm birth - epidemiology of preterm birth. *Br Med J* 2004;**329**(7467):675-78.
2. Gates S, Brocklehurst P. How should randomised trials including multiple pregnancies be analysed? *BJOG* 2004;**111**(3):213-19.
3. Marston L, Peacock JL, Yu KM, *et al.* Comparing methods of analysing datasets with small clusters: case studies using four paediatric datasets. *Paediatr Perinat Epidemiol* 2009;**23**(4):380-92.
4. Liang KY, Zeger SL. Longitudinal data analysis using generalized linear models. *Biometrika* 1986;**73**(1):13-22.
5. Donner A, Klar N. *Design and Analysis of Cluster Randomization Trials in Health Research*. London: Arnold, 2000.
6. Diggle P. *Analysis of Longitudinal Data*. 2nd ed. Oxford: Oxford University Press, 2002.
7. Stiratelli R, Laird N, Ware JH. Random-effects models for serial observations with binary response. *Biometrics* 1984;**40**(4):961-71.
8. Laird NM, Ware JH. Random-effects models for longitudinal data. *Biometrics* 1982;**38**(4):963-74.
9. Ananth CV, Platt RW, Savitz DA. Regression models for clustered binary responses: implications of ignoring the intracluster correlation in an analysis of perinatal mortality in twin gestations. *Ann Epidemiol* 2005;**15**(4):293-301.
10. Shaffer ML, Kunselman AR, Watterberg KL. Analysis of neonatal clinical trials with twin births. *BMC medical research methodology* 2009;**9**:12.

11. Hibbs AM, Black D, Palermo L, *et al.* Accounting for multiple births in neonatal and perinatal trials: systematic review and case study. *J Pediatr* 2010;**156**(2):202-08.
12. Yelland LN, Salter AB, Ryan P, *et al.* Analysis of binary outcomes from randomised trials including multiple births: when should clustering be taken into account? *Paediatr Perinat Epidemiol* 2011;**25**:283-97.
13. Sauzet O, Wright KC, Marston L, *et al.* Modelling the hierarchical structure in datasets with very small clusters: a simulation study to explore the effect of the proportion of clusters when the outcome is continuous. *Stat Med* 2013;**32**(8):1429-38.
14. Shaffer ML, Hiriote S. Analysis of time-to-event and duration outcomes in neonatal clinical trials with twin births. *Contemp Clin Trials* 2009;**30**(2):150-54.
15. Killip S, Mahfoud Z, Pearce K. What is an intracluster correlation coefficient? Crucial concepts for primary care researchers. *Ann Fam Med* 2004;**2**(3):204-08.
16. DeMauro SB, Giaccone A, Kirpalani H, *et al.* Quality of reporting of neonatal and infant trials in high-impact journals. *Pediatr* 2011;**128**(3):e639-44.
17. Zachariassen G, Faerk J, Grytter C, *et al.* Nutrient enrichment of mother's milk and growth of very preterm infants after hospital discharge. *Pediatr* 2011;**127**(4):e995-1003.
18. Carlo WA, Finer NN, Walsh MC, *et al.* Target ranges of oxygen saturation in extremely preterm infants. *N Engl J Med* 2010;**362**(21):1959-69.
19. Gopel W, Kribs A, Ziegler A, *et al.* Avoidance of mechanical ventilation by surfactant treatment of spontaneously breathing preterm infants (AMV): an open-label, randomised, controlled trial. *Lancet* 2011;**378**(9803):1627-34.
20. Sandri F, Plavka R, Ancora G, *et al.* Prophylactic or early selective surfactant combined with nCPAP in very preterm infants. *Pediatr* 2010;**125**(6):e1402-09.

21. Makrides M, Gibson RA, McPhee AJ, *et al.* Neurodevelopmental outcomes of preterm infants fed high-dose docosahexaenoic acid: a randomized controlled trial. *J Am Med Assoc* 2009;**301**(2):175-82.

22. Koldewijn K, Wolf MJ, van Wassenaer A, *et al.* The infant behavioral assessment and intervention program for very low birth weight infants at 6 months corrected age. *J Pediatr* 2009;**154**(1):33-38.

23. Schmidt B, Whyte RK, Asztalos EV, *et al.* Effects of targeting higher vs lower arterial oxygen saturations on death or disability in extremely preterm infants: a randomized clinical trial. *J Am Med Assoc* 2013;**309**(20):2111-20.

24. Dunn MS, Kaempf J, de Klerk A, *et al.* Randomized trial comparing 3 approaches to the initial respiratory management of preterm neonates. *Pediatr* 2011;**128**(5):e1069-76.

25. Garite TJ, Kurtzman J, Maurel K, *et al.* Impact of a 'rescue course' of antenatal corticosteroids: a multicenter randomized placebo-controlled trial. *Am J Obstet Gynecol* 2009;**200**(3):248 e1-9.

26. Murphy KE, Hannah ME, Willan AR, *et al.* Multiple courses of antenatal corticosteroids for preterm birth (MACS): a randomised controlled trial. *Lancet* 2008;**372**(9656):2143-51.

27. Rouse DJ, Hirtz DG, Thom E, *et al.* A randomized, controlled trial of magnesium sulfate for the prevention of cerebral palsy. *N Engl J Med* 2008;**359**(9):895-905.

28. Ramanathan R, Sekar KC, Rasmussen M, *et al.* Nasal intermittent positive pressure ventilation after surfactant treatment for respiratory distress syndrome in preterm infants <30 weeks' gestation: a randomized, controlled trial. *J Perinatol* 2012;**32**(5):336-43.

29. Roos C, Spaanderman ME, Schuit E, *et al.* Effect of maintenance tocolysis with nifedipine in threatened preterm labor on perinatal outcomes: a randomized controlled trial. *J Am Med Assoc* 2013;**309**(1):41-47.

30. Vierron E, Giraudeau B. Design effect in multicenter studies: gain or loss of power? *BMC medical research methodology* 2009;**9**:39.

31. Campbell MK, Piaggio G, Elbourne DR, *et al.* Consort 2010 statement: extension to cluster randomised trials. *Br Med J* 2012;**345**:e5661.

32. Murray DM, Varnell SP, Blitstein JL. Design and analysis of group-randomized trials: a review of recent methodological developments. *Am J Epidemiol* 2004;**94**(3):423-32.

33. Cannon MJ, Warner L, Taddei JA, *et al.* What can go wrong when you assume that correlated data are independent: an illustration from the evaluation of a childhood health intervention in Brazil. *Stat Med* 2001;**20**(9-10):1461-67.

34. Kernan WN, Viscoli CM, Makuch RW, *et al.* Stratified randomization for clinical trials. *J Clin Epidemiol* 1999;**52**(1):19-26.

35. Neuhaus JM. Estimation efficiency with omitted covariates in generalized linear models. *J Am Stat Assoc* 1998;**93**(443):1124-29.

36. Committee for Proprietary Medicinal Products. Points to consider on adjustment for baseline covariates. *Stat Med* 2004;**23**(5):701-9.

37. Andridge RR. Quantifying the impact of fixed effects modeling of clusters in multiple imputation for cluster randomized trials. *Biom J* 2011;**53**(1):57-74.

38. Taljaard M, Donner A, Klar N. Imputation strategies for missing continuous outcomes in cluster randomized trials. *Biom J* 2008;**50**(3):329-45.

39. Schulz KF, Altman DG, Moher D. CONSORT 2010 statement: updated guidelines for reporting parallel group randomised trials. *Br Med J* 2010;**340**:c332.

**LICENCE STATEMENT**

**Table 1: Characteristics of trials including infants from multiple births**

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

* 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**

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

## 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


1 Blanken MO, Rovers MM, Molenaar JM, *et al.* Respiratory syncytial virus and recurrent wheeze in healthy preterm infants. *N Engl J Med* 2013;**368**(19):1791-99.

2 Schmidt B, Whyte RK, Asztalos EV, *et al.* Effects of targeting higher vs lower arterial oxygen saturations on death or disability in extremely preterm infants: a randomized clinical trial. *J Am Med Assoc* 2013;**309**(20):2111-20.

3 Roos C, Spaanderman ME, Schuit E, *et al.* Effect of maintenance tocolysis with nifedipine in threatened preterm labor on perinatal outcomes: a randomized controlled trial. *J Am Med Assoc* 2013;**309**(1):41-47.

4 Fergusson DA, Hebert P, Hogan DL, *et al.* Effect of fresh red blood cell transfusions on clinical outcomes in premature, very low-birth-weight infants: the ARIPI randomized trial. *J Am Med Assoc* 2012;**308**(14):1443-51.

5 van der Ham DP, van der Heyden JL, Opmeer BC, *et al.* Management of late-preterm premature rupture of membranes: the PPROMEXIL-2 trial. *Am J Obstet Gynecol* 2012;**207**(4):276 e1-10.

6 van der Ham DP, Vijgen SM, Nijhuis JG, *et al.* Induction of labor versus expectant management in women with preterm prelabor rupture of membranes between 34 and 37 weeks: a randomized controlled trial. *PLoS Med* 2012;**9**(4):e1001208.

7 Leaf A, Dorling J, Kempley S, *et al.* Early or delayed enteral feeding for preterm growth-restricted infants: a randomized trial. *Pediatr* 2012;**129**(5):e1260-68.

8 Ramanathan R, Sekar KC, Rasmussen M, *et al.* Nasal intermittent positive pressure ventilation after surfactant treatment for respiratory distress syndrome in preterm infants <30 weeks' gestation: a randomized, controlled trial. *J Perinatol* 2012;**32**(5):336-43.

9 Cignacco EL, Sellam G, Stoffel L, *et al.* Oral sucrose and "facilitated tucking" for repeated pain relief in preterms: a randomized controlled trial. *Pediatr* 2012;**129**(2):299-308.

10 Dunn MS, Kaempf J, de Klerk A, *et al.* Randomized trial comparing 3 approaches to the initial respiratory management of preterm neonates. *Pediatr* 2011;**128**(5):e1069-76.

11 Gopel W, Kribs A, Ziegler A, *et al.* Avoidance of mechanical ventilation by surfactant treatment of spontaneously breathing preterm infants (AMV): an open-label, randomised, controlled trial. *Lancet* 2011;**378**(9803):1627-34.

12 New K, Flint A, Bogossian F, *et al.* Transferring preterm infants from incubators to open cots at 1600 g: a multicentre randomised controlled trial. *Arch Dis Child Fetal Neonatal Ed* 2012;**97**(2):F88-92.

13 Rubin LP, Chan GM, Barrett-Reis BM, *et al.* Effect of carotenoid supplementation on plasma carotenoids, inflammation and visual development in preterm infants. *J Perinatol* 2012;**32**(6):418-24.

14 Zachariassen G, Faerk J, Grytter C, *et al.* Nutrient enrichment of mother's milk and growth of very preterm infants after hospital discharge. *Pediatr* 2011;**127**(4):e995-1003.

15 Mintz-Hittner HA, Kennedy KA, Chuang AZ. Efficacy of intravitreal bevacizumab for stage 3+ retinopathy of prematurity. *N Engl J Med* 2011;**364**(7):603-15.

16 Mercier JC, Hummler H, Durrmeyer X, *et al.* Inhaled nitric oxide for prevention of bronchopulmonary dysplasia in premature babies (EUNO): a randomised controlled trial. *Lancet* 2010;**376**(9738):346-54.
17 Carlo WA, Finer NN, Walsh MC, *et al.* Target ranges of oxygen saturation in extremely preterm infants. *N Engl J Med* 2010;**362**(21):1959-69.
18 Sandri F, Plavka R, Ancora G, *et al.* Prophylactic or early selective surfactant combined with nCPAP in very preterm infants. *Pediatr* 2010;**125**(6):e1402-09.
19 Ortenstrand A, Westrup B, Brostrom EB, *et al.* The Stockholm Neonatal Family Centered Care Study: effects on length of stay and infant morbidity. *Pediatr* 2010;**125**(2):e278-85.
20 Carbonell-Estrany X, Simoes EA, Dagan R, *et al.* Motavizumab for prophylaxis of respiratory syncytial virus in high-risk children: a noninferiority trial. *Pediatr* 2010;**125**(1):e35-51.
21 Garite TJ, Kurtzman J, Maurel K, *et al.* Impact of a 'rescue course' of antenatal corticosteroids: a multicenter randomized placebo-controlled trial. *Am J Obstet Gynecol* 2009;**200**(3):248 e1-9.
22 Carr R, Brocklehurst P, Dore CJ, *et al.* Granulocyte-macrophage colony stimulating factor administered as prophylaxis for reduction of sepsis in extremely preterm, small for gestational age neonates (the PROGRAMS trial): a single-blind, multicentre, randomised controlled trial. *Lancet* 2009;**373**(9659):226-33.
23 Makrides M, Gibson RA, McPhee AJ, *et al.* Neurodevelopmental outcomes of preterm infants fed high-dose docosahexaenoic acid: a randomized controlled trial. *J Am Med Assoc* 2009;**301**(2):175-82.
24 Murphy KE, Hannah ME, Willan AR, *et al.* Multiple courses of antenatal corticosteroids for preterm birth (MACS): a randomised controlled trial. *Lancet* 2008;**372**(9656):2143-51.
25 Koldewijn K, Wolf MJ, van Wassenaer A, *et al.* The Infant Behavioral Assessment and Intervention Program for very low birth weight infants at 6 months corrected age. *J Pediatr* 2009;**154**(1):33-38.
26 Rouse DJ, Hirtz DG, Thom E, *et al.* A randomized, controlled trial of magnesium sulfate for the prevention of cerebral palsy. *N Engl J Med* 2008;**359**(9):895-905.
27 Rojas MA, Lozano JM, Rojas MX, *et al.* Prophylactic probiotics to prevent death and nosocomial infection in preterm infants. *Pediatr* 2012;**130**(5):e1113-20.
28 Moya F, Sisk PM, Walsh KR, *et al.* A new liquid human milk fortifier and linear growth in preterm infants. *Pediatr* 2012;**130**(4):e928-35.
29 Todd DA, Wright A, Broom M, *et al.* Methods of weaning preterm babies <30 weeks gestation off CPAP: a multicentre randomised controlled trial. *Arch Dis Child Fetal Neonatal Ed* 2012;**97**(4):F236-40.
30 Dani C, Vangi V, Bertini G, *et al.* High-dose ibuprofen for patent ductus arteriosus in extremely preterm infants: a randomized controlled study. *Clin Pharmacol Ther* 2012;**91**(4):590-96.
31 Dani C, Lori I, Favelli F, *et al.* Lutein and zeaxanthin supplementation in preterm infants to prevent retinopathy of prematurity: a randomized controlled study. *J Matern Fetal Neonatal Med* 2012;**25**(5):523-27.
32 Bober K, Swietlinski J, Zejda J, *et al.* A multicenter randomized controlled trial comparing effectiveness of two nasal continuous positive airway pressure devices in very-low-birth-weight infants. *Pediatr Crit Care Med* 2012;**13**(2):191-96.

33 Marcatto JO, Vasconcelos PC, Araujo CM, *et al.* EMLA versus glucose for PICC insertion: a randomised triple-masked controlled study. *Arch Dis Child Fetal Neonatal Ed* 2011;**96**(6):F467-68.
34 Acolet D, Allen E, Houston R, *et al.* Improvement in neonatal intensive care unit care: a cluster randomised controlled trial of active dissemination of information. *Arch Dis Child Fetal Neonatal Ed* 2011;**96**(6):F434-39.
35 Standley JM, Cassidy J, Grant R, *et al.* The effect of music reinforcement for non-nutritive sucking on nipple feeding of premature infants. *Pediatr Nurs* 2010;**36**(3):138-45.
36 Modi N, Uthaya S, Fell J, *et al.* A randomized, double-blind, controlled trial of the effect of prebiotic oligosaccharides on enteral tolerance in preterm infants. *Pediatr Res* 2010;**68**(5):440-45.
37 Wang Y, Tao YX, Cai W, *et al.* Protective effect of parenteral glutamine supplementation on hepatic function in very low birth weight infants. *Clin Nutr* 2010;**29**(3):307-11.
38 Fujii AM, Patel SM, Allen R, *et al.* Poractant alfa and beractant treatment of very premature infants with respiratory distress syndrome. *J Perinatol* 2010;**30**(10):665-70.
39 Sullivan S, Schanler RJ, Kim JH, *et al.* An exclusively human milk-based diet is associated with a lower rate of necrotizing enterocolitis than a diet of human milk and bovine milk-based products. *J Pediatr* 2010;**156**(4):562-67.
40 Manzoni P, Rinaldi M, Cattani S, *et al.* Bovine lactoferrin supplementation for prevention of late-onset sepsis in very low-birth-weight neonates: a randomized trial. *J Am Med Assoc* 2009;**302**(13):1421-28.
41 Lee SK, Aziz K, Singhal N, *et al.* Improving the quality of care for infants: a cluster randomized controlled trial. *CMAJ* 2009;**181**(8):469-76.
42 Kumar P, Murki S, Malik GK, *et al.* Light emitting diodes versus compact fluorescent tubes for phototherapy in neonatal jaundice: a multi center randomized controlled trial. *Indian J Pediatr* 2010;**47**(2):131-37.
43 Kuhn P, Messer J, Paupe A, *et al.* A multicenter, randomized, placebo-controlled trial of prophylactic recombinant granulocyte-colony stimulating factor in preterm neonates with neutropenia. *J Pediatr* 2009;**155**(3):324-30.
44 Rojas MA, Lozano JM, Rojas MX, *et al.* Very early surfactant without mandatory ventilation in premature infants treated with early continuous positive airway pressure: a randomized, controlled trial. *Pediatr* 2009;**123**(1):137-42.
45 Aranda JV, Clyman R, Cox B, *et al.* A randomized, double-blind, placebo-controlled trial on intravenous ibuprofen L-lysine for the early closure of nonsymptomatic patent ductus arteriosus within 72 hours of birth in extremely low-birth-weight infants. *Am J Perinatol* 2009;**26**(3):235-45.
46 Lin HC, Hsu CH, Chen HL, *et al.* Oral probiotics prevent necrotizing enterocolitis in very low birth weight preterm infants: a multicenter, randomized, controlled trial. *Pediatr* 2008;**122**(4):693-700.
47 Maguire CM, Walther FJ, van Zwieten PH, *et al.* No change in developmental outcome with incubator covers and nesting for very preterm infants in a randomised controlled trial. *Arch Dis Child Fetal Neonatal Ed* 2009;**94**(2):F92-97.
48 Picaud JC, Decullier E, Plan O, *et al.* Growth and bone mineralization in preterm infants fed preterm formula or standard term formula after discharge. *J Pediatr* 2008;**153**(5):616-21.

49 Rees CM, Eaton S, Kiely EM, *et al.* Peritoneal drainage or laparotomy for neonatal bowel perforation? A randomized controlled trial. *Ann Surg* 2008;**248**(1):44-51.
50 Jegatheesan P, Ianus V, Buchh B, *et al.* Increased indomethacin dosing for persistent patent ductus arteriosus in preterm infants: a multicenter, randomized, controlled trial. *J Pediatr* 2008;**153**(2):183-89.
51 Grobman WA, Thom EA, Spong CY, *et al.* 17 alpha-hydroxyprogesterone caproate to prevent prematurity in nulliparas with cervical length less than 30 mm. *Am J Obstet Gynecol* 2012;**207**(5):390 e1-8.
52 Adzick NS, Thom EA, Spong CY, *et al.* A randomized trial of prenatal versus postnatal repair of myelomeningocele. *N Engl J Med* 2011;**364**(11):993-1004.
53 Macones GA, Parry S, Nelson DB, *et al.* Treatment of localized periodontal disease in pregnancy does not reduce the occurrence of preterm birth: results from the Periodontal Infections and Prematurity Study (PIPS). *Am J Obstet Gynecol* 2010;**202**(2):147 e1-8.
54 Harper M, Thom E, Klebanoff MA, *et al.* Omega-3 fatty acid supplementation to prevent recurrent preterm birth: a randomized controlled trial. *Obstet Gynecol* 2010;**115**(2 Pt 1):234-42.
55 Landon MB, Spong CY, Thom E, *et al.* A multicenter, randomized trial of treatment for mild gestational diabetes. *N Engl J Med* 2009;**361**(14):1339-48.
56 Salvo V, Zimmermann LJ, Gavilanes AW, *et al.* First intention high-frequency oscillatory and conventional mechanical ventilation in premature infants without antenatal glucocorticoid prophylaxis. *Pediatr Crit Care Med* 2012;**13**(1):72-79.

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

| **Trial** | **Interventions** | **Primary Outcome*** |
|---|---|---|
| Blanken et al.[1] | Monthly palivizumab injections vs placebo | Total number of parent reported wheezing days in the first year of life |
| Schmidt et al.[2] | Pulse oximiters displaying oxygen saturations 3% higher vs lower than the true value | Composite of death, gross motor disability, cognitive or language delay, severe hearing loss or bilateral blindness at 18 months corrected age |
| Roos et al.[3] | Maintenance tocolysis with nifedipine orally vs placebo | Composite of perinatal death, chronic lung disease, neonatal sepsis, intraventricular haemorrhage >grade 2, periventricular leukomalacia >grade 1 or necrotizing enterocolitis up to 6 months of age |
| Fergusson et al.[4] | Transfusion of red blood cells stored 7 days or less vs standard issue red blood cells | Composite of necrotizing enterocolitis, retinopathy of prematurity, bronchopulmonary dysplasia, intraventricular haemorrhage or death while in the neonatal intensive care unit up to |

| | | |
|---|---|---|
| | | 90 days |
| van der Ham et al.[5] | Expectant management vs induction of labour | Neonatal sepsis |
| van der Ham et al.[6] | Expectant management vs induction of labour | Neonatal sepsis |
| Leaf et al.[7] | Early vs late commencement of enteral feeds | Time to achieve full enteral feeding sustained for 72 hours |
| Ramanathan et al.[8] | Nasal intermittent positive pressure ventilation vs nasal continuous positive airway pressure | Need for mechanical ventilation via endotracheal tube at 7 days of age |
| Cignacco et al.[9] | Oral sucrose vs facilitated tucking vs combined oral sucrose and facilitated tucking | Pain response measured on the Bernese Pain Scale for Neonates |
| Dunn et al.[10] | Prophylactic surfactant vs intubate-surfactant-extubate vs nasal continuous positive airway pressure | Death or bronchopulmonary dysplasia at 36 weeks postmenstrual age |
| Gopel et al.[11] | Surfactant without ventilation vs standard care | Need for mechanical ventilation or not being ventilated but having $pCO_2$>65mmHg or $FiO2_2$>0.60 or both for more than 2 hours between 25 and 72 hours of age |
| New et al.[12] | Transfer to an open cot at 1600 vs 1800g | Average daily weight gain over the first 14 days following transfer to an open cot |
| Rubin et al.[13] | Carotenoid supplemented formula vs control formula | Plasma lutein concentration |
| Zachariassen et al.[14] | Unfortified vs fortified mother's milk | Weight |
| Mintz-Hittner et al.[15] | Conventional laser therapy vs intravitreal bevacizumab monotherapy | Recurrence of neovascularization in one or both eyes arising from the retinal vessels and requiring retreatment by 54 weeks postmenstrual age |
| Mercier et al.[16] | Inhaled nitric oxide vs placebo gas | Survival without development of bronchopulmonary dysplasia at 36 weeks postmenstrual age |
| Carlo et al.[17] | Target range of oxygen saturation of 85-89% vs 81-95% | Composite of severe retinopathy of prematurity or death before hospital discharge |
| Sandri et al.[18] | Prophylactic surfactant followed by nasal continuous positive airway pressure vs early nasal | Need for mechanical ventilation in the first 5 days of life |

| | | |
|---|---|---|
| | continuous positive airway pressure followed by early selective surfactant | |
| Ortenstrand et al.[19] | Family care ward vs standard care ward | Total length of hospital stay |
| Carbonell-Estrany et al.[20] | Monthly motavizumab vs palivizumab injections | 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 |
| Garite et al.[21] | Single rescue course of betamethasone vs placebo | Composite of respiratory distress syndrome, bronchopulmonary dysplasia, severe intraventricular haemorrhage, periventricular leukomalacia, blood culture proven sepsis, necrotizing enterocolitis or perinatal death |
| Carr et al.[22] | Granulocyte-macrophage colony stimulating factor vs standard management | Survival without an episode of culture positive systemic sepsis to 14 days from trial entry |
| Makrides et al.[23] | High vs standard docosahexaenoic acid enteral feeds | Bayley Mental Development Index at 18 months corrected age |
| Murphy et al.[24] | Multiple courses of antenatal corticosteroids vs placebo | Composite of perinatal or neonatal mortality, severe respiratory distress syndrome, intraventricular haemorrhage >grade 2, periventricular leukomalacia, bronchopulmonary dysplasia or necrotizing enterocolitis |
| Koldewijn et al.[25] | Infant Behavioural Assessment and Intervention Program vs standard care | Bayley Mental Development Index at 6 months corrected age |
| Rouse et al.[26] | Magnesium sulfate vs placebo | Composite of stillbirth or infant death by 1 year corrected age or moderate or severe cerebral palsy at or beyond 2 years corrected age |

* According to the definition described in the Methods. All outcomes relate to the infant.

**Figure 1: Flow diagram**