Early parenteral nutrition and growth outcomes in preterm infants: a systematic review and meta-analysis1–4

Helen E Moyses, Mark J Johnson, Alison A Leaf, and Victoria R Cornelius

ABSTRACT

Background: The achievement of adequate nutritional intakes in preterm infants is challenging and may explain the poor growth often seen in this group. The use of early parenteral nutrition (PN) is one potential strategy to address this problem, although the benefits and harms are unknown.

Objective: We determined whether earlier administration of PN benefits growth outcomes in preterm infants.

Design: We conducted a systematic review of randomized controlled trials (RCTs) and observational studies.

Results: Eight RCTs and 13 observational studies met the inclusion criteria (n = 553 and 1796 infants). The meta-analysis was limited by disparate growth-outcome measures. An assessment of bias was difficult because of inadequate reporting. Results are given as mean differences (95% CIs). Early PN reduced the time to regain birth weight by 2.2 d (1.1, 3.2 d) for RCTs and 3.2 d (2.0, 4.4 d) in observational studies. The maximum percentage weight loss with early PN was lower by 3.1 percentage points (1.7, 4.5 percentage points) for RCTs and by 3.5 percentage points (2.6, 4.3 percentage points) for observational studies. Early PN improved weight at discharge or 36 wk postmenstrual age by 14.9 g (5.3, 24.5 g) (observational studies only), but no benefit was shown for length or head circumference. There was no evidence that early PN significantly affects risk of mortality, necrotizing enterocolitis, sepsis, chronic lung disease, intraventricular hemorrhage, or cholestasis.

Conclusions: The results of this review, although subject to some limitations, show that early PN provides a benefit for some short-term growth outcomes. No evidence that early PN increases morbidity or mortality was found. Neonatal research would benefit from the development of a set of core growth outcome measures.


INTRODUCTION

Infants born at extremes of prematurity often fail to achieve rates of growth similar to those seen in utero at the equivalent gestation. This failure of growth has been shown in several studies, with preterm infants often being discharged at significantly lower percentiles for weight, head circumference, and length than those at which they were born (1, 2). Growth is a complex process, and the reasons for the poor growth seen in this population are almost certainly multifactorial. However, one important factor is that these infants receive inadequate nutrition, particularly early in life, and fail to achieve recommended nutrient targets. Embleton et al (3) looked at the provision of energy and protein to preterm infants and showed the accumulation of significant deficits early in life that were never recovered. These deficits were directly related to the degree of postnatal growth failure. Similarly, Martin et al (4) also showed that the early provision of nutrients was an important determinant of postnatal growth. This provision is important because improved postnatal growth has been shown to reduce risk of neurodevelopmental impairment at 18 mo (5). Furthermore, a higher early provision of protein and energy has been shown to be associated with improved neurodevelopmental outcomes at 18 mo (6).

The achievement of adequate nutritional intakes in preterm infants during the first few weeks of life is challenging, and one potential strategy to address this task is the use of early parenteral nutrition (PN)5 initiated soon after birth (7, 8). However, practice in this area is variable as highlighted by a recent confidential inquiry into the use of PN in the United Kingdom that showed frequent shortfalls in the nutritional adequacy of neonatal PN and delays in its initiation (9). This inquiry showed the variability in practice and need for clearer guidelines on the basis of recommended nutrient intakes.

Our research group is currently undertaking a study to improve nutrient delivery to preterm infants through a program aimed at standardizing nutritional care. To inform the evidence base for this study, we carried out a systematic review to investigate the effect of an early compared with late initiation of PN on growth, mortality, and neurodevelopment.

1 From the National Institute for Health Research (NIHR) Southampton Biomedical Research Centre (HEM, MJJ, and AAL) and the NIHR Southampton Respiratory Biomedical Research Unit (HEM), University of Southampton and University Hospital Southampton National Health Service (NHS) Foundation Trust, Southampton, United Kingdom; the Department of Neonatal Medicine, University Hospital Southampton NHS Foundation Trust, Southampton, United Kingdom (MJJ and AAL); and the Department of Primary Care and Public Health Sciences, Kings College, London, United Kingdom (VRC).

2 Supported by the NIHR Southampton Nutrition, Diet and Lifestyle Biomedical Research Unit (HEM, MJJ, AAL, and VRC); the NIHR Southampton Respiratory Biomedical Research Unit (HEM and VRC); and the NIHR Southampton Biomedical Research Centre (HEM, MJJ, and AAL).

3 HEM and MJJ contributed equally to this work and are joint first authors.

4 Address correspondence to HE Moyses, Biomedical Research Unit, University Hospital Southampton National Health Service Foundation Trust, Mailpoint 218, Southampton General Hospital, Tremona Road, Southampton SO16 6YD, United Kingdom. E-mail: h.e.moyses@soton.ac.uk.

5 Abbreviations used: CLD, chronic lung disease; IVH, intraventricular hemorrhage; NEC, necrotizing enterocolitis; PMA, postmenstrual age; PN, parenteral nutrition; RCT, randomized controlled trial.

Received April 27, 2012. Accepted for publication January 9, 2013.
First published online February 27, 2013; doi: 10.3945/ajcn.112.042028.
morbidity, and mortality in preterm infants. A previous systematic review (10) that looked specifically at the early introduction of lipids showed no benefit in the introduction of lipids <5 d of age. We aimed to systematically analyze data from all available observational studies and randomized controlled trials (RCTs) on the effect of the timing of commencement of PN on growth and risk of morbidity and mortality in preterm infants.

METHODS

See “Supplemental data” in the online issue for a protocol for systematic review that we produced.

Eligibility criteria

Studies were eligible if they were RCTs or observational studies that compared early with late PN or intravenous amino acids and had at least one growth outcome. Study participants were preterm infants. The definition of early was ≤48 h of birth, and studies in which the early feeding intervention was >48 h were excluded. There was no restriction on language.

Searches and information sources

An all-language literature search was carried out by using the key words and search strategy detailed in Table 1. The electronic databases MEDLINE (1947 to present), EMBASE (1947 to present), the Cumulative Index to Nursing and Allied Health Literature (1981 to present), and the Health Management Information Consortium were searched by using OvidSP (Wolters Kluwer Health; http://www.ovid.com). The last search was run on 10 July 2012. Conference abstracts and other citations were identified by searching the websites Web of Science (http://wok.mimas.ac.uk) and CSA Conference Papers Index (http://www.csa.com) by using the same search strategy outlined in Table 1 but adapted for use in the Web interface. Citation and reference searching was performed on articles that were selected for review.

Study selection

Studies were independently assessed for eligibility by 2 reviewers (MJJ and VRC). Reviewers were not blinded to identities of study authors or institutions.

Data-collection process

Data extraction was carried out by 2 reviewers (MJJ and VRC) by using a specially designed spreadsheet, which collected data on study characteristics (ie, methods, design, nature of PN intervention, gestational age at birth, birth weight, and definitions of early and late initiation of feeding).

Outcome measures collected were physical growth (including weight, length, and head circumference), morbidity [including necrotizing enterocolitis (NEC), sepsis, chronic lung disease (CLD), intraventricular hemorrhage (IVH) and cholestasis], and mortality.

Risk of bias in individual studies

The quality of observational studies was assessed by using items from the Newcastle Ottawa Scale (11), whereas RCTs were assessed by using the Cochrane risk of bias assessment tool (12).

Synthesis of results

Outcome measures available for each study were tabulated, and the study methodology was assessed, to ascertain the potential for synthesis. RCTs and observational studies were considered separately. Heterogeneity was checked by using the I² statistic, which reports the percentage of variation attributable to heterogeneity. If sufficient data were available, and it was deemed appropriate, summary estimates for effects of early compare with late feeding were produced. For continuous outcomes, the weighted mean difference was calculated by using the inverse-variance method. When the I² statistic was >50%, a random-effects model by using the method of DerSimonian and Laird (13) was used. For binary-outcome data, Peto’s method was used to calculate a pooled OR because this has been shown to provide the least-biased estimated when the outcome is rare (14). In cases where SDs were not given, these were calculated from CIs if available or from the median and range by using the method described by Hozo et al (15). The analysis was performed with Stata version 11.0 software (StataCorp LP).

RESULTS

Study selection

Searches identified 3340 potentially relevant studies, with 2166 studies that remained after the removal of duplicates. After the screening of titles and abstracts, 86 full-text articles were
obtained. Of these articles, 8 RCTs (10 articles) and 13 observational studies (15 articles) met the inclusion criteria (Figure 1).

Included studies were published between 1989 and 2012.

Study characteristics

Included studies had a total of 2349 participants (553 participants in RCTs and 1796 participants in observational studies). The timing of early initiation ranged from immediately after delivery to ≤48 h after birth. The timing of introduction in the late group ranged from 12 h to the sixth day of life. The duration of PN between early and late groups was not significantly different in RCTs [mean difference: 3.24 d; 95% CI: −0.99, 7.47 d; P = 0.133] or observational studies (mean difference: 0.29 d; 95% CI: −3.16, 3.74 d; P = 0.868). Of included studies, one RCT and 5 observational studies assessed complex interventions that involved the use of other measures in addition to early PN such as early enteral feeding or nutritional guidelines. Because of the potential impact that such measures might have had on outcomes, a sensitivity analysis was carried out, with these studies coded as complex and considered as subgroups. A summary of characteristics of included studies is shown in Table 2.

Risk of study bias

Quality-assessment criteria of included RCTs and observational studies are summarized in Tables 3 and 4, respectively. Inadequate reporting by authors meant that it was not possible to fully assess risk of bias in RCTs. Six of 8 studies reported such that ≤3 items on the bias-assessment scale could be scored. In general, observational studies were assessed as having included a representative cohort of infants, with data collected from reliable records. However, there was inadequate reporting of items, such as the assessment of outcome, which are crucial to understanding the potential bias. Of note, only 2 observational studies adjusted the estimated intervention effects for potential confounding factors between groups (17, 29). Some observational studies did not report follow-up information for a significant (or, in some cases, unknown) number of infants who were initially entered into the study. The exclusion of these infants may have introduced bias because reasons for exclusion may have been associated with poor outcomes. For all studies, it was difficult to assess whether they were subject to a selective reporting bias.

Growth outcomes

Reported measures throughout this section are presented as mean differences (95% CIs).

Time to regain birth weight

Ten observational studies and 4 RCTs reported the time to regain birth weight. Both observational studies and RCTs showed a significant reduction in the time to regain birth weight for early PN. Reductions were 3.19 d (1.95, 4.44 d) (P < 0.001) and 2.24 d (1.29, 3.19 d) (P < 0.001), respectively (Figure 2A, forest plot). A sensitivity analysis in which observational studies were split by whether or not they used a complex intervention gave a similar and significant result.

Maximum weight loss

Six observational studies and 2 RCTs reported the maximum percentage weight loss. Both observational studies and RCTs showed a significant reduction in the maximum weight loss with early PN by 3.48 percentage points (2.63, 4.33 percentage points) (P < 0.001) and 3.09 percentage points (1.70, 4.48 percentage points) (P < 0.001), respectively (Figure 2B, forest plot).

Growth at discharge or 36 wk postmenstrual age

Studies that reported growth outcomes at either discharge or 36 wk postmenstrual age (PMA) were pooled for a meta-analysis. For weight, although the 6 included observational studies showed a small but significant improvement of 14.94 g (5.31, 24.5 g) (P = 0.002) in weight that favored early PN, this result was not corroborated by the RCTs (2 studies), in which no evidence of a significant difference was shown (Figure 3A). A sensitivity analysis that split observational studies with the use of a complex intervention showed a smaller improvement of 8.91 g (2.18, 17.99 g) (P = 0.055) for noncomplex studies that used early PN. Complex studies showed an increase of 24.25 g (15.27, 33.24 g) (P < 0.001) with early PN.

The meta-analysis showed no evidence of a significant difference in head circumference at discharge or 36 wk PMA for either observational studies or RCTs (Figure 3B). For length (Figure 3C), the single observational study by Dinerstein et al (17) showed a significant improvement in length at discharge or 36 wk PMA for early PN, but no evidence for a significant difference was shown for RCTs (2 studies).

FIGURE 1. Flow diagram showing the results of the literature search and selection of studies for the systematic review.
### Characteristics of included studies

<table>
<thead>
<tr>
<th>First author, year of publication (reference)</th>
<th>Study type</th>
<th>Complex</th>
<th>No. of subjects</th>
<th>Gestational age</th>
<th>Birth weight</th>
<th>Timing of PN</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Early</td>
<td>Late</td>
<td>Early</td>
<td>Late</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>wk</td>
<td>kg</td>
<td>wk</td>
<td>kg</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aroor, 2012 (16)</td>
<td>Retrospective</td>
<td>No</td>
<td>49</td>
<td>54</td>
<td>28.0</td>
<td>27.0</td>
</tr>
<tr>
<td>Dinerstein, 2006 (17)</td>
<td>Retrospective</td>
<td>Yes</td>
<td>117</td>
<td>65</td>
<td>30.0</td>
<td>30.0</td>
</tr>
<tr>
<td>Donovan, 2006 (18)</td>
<td>Retrospective</td>
<td>Yes</td>
<td>57</td>
<td>23</td>
<td>27.3</td>
<td>27.1</td>
</tr>
<tr>
<td>García, 2012 (19)</td>
<td>Prospective</td>
<td>No</td>
<td>29</td>
<td>29</td>
<td>30.1</td>
<td>30.1</td>
</tr>
<tr>
<td>Ho, 2003 (20)</td>
<td>Retrospective</td>
<td>Yes</td>
<td>17</td>
<td>19</td>
<td>30.9</td>
<td>30.7</td>
</tr>
<tr>
<td>Elstgeest, 2010 (21)</td>
<td>Retrospective</td>
<td>No</td>
<td>73</td>
<td>70</td>
<td>26.0</td>
<td>26.0</td>
</tr>
<tr>
<td>Geary, 2008 (22)</td>
<td>Retrospective</td>
<td>Yes</td>
<td>76</td>
<td>87</td>
<td>25.7</td>
<td>25.4</td>
</tr>
<tr>
<td>Han, 2012 (23)</td>
<td>Retrospective</td>
<td>No</td>
<td>112</td>
<td>52</td>
<td>31.3</td>
<td>31.3</td>
</tr>
<tr>
<td>Janeiro, 2010 (24)</td>
<td>Retrospective and prospective</td>
<td>Yes</td>
<td>28</td>
<td>28</td>
<td>29.8</td>
<td>29.1</td>
</tr>
<tr>
<td>Kotsopoulos, 2006 (25)</td>
<td>Prospective</td>
<td>No</td>
<td>54</td>
<td>54</td>
<td>26.1</td>
<td>26.3</td>
</tr>
<tr>
<td>Radmacher, 2009 (26, 27)</td>
<td>Retrospective</td>
<td>No</td>
<td>20</td>
<td>70</td>
<td>26.0</td>
<td>26.2</td>
</tr>
<tr>
<td>Trinitis, 2012 (28)</td>
<td>Retrospective and prospective</td>
<td>Yes</td>
<td>85</td>
<td>88</td>
<td>27.6</td>
<td>27.8</td>
</tr>
<tr>
<td>Valentine, 2009 (29)</td>
<td>Retrospective and prospective</td>
<td>No</td>
<td>308</td>
<td>132</td>
<td>29.1</td>
<td>29.4</td>
</tr>
<tr>
<td>Bai, 2005 (30)</td>
<td>RCT</td>
<td>No</td>
<td>20</td>
<td>20</td>
<td>32.0</td>
<td>32.3</td>
</tr>
<tr>
<td>Brownlee, 1993 (31)</td>
<td>RCT</td>
<td>No</td>
<td>63</td>
<td>66</td>
<td>29.0</td>
<td>29.0</td>
</tr>
<tr>
<td>Heinler, 2010 (32)</td>
<td>RCT</td>
<td>No</td>
<td>8</td>
<td>9</td>
<td>29.6</td>
<td>30.2</td>
</tr>
<tr>
<td>Ibrahim, 2004 (33)</td>
<td>RCT</td>
<td>No</td>
<td>16</td>
<td>16</td>
<td>27.0</td>
<td>26.8</td>
</tr>
<tr>
<td>Wilson, 1997 (34)</td>
<td>RCT</td>
<td>Yes</td>
<td>64</td>
<td>61</td>
<td>27.0</td>
<td>27.4</td>
</tr>
<tr>
<td>Blanco, 2008–2012 (35–37)</td>
<td>RCT</td>
<td>No</td>
<td>30</td>
<td>31</td>
<td>25.7</td>
<td>26.3</td>
</tr>
<tr>
<td>te Braake, 2005 (38)</td>
<td>RCT</td>
<td>No</td>
<td>66</td>
<td>69</td>
<td>28.4</td>
<td>28.4</td>
</tr>
<tr>
<td>Weiler, 2006 (39)</td>
<td>RCT</td>
<td>No</td>
<td>7</td>
<td>7</td>
<td>26.1</td>
<td>26.7</td>
</tr>
</tbody>
</table>

1 PN, parenteral nutrition; RCT, randomized controlled trial.

**Binary growth outcomes**

Effect measures throughout this section and the Harms section are reported as Peto’s ORs (95% CIs). Several studies reported binary growth outcomes, which are summarized in Table 5. Meta-analyses of 2 observational studies and a single RCT that reported the number of infants with weights below the 10th percentile at 36 wk PMA, discharge, or death showed a significant benefit in favor of early PN. ORs were 0.43 (0.26, 0.71) ($P = 0.001$) and 0.34 (0.16, 0.73) ($P = 0.006$), respectively. There was a similar result for the same studies when reporting the number of infants weights below the third percentile at 36 wk PMA, discharge, or death, although this result was NS at the 5% level for observational studies. In general, analyses indicated in favor of early PN in terms of a reduction of the probability of an infant having a length or head circumference below the 10th or third percentiles at 36 wk PMA, discharge, or death (Table 5). However, in most cases, only one study was included in each analysis, and we could not assess whether these results were vulnerable to a selective-outcome reporting bias. See Table S1 under “Supplemental data” in the online issue for a summary of all growth outcomes.

**Harms**

**Mortality**

Eleven studies reported mortality. Meta-analyses of 7 observational studies and 4 RCTs showed no evidence of a significant difference between early and late PN with ORs of 0.87 (0.52, 1.45) ($P = 0.590$) and 0.96 (0.55, 1.66) ($P = 0.882$), respectively (Figure 4). A sensitivity analysis that split observational studies by whether or not they used a complex intervention produced a similar result.

**CLD**

Nine observational studies and 5 RCTs reported CLD as an outcome. Meta-analyses of both study types showed no evidence of a difference in CLD rates between early and late PN groups, with ORs of 1.07 (0.82, 1.39) ($P = 0.616$) and 1.34 (0.87, 2.06) ($P = 0.189$) for observational studies and RCTs, respectively. A compound forest plot of all morbidity outcomes is shown in Figure 5. (See Figures S1–S5 under “Supplemental data” in the online issue for full forest plots for all morbidity outcomes.) A sensitivity analysis that split observational studies by whether they used a complex intervention or not produced a lower and nonsignificant estimated risk of complex studies that involved early PN (OR: 0.65; 95% CI: 0.41, 1.04; $P = 0.067$). Studies that purely looked at early PN showed higher nonsignificant estimated risk of CLD (OR: 1.34; 95% CI: 0.98, 1.84; $P = 0.074$).

**NEC**

Only one RCT reported the number of NEC events. There were less events in the early PN group, although this was NS with an OR of 0.52 (0.12, 2.16) ($P = 0.368$). Eleven observational studies reported NEC. There was a nonsignificant increased risk of NEC with early PN with an OR of 1.52 (0.97, 2.38) ($P = 0.065$; Figure 5, forest plot). A sensitivity analysis that split observational studies by the use of complex intervention showed no evidence of a significant difference between early and late groups when a complex intervention was used (5 studies; OR: 1.04; 95% CI: 0.51, 2.10; $P = 0.914$). The 6 studies that purely used early PN
showed a significant increase in risk of NEC with an OR of 1.95 (1.10, 3.46) ($P = 0.022$).

**Sepsis**

Both meta-analyses of 8 observational studies and 4 RCTs that reported sepsis as an outcome showed a nonsignificant reduction in risk of sepsis with early PN with ORs of 0.74 (0.54, 1.01) ($P = 0.050$) and 0.65 (0.37, 1.15) ($P = 0.136$), respectively (Figure 5). A sensitivity analysis that split observational studies by the use of complex intervention showed no evidence of a significant difference in risk of sepsis between early and late groups for complex interventions (OR: 0.96; 95% CI: 0.54, 1.73; $P = 0.898$), whereas studies that used PN alone showed a significant reduction in risk of sepsis with early PN (OR: 0.66; 95% CI: 0.46, 0.95; $P = 0.026$).

**IVH**

A meta-analysis of the 5 observational studies that reported IVH showed a nonsignificant reduction in risk of IVH with early PN (OR: 0.74; 95% CI: 0.51, 1.07; $P = 0.112$). A sensitivity analysis, again with observational studies split by the use of a complex intervention, gave a similar result, with ORs of 0.82 (0.53, 1.26) ($P = 0.363$) and 0.53 (0.25, 1.14) ($P = 0.106$) for noncomplex and complex studies, respectively. Meta-analysis of the 2 RCTs reporting IVH showed no evidence of a significant difference between the early and late groups with an OR of 1.05 (0.32, 3.40) $P = 0.930$.

**Cholestasis**

Only 5 observational studies reported rates of cholestasis. Risk of cholestasis with early compared with late PN was not shown to be different with an OR of 0.92 (0.58, 1.47) ($P = 0.713$; Figure 5).

### DISCUSSION

The use of early PN is becoming increasingly common practice in neonatal intensive care units, but to our knowledge, this is the first systematic review and meta-analysis that considered its effects on growth and any potential risks. For growth outcomes, this study showed significant benefits with early PN for initial weight loss and time to regain birth weight. These benefits were consistent across observational studies and RCTs. In addition, observational studies suggested a benefit for weight and length at discharge or 36 wk PMA, although these effects were not borne out by RCTs. Binary growth outcomes also suggested that early PN use significantly reduced the number of infants below the 10th and third percentiles at discharge or 36 wk.

It may also be argued that growth, particularly weight, may not be the best measure of clinical outcome in nutritional studies in preterm infants. Other criteria such as neurodevelopmental status at 2 y or other morbidities may be preferred, especially because of the association between poor growth and these outcomes (4). However, only one study, the RCT by Blanco et al (37), reported long-term effects on neurodevelopmental outcomes at 2 y of age. Although growth
### Table 4

<table>
<thead>
<tr>
<th>First author, year of publication (reference)</th>
<th>Representativeness of exposed cohort</th>
<th>Selection of the nonexposed cohort</th>
<th>Ascertainment of exposed cohort</th>
<th>Outcome not present at start of study</th>
<th>Comparability of cohorts</th>
<th>Assessment of outcome</th>
<th>Follow-up long enough for outcome to occur</th>
<th>Adequacy of follow-up</th>
</tr>
</thead>
<tbody>
<tr>
<td>Aroor, 2012 (16)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Dinerstein, 2006 (17)</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Donovan, 2006 (18)</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Elstgeest, 2010 (21)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>In Methods but not Results</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Garcia, 2012 (19)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Geary, 2008 (22)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Han, 2011 (23)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Ho, 2003 (20)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Janeiro, 2010 (24)</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>–</td>
</tr>
<tr>
<td>Kotsopoulos, 2006 (25)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Radmacher, 2009 (26, 27)</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Triniis, 2010 (28)</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Valentine, 2009 (29)</td>
<td>+</td>
<td>+</td>
<td>–</td>
<td>+</td>
<td>–</td>
<td>–</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

1 NOS, Newcastle-Ottawa Scale.
2 –, selected group of users or no description; +, truly or somewhat representative.
3 –, drawn from different source or no description; +, drawn from same community as the exposed.
4 –, written self-report or no description; +, secure record or structured interview.
5 –, no; +, yes. Item was true by default for each study as a result of the use of a growth outcome.
6 –, does not control for any factor; +, controls for baseline weight or other factors.
7 –, self-report or no description; +, independent blind assessment or record linkage.
8 –, no; +, yes. Item was true by default for each study as a result of the use of growth outcomes.
9 –, no description or statement; +, complete follow-up of all subjects accounted for or minimal loss unlikely to introduce bias.
outcomes at 2 y appeared worse in the early PN group, there were no differences in neurodevelopmental outcomes (Bayley II mental and psychomotor developmental indexes).

This review has not raised any concerns that early PN causes harm, and this is particularly important because of the increasing prevalence of early PN use in clinical practice. There was no significant increase observed in mortality or morbidity in terms of risk of CLD, NEC, or IVH. Of note, sensitivity analysis did reveal a small but significant increase in risk of NEC associated with early PN in noncomplex observational studies. However, whether this observation is true or was as a result of bias introduced by study design needs additional clarification. One ongoing concern associated with early PN use is that infants are exposed to PN for a more prolonged period, which increases their risk of sepsis (secondary to an increased duration of central venous catheter use) and cholestasis. In this review, the duration of PN was not shown to be significantly different between early and late PN groups, and estimated risk of sepsis, although nonsignificant, was

---

**FIGURE 2.** A: Time to regain birth weight (d). B: Maximum weight loss (in percentage points of birth weight). Open diamonds represent pooled mean differences (95% CIs) for each study type. Black circles represent the study mean differences, and the black bars are the 95% CIs. The size of the gray box is proportional to the weight of the study estimate in the meta-analysis. I-squared represents the percentage of variation attributable to heterogeneity. RCT, randomized controlled trial.
lower with early PN use. No significant difference was shown in the incidence of cholestasis.

This review was subject to several limitations. Although RCTs are the gold standard and produce an essentially unbiased estimate of combined studies (40), there was a lack of suitably powered RCTs in this area. RCTs that were included were small, which meant that, for some outcomes, the pooled sample would have still lacked power to detect significant differences even if

**FIGURE 3.** A: Weight at 36 wk or discharge (g). B: Head circumference (cm). C: Length (cm). Open diamonds represent pooled mean differences (95% CIs) for each study type. Black circles represent the study mean differences, and the black bars are the 95% CIs. The size of the gray box is proportional to the weight of the study estimate in the meta-analysis. I-squared represents the percentage of variation attributable to heterogeneity. RCT, randomized controlled trial.
present. We have reported RCTs and observational studies separately because confounding and selection bias often distort findings of observational studies. The size of observational studies ranged from 36 to 440 participants. However, smaller sample sizes can be preferable if it means that more attention is given to characterizing confounding factors (40). The observational studies included some with a before-and-after study design (24, 26, 29). With these studies, it was not possible to exclude the possibility that any changes in outcome may have occurred as a result of other factors that changed over time rather than the intervention.

A reporting bias and publication bias are potential limitations of any systematic review. We used the Newcastle Ottawa Scale (11) to assess the quality of observational studies and the Cochrane risk of bias assessment tool for RCTs (12). The reporting of items by RCTs was inadequate, which made it impossible to assess risk of bias or explore the potential effect on study results. Some aspects of observational studies were consistent with quality components; however, only 2 studies attempted to adjust for confounding variables. Moher et al (41) completed an extensive search to identify studies that examined the general quality of reporting of systematic reviews in nutrition but showed that few such studies exist. Issues highlighted were difficulties in sorting through numerous endpoints and variations in the intervention, study design, and study duration, which were similar to the difficulties encountered in this review.

In addition, there was considerably heterogeneity between included studies in terms of the nature of the intervention and exact groups of preterm infants studied. The composition of PN present. We have reported RCTs and observational studies separately because confounding and selection bias often distort findings of observational studies. The size of observational studies ranged from 36 to 440 participants. However, smaller sample sizes can be preferable if it means that more attention is given to characterizing confounding factors (40). The observational studies included some with a before-and-after study design (24, 26, 29). With these studies, it was not possible to exclude the possibility that any changes in outcome may have occurred as a result of other factors that changed over time rather than the intervention.

A reporting bias and publication bias are potential limitations of any systematic review. We used the Newcastle Ottawa Scale (11) to assess the quality of observational studies and the Cochrane risk of bias assessment tool for RCTs (12). The reporting of items by RCTs was inadequate, which made it impossible to assess risk of bias or explore the potential effect on study results. Some aspects of observational studies were consistent with quality components; however, only 2 studies attempted to adjust for confounding variables. Moher et al (41) completed an extensive search to identify studies that examined the general quality of reporting of systematic reviews in nutrition but showed that few such studies exist. Issues highlighted were difficulties in sorting through numerous endpoints and variations in the intervention, study design, and study duration, which were similar to the difficulties encountered in this review.

In addition, there was considerably heterogeneity between included studies in terms of the nature of the intervention and exact groups of preterm infants studied. The composition of PN present. We have reported RCTs and observational studies separately because confounding and selection bias often distort findings of observational studies. The size of observational studies ranged from 36 to 440 participants. However, smaller sample sizes can be preferable if it means that more attention is given to characterizing confounding factors (40). The observational studies included some with a before-and-after study design (24, 26, 29). With these studies, it was not possible to exclude the possibility that any changes in outcome may have occurred as a result of other factors that changed over time rather than the intervention.

A reporting bias and publication bias are potential limitations of any systematic review. We used the Newcastle Ottawa Scale (11) to assess the quality of observational studies and the Cochrane risk of bias assessment tool for RCTs (12). The reporting of items by RCTs was inadequate, which made it impossible to assess risk of bias or explore the potential effect on study results. Some aspects of observational studies were consistent with quality components; however, only 2 studies attempted to adjust for confounding variables. Moher et al (41) completed an extensive search to identify studies that examined the general quality of reporting of systematic reviews in nutrition but showed that few such studies exist. Issues highlighted were difficulties in sorting through numerous endpoints and variations in the intervention, study design, and study duration, which were similar to the difficulties encountered in this review.

In addition, there was considerably heterogeneity between included studies in terms of the nature of the intervention and exact groups of preterm infants studied. The composition of PN present. We have reported RCTs and observational studies separately because confounding and selection bias often distort findings of observational studies. The size of observational studies ranged from 36 to 440 participants. However, smaller sample sizes can be preferable if it means that more attention is given to characterizing confounding factors (40). The observational studies included some with a before-and-after study design (24, 26, 29). With these studies, it was not possible to exclude the possibility that any changes in outcome may have occurred as a result of other factors that changed over time rather than the intervention.

A reporting bias and publication bias are potential limitations of any systematic review. We used the Newcastle Ottawa Scale (11) to assess the quality of observational studies and the Cochrane risk of bias assessment tool for RCTs (12). The reporting of items by RCTs was inadequate, which made it impossible to assess risk of bias or explore the potential effect on study results. Some aspects of observational studies were consistent with quality components; however, only 2 studies attempted to adjust for confounding variables. Moher et al (41) completed an extensive search to identify studies that examined the general quality of reporting of systematic reviews in nutrition but showed that few such studies exist. Issues highlighted were difficulties in sorting through numerous endpoints and variations in the intervention, study design, and study duration, which were similar to the difficulties encountered in this review.

In addition, there was considerably heterogeneity between included studies in terms of the nature of the intervention and exact groups of preterm infants studied. The composition of PN present. We have reported RCTs and observational studies separately because confounding and selection bias often distort findings of observational studies. The size of observational studies ranged from 36 to 440 participants. However, smaller sample sizes can be preferable if it means that more attention is given to characterizing confounding factors (40). The observational studies included some with a before-and-after study design (24, 26, 29). With these studies, it was not possible to exclude the possibility that any changes in outcome may have occurred as a result of other factors that changed over time rather than the intervention.

A reporting bias and publication bias are potential limitations of any systematic review. We used the Newcastle Ottawa Scale (11) to assess the quality of observational studies and the Cochrane risk of bias assessment tool for RCTs (12). The reporting of items by RCTs was inadequate, which made it impossible to assess risk of bias or explore the potential effect on study results. Some aspects of observational studies were consistent with quality components; however, only 2 studies attempted to adjust for confounding variables. Moher et al (41) completed an extensive search to identify studies that examined the general quality of reporting of systematic reviews in nutrition but showed that few such studies exist. Issues highlighted were difficulties in sorting through numerous endpoints and variations in the intervention, study design, and study duration, which were similar to the difficulties encountered in this review.

In addition, there was considerably heterogeneity between included studies in terms of the nature of the intervention and exact groups of preterm infants studied. The composition of PN present. We have reported RCTs and observational studies separately because confounding and selection bias often distort findings of observational studies. The size of observational studies ranged from 36 to 440 participants. However, smaller sample sizes can be preferable if it means that more attention is given to characterizing confounding factors (40). The observational studies included some with a before-and-after study design (24, 26, 29). With these studies, it was not possible to exclude the possibility that any changes in outcome may have occurred as a result of other factors that changed over time rather than the intervention.

A reporting bias and publication bias are potential limitations of any systematic review. We used the Newcastle Ottawa Scale (11) to assess the quality of observational studies and the Cochrane risk of bias assessment tool for RCTs (12). The reporting of items by RCTs was inadequate, which made it impossible to assess risk of bias or explore the potential effect on study results. Some aspects of observational studies were consistent with quality components; however, only 2 studies attempted to adjust for confounding variables. Moher et al (41) completed an extensive search to identify studies that examined the general quality of reporting of systematic reviews in nutrition but showed that few such studies exist. Issues highlighted were difficulties in sorting through numerous endpoints and variations in the intervention, study design, and study duration, which were similar to the difficulties encountered in this review.

In addition, there was considerably heterogeneity between included studies in terms of the nature of the intervention and exact groups of preterm infants studied. The composition of PN present. We have reported RCTs and observational studies separately because confounding and selection bias often distort findings of observational studies. The size of observational studies ranged from 36 to 440 participants. However, smaller sample sizes can be preferable if it means that more attention is given to characterizing confounding factors (40). The observational studies included some with a before-and-after study design (24, 26, 29). With these studies, it was not possible to exclude the possibility that any changes in outcome may have occurred as a result of other factors that changed over time rather than the intervention.

A reporting bias and publication bias are potential limitations of any systematic review. We used the Newcastle Ottawa Scale (11) to assess the quality of observational studies and the Cochrane risk of bias assessment tool for RCTs (12). The reporting of items by RCTs was inadequate, which made it impossible to assess risk of bias or explore the potential effect on study results. Some aspects of observational studies were consistent with quality components; however, only 2 studies attempted to adjust for confounding variables. Moher et al (41) completed an extensive search to identify studies that examined the general quality of reporting of systematic reviews in nutrition but showed that few such studies exist. Issues highlighted were difficulties in sorting through numerous endpoints and variations in the intervention, study design, and study duration, which were similar to the difficulties encountered in this review.

In addition, there was considerably heterogeneity between included studies in terms of the nature of the intervention and exact groups of preterm infants studied. The composition of PN present. We have reported RCTs and observational studies separately because confounding and selection bias often distort findings of observational studies. The size of observational studies ranged from 36 to 440 participants. However, smaller sample sizes can be preferable if it means that more attention is given to characterizing confounding factors (40). The observational studies included some with a before-and-after study design (24, 26, 29). With these studies, it was not possible to exclude the possibility that any changes in outcome may have occurred as a result of other factors that changed over time rather than the intervention.

A reporting bias and publication bias are potential limitations of any systematic review. We used the Newcastle Ottawa Scale (11) to assess the quality of observational studies and the Cochrane risk of bias assessment tool for RCTs (12). The reporting of items by RCTs was inadequate, which made it impossible to assess risk of bias or explore the potential effect on study results. Some aspects of observational studies were consistent with quality components; however, only 2 studies attempted to adjust for confounding variables. Moher et al (41) completed an extensive search to identify studies that examined the general quality of reporting of systematic reviews in nutrition but showed that few such studies exist. Issues highlighted were difficulties in sorting through numerous endpoints and variations in the intervention, study design, and study duration, which were similar to the difficulties encountered in this review.
was extremely variable, and there was some degree of overlap in timing between early and late groups. There was also a fairly broad range of gestational ages and birth weights within each group. Although it remains reasonable to compare the early and late groups, it is important to consider these limitations when reviewing the results in relation to practice, particularly when different compositions of PN and different populations of preterm infants are considered.

There was an inconsistency in reported growth outcomes and variation in assessment time points. Although short-term growth outcomes, such as the maximum percentage weight loss, were fairly consistently reported, there was considerable variability in the reporting of long-term growth outcomes, such as weight, length, and head circumference at discharge or beyond. This limited the potential for a meta-analysis on long-term growth. In relation to this, most studies reported growth outcomes as absolute values. The use of SD scores for weight, length, or head circumference at birth and at selected endpoints would allow more-objective comparisons to be made between groups. Several studies reported binary growth outcomes, and it may be more appropriate to measure an intervention effect in terms of whether it reduces the proportion of babies who are less than a certain percentile rather than detecting a difference in average values. However, both the third and 10th percentiles were used and at several different study time points, which made multiplicity an issue. It was not stated in articles whether these percentiles and time points were chosen pre hoc or post hoc.

The disparate nature of measures reported for growth and nutritional outcomes in neonatal research shows the need for defined core outcome measures in this area to facilitate comparison and enable the synthesis of study data in future systematic reviews. Following the Core Outcome Measures in Effectiveness Trials initiative (42), a CORE set of agreed standardized outcomes could be developed.

In conclusion, the results of this review, although subject to some limitations, provide evidence that the use of early PN reduces early weight loss and the time to return to birth weight. The review did not provide any evidence that early PN increases morbidity or mortality. There is also some evidence that growth outcomes at discharge or 36 wk PMA are improved, although these data were mainly from observational studies or single RCTs. Because of these findings and because the use of early PN is becoming part of the normal practice of many clinicians, it is unlikely that there is sufficient clinical equipoise to conduct additional RCTs that address the short-term effects of early compared with late PN. However, additional research efforts should be directed toward ascertaining the optimum composition and approach for the use of PN in preterm infants and understanding the longer-term effects of early nutrition on growth and neurodevelopment. In this context, neonatal research would benefit from the development of a set of core outcome measures to facilitate the comparison of data between studies.

The authors’ responsibilities were as follows—HEM and VRC: identified and extracted data initially; MJJ reviewed disagreements regarding the identification and extraction of data and repeated searches and identified articles in response to initial reviewer comments; MJJ and VRC: repeated the statistical analysis after the repeated search; HEM, MJJ, VRC.

**TABLE 5.** Compound plot of morbidity outcomes. Point estimates and bars represent pooled Peto’s ORs (95% CIs) for the comparison between early and late parenteral nutrition. CLD, chronic lung disease; IVH, intraventricular hemorrhage; NEC, necrotizing enterocolitis; RCT, randomized controlled trial.
and AAL: designed the research, wrote the article, read and approved the final manuscript, and had primary responsibility for the final content of the article. None of the authors had a conflict of interest.

REFERENCES


