Overview of Cochrane systematic reviews of early enteral feeding strategies for improving growth and reducing necrotising enterocolitis in very preterm or very low birth weight infants

Dr Verena Walsh BSc, MBChB, MRCPCH

Thesis submitted for MSc in Medical Sciences (by thesis)

The University of Hull and the University of York

Hull York Medical School

March 2021

1 Abstract

Background:

Balancing nutrition for healthy growth and development of preterm infants while avoiding necrotising enterocolitis (NEC) and consequences of prolonged parenteral nutrition remains one of the greatest challenges in neonatal care. Early enteral nutrition is important for gastrointestinal maturation but is feared to create a physiological strain contributing to NEC. Despite vast research into best practice, large variations in feeding practices continue to exist.

Methods:

Using Cochrane methodology, we conducted a search for Cochrane systematic reviews evaluating the effects of early enteral feeding strategies on NEC and growth in very preterm or very low birth weight (VLBW) infants. Eligible reviews were assessed for certainty of the evidence and quality of the systematic review methodology, following which the certainty of the evidence base was summarised.

Results:

Review quality was generally high, however some methodological areas for improvement were highlighted. Of the thirty completed eligible reviews found, no interventions provided high certainty or probable evidence for improvement in short term growth while reducing or having no effect on NEC. The evidence base for long term growth was sparse. Only donor breast milk compared with formula was shown to be a promising intervention for reducing NEC, but probably harmful for short term growth. Probable evidence for no effect of different rates of feed advancement was concluded for both outcomes.

Conclusion:

Evidence for effects of many common early enteral feeding interventions on growth and NEC is uncertain. Investigation of NEC requires large improvements focussing on precision of the effect estimate and reducing risk of performance and detection bias. Evidence for growth would benefit from consistency in measurements between trials. Future trials should focus on infants at greatest risk of necrotising enterocolitis and subsequent feeding of infants affected by this early in their stay.

2 List of contents

3 List of Tables

4 List of Figures

5 Acknowledgements

I would like to thank my supervisor, Professor William McGuire, for his guidance and support throughout my work on this overview, including contribution as second overview author to the search, data extraction and evaluation of quality of the included trials. Through many constructive discussions, you have deepened my understanding both in neonatal evidence base and the critical review of the evidence through systematic review and allowed me to develop the curiosity and enthusiasm to explore this through future work.

Thank you to Jennifer Brown, for all you taught me at the start of my journey with Cochrane Neonatal, your contribution to my work as co-supervisor through the Thesis Advisory Panel, and always being there when I needed guidance and support with the review methodology. Your support all the way through the project has been invaluable. I would also like to thank Dr Sam Oddie, for his guidance and advice as Thesis Advisory Panel chair and his contribution as third review author assessment of reviews affected by dual authorship.

Finally, I would also like to thank my partner, my family and friends for supporting me through the more challenging parts of the journey and encouraging me to remain motivated and enthusiastic about my work throughout the project.

6 Declaration

I confirm that this work is original and that if any passage(s) or diagram(s) have been copied from academic papers, books, the internet or any other sources these are clearly identified by the use of quotation marks and the reference(s) is fully cited. I certify that, other than where indicated, this is my own work and does not breach the regulations of HYMS, the University of Hull or the University of York regarding plagiarism or academic conduct in examinations. I have read the HYMS Code of Practice on Academic Misconduct, and state that this piece of work is my own and does not contain any unacknowledged work from any other sources.

7 Background

7.1 Description of the condition

Extremely preterm and very preterm infants are defined as an infant born before 28 weeks and 28-31 gestational weeks, respectively. Low, very low and extremely low birthweight are defined as 1500 to 2499g, 1000-1499g and 999g or less, respectively[1] and result from shortened gestation, intrauterine growth restriction, or both. After major improvement in survival of infants born preterm, the requirement for research and management has shifted to optimising their disability-free survival. One of the main elements required for this is adequate nutrition to allow optimal growth and development.

Rate of appropriate growth of preterm infants is assumed to follow a similar trajectory as a foetus of equal postconceptional age and recommendations for their nutritional requirements have been established by international consensus[2]. Most preterm infants, however, do not achieve this growth and are growth restricted on discharge from the neonatal unit[3]. Preterm, and especially growth restricted infants have not developed a buffer in nutrient reserves which term infants are said to benefit from while establishing feeding in the first days of life, therefore require calories early to fuel their metabolic demands. Very preterm infants develop in an extrauterine environment during the third trimester, a critical period for brain development[4,5], fragile period of saccular lung development[6]; and rapid period of intrauterine weight gain[2]. Undernutrition during this time is shown to exert a detrimental effect on the infants both in short term clinically important outcomes including infection and necrotising enterocolitis, as well as affecting long term growth and neurodevelopmental outcomes[5,7].

The ability for infants born before 34 weeks gestation to establish breast or bottle feeding with coordination of sucking, swallowing and breathing is compromised by neurological immaturity and respiratory distress[8]. In addition, preterm infants may not able to tolerate the increases in enteral feed volumes required to match nutritional requirements due to delayed gastric emptying and intestinal peristalsis[8].

A significant barrier within clinical practice against establishing enteral nutritional intake is the fear of necrotising enterocolitis (NEC), its potential mortality and

serious long term neurodevelopmental consequences[9]. Infants who have developed NEC, when compared with infants who have not developed NEC, have greater mortality and disability, slower growth, lower nutrient intake, experience more infections and have a longer intensive care and hospital stay[9]. The multifactorial pathogenesis of NEC is believed to include a compromised physiological and immunological intestinal defence, and abnormal vascular development and tone[10]. Introducing enteral feeds to the immature gastrointestinal tract of the very low birthweight (VLBW) or very preterm is feared to create a physiological strain which can lead to development of this acute intestinal necrosis syndrome[10]. Avoiding enteral nutrition involves parenteral nutrition, which comes with its own challenges, including avoiding bloodstream or line infection from the required presence of central venous access, and the conundrum of what nutrition is best to give. An important aim, therefore, of preterm care is a balance of providing nutrition for healthy growth and development of important organ systems including the gastrointestinal system but avoiding development of necrotising enterocolitis and consequences of prolonged parenteral nutrition.

Outcomes traditionally considered important are acknowledged to have been chosen by health care professionals, including researchers and clinicians, and fixed in routine practice without further evaluation of their importance to parents and patients[11]. Therefore there has been a recent emphasis on involving parents and patients in determining important outcomes and designing research trials[12,13]. The high clinical importance of the outcome of NEC is agreed, including consensus among parents, healthcare providers and researchers[14]. Monitoring of growth through anthropometric measurements continues to be an important consideration in routine clinical practice as part of ongoing nutritional assessment[15] and aiming to improve neurodevelopment[7]. Yet what constitutes a clinically meaningful difference between interventions for each outcome has not been clearly defined.

A statistically significant difference may not necessarily imply a clinically meaningful difference. What is recognised as a clinically meaningful effect size may vary between individuals or groups of individuals (clinicians, parents, service providers), between settings (such as those resource-limited compared with resource rich) and depending on the nature of the intervention (including effect on other outcomes)[16]. For example, a low cost, easily delivered and low risk intervention,

such as breastfeeding would tolerate a much higher number needed to treat for benefit (NNTB) than an intervention with high risk, high cost, and high toxicity or workload. Therefore what constitutes a clinically meaningful effect size is variable and will need to be interpreted for each intervention individually in the context a policy decision is applied to. Cochrane reviews authors may report the number needed to treat for an additional beneficial (NNTB) or harmful (NNTH) outcome for dichotomous outcomes, or indicate the minimal important difference (MID) for continuous outcomes to guide decisions[16].

7.2 Description of the interventions and how the interventions might work

Research over recent decades has provided a wealth of knowledge to inform how to best provide nutrition to very preterm infants to allow good growth and development but avoiding infection and NEC. Yet large variations in feeding practices both within countries and globally continue[17].

'Early' refers to feeding strategies during initiation and establishment of enteral feeds, typically during the first few weeks after birth. Early enteral feeding interventions for preterm infants can be grouped into the categories of 'what is given', 'when it is given' and 'how it is given'[18]. We have additionally considered a further category of 'adjunctive strategies' which include early strategies that are not directly involved with the delivery of the nutrition but aim to improve the ability of preterm infants to achieve good nutrient intake, absorption and passage through the bowel; or by early implementation shorten the time for infants to develop the means required to improve nutrient intake and absorption.

7.2.1 What to feed

7.2.1.1 Maternal milk

Maternal milk is regarded as the enteral feed of choice, providing both immunological and nutritional content. Its composition changes with changing infant requirements, producing colostrum in the initial days after birth which containing a higher concentration of protein; immuno-nutrients such as secretory Immunoglobulin A (IgA), lactoferrin, cytokines, enzymes and growth factors; and lower concentrations of lactose[19–21]. After a few days, maternal milk composition changes with the transitional milk gradually increasing in nutritional content and becoming mature between five days to six weeks postpartum[21].

Preterm milk differs in macronutrient composition, providing a relatively higher concentrations of protein, fat, lactose and energy compared with milk from a mother of a term infant[22]. Composition of human milk also differs between the initial foremilk, and the higher fat content of the last milk of the feed, hindmilk[21]. The true energy, protein and fat intake from own mothers' milk is variable and often overestimated, and at a feeding volume of 150ml/kg/day maternal breast milk alone is believed to be insufficient to meet energy and nutritional requirements of preterm infants[23]. In clinical practice, the nutritional intake is increased either by supplementing the breastmilk with milk fortifiers or giving the milk in higher volumes. Exclusive maternal breastmilk provision relies on maternal supply, where adequate provision is often more challenging for mothers of non-nursing preterm mothers and can lead to delays in starting and level of this beneficial component within their diet[24].

7.2.1.2 Donor breast milk

When maternal milk is not available or insufficient, donor breast milk (DBM) can be given[17]. This milk is generally pooled and pasteurised milk donated and therefore is primarily obtained from mothers of term infants, providing milk surplus to their own infant's requirements and later in lactation. Neonatal units that use donor breastmilk have specific criteria for their use and therefore donor milk is often used in extremely and very preterm infants and then discontinued once infants have achieved a defined weight or corrected gestational age. DBM is thought to provide an immunonutritional advantage and reduce feed intolerance, however the nutritional

content is regarded as insufficient to support adequate growth of the preterm infant. and requiring fortification to meet their requirements. Observational studies have reported concerns of slower initial weight gain[25,26] and possible increased rate of early cognitive delays compared with infants fed their own mother's milk or formula[26].

Unlike maternal breast milk, the composition of DBM is not unique to the infant in macronutrient and immunonutrient composition. Milk banking practices vary between countries, including in screening and pasteurisation. Due to risk of infection and transmission of harmful substances through breastmilk, in several countries breastmilk donors are screened for alcohol, nicotine and drug use, and milk is pasteurised to inactivate or kill viruses and bacteria[27–29]. The pasteurisation process lowers the immunonutritional content of the milk and there is concern about pasteurisation altering how well donor milk can be digested and absorbed[27].

Sparse evidence from an early trial[30,31] suggested that unfortified donor breast milk when compared with preterm formula was associated with lower mean blood pressure and a more favourable plasma lipid profile. It is questionable if bovine fortification of the donor milk which aims to lead to more rapid growth would reduce or negate such a benefit[27].

7.2.1.3 Formula milk

Where human breast milk is unavailable, insufficient, or by parental choice, cow's milk based formulas may be used for feeding preterm infants either as a sole diet or as a supplement to human breast milk[17].

Formulas can be grouped as standard formulas and preterm formulas. Preterm formulas have been designed with higher macronutrient and micronutrient content to respond to a recognised increased requirement to support growth and development of a preterm infant[32]. The composition of standard term formula is based on mature breast milk, providing energy at about 67-70 kCal/100ml, protein at 1.4-1.7 g/100mL, calcium at around 50mg/100mL and phosphate at 30mg/100mL. In comparison, preterm formulas provide a higher content of these nutrients, targeting an energy content of between 75 to 80 kCal/100ml and protein of 2 to 2.4 g/100mL. These formulas also provide variably higher concentrations of minerals, most commonly a higher calcium and phosphate to support bone mineralisation and

prevent development of osteopenia of prematurity, and sodium and potassium to counter for the increased renal losses of the preterm kidneys[32].

Cow's milk protein provided within formula is a ratio, mostly between whey protein and casein protein. Casein coagulates in the acidic environment of the stomach and therefore is slower to digest and empty through the pylorus[32–34]. Whey protein is more soluble, allowing faster gastric transit and facilitating faster amino acid absorption. The ratio of these proteins in human milk is variable and changes with the stage of lactation, with the 90:10 whey to casein ratio seen initially reducing to 60:40 in mature milk[35]. The ratio of whey to casein within infant formulas change to reflect this transition. As concerns with preterm infant feeding include their tolerance to feed and impact on gastrointestinal function including absorption, preterm formulas have been designed with a higher whey to casein ratio[32].

Infant formulas can also differ to the degree to which the proteins are broken down. These are broadly classified as ‘intact’, ‘extensively hydrolysed’ and ‘partially hydrolysed’ formulas. Hydrolysed formulas were initially developed for infants with cow’s milk allergy or intolerance, but as the cow’s milk derived protein is broken down to shorter peptide chains, it is also believed to be better tolerated by the immature preterm gut, allowing a faster intestinal transit time and achieving full enteral feeds earlier [36,37]. Uncertainty exists, however, regarding whether these formulas provide adequate nutrition, bioavailability of protein and important minerals, and benefit growth and neurodevelopment[37].

Lactose and milk oligosaccharides form the primary source of carbohydrates in human milk[21]. Lactose also forms the major carbohydrate source in term formulas, however preterm formulas have a combination of lactose and either sucrose or low osmolar glucose polymers[32]. Breakdown of sucrose produces glucose and fructose which are easily transported across the intestinal mucosa, but lactose produces galactose instead of fructose which is more efficient for glycogen production in the liver. Lactose is thought to be beneficial for a healthier microbiome and development of colonic butyrate involved in colonic maturation and repair[32].

Preterm infants are developmentally lactase deficient and have been shown to have incomplete digestion of lactose[38]. Maternal milk increases in lactose as lactation progresses from the low concentrations found in colostrum[21,39]. Lactose intolerance

in term infants can be managed with reduced lactose or lactose-free formula, and treating preterm feeds with the addition of lactase has also been explored[40]. Colonic fermentation of undigested carbohydrates leads to reduced pH and production of gas, and there has been concern of its relevance to NEC[41]. Yet, the reduced pH also promotes beneficial gastrointestinal flora to develop[38]. Lactase activity increases with gestational age and with earlier enteral feeds and human milk fed infants in whom lactose is the only carbohydrate have a higher lactase activity[39,42]. A more recent study on preterm pigs, showed a higher incidence and severity of NEC in preterm piglets enterally fed maltodextrose-based feeds than lactose-based feeds[43,44], raising the importance of this carbohydrate source in preterm formulas.

The lipid content of formula is a significant energy source, provides an essential component of the cell membranes, and is important for development of the brain. Neonates have a transient exocrine pancreatic insufficiency with reduced amylase and lipase secretion[45]. Unlike human milk, which contains lipase which can compensate for the infant's deficiency, most formula does not. As an aim to improve fat absorption, some formulas have been adapted to provide shorter lipid chains that are more easily digested and absorbed. Preterm infants fed formula have been shown to have reduced absorption of fatty acids with increasing chain length[45]. The absorption, transport and utilisation of long chain triglycerides (LCT) require a complicated process. LCTs need emulisification prior to intraluminal digestion by pancreatic lipase and mixing with bile salts to form mixed micelles which can diffuse into the enterocytes. Their absorption can therefore be affected by the reduced levels of pancreatic lipase and bile salts in early life[46]. Medium chain triglycerides (MCT) have a shorter fatty acid chain length of 6-12 carbons. They are more water soluble and are more rapidly hydrolysed by gastric, lingual and pancreatic lipases and absorbed through the gastric and intestinal mucosa. Unlike the long-chain triglycerides, MCTs do not need to travel as chylomicrons in the lymphatic system and can be transported in the blood bound to serum albumin[46]. They also do not need conversion into acyl-carnitines for diffusion across the mitochondrial membrane[46]. With easier absorption, transport and utilisation of MCTs, formulas with a higher proportion of fats in this form have the theoretical potential to improve feed tolerance and lipid absorption in the neonate, therefore improve growth and development despite a lower caloric density. Yet the theoretical

benefit needs to be reflected in demonstrable effect on clinical outcomes in infants, which is still felt to be uncertain[47]. Furthermore, long-chain polyunsaturated fatty acids, such as arachidonic acid and docosahexanoic acid, are important components of human milk and beneficial for visual and central nervous system development. These are not present in bovine milk, therefore are added to some formulas. A retrospective study by Martin *et al.*[45] found a reduced absorption of docohexanoic acid in formula fed infants compared with breastfed infants, hypothesising this difference to be due to reduced ability for the formula fed infants to digest these long chained lipids without additional lipase in formula.

7.2.1.4 Breast milk fortification

The aim of milk fortification is to provide the nutritional requirements for equivalent foetal growth in volumes recommended for feeding. Despite the consensus of preference for breastmilk, exclusive feeding with only mother's own breastmilk and donor breast milk has been associated with slow growth in very low birth weight infants and therefore the European Milk Bank Association Working Group on Human Milk Fortification state that human milk should be supplemented, especially with protein, calcium and phosphate, for which preterm infants have a higher demand[48].

Available fortifiers are derived from either cow, donkey or human breast milk concentrated to allow additional calories and nutrients, without the additional volumes otherwise required. Multi-nutrient fortification of breastmilk aims to achieve the infants' nutritional requirements by simultaneously increasing a variety of nutrients, including carbohydrates and protein, calcium, phosphate, electrolytes, minerals and vitamins[49]. The greatest concern in fortification is meeting the infants' protein requirements to allow good postnatal growth and neurodevelopment. A recent prospective observational study[19] comparing the macronutrient composition of preterm milk in the first two months of lactation found that almost half of the preterm milk samples did not achieve an energy content of 67 kcal/100ml, and over three quarters (78.9%) had a protein content of less than 1.5 g/100ml. After the first week of lactation, carbohydrate and fat content remained stable, however protein content declined until the third week of lactation, after which it remained stable.

The most common approach for fortification, 'standard human milk fortification', is by adding a fixed amount of fortifier to a specified volume of the expressed mother's or donated breastmilk. This provides a standard extra amount of protein, fats, carbohydrates and micronutrients regardless of variable milk content or infant growth and requirements, assuming a fixed amount of unfortified breastmilk protein content of 1.5g/dL[50]. This approach does not consider variation in milk both from the same mother over the course of lactation and between different mothers[50], nor does it consider the variability in infant nutritional requirements[48]. Standard fortification has been shown to have the greatest discrepancy between required and actual nutritional content in the protein content, leading to protein undernutrition and therefore potentially suboptimal growth[48].

Alternative strategies for fortification have been proposed, which are termed 'individualised human milk fortification'. These include adjusting the fortification given depending on either the levels of macronutrients from analysis of the breastmilk the infant receives (Targeted Fortification) or according to the blood urea nitrogen measured in the blood (Adjustable Fortification)[48,50].

The concept of 'targeted human milk fortification' is that the infants' requirements vary and therefore fortification is dosed to achieve recommended target intake according to population based recommendations for post conceptual age and according to analysis of the human milk composition[51,52]. This approach requires frequent testing of milk due to variability between milk. This involves additional time and ability of staff to interpret and correctly amend fat and protein content, introducing potential of human error, and the expensive purchase of equipment for the frequent testing of the human milk given[51]. In a survey of neonatal care units in high income countries, performed in 2010[17], routine human milk analysis was most common in Scandinavian intensive care units, occurring only in one unit studied outside this region. The infra-red analysers used were originally developed for the dairy industry and with calibration can indicate protein and fat concentration, however is unable to inform about lactose and therefore energy content[48]. In a matched pair analysis of infants supplemented by targeted fortification with infants with standard fortification[53] there was a linear relationship of milk volume to weight gain seen in the infants who were target fortified, that was not seen in infants with routine supplementation. The mean weight gain observed was similar between target

fortified and matched controls, however the results of this were likely confounded by an observed reduction in feeding volumes in infants who were target fortified reported to be due to bedside staff reducing feeding volumes due to an unusually high weight gain[53].

'Adjustable fortification' uses measurements of blood urea nitrogen to determine the need for initiating, increasing, and reducing protein supplementation[54]. It uses blood urea nitrogen, a routinely used investigation, as a surrogate marker for assessing if the level of protein in an infant's nutrition matches its requirements. The fortification begins as standard fortification with a multi-nutrient fortifier. The blood urea nitrogen level is used to assess nutritional protein adequacy, twice weekly. Where the infants' renal function is normal, a low and high urea is said to reflect inadequate and excessive protein intakes, respectively[54]. In the protocol by Arslanoglu *et al.*[54], these thresholds are stated as <9 mg/dl at which point the protein content is increased, and >14 mg/dl at which point the level of protein fortification is reduced. This method was felt to be less labour intensive than milk analysis and is directed more at the infant according to its metabolic response to protein, avoiding excessive protein intake and not making assumptions on an average infant's protein requirements[49].

Yet, fortification of human milk does not come without challenges and concerns. Increasing protein in the diet is one of the main aims of fortification, however when comparing discrepancies between assumed and actual protein content, it has been shown that actual protein content is consistently significantly lower in both standard and adjustable fortification strategies[55].

Whole protein bovine milk fortifier is the most widely used human milk fortifier, however other avenues are being explored. Fortifier derived from human breast milk has been developed attempting to achieve similar benefits of an exclusively human milk-based diet, with reduction of NEC, feeding intolerance and morbidity. However, human milk-based fortifier requires concentration of a large volume of human milk, the practicality, and ethics of sourcing of which may prove challenging if more widely used. Manufacturers have also produced fortification using a hydrolysed protein source, to improve feed tolerance similarly to differing protein contents in formula milk[48].

Concentrating nutrients into defined volumes of milk increases the osmolality of their nutrition and therefore concerns exist regarding its effect on the preterm gut. Osmolality has been thought to be linked to feed intolerance and the pathogenesis of necrotising enterocolitis. Many clinicians balance the theoretical risk of increased NEC and feed intolerance against the benefit of fortification, delaying introduction of fortification until a specific intake is met. Osmolality can affect gastrointestinal motility, therefore affecting tolerance to feeds and transit of higher volumes of bolus feeds. There have also been concerns of cases of bowel obstruction from lactobezoar after introduction of fortifier in preterm infants of low birth weight[56] . Yet, no evidence for a causal relationship between hyperosmolar feeds and necrotising enterocolitis has been found[57], and a recent systematic review of animal and human studies showed no consistent evidence that feed osmolalities between 300-500 mOsm/kg are associated with gastrointestinal symptoms in neonates[58].

7.2.1.5 Target volume

Once feeds have been established, the majority of neonatal units target an enteral feeding volume of 140-180 ml/kg/day[17]. Current recommendations advise a feeding volume between 135-200ml/kg/day[2]. Unfortified breastmilk (which is assumed to have 1.5g protein/100ml and 67 kcal/100ml) and standard formula would provide approximately 100-126 kcal/kg/day, and therefore nutrition for preterm infants is routinely concentrated into lower volumes, either as nutrient-enriched formula or as fortified breastmilk. An alternative approach to improving the enteral nutrition provided is to increase the volume of feed given. Assuming standard protein and calorie concentrations of breastmilk and term formula, a volume of 200ml/kg/day would provide 134 kcal/kg/day and 3g protein/kg/day. Although this meets the recommended calorie requirement in preterm infants, volumes in excess of this are needed to reach the protein requirement for preterm infants of 3.5-4.5 g/kg/day protein[2]. This approach, if safely tolerated, could allow an exclusive human milk diet to be preserved. It may also be a cheaper and more accessible alternative avoiding the additional cost of milk fortifier and nutrient enriched preterm formula, which may be especially beneficial in low- and middle-income countries. Yet, even in high-income countries a higher feeding volume without breastmilk fortification could potentially improve growth and development by continuing to meet nutritional requirements where there is hesitancy to start or continue milk fortification such as

following need for gastrointestinal surgery or obstruction felt to be related to fortification. Nevertheless, there are concerns that high volume feeds may cause complications related to fluid overload such as hyponatraemia, patent ductus arteriosus and bronchopulmonary dysplasia, as well as increasing feed intolerance and necrotising enterocolitis[59].

7.2.2 When to feed

7.2.2.1 Introducing and advancing enteral feeds

The immature preterm gut must quickly adapt to being able to mobilise milk through the gut and absorb the required nutrients. Hesitancy with initiating and advancing feeds is fuelled by concerns about feeding intolerance and necrotising enterocolitis[60]. Yet, this careful approach to introducing enteral nutrition may also adversely affect these outcomes by influencing development of the microbiome and maturation of gastrointestinal motility and function.

Minimal enteral nutrition (trophic feeding/gut priming/hypocaloric feeding) starts with nutritionally insignificant volumes of milk to stimulate postnatal development of the immature gastrointestinal tract of the preterm infant[61]. This exposes the preterm gut to milk while parenteral nutrition is being given to provide the infant's nutritional needs[62,63]. Exposure to even low quantities of milk induce gastrointestinal hormone secretion stimulating gastrointestinal growth, function and motility and alter relative gastrointestinal disaccharidase activity to a higher lactase to sucrase ratio[62,63]. Minimal enteral nutrition is believed to improve tolerance to milk, postnatal growth and reduce systemic sepsis and length of hospital stay[61].

Yet delaying introduction of nutritionally significant feeds, prolongs the time the infant is reliant on a parenteral mode of nutrition. Parenteral nutrition increases the risk of complications including line-related infection, parenteral nutrition associated liver disease (PNALD) and metabolic bone disease with increasing duration. It is therefore beneficial for the infant to achieve and tolerate full enteral feeds as early as possible. To achieve full feeding volumes earlier, enteral feeds can be increased to provide nutritionally significant volumes earlier and can be increased by larger increments per day. Yet there are concerns about how well the immature preterm gut can accept the associated functional demands, with previous evidence from a case-control study by McKeown *et al.*[64] showing higher incidence of necrotising

enterocolitis in infants fed earlier, with greater increments of volume advancement and with higher volumes.

7.2.2.2 Time to pause advancement or holding feeds

Feed intolerance is a marker which may influence a clinical decision about withholding, reducing or deciding not to advance feeds. Decisions are subjective and variable. Clinical signs in determining feed intolerance include gastric residual volume, abdominal distension, and vomiting.

One of the main measures of feed intolerance used is measuring gastric residuals. If the infant has increased gastric residuals, then decisions such as reducing or pausing feeds may be made. When assessing gastric residuals, decisions are also made with changes in colour of the milk to green or blood-stained. While increased gastric residuals may increase with impending or current NEC, these cases are accompanied by more striking local and systemic features such as a tense abdomen with abdominal wall discolouration, absent bowel sounds, acidosis, and temperature and cardiorespiratory instability[65]. Meanwhile increased gastric residual volume and assessing colour may be overcautious and harmful in preterm infants both from trauma and from preventing an otherwise healthy infant from achieving adequate nutrition. When feeds are withheld, the infant is often placed on IV fluids therefore receiving glucose and sodium, but not receiving the protein and lipid required for growth. If the feed is withheld for longer, a decision may be made for a central venous access for parenteral nutrition, however there may also be a delay in this decision being made. A recent trial[66] comparing extremely preterm infants who did not have their gastric residuals routinely measured, with infants who had their gastric residuals measured pre-feeds, showed that infants in whom the gastric residuals were not measured received more enteral nutrition, had improved weight gain and left hospital sooner than infants in whom gastric residuals were measured. The question is raised whether not utilising this marker of feeding intolerance and gastrointestinal dysfunction would affect outcomes by delayed diagnosis and increased risk and severity of necrotising enterocolitis when problems are recognised.

7.2.3 How to feed

7.2.3.1 Intermittent feed frequency and continuous feeds

Milk can be given either through a continuous slow administration of feeds, or by giving an equivalent volume more rapidly over a shorter defined time, also known as 'bolus' or 'intermittent gavage' feeding. The feed frequency can range, most often between every one to four hours.

An intermittent feed can be given either by allowing the milk to enter by gravity, or by delivering the volume through compression of a syringe. Bolus feeding reflects the physiological feeding pattern of term infants and adults, with the duration corresponding to the duration of a breast or bottle feed. This modality demands the ability for preterm infants to adapt from a continuous in utero supply to that of maintaining metabolic homeostasis despite a fluctuating enteral supply. Mizumoto *et al.*[67] demonstrated initial large variability in serum glucose to severe hyperglycaemic and hypoglycaemic levels in preterm infants in response to intermittent bolus feedings. This improved over a period of weeks with acquired ability to maintain glucose homeostasis[67]. The clinical significance of these fluctuations remain undetermined[67]. In term infants, exposure to their first feeds cause a rise in glucose and significant surges in gastrointestinal related hormones. In preterm infants this cyclical response does not initially present, but develops after regular intermittent milk feeds[68,69]. This cyclical secretion of gastrointestinal hormones and changes in metabolites resulting from the fluctuating feeding fasting pattern stimulates is believed to be important for the development of the gastrointestinal tract and stimulates protein accretion[63,69–71]. During continuous feeds, these hormones and metabolites remain at a steady state[69]. Due to the slow infusion of the milk, there are concerns that continuous feeding may lead to large changes in fat concentration given to the infant[72]. Similarly to the continuous administration of parenteral feeds, continuous enteral feeds have been associated with extrahepatic biliary stasis which were shown to resolve following a bolus feed[73].

A concern with intermittent feeding is through the physical impact of the rapid increase in volume. By slow administration of the continuous feed, clinicians aim to improve gastro-oesphageal reflux (GOR) and feeding tolerance by limiting gastric distension, reducing pressure on the lower oesophageal sphincter, and allowing

gastric emptying to be faster and more complete[74]. Improved feeding tolerance allows improved absorption of nutrients and therefore potential for improved growth. Yet, GOR is very common in preterm infants, and the benefit of management is debated[75]. In addition to the potential impact on cardiorespiratory events associated with GOR, gastric distension following a bolus feed reduces the infant's functional residual capacity which has the potential to result in hypoxaemia and desaturation associated with feeding[76,77]. Studies investigating the impact of cardiorespiratory events, however have not reflected this, with Corvaglia *et al.*[78] finding that episodes of prolonged apnoeas and apnoea-related hypoxic episodes was greater with continuous feeds. Yet this study[78] had the confounding difference of the NG tubes removed between bolus feeding but not continuous feeding. Indeed, the constant presence of contents in the stomach in continuous feeding has the potential to predispose to more reflux and aspiration.

The splanchnic blood flow is compromised in infants with intrauterine growth restriction, and increased superior mesenteric artery blood flow and reduced resistance index have been shown to correlate with early tolerance of enteral feeding in preterm infants[79–81]. The length of feed interval has been shown to affect splanchnic perfusion, with a longer interval resulting in a larger postprandial response[82]. This study[82] also showed that when the splanchnic blood flow of infants fed hourly compared to infants fed 3 hourly or more, these infants had a persistent hyperaemia. Bolus feeding is followed by increase splanchnic perfusion to improve oxygenation. Dani *et al.*[81] suggest that the gut of small for gestational age infants has a low intestinal oxygenation due to prenatal haemodynamic compromise which is less able to meet the additional metabolic demand of feeding and that a lower oxygen requirement of continuous feeding may in these cases limit the potential for development of necrotising enterocolitis from hypoxic-ischaemic damage.

Feed frequency is a further extension of the argument between continuous and intermittent feeds. Shorter intervals allow a smaller volume per feed which potentially are better tolerated, with less gastric distension and pressure on the lower oesophageal sphincter, and more complete gastric emptying. If better tolerated it may allow a higher feed volume to be achieved per day. With more frequent feeds, however, the infant has less time to settle between disturbances from healthcare interventions. Similarly, what effect the increased more persistent shunting of blood

and hyperaemia associated with shorter intervals has on the gut and its motility is not known.

7.2.3.2 Feeding tube placement

As the suck and swallow reflexes have not fully developed in very preterm or very low birth weight infants, in clinical practice enteral nutrition is given by a nasogastric (NG) or orogastric (OG) tube. Placement of an NG tube is felt to be easier to secure in place with reduced movement, however its location may interfere with the available airway of the obligate nose breathing infant and therefore increase work of breathing[83]. Although the orogastric route does not obstruct the nasal airway, the increased mobility has the potential for increased mucosal injury and vagal stimulation provoking apnoea and bradycardias[83].

The tube passing via nasal or oral passages can be positioned with the tip sitting in the stomach or in the duodenum/ileum. GOR and vomiting are major concerns in the neonatal unit. If considered problematic, this may lead to treatment decisions including reducing total feeding volumes and reduced feeding intervals. It may also be associated with increased bradycardia and apnoea, and potential milk aspiration with its complications[84]. Clinicians may therefore decide to place the feeding tube more distal than the pylorus and in the duodenum or jejunum. Despite the theoretical advantage of increasing enteral feeds reaching the main site of nutrient absorption, this mode of feeding has its disadvantages. The positioning of the tube is more challenging and is associated with increased imaging[85]. As the milk does not pass through the stomach, this may also reduce the stimulation of gastrointestinal hormone and growth factor secretion, and the antimicrobial protection from exposure to gastric acid[83]. Furthermore, gastric acid secretion following stimulation of gastrin by milk in the small intestine may lead to increased gastric bleeding without neutralisation by milk within the stomach[85].

7.2.3.3 Cup feeding

Cup feeding is the practice of allowing the preterm infant to lap up the milk from a cup. This requires coordination of swallow and breathing, however does not rely on the infant to have a strong suck. The tilt of the cup allows the milk to touch, but not pour into the infant's mouth. The infant uses the tongue to lap up small boluses into

the mouth and through use of the tongue is thought to promote the tongue mobility required for breastfeeding[86]. An observational study described in 1994[86], noted that cup feeding was only possible at a postconceptional age of 30 weeks and that oral feeding in very preterm infants (>28 weeks) was delayed due to the ventilatory assistance required. As respiratory care of preterm infants improves, we need to question if cup feeding is possible to initiate prior to this time. Yet, cup feeding needs the infant to actively feed and can take longer for each feed. This may result in feeding for each infant to be more time intensive for neonatal staff, and with increasing time of active feeding may be more tiring for the infant and therefore impact ability for further feeding. Furthermore, if milk is poured rather than lapped, this has the potential for milk aspiration and its respiratory complications.

7.2.4 Adjunctive strategies

Several adjunctive strategies have been tried to increase success of early enteral feeding of preterm infants. These include commencing pharmacological agents such as prokinetic agents to improve transit of enteral feed through the intestine, and suppositories to facilitate continuous passage through the bowel. Other strategies are started within the early feeding period to improve and achieve earlier successful oral feeding of preterm infants, such as programmes of oral stimulation and non-nutritive sucking.

7.2.4.1 Stimulation of oropharyngeal motor skills

Despite sucking being demonstrated in utero from early in gestation, the sucking demonstrated in very preterm infants is inadequate to allow sufficient milk to be withdrawn and to be coordinated with a safe swallowing mechanism[87]. Non-nutritive sucking occurs in the absence of oral feed, occurring in bursts of a more rapid suck of two sucks/second between brief pauses[87,88]. Infants of an earlier postmenstrual age have more irregular patterns of sucking, which develops as they mature to become more rhythmic, faster and intense with longer duration of sucking and shorter pauses between[87]. Nutritive sucking is more settled and constant with one suck/second and much greater movement of the jaw, tongue, and larynx to achieve nutrient flow[87,88]. A non-nutritive suck can be stimulated in a variety of ways including by a pacifier/dummy, nipple of an empty breast, or sucking of a gloved finger. It is

thought to promote establishment of oral feeding with development of sucking patterns and preventing oral aversion after prolonged tube feeding[89]. It is also believed to encourage growth and development through reducing energy use with more settled breathing and reducing distress with pain or during procedures[87]. The effects of non-nutritive sucking on digestion, gastrointestinal transit and tolerance are not clear, however is claimed to be improved by the neuroendocrine effect of non-nutritive sucking on gastrointestinal hormones by stimulating the vagal nerve[90,91].

7.2.4.2 Pharmacological agents to facilitate gastric transit

The macrolides erythromycin and clarithromycin are used in the clinical setting to improve gastrointestinal motility through its agonistic effect on the motilin receptor. Along with improving feeding tolerance and therefore facilitating more nutrition to be given enterally, it has also been shown to affect parenteral nutrition associated liver disease and sepsis[92]. However, in a review of the literature, Ng *et al.*[92] felt that its action is less effective at lower doses and infant gestation. Concerns continue to exist regarding the association of macrolide use on development of infantile hypertrophic pyloric stenosis [93], and development of resistance to the antibiotic action of the medication[92].

Delayed passage of meconium and infrequent passing of stool is a common problem in preterm infants and may contribute to feed intolerance. To prevent and manage this, rectal suppositories such as glycerin laxatives can be given to encourage passage of stool through its irritation and hyperosmolar effects, and therefore facilitating more proximal gastrointestinal contents to mobilise. These may be used to treat discrete episodes when identified or could be used prophylactically to prevent the effects on enteral feeding due to feed intolerance from building up of stool from occurring.

7.3 Why it is important to do this overview

Although the importance of promoting early enteral feeding has been recognised, there is large variation and little consensus on the best approach to the way in which this should occur[17].

Cochrane systematic reviews have evaluated a broad range of early feeding strategies in the VLBW and very preterm infant. They are regarded as at the high level of the evidence pyramid and inform guidelines and recommendations in neonatal care. The evidence generated by these systematic reviews exist only as isolated components of evidence, each from a specific comparison of two or more interventions, without providing an indication of the context of what alternative or synergistic interventions exist within the decision of when, what and how enteral feeds are given to these infants.

Part of a systematic review is to evaluate what population the evidence applies to, by presenting what the characteristics of the included population are. In neonatal trials, the population who participate may be primarily towards the higher end of a specified range of gestational age and birthweight. Additionally, exclusion criteria may result in low or absent participation for infants who have the highest morbidity and mortality, and therefore who may benefit the most from research. Pre-defined subgroup analyses can be conducted to further investigate differences in effect with a specific differentiating feature within a population, however this is not always possible or done. Although the limitations from the studied population is noted during the data extraction and analysis, this may not be highlighted in the conclusions. The conclusions about evidence from trials and systematic reviews may be extrapolated to be implemented in these infants, despite evidence of benefit or harm.

Guidelines and policy recommendations informed by systematic reviews and meta-analyses are dependent on the quality of the included reviews. The methodological quality of Cochrane reviews within neonatal care is generally felt to be good, yet variation across reviews published and improvements to methodology have been reported[94,95]. The quality of the evidence provided by the included study data, as well as variation in the methodological quality of the Cochrane reviews themselves can affect whether clinicians involve the systematic review evidence to directly inform

practice or primarily influence further research. Guidelines or policy recommendations based on flawed evidence without appreciation of its limitations, may result in poor practice with adverse effects on outcomes for infants and families[96].

An overview of systematic reviews aims to collect the evidence from multiple systematic reviews of interventions and present this into a document that can be easily used to inform the decision makers of the best current evidence surrounding a particular topic. There are several potential approaches to an overview, in this case we will "summarize evidence from more than one systematic review of different interventions for the same condition or problem"[97], namely, to summarise the current evidence from the systematic reviews concerning different early enteral nutritional interventions to improve growth and reduce necrotising enterocolitis in very low birth weight and very preterm infants on the neonatal unit.

7.4 Is an overview the correct approach?

The "Editorial Decision Tree for Overviews" produced by the Cochrane Methods Comparing Multiple Interventions Group suggests that an overview is an appropriate format to meet the aims of this study[98]:

1. This study will be a review of reviews, analysing the results at the review level.
2. Early enteral nutrition in very low birth weight and very preterm infants has been identified as a priority for the Cochrane Neonatal Review Group, with a need for a "friendly front end" document for users to access the evidence-base related to this topic.
3. The review will not aim for a direct comparison of the effect of multiple interventions within the different systematic reviews but will map and summarise the evidence from multiple systematic reviews on the same group of interventions (early enteral feeding interventions) for the same condition (being born at very low birthweight or very preterm).

8 Objectives

The aim of this overview will be to:

1. Identify and summarise the existing Cochrane reviews of early enteral feeding strategies for improving growth and reducing necrotising enterocolitis in very and extreme preterm or very and extremely low birth weight infants.
2. Assess the quality of the methods used in the systematic review process.
3. Assess the quality of the evidence available within the studies included in each systematic review and the validity of their findings.
4. Map the existing evidence from Cochrane reviews.
5. Identify where there are important gaps in the current coverage of the topic through Cochrane reviews.
6. Identify where there are important gaps in the current evidence base, to identify the priorities for new primary research in this field.

9 Methods

9.1 Inclusion and exclusion criteria

Criteria for reviews to be included in the overview are illustrated in table 1.

Table 1: Inclusion and Exclusion criteria for selection of systematic reviews for the overview.

Inclusion criteria	
Type of study:	All systematic reviews published in the Cochrane Database of Systematic Reviews (CDSR). Reviews will be eligible for inclusion regardless of number, methodological quality, and type of studies included.
Type of participants:	Preterm infants (<37 weeks gestation) where the group studied includes: • Very preterm infants (born <31+6 weeks gestation) • And/or very low birthweight infants (<1500g)
Types of interventions:	All early enteral feeding interventions including, but not limited to: • What to feed: Type of feeds (source and level of nutritional content). • When to feed: Timing of starting, advancing and pausing of feeds. • How to feed: Route of feeding; volume of feeding; frequency of feeding. • Adjunctive strategies: Non-nutritive and pharmaceutical interventions to improve physiological ability and tolerance to enteral feeding.
Types of comparator:	Standard care, existing or alternative intervention, placebo, no treatment, or any other comparator.
Types of outcome:	Eligibility will be restricted to reviews whose outcomes include both necrotising enterocolitis and growth.

We did not include reviews of interventions that would only be delivered to very preterm or VLBW infants until after the first few weeks of life, or around/after the time of discharge from the neonatal unit. When stating 'very low birthweight' or 'VLBW', we will be referring to at least very low birthweight but this may also include extremely low birthweight infants.

9.2 Search methods for identification of reviews

This overview only included systematic reviews published in the Cochrane Database of Systematic Reviews (CDSR).

Cochrane systematic reviews are peer-reviewed systematic reviews conducted and supervised within a Cochrane Review Group. They are traditionally regarded highly due to the rigour by adherence to the well-described and validated standardised methodology of the Cochrane handbook and standard reporting decisions made by the Cochrane Neonatal review group, a priori publication of a protocol, peer review process by individuals with methodological expertise, and policy for keeping reviews up to date[99–104]. By excluding non-Cochrane systematic reviews we aim to reduce the complexity introduced by inconsistent data reporting[100], overlapping reviews[97,105] and introduction of bias by inclusion of non-randomised studies[97,106]. We have concluded that the relevant Cochrane reviews of interventions are sufficiently comprehensive to exclude non-Cochrane reviews[97].

As majority of reviews by Cochrane neonatal exclude non-randomised trials, by excluding broader overviews we may concluding no evidence where there may be very low certainty evidence from lower levels of the evidence pyramid, such as from observational studies[107]. We also acknowledge the risk of excluding potentially important information from studies included in reviews with a different focus or may be more up-to-date, or included additional analyses such as sub-groups of interest[108].

Two overview authors (VW and WM) independently conducted a search of the lists of reviews published by Cochrane Neonatal on 18th November 2019. We first scrutinized the list of completed reviews, protocols and titles on the Cochrane Neonatal Website for relevant records. In addition, we searched the CDSR on the same day in case the manual search had missed any records. This search did not reveal any additional relevant results.

No other databases were searched.

9.3 Data collection and analysis

The methodology for data collection and synthesis was conducted as per the Cochrane Handbook for Systematic Reviews of Interventions[97].

9.3.1 Selection of reviews:

All systematic reviews within Cochrane Neonatal were assessed independently for eligibility for inclusion by two overview authors (VW and WM). Any conflicts were resolved through discussion on 26th November 2019 until a consensus was reached.

The Cochrane Neonatal website contains not only completed reviews, but also protocols of ongoing reviews and registered titles of planned reviews. Where protocols were found to be eligible for inclusion, we categorised them as 'ongoing reviews' and established the anticipated completion date by contact with the relevant corresponding author or Cochrane neonatal editorial team.

As titles will not have outlined their inclusion/exclusion criteria, where it is not clear from the title if the future review would be eligible, we contacted the Cochrane neonatal editorial team to confirm likely eligibility, and these were classified into 'Characteristics of reviews awaiting assessment'.

As the evidence base in neonatal nutrition is a dynamic field, with continuous new emerging evidence, it is important that a systematic overview addresses the most up-to-date evidence. Yet, we may find that a systematic review may not have been updated recently. To address this problem, we will take a similar approach to that taken in two recent Cochrane Neonatal Overview protocols[109,110]. All reviews were assessed for eligibility, and if 'up to date' and published within the past 5 years, it was included in our review. Any reviews that have not been updated in the past five years (published in 2013 or earlier) had their status assessed and classified into one of four categories (Figure 1).

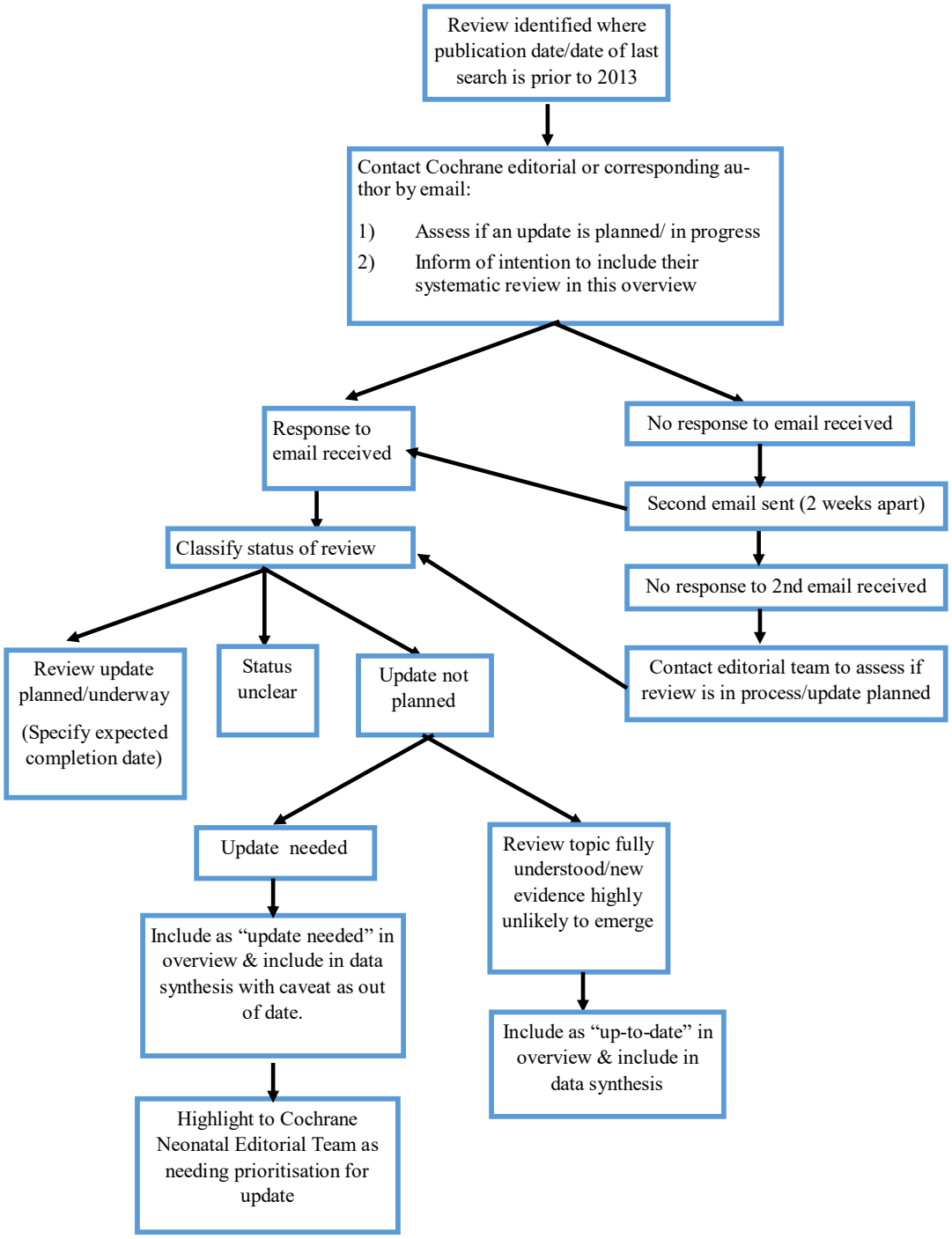


Figure 1: Flowchart of assessing "out of date" reviews

The Cochrane handbook identifies two anticipated complications which require a plan prior to conducting the search and analysis, namely 'out of date' and 'overlapping" reviews'[97].

'Overlapping' reviews are reviews which may address the same or similar questions and therefore which may include the same primary study data. This may lead to the data being over-represented by double counting in the systematic review, which if used in re-analysis of the data could lead to an excessively precise estimate of the true treatment effect[111]. Yet, as this overview will only include Cochrane intervention reviews, we do not expect that overlapping reviews will be a large problem, with overlapping reviews only occurring when an 'out of date' review has been revised and a new protocol written and published as a separate review with the previous review withdrawn. If overlapping reviews are found, only the most recent review will be included.

This overview concentrates on very low birth weight infants and very preterm infants, but we anticipate that the inclusion criteria of many of the reviews and primary studies will encompass a much wider range of infants, in particular groups of <37 weeks gestation. If a clear subset of the primary studies contained within the review only include infants above 32+0 weeks gestation, we planned not to include these results in our overview. This was not possible for the reviews included. For studies which include a range of infants <37 weeks without a clear distinction in eligibility of primary studies, we included all this data in our analysis.

9.3.2 Data extraction

Data extraction was performed by one review author (VW) on a pre-designed data collection form (Appendix 1) and independently checked by a second review author (WM or SO).

Initial extracted data and independent second review of extracted data were compared and disagreements were resolved by discussion. Where data was missing, we planned to contact the authors of the eligible reviews but did not plan to contact the primary study authors to retrieve unpublished data.

9.3.3 Dual authorship

The Cochrane Reviews relevant to this overview may include reviews that were authored by members of the Overview team, introducing a potential source of bias[112]. We identified these reviews and ensured that the eligibility and quality

assessment of each of the affected reviews is additionally checked by a reviewer who is not affiliated with the Cochrane review in question. We have highlighted these in our results and discussed the potential effect of inclusion of the Cochrane Reviews affected by dual authorship in the discussion of our Overview.

9.4 Assessment of methodological quality of included reviews

The methodological quality of the included reviews was assessed using the AMSTAR 2 tool[106]. Risk of bias of the systematic reviews was assessed using the ROBIS tool[113]. Quality assessment will be carried out by one author and checked by another, with any disagreements discussed until consensus is reached. We will use a high/moderate/low/critically low assessment of methodological quality, as per the guidance provided by the developers of the AMSTAR2 tool[106].

To allow an assessment methodological transparency and rigour we have checked included reviews against their protocols, specifically noting differences between the outcomes pre-specified in the review protocol versus outcomes reported in the published review. We reported discrepancies between review protocols and the completed published review that were not reported as amendments to the protocol in the publication.

We have reported study quality according to the assessment by the review authors within their published systematic review. We have not reassessed the quality of included primary studies within reviews, however during the collection of this information during the data extraction process, we have collected data including the quality assessment tool used, the overall conclusions of the authors and the justification they gave for their assessment. We have discussed any variation between reviews on the assessment and justification provided in the reviews.

AMSTAR2 and ROBIS are both validated assessment tools for critical appraisal of systematic reviews which each generate an overall rating[106,113–117]. Despite significant overlap between the areas assessed by the tools, there are important differences[114]. AMSTAR2 focusses more on methodological quality, including consideration whether exclusion of studies was justified, and whether authors reported conflict of interest both from funding of included studies but also declaration of funding and conflict of interest affecting the systematic review authors

themselves[106,114]. ROBIS places more consideration on risk of bias introduced during the review process and the results section itself, including scrutinising eligibility criteria, considering whether all relevant results and studies are included, and reporting of all predefined analyses or explanation of their departure[113,114]. This, as noted by Pieper et al.[114], requires a higher degree of subject knowledge from the review authors themselves to answer the questions in ROBIS. Inter-rater reliability (IRR) for both tools is variable between studies and individual components of each assessment, but the overall scoring assessed as fair to moderate for AMSTAR2[114,117] and fair to substantial for ROBIS[114–117]. The IRR for ROBIS has been found to be higher where reviewers have prior experience of working together[114,116], as is the case in this overview. Some disagreement is expected and should generate discussion of causes and importance between reviewers and consultation with experts[106]. In the presence of the discrepancy in focus and coverage of important considerations between the two tools, our review has used both assessment tools to achieve a more comprehensive critical appraisal.

9.5 Data synthesis

We have provided a narrative description of the characteristics of the included Cochrane Reviews. We have organised review findings by the subcategories specified within the review: ‘What to feed’, ‘When to feed’, ‘How to feed’ and ‘Adjunctive Strategies’.

The main results of the included reviews are first presented by each outcome assessed. We have then summarised the main results by categorising their findings building on the framework applied by the Cochrane overview of interventions to prevent cerebral palsy[118] and overview protocols for Birth Room transition support for preterm, near term and term infants[109,110]. We have also added two further categories of harmful interventions and probably harmful interventions as seen in the overview assessing interventions to prevent preterm birth[119]

1. Effective interventions: the review found high-quality evidence of effectiveness for an intervention.

2. Promising interventions (more evidence needed): the review found moderate-quality evidence of effectiveness for an intervention, but more evidence is needed.
3. Ineffective interventions: the review found high-quality evidence of lack of effectiveness for an intervention.
4. Probably ineffective interventions (more evidence needed): the review found moderate-quality evidence suggesting lack of effectiveness for an intervention, but more evidence is needed.
5. Harmful interventions: the review found high-quality evidence of harm for an intervention
6. Possibly harmful interventions: (more evidence needed): the review found moderate-quality evidence of harm for an intervention, but more evidence is needed.
7. No conclusions possible: the review found low- or very low-quality evidence, or insufficient evidence to comment on the effectiveness of an intervention.

Where review authors have assessed Grading of Recommendations Assessment, Development and Evaluation (GRADE) for our outcomes of interest, we will use their assessment of the quality of evidence to categorise findings into the above criteria. Where review authors have not assessed GRADE for our outcomes of interest, GRADE was assessed by two overview authors (VW and WM or SO) independently applying GRADE criteria and resolving disagreement through discussion. We had not planned to perform indirect or mixed treatment comparisons as part of this overview but planned to assess if there is a need for a network meta-analysis to be undertaken in the future.

10 **Results**

Five hundred and ninety-eight Cochrane systematic reviews, protocols and titles were screened by Title and Abstract. Forty results were included in full text screening (Figure 2). Five reviews were excluded after full text as their pre-determined outcomes did not include necrotizing enterocolitis (Figure 2).

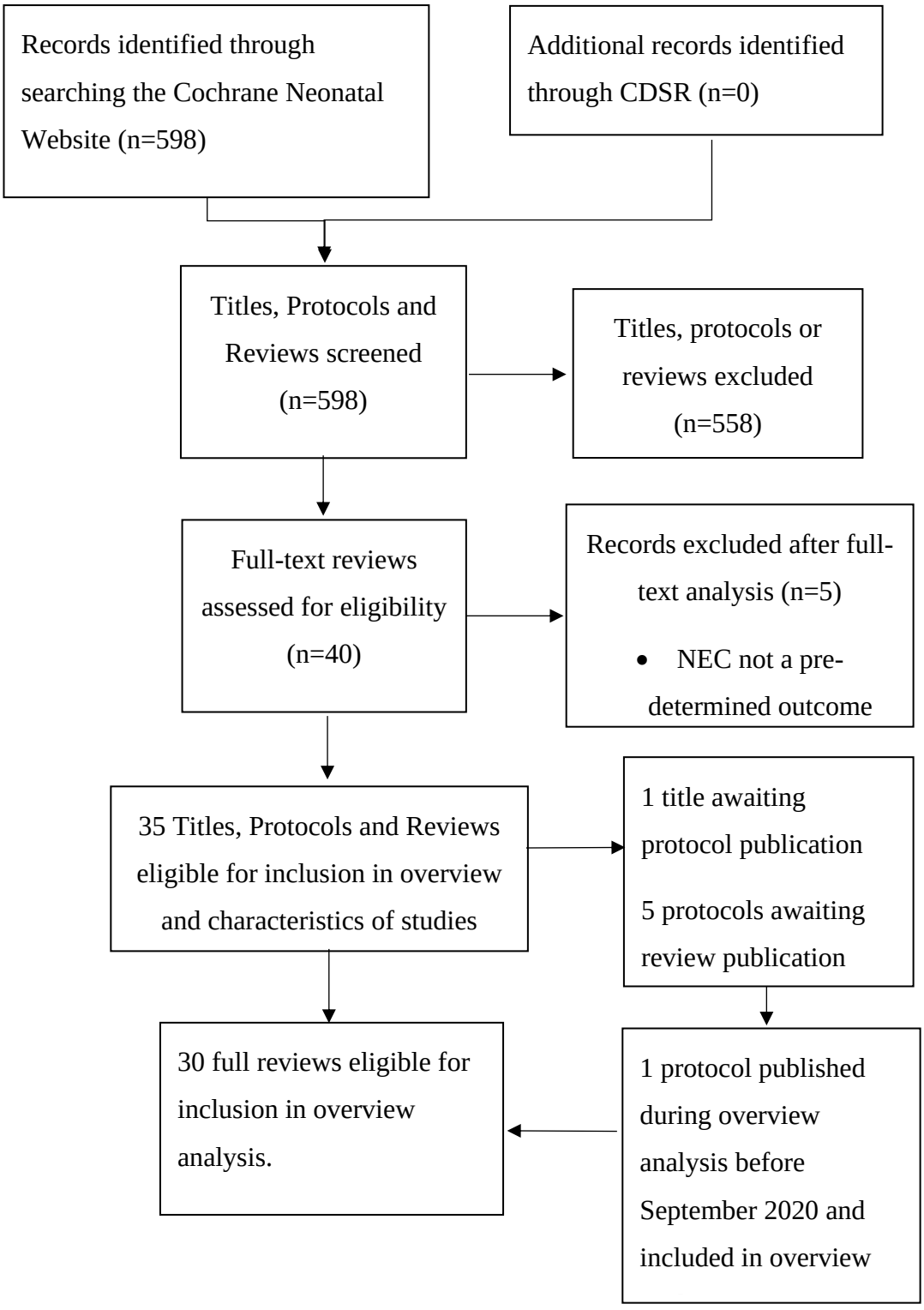


Figure 2: Study Flow Chart

Table 2: Characteristics of excluded reviews

Author	Title	Reason for exclusion
Amari *et al.* 2020[120]	Branched-chain amino acid supplementation for improving growth and development in term and preterm neonates	NEC not included as a pre-determined outcome.
Moon *et al.* 2019[121]	Longchain polyunsaturated fatty acid supplementation in preterm infants	
Watson *et al.* 2016[122]	Responsive versus scheduled feeding for preterm infants	
Flint *et al.* 2016[123]	Cup feeding versus other forms of supplemental enteral feeding for newborn infants unable to fully breastfeed	
Foster *et al.* 2016[124]	Non-nutritive sucking for increasing physiologic stability and nutrition in preterm infants	
Watson *et al.* 2013b[125]	Nasal versus oral route for placing feeding tubes in preterm or low birth weight infants	

Two reviews were identified which included both overview authors as review authors[126,127]. These were additionally assessed by a third assessor (SO) to mitigate risk of bias with conflicts resolved by consensus.

No overlapping reviews were found; however, we acknowledge that the topic of more recently updated reviews may have stemmed from the results of previous reviews.

10.1 Included reviews

Of the 40 included titles, 17 full reviews and 2 protocols assessed ‘what was fed’; 5 full reviews, 1 protocol and 1 title assessed ‘when to feed’, 4 full reviews and one protocol assessed ‘how to feed’, and four full reviews assessed ‘Adjunctive strategies’.

One review, found as a protocol on the original search, was published in August 2020 during the overview process, therefore has been included in our analysis and overview[128].

10.1.1 Out of date reviews

Seven reviews were assessed as out of date[129–135]. These reviews were published between 2007-2013.

Three reviews[129–131] concluded that further research was needed to address the review question, three reviews[132–134] concluded that there was insufficient evidence and one study[135] concluded that a cautious approach to the evidence was required due to methodological weaknesses. We assessed these as requiring an update.

Three reviews were identified as 'update planned or underway'[129,133,135]. Updates for these reviews are expected to be published in 2021.

Two reviews were identified as 'update not planned but update needed'[130,134]. We were informed that these reviews are low priority but an update may be needed if new evidence emerges.

For the review assessing early trophic feeding, the topic is felt to be fully understood with new evidence highly unlikely to emerge [131]. This review should therefore be treated as 'up to date'. More recent reviews have continued to develop on the evidence concluded by this review[136,137].

10.1.2 Ongoing reviews

One ongoing review was found only as a registered title on original search with the protocol published during the overview process[138]. Review authors have confirmed this review is near completion and aim for the completed review to be published towards the end of 2020.

Two reviews found as protocols have been completed and are expected to be published in early 2021[139,140]. One review is in progress and is expected to be published in 2021[141].

10.1.3 Reviews requiring assessment

One registered title was found on the Cochrane Neonatal Website. The protocol for this review on "Feeding during treatment of patent ductus arteriosus" has been completed and is expected to be published in early 2021.

10.2 **Description of Included Reviews**

10.2.1 ***What to feed***

Seventeen completed systematic reviews addressed "what to feed", of which 15 were assessed as "up to date" (Table 3).

The reviews included between 0[126,142] to 18[143] studies, with two reviews finding no eligible studies and four reviews only included a single study. The reviews where eligible studies were identified included between 14[144] and 1879[145] participants with 3 reviews including less than 100 participants[59,144,146] and two studies with more than 1000 participants[143,145]

Four reviews did not find any eligible studies conducted after 2000[127,130,144,147].

Table 3: Characteristics of included reviews (What to feed)

Review	Last search:	Population & Setting (within included studies) • *Gestational age at birth* • *Birthweight* • *Significant Exclusion factors* • *Setting of included studies*	Intervention	Comparison	Primary outcomes	No. studies included	Sample size (range)
Completed reviews "up to date"							
Thanigainathan *et al.* 2020[128]	15th August 2019	"preterm" 500-1500g Exclusions included chromosomal or congenital anomalies (all); death expected within 72 hours; no maternal milk/DBM refused; high likelihood of transfer during study period. USA and Austria, previous 20 years (published 2010 and 2016).	Early fortification of human milk (commenced <100 ml/kg/day or <7 days postnatal age) - in studies at 20 & 40 ml/kg/day	Later fortification of human milk (commenced ≥100 ml/kg/day or ≥ 7 days postnatal age - - in studies at 100 ml/kg/day	Short-term growth. Necrotising enterocolitis *(Long-term growtb in secondary outcomes)*	2	237 (100-138)
Brown *et al.* 2020[143]	30th September 2019	<37 weeks gestation at birth <2000g Exclusions included major congenital malformations, prolonged mechanical ventilation. Some excluding supplemental oxygen, diuretics, steroids, unable to achieve target feed volume. 1 study excluded NEC stage 2/3. 14 single centre trials, 4 multicentre. 5 in Europe, 4 in North America, 5 in Asia, 3 in Africa, 1 in South America. Conducted 1986-2019.	Energy (carbohydrate or fat) and protein fortification of human breast milk.	No fortification of human milk	Growth (short-term and long-term). Neurodevelopment *(NEC included in secondary outcomes)*	18	1456 (14-275)

Review	**Last search:**	**Population & Setting (within included studies)** • *Gestational age at birth* • *Birthweight* • *Significant Exclusion factors* • *Setting of included studies*	**Intervention**	**Comparison**	**Primary outcomes**	**No. studies included**	**Sample size (range)**
Amissah *et al.* 2020a[146]	22 August 2019	≤ 34 gestational weeks at birth Birthweights not stated (≤ 1500 grams in methods) Exclusion criteria included no asphyxia, major congenital anomalies, congenital cyanotic heart disease, gastrointestinal system anomalies, proven sepsis, or infection before the start of the study and were not transferred to other departments Iran. Conducted 2012-2013.	Human milk with carbohydrate supplementation	Human milk without carbohydrate supplementation	Growth (short-term and long-term). Neurodevelopment *(NEC included in secondary outcomes)*	1	75
Amissah *et al.* 2020b[148]	2 February 2018	≤36 weeks gestation at birth: <1990g at birth Excluded major congenital malformations. Some excluded septicaemia, chronic intrauterine infection, IVH, seizures, oxygen therapy. USA and Europe. Most conducted in 1980s & 1 early 2000s. Included 2 multicentre trials.	Human milk with additional protein supplementation	Human milk without additional protein supplementation	Growth (short-term and long-term). *(NEC included in secondary outcomes)*	6	204 (14-103)
Amissah *et al.* 2020c[144]	23 August 2019	<32 weeks gestation at birth <1500g Exclusion criteria not stated in the included trial. Sweden. Late 90s.	Human milk with additional fat supplementation	Human milk without additional fat supplementation	Growth (short-term and long-term). Neurodevelopment *NEC included in secondary outcomes.*	1	14
Fenton *et al.* 2020[149]	2 August 2019	<37 weeks PMA <2500g Excluded congenital anomalies, gastrointestinal disease, respiratory pathology (2 studies excluded supplemental oxygen). Country of trial not stated. During initial hospital stay only. Conducted 1980s-early 2000s.	*2 main comparisons (a priori inclusion criteria):* Low protein intake (<3 g/kg/day) High protein intake 3.0 g/kg/d≤x<4.0 g/kg/d	High protein intake 3.0 g/kg/d≤x<4.0 g/kg/d Very high protein intake ≥4.0 g/kg/d	Short-term growth Nitrogen utilisation and accretions Abnormal phenylalanine levels *(NEC included in secondary outcomes)*	6	228 (16-77)

Review	**Last search:**	**Population & Setting (within included studies)** • *Gestational age at birth* • *Birthweight* • *Significant Exclusion factors* • *Setting of included studies*	**Intervention**	**Comparison**	**Primary outcomes**	**No. studies included**	**Sample size (range)**
Quigley *et al.* 2019[145]	3 May 2019	<32 weeks gestational age <1800g birthweight Most excluded infants small for gestation age,with congenital anomalies, gastrointestinal or neurological problems. Europe and North America in 1970s-2016. 7 of 12 trials conducted >30 years ago, 5 trials since 2000.	Enteral feeding with formula either as sole diet or a supplement to maternal breast milk	Enteral feeding with donor breast milk either as sole diet or a supplement to maternal breast milk	Growth (short-term and long-term). Neurodevelopment *(NEC included in secondary outcomes)*	12	1879 (53-373)
Ng *et al.* 2019[150]	28 January 2019	<37 weeks gestational age <2500g birthweight Most trials excluded infants with congenital anomalies, gastrointestinal or neurological problems. Europe and North America from 1990s to present n.b. Majority funded by formula manufacturers	Hydrolysed cow's milk formula	Standard cow's milk formula (equivalent energy and protein content to intervention)	Feed intolerance NEC *(Short and long-term growtb in secondary outcomes)*	11	665 (16-108)
Walsh *et al.* 2019[127]	12 November 2018	<37 gestational age <2000g birthweight Excluded infants with congenital anomalies, respiratory, gastrointestinal or neurological problems. UK, USA, Turkey, Thailand and South Africa in 1980s & 90s.	Nutrient-enriched formula (> 72 kcal/100 mL and > 1.7 g protein/100 mL)	Standard formula: (≤ 72 kcal/100 mL and ≤ 1.7 g protein/100 mL	Growth (short-term and long-term). Neurodevelopment *(NEC included in secondary outcomes)*	7	590 (22-264)
Dempsey *et al.* 2019[142]	23 October 2018	*No studies found* Gestation at birth criteria not specified Birthweight criteria <1500g	Banked donor preterm milk with or without fortification fed either as sole enteral diet or as a supplement to mother's own milk	Banked donor term milk with or without fortification fed either as sole enteral diet or as a supplement to mother's own milk	Growth (short-term and long-term). Neurodevelopment *(NEC included in secondary outcomes)*	0	n/a

Review	**Last search:**	**Population & Setting (within included studies)** • *Gestational age at birth* • *Birthweight* • *Significant Exclusion factors* • *Setting of included studies*	**Intervention**	**Comparison**	**Primary outcomes**	**No. studies included**	**Sample size (range)**
Brown *et al.* 2019[126]	1 October 2018	*No studies found* Gestation at birth criteria: <37 Birthweight criteria <2500g	Feeding with formula milk	Feeding with maternal expressed milk	Growth (short-term and long-term). Neurodevelopment NEC	0	n/a
Basuki *et al.* 2019[147]	1 October 2018	<37 weeks gestational age (Mean 29-33 weeks) <1750g at birth Excluded major congenital malformations and serious gastrointestinal problems (including NEC). USA & India. Late 80s-early 90s.	Diluted preterm/term formula during feeding advancement (total enteral nutrient intake equivalent to comparator, therefore at increased volume)	Full-strength preterm/term formula during feeding advancement	NEC Feed intolerance *(Short and long-term growtb in secondary outcomes)*	3	102 (14-50)
Premkumar *et al*. 2019[151]	20 September 2018	"Preterm": (mean 27.7 weeks) <1250g (mean 888g) at birth Excluded congenital or chromosomal anomalies, fed any cow's milk, or not able to feed enterally ≤14 days of life. North America in the previous decade (2018 published).	Human milk fortified with human-milk derived multi-nutrient fortifier.	Human milk fortified with bovine milk-derived multi-nutrient fortifier or formula.	NEC Mortality *(Short and long-term growtb in secondary outcomes)*	1	127
Nasuf *et al.* 2018[152]	August 2017	<33 weeks gestation at birth 410-2500g at birth Excluded major congenital anomalies, infection, blood borne virus, maternal infection or early infection. Single centre in North America and Brazil. Conducted 2011-2016.	Early oropharyngeal colostrum	Control (water, oral formula, donor breast milk or no intervention)	NEC Late onset infection Mortality before discharge. *(Short-term growtb in secondary outcomes)*	6	335 (12-149)
Abiramalatha *et al*. 2017[59]	14 November 2016	Not specified (review all <37 weeks accepted) <1500g No exclusion criteria stated. India in early 2010s.	high-volume enteral feeds: > 200 mL/kg/d	standard-volume enteral feeds: ≤ 200 mL/kg/d	Growth (short-term and long-term). *(NEC included in secondary outcomes)*	1	64

Review	Last search:	Population & Setting (within included studies) • *Gestational age at birth* • *Birthweight* • *Significant Exclusion factors* • *Setting of included studies*	Intervention	Comparison	Primary outcomes	No. studies included	Sample size (range)
Out of date reviews							
Tan-Dy *et al.* 2013[130]	June 2012	26-34 weeks gestation at birth: (mean 31.4 weeks in both groups) Majority <1500g at birth No exclusion criteria reported in review: North America in 1997-2000.	Addition of lactase to milk	Placebo/no intervention	Short-term growth Time to achieve full enteral feeds *(NEC included in secondary outcomes)* n.b. no included trials reported the defined primary outcomes	1	130
	Status: Update not planned. Update may be needed if new evidence emerges. Low priority.						
Nehra *et al.* 2002[129]	11 September 2007	<37 weeks gestation at birth:. (Mean 29-32 weeks) n.b. infant age at start of trial 1-6 weeks. Mean 1010-1476g at birth Excluded major congenital anomalies, surgery, major illnesses, respiratory support or supplementary oxygen, or fed human milk. Setting not described. Conducted in 1980s and early 1990s.	High MCT formula (40% or more by weight) over at least one week	Low MCT formula (20% or less by weight) over at least one week	Short-term growth *(NEC included in secondary outcomes as part of "adverse events") Long-term growtb in secondary outcomes)*	8	182 (14-30)
	Status: Review update planned/underway. Expected completion date 2021. Low priority.						

Review	Last search:	Population & Setting (within included studies) • *Gestational age at birth* • *Birthweight* • *Significant Exclusion factors* • *Setting of included studies*	Intervention	Comparison	Primary outcomes	No. studies included	Sample size (range)
Ongoing reviews							
Fabrizio et al. 2019[139] *Protocol*	*Protocol published 4th November 2019.* *Protocol published 4th November 2019.*	*Status: title changed to "Individualized versus standard diet fortification for growth and development in preterm infants receiving human milk".* *Gestation at birth: -* *Birthweight: <1500g* *Setting:-* *Significant Exclusion factors: -*	*3 comparisons:* *Targeted fortification* *Adjustable fortification* *Targeted fortification*	*Adjustable fortification* *Standard fortification* *Standard fortification*	*Growth* *(NEC included in secondary outcomes)*	-	-
	Status:In progress. *Expected publication early 2021*						
Gao et al. 2008[141] *Protocol*	*Protocol published 16th April 2008*	*Protocol.* *Gestation at birth: 37 weeks* *Birthweight:* *Setting:-* *Significant Exclusion factors: -*	*High level of protein (=/>1.4g/100ml of human milk)* *Moderate level of protein content of human milk fortifier (1<x<1.4g protein per 100 ml of human milk)* *Low level of protein content of human milk fortifier (<1 g protein per 100 ml of human milk)*		*Growth* *Neurodevelopment* *Mortality* *(NEC included in secondary outcomes)*	-	-
	Status: In progress. *Expected publication 2021.*						

10.2.1.1 *Included Participants*

Only two reviews included only very preterm infants (<32 weeks)[144,145]. Other reviews included infants who were preterm (<37 weeks), with three reviews only including infants ≤34 weeks[130,146,152]. Only seven reviews had at least the majority of infants as very low birth weight (<1500g)[59,128–130,144,146,151]. Three reviews included infants up to 2500g[149,150,152] and five reviews up to 2000g[127,143,145,147,148].

The included studies were conducted mostly across North America and Europe, with eligible studies in five reviews including studies from low and middle income countries in Asia, South America, Africa[59,127,143,147,152] and one in Iran[146]. Two reviews did not report where the included studies were conducted[129,149].

The majority of reviews included studies which excluded congenital anomalies[127,128,152,129,143,145–150] gastrointestinal[127,145–147,149,150] or neurological problems[127,145,150]. Four reviews included studies which excluded infants with respiratory problems[127,149], with three reviews including studies excluding infants on supplemental oxygen [129,148,149], and one review included studies excluding infants receiving prolonged mechanical ventilation[143]. Two reviews[143,147] included studies excluding infants with necrotising enterocolitis. One review included studies excluding infants who received surgery[129]. Three reviews included studies which excluded infants where maternal or early neonatal infection were identified[146,148,152]. Two reviews evaluating multi-nutrient fortification excluded infants who were unable to enterally feed by 2 weeks of life[151] or were unable to achieve the study target volume[143]. Two reviews investigating components of formula or breastmilk fortifier specifically, included studies which excluded infants not exclusively fed formula[129] or breastmilk[151], respectively. A review on breastmilk fortification[128] included a study where infants with no maternal breastmilk or consent for DBM was refused, and a trial where transfer during the study period was likely. Two reviews did not describe exclusion criteria[59,130], and Amissah *et al.*[144] stated that the included trial did not describe their exclusion criteria.

10.2.1.2 Interventions and comparators

The eligible systematic reviews compared the nutritional content and source of the enteral nutrition. They compared the source of nutrition with both maternal breastmilk[126] and donor breastmilk[145] compared to formula milk, human versus bovine milk derived fortification[151], and compared early colostrum[152] to other early feed exposures. Reviews investigated the structural qualities of the nutrients with addition of lactase[130], hydrolysed protein[153], and the degree of medium chain triglyceride content[129] within cow's milk formula. The reviews investigated the concentration of nutrients within a specified volume by investigating supplementation of lipids[144], proteins [148] and carbohydrates[146], or combination of these facilitated by different artificial formulas [127] and addition of fortifier[143,151], but also investigating how dilution of formula [147] could deliver similar degree of nutrition but at a reduced osmolality, or and increased volume of milk[59] could be used to provide additional nutrition. One recently completed review explored the age or stage of feeding at which fortification is introduced[128].

The protocols plan to further investigate fortification with different strategies of tailoring the fortification to the infant. One protocol published in 2008[141], planned to investigate the level of protein in the fortified human milk given. This differed by the thresholds of protein per volume of milk in the three comparison groups, whereas the two published reviews compared the protein levels by g/kg/day in only formula milk[149] and presence or absence of additional protein supplementation in human milk, with no overall level of protein specified[148].

The comparator groups for these interventions were either another or "standard" form or volume of milk or fortifier[59,126,130,142,143,145,151,152], the "standard" or "normal nutritional content"[127,144,146,148], or fortification introduced at a later stage[128]. Where the structural quality of the nutritional components or the dilution of the milk were tested, this was done at an equivalent energy and protein content to the intervention[147,150]. In one completed review and two protocols, there were three intervention groups split into more than two groups, with levels of protein intake measured by g/kg/day[149], protein content in g/100ml of human milk[141], or mode of fortification[139]. One review compares two intervention groups split by percentage by weight of MCT in formula[129].

10.2.1.3 Primary outcomes assessed

Short-term growth was assessed as a primary outcome by 13 of 17 completed reviews[59,126,146,148,149,127–130,142–145]. Four further reviews assessed short term growth as a secondary outcome[147,150–152].

Nine reviews assessed long-term growth as a primary outcome[59,126,127,142–146,148], with five reviews assessing this as a secondary outcome[128,129,147,150,151]. Three reviews did not include long-term growth as an a priori outcome[130,149,152].

Necrotising enterocolitis was assessed as a primary outcome in six reviews[126,128,147,150–152] and as a secondary outcome in eleven reviews [59,127,149,129,130,142–146,148].

10.2.2 When to feed

Five completed reviews addressed the question of "when to feed", of which four were assessed as 'up to date' (Table 4). The reviews included one[154] to 10[136] studies, with the number of participants ranging from 22[154] to 3757[136] infants.

Table 4: Characteristics of included reviews (When to feed)

Review	**Last search:**	**Population & Setting (within included studies)** • ***Gestational age at birth*** • ***Birthweight*** • ***Significant Exclusion factors*** • ***Setting of included studies***	**Intervention**	**Comparison**	**Primary outcomes**	**No. studies included**	**Sample size (range)**
Completed reviews "up to date"							
Yeo *et al.* 2019[154]	14th November 2018	≤32 weeks gestation Birthweight not reported Congenital anomalies, higher ordered multiples, previous NEC, current sepsis & feed intolerance, evidence of altered mesenteric blood flow velocity (IUGR/PDA) USA	Temporary stopping of feeds before, during or after transfusion of all blood products.	Continuation of feeding s per schedule	NEC within 48 hours after transfusion NEC at any point after first blood transfusion Mortality *(Growth in secondary outcomes)*	1	22
Abiramalatha *et al*. 2019a[155]	19th February 2018	Majority <34 weeks gestation at birth 2 trials <1500g, 1 trial 1500-2000g. Excluded major congenital abnormalities, absent/reversed EDF, perinatal asphyxia/poor APGAR scores. 1 trial excluded growth restricted infants. Setting not stated in review.	Routine monitoring of gastric residual for decisions on advancement of enteral feeds in absence of signs of feed intolerance/NEC	No monitoring of gastric residual until sign of feed intolerance/NEC.	NEC Time to establish full enteral feeds *(Growth in secondary outcomes)*	3	228 (61-87)

Review	**Last search:**	**Population & Setting (within included studies)** • ***Gestational age at birth*** • ***Birthweight*** • ***Significant Exclusion factors*** • ***Setting of included studies***	**Intervention**	**Comparison**	**Primary outcomes**	**No. studies included**	**Sample size (range)**
Abiramalatha *et al*. 2019a[155] (continued)	19th February 2018			2nd comparison: monitoring of gastric residual in both groups, by decisions made on two different predefined criteria.			
Oddie *et al*. 2017[136]	20th June 2017	<37 weeks gestation at birth:, 1 large trial with majority 23-32 weeks. Majority <1500g at birth Majority excluded severe congenital anomalies. 2 trials formula only, 2 trials breastmilk/donor breastmilk only. 2 excluded respiratory distress syndrome. 1 excluded infants with umbilical vessel catheters present. 1 excluded infants with absent end-diastolic flow. UK & Ireland*, USA, India, Turkey, South Africa.	Enteral feed advancement ≤24 ml/kg/day	Enteral feed advancement >4 ml/kg/day	NEC All-cause mortality *(Short term and Long term growth in secondary outcomes)*	10	3757 (30-2804) *1 trial = 75% of participants

Review	**Last search:**	**Population & Setting (within included studies)** • ***Gestational age at birth*** • ***Birthweight*** • ***Significant Exclusion factors*** • ***Setting of included studies***	**Intervention**	**Comparison**	**Primary outcomes**	**No. studies included**	**Sample size (range)**
Morgan *et al.* 2014[137]	September 2014	<37 weeks gestation at birth Majority <1500g All excluded major congenital anomalies. Majority excluded infants with multi-organ failure, intrauterine transfusion or exchange transfusions, or requiring inotropic support. 1 trial excluded infants with intrauterine growth restriction. USA, UK/Ireland, Greece, Columbia, Qatar, Iran, Israel.	Delayed introduction (≥4 days after birth) of progressive enteral feeds	Early introduction (<4 days after birth) of progressive enteral feeds	NEC All-cause mortality *(Short term and Long term growth in secondary outcomes)*	9	1106 (12-404)
Out of date Reviews							
Morgan *et al.* 2013[131]	December 2012 Status: Update not planned. Review topic fully understood/new evidence highly unlikely to emerge**.** **"Up to date".**	Majority <32 weeks gestation at birth: Majority <1500g, 1 trial <2000g at birth Majority excluded major congenital abnormalities. Unstable infants despite respiratory support/inotropes. Included mechanically ventilated neonates. Infants were fed breastmilk, formula or both. USA, UK, Spain, some not reported.	Early trophic feeding (up to 24 ml/kg/day) beginning within first 4 days after birth and continued for at least 5 days/at least one week after birth.	Enteral fasting for at least 5 days or until at least one week after birth.	Feeding intolerance NEC *(Short term and Long term growth in secondary outcomes)*	9	754 (29 to 190)

Review	**Last search:**	**Population & Setting (within included studies)** • ***Gestational age at birth*** • ***Birthweight*** • ***Significant Exclusion factors*** • ***Setting of included studies***	**Intervention**	**Comparison**	**Primary outcomes**	**No. studies included**	**Sample size (range)**
Ongoing reviews							
Walsh et al. 2020[138]	*Protocol published 12th March 2020*	*Gestation at birth: <37 weeks' gestation* *Birthweight: <2500g* *Setting: unspecified* *Significant exclusion: none stated*	*Full enteral feeds from birth without parenteral fluids or nutrition.*	*Any other feeding regimen, such as delayed initiation of full milk feeds and gradual advancement of feed volumes while receiving supplemental fluid or nutrients parenterally.*	*In hospital growth (weight gain, head circumference)* *Growth restriction* *Necrotising enterocolitis*	*n/a*	*n/a*
	Status: near completion. *Expected publication: early 2021*						
Reviews awaiting assessment (Reviews at title stage):							
Feeding during treatment of patent ductus arteriosus *Status: Protocol awaiting publication..* *Expected publication: early 2021*							

10.2.2.1 Included participants

The included infants were preterm (<37 weeks), with 2 reviews declaring the majority of infant to be <32 weeks gestation at birth[137,154], and one review with the majority of infants <34 weeks[155].

The majority of infants in four reviews were very low birth weight (<1500g) [131,136,137,155], with some infants <2000g[131,155]. Yeo *et al.*[154] did not state the birthweight of included participants.

All reviews included studies excluding infants with congenital anomalies. Most reviews included studies excluding infants with a risk of disrupted gut flow such as intrauterine growth restriction (IUGR)[137,154], patent ductus arteriosis (PDA)[154], presence of umbilical vessel catheters or absent end-diastolic flow[136,155]. Three reviews included studies excluding unstable participants requiring inotropes[131,137], or current sepsis[154]. The trial included in the review assessing feeding during blood transfusion also excluded infants with previous NEC or feed intolerance[154].

10.2.2.2 *Interventions and comparators*

The eligible systematic reviews in this category compared decisions about feeding during blood transfusion[154], decisions affected by routine monitoring of gastric residuals [155], the time of introduction of trophic[131] and progressive[137] feeding and speed of feeding advancement[136].

The comparisons were split by a defined time (day) after birth[131,136,137] or by intervention versus absence of intervention or routine care[154,155]. One review had two planned comparisons, comparing routine monitoring of gastric residuals against each no monitoring or monitoring against different predefined criteria[155].

10.2.2.3 Primary outcomes assessed

All five reviews assessed necrotising enterocolitis as a primary outcome[131,136,137,154,155]. The review assessing feeding around blood transfusion[154] assessed necrotising enterocolitis at two time points, both within 48 hours after transfusion and at any point after the first blood transfusion.

Growth was assessed as a secondary outcome in all reviews, with short term growth assessed by all reviews[131,136,137,154,155], and long-term growth assessed by all but one review[131,136,137,155].

10.2.3 How to feed

Four completed reviews assessed 'how to feed' (Table 5). Only one review was assessed as "up to date", with the remaining reviews last published in 2011-2013. Two reviews found only a single eligible study[132,156]. The number of included participants ranged from 31[132] to 511[133].

Table 5:Characteristics of included reviews (How to feed)

Review	**Last search:**	**Population (within included studies)** • ***Gestational age at birth*** • ***Birthweight*** • ***Significant Exclusion factors*** • ***Setting of included studies***	**Intervention**	**Comparison**	**Primary outcomes**	**No. studies included**	**Sample size (range)**
Completed reviews "up to date"							
Abiramalatha *et al*. 2019b[156]	19 February 2018	23-28 weeks gestation at birth Birthweight not reported Excluded major congenital/chromosomal anomalies, low likelihood of survival. Performed in the past 10 years. Country of trials performed in not reported.	Re-feeding the gastric residuals unless the predefined quality parameters were not satisfied.	Discarding the gastric residuals, irrespective of quantity and quality.	Time to regain birthweight Necrotising enterocolitis stage 2/3	1	72
Out of date reviews							
Watson *et al.* 2013[135]	June 2012	Gestation at birth mostly unspecified. 2 trials ≤32 weeks. Majority <1500g at birth, appropriate for gestational age. 2 trials excluded infants requiring "assisted ventilation" USA, Canada, UK, Australia. 1970s to early 90s.	Transpyloric tube feeding (passed via nose or mouth)	Gastric tube feeding (passed via nose or mouth)	Days to establish full enteral feeds Growth *(Necrotising enterocolitis in secondary outcomes)*	9	316 (11-80)
	Status: Review update planned/underway. Expected completion: 2021						

Review	**Last search:**	**Population (within included studies)** • ***Gestational age at birth*** • ***Birthweight*** • ***Significant Exclusion factors*** • ***Setting of included studies***	**Intervention**	**Comparison**	**Primary outcomes**	**No. studies included**	**Sample size (range)**
Dawson *et al.* 2012[132]	1 May 2012	≤32 weeks gestation at birth: ≤1750g at birth Exclusion factors not stated. Setting not described. Crossover trial. Early 90s	Push gavage feeding	Gravity gavage feeding	Time to achieve full enteral feeding Feeding intolerance *(Necrotising enterocolitis included in secondary outcomes)*	1	31
	Status: Update not planned. Update may be needed if new evidence emerges. Low priority.						
Premji *et al.*2011[133]	25 July 2011	Gestation at birth: Preterm, 2 trials ≤30 weeks. Birthweight: <1500g Setting: Not stated in the review. Late 80s-early 2000s. Significant Exclusion factors: Excluded major congenital anomalies, Some excluded infants with low initial APGAR scores, unstable respiratory status, intrauterine infections, sepsis or NEC. Trials fed human milk only, formula only or both.	Continuous nasogastric feeding for initiation of feeds and advancement to full enteral feeds.	Intermittent bolus nasogastric feeding for initiation of feeds and advancement to full enteral feeds.	Feeding intolerance Days to regain birthweight Age at full enteral feedings Age at discharge Growth NEC	7	511 (23-171)

Review	**Last search:**	**Population (within included studies)** • ***Gestational age at birth*** • ***Birthweight*** • ***Significant Exclusion factors*** • ***Setting of included studies***	**Intervention**	**Comparison**	**Primary outcomes**	**No. studies included**	**Sample size (range)**
Premji *et al.*2011[133] (continued)	Status: Review update planned/underway. Expected completion date: 2021 Low priority						
Ongoing reviews							
Ibrahim et al. 2016 (protocol)[140]	*Protocol published 17th August 2016*	*Planned:* *Gestational age at birth: <32 weeks gestation* *Birthweight: any* *Setting: unspecified* *Exclusion: gastrostomy/jejunostomy/transpyloric feeding/pre-existing feeding problems.*	*Short feeding interval for bolus (breastmilk/formula): <3 every hours*	*Short feeding interval for bolus (breastmilk/formula): ≥3 hours*	*Time to achieve full enteral feeding* *Time to regain birthweight* *(Necrotising enterocolitis included in secondary outcomes)*	*n/a*	*n/a*
	Status: In progress. *Expected publication: early 2021*						

10.2.3.1 Included participants

The majority of included infants were very preterm (<32 weeks gestation), but two reviews including trials where gestation at birth were unspecified [133,135]. The review investigating refeeding of gastric residuals[156] only included infants ≤28 weeks. Birthweight was only reported in 3 reviews, with the majority of infants very low birth weight, reported as ≤1750g[132] and <1500g[133,135].

Two reviews stated that infants with congenital anomalies were excluded[133,156]. The review assessing site of feeding tube placement included some trials excluding infants on "assisted ventilation"[135]. The review assessing continuous and intermittent bolus feeding had some infants excluded with low initial APGAR scores, necrotising enterocolitis, intrauterine infection or current sepsis[133].

10.2.3.2 Interventions and comparators

Each of the four eligible reviews compared two defined mode of feeding interventions. These compared re-feeding versus discarding gastric residuals[156], the transpyloric or gastric site of a feeding tube placement[135], push versus gravity gavage feeding[132], continuous versus intermittent bolus feeding[133]. A protocol[140] plans to assess the length of feeding interval of less than three hours compared with 3 hours or more.

10.2.3.3 *Primary outcomes assessed*

Two reviews assessed necrotising enterocolitis as a primary outcome[133,156] and two reviews assessed this as a secondary outcome[132,135].

Three reviews assessed growth as a primary outcome[133,135,156], with two reviews assessing time to regain birthweight[133,156], and three reviews assessing a variety of growth markers[133,135,156]. Premji *et al.*[133] and Watson *et al.*[135] do not specify timescale of the growth outcomes in their protocols.

Dawson *et al.*[132] assessed time to regain birthweight in their secondary outcomes and Abiramalatha *et al.*[156] assesses long term growth in their secondary outcomes.

Adjunctive strategies

Of the four reviews found assessing adjunctive strategies to improve feeding in preterm infants, 3 reviews were assessed as 'up to date' (Table 6). The reviews included between 3[157,158] and 16[159] studies, with between 161[157] and 825[159] participants. The review on oral stimulation for promoting oral feeding[159] was split into three comparisons including 9, 7 and zero studies in each comparison.

Table 6: Characteristics of included studies (Adjunctive strategies)

Review	**Last search:**	**Population (within included studies)** • ***Gestational age at birth*** • ***Birthweight*** • ***Significant Exclusion factors*** • ***Setting of included studies***	**Intervention**	**Comparison**	**Primary outcomes**	**No. studies included**		**Sample size (range)**
Completed review "up to date"								
Muelbert *et al.* 2019[157]	1st June 2018	Gestation at birth: Infants <34 weeks, 1 trial <29 weeks. Birthweight: Not stated in 2 trials. 1 trial ~1kg. Setting: Tertiary NICU in Turkey, Austrialia and 1 not stated. Significant Exclusion factors: • Most excluded congenital anomalies • 1 trial excluded growth restricted infants • 1 trial excluded sepsis and commonly seen neonatal neurological, gastrointestinal and respiratory comorbidities. • One trial only included mothers who planned to breastfeed.	Delivery of smell/taste of breast milk or formula milk immediately before or at time of tube feedings	No exposure to the smell/taste of milk prior/during milk feeding	Time to reach full sucking feeds Adverse effects related to the intervention. *(Necrotising enterocolitis and short-term growth included in secondary outcomes)*	3		161 (30-80)
Greene *et al.* 2016[159]	25th February 2016	Gestation at birth: <37 weeks Birthweight: 1 <2000g, 1 <1000g, 3 not described, 8 with mean birthweight <1500g.	Oral stimulation	No intervention/standard care/sham treatment Non-oral intervention	Time to achieve exclusive oral feeding. Time (days) spent in NICU	16	9 7	825 (14-108)

Review	**Last search:**	**Population (within included studies)** • ***Gestational age at birth*** • ***Birthweight*** • ***Significant Exclusion factors*** • ***Setting of included studies***	**Intervention**	**Comparison**	**Primary outcomes**	**No. studies included**		**Sample size (range)**
Greene *et al.* 2016[159] (continued)		Setting: 5 USA, 2 UK, 2 Iran, 1 India, 1 France, 2 China, 3 Brazil. Significant exclusion factors: Excluded defined respiratory disease and significant comorbid conditions that preclude the introduction of oral feeding.		Oral stimulation delivered by a different method	Total Hospital stay (days) Duration (days of parenteral nutrition. *(Necrotising enterocolitis and short-term growth included in secondary outcomes)*		0	
Anabrees *et al*. 2015[158]	1st April 2015	Gestation at birth: ≤32 weeks Birthweight: Majority <1500g Setting: Austria, UK, India Significant Exclusion factors: Majority excluded infants with major congenital malformations and gastrointestinal abnormalities. 1 excluded hypoxic ischaemic encephalopathy.	Prophylactic or therapeutic glycerin laxatives	Placebo or no treatment	Time to full enteral feeds *(Necrotising enterocolitis and short-term growth included in secondary outcomes)*	3		177 (42-81)
Out of date reviews								
Ng *et al.* 2008[134]	1st December 2007	Gestation at birth: <37 weeks, 3 trials only including <32 weeks, 9 trials with median/mean <32 weeks. Birthweight: Most <2000g, 2 trials <1500g	Prevention: Erythromycin either 3-12 mg/kg/day or >12 mg/kg/day for up	Prevention: Placebo for up to 2 weeks to promote gastrointestinal motility once enteral feeding has begun.	Days to achieve full enteral feeding.	10	3	199 (50-76)

Review	**Last search:**	**Population (within included studies)** • ***Gestational age at birth*** • ***Birthweight*** • ***Significant Exclusion factors*** • ***Setting of included studies***	**Intervention**	**Comparison**	**Primary outcomes**	**No. studies included**		**Sample size (range)**
Ng *et al.* 2008[134]		Setting: Not stated in systematic review. Significant Exclusion factors: • Mainly included clinically stable infants >5 days of life at study entry. • Most excluding congenital anomalies or chromosomal abnormalities, previous/current NEC of gastrointestinal surgery, sepsis, perinatal hypoxia. • 1 excluded infants on continuous milk feeding, 4 studies excluded infants for whom them were unable to initially advance feeds. • 1 study excluded infants with growth restriction or absent/reversed antenatal end diastolic flow. Receiving breastmilk, formula milk or both.	to 2 weeks to promote gastrointestinal motility once enteral feeding has begun. Treatment: oral/IV erythromycin 3-12 mg/kg/day or >12 mg/kg/day for up to 2 weeks to promote gastrointestinal motility commenced once feeding intolerance is diagnosed.	Treatment: Placebo for up to 2 weeks to promote gastrointestinal motility commenced once feeding intolerance is diagnosed.	*(Necrotising enterocolitis and short-term growth included in secondary outcomes)*		7	330 (24-60)
	Status: Update not planned. Update may be needed if new evidence emerges. Low priority.							

10.2.3.4 Included participants

The included infants were preterm, with two reviews reporting infants to be <34 weeks[157] and ≤32 weeks gestation at birth[158], and one further review specifying that the mean/median gestation in the majority of trials were <32 weeks [134]. The majority of infants were <2000g, excluding infants with significant congenital malformations[134,157,158]. Two reviews included trials excluding infants with previous/current necrotising enterocolitis, gastrointestinal morbidity or surgery, respiratory morbidity or sepsis. Two reviews had trials excluding perinatal asphyxia or hypoxic ischaemic encephalopathy[134,158]. Ng *et al.* [134] had studies excluding continuous milk feeding, inability to advance feeds initially, and infants with growth restriction or absent/reversed end-diastolic flow.

10.2.3.5 Interventions and comparators

The four eligible reviews in this category compared an intervention against absence[157–159] or placebo[134,158,159]. One trial compared the intervention of oral stimulation against three different comparators[159]. These were firstly against the absence or placebo of the oral intervention, against a non-oral intervention, or against an oral stimulation intervention delivered by a different method. The review by Ng *et al.*[134] had two comparisons. Firstly, comparing the intervention against a placebo when used as prevention, and secondly comparing the intervention when used as a treatment when feeding intolerance is diagnosed.

10.2.3.6 Primary outcomes assessed

All four reviews assessed necrotising enterocolitis and short term growth as secondary outcomes[134,157–159].

10.3 Review author assessment of Risk of Bias in the included studies

A summary table and discussion of the review author assessments of the included reviews is presented in Appendix 2.

Sixteen reviews assessed all risk of bias domains[59,127,151,154–157,159,128,132,144–146,148–150]. Nine reviews did not assess 'other bias'[129–131,133,135–137,143,152], with one review only assessing it for one of its three studies[147]. Seven reviews did not assess for reporting bias[129,131,133,135–137,143]. Anabrees *et al.*[158] only assessed detection bias for one of the three trials.

Three reviews only reported some of the above risk of bias outcomes, but instead reported 'sequence generation', 'allocation concealment', 'blinding of intervention' 'blinding of outcomes' and 'complete follow-up' as 'yes', 'no' and 'can't tell' within the 'characteristics of included studies'[129,133,134] . Nehra *et al.*[129] and Ng *et al.*[134] additionally only reported risk of bias assessment of selection bias[129,134]. Premji *et al.*[133] reported selection bias from allocation concealment and performance bias for all studies, while only reporting detection bias and attrition bias for some[133].

The majority of systematic reviews included evidence with a combination of reviews mostly at low risk or unclear risk of selection bias from randomisation sequence generation and allocation concealment.

Nine reviews identified at least one included study as high risk of random sequence generation[135,143,145–147,149,155,157,159] and thirteen reviews identified at least one study at high risk of bias from allocation concealment[131,132,155,157,159,133,135,143,145–147,149,152]. Trials at high risk of selection bias from random sequence generation or allocation concealment were mostly quasi-randomised trials where randomisation or allocation was affected by alternate or at least in part predictable allocation [133,143,145,147,159], allocation was not used or 'not applicable'[143,152]. Other reasons included where it was felt that the sequence was able to be predicted due to caregivers not being masked [146,159] or block size being fixed[155,159]. Most reviews included studies at unclear risk of randomisation or selection bias due to insufficient descriptions of their methodology, although Dawson *et al.*[132] assessed the only included study as

unclear risk for randomisation as it had only stated "allocated randomly" but and high risk for selection bias as no information was provided.

Most reviews assessed performance and detection bias separately, while four reviews only assessed it as a combined outcome[135–137,145]. Two reviews assessed the combined performance and detection bias outcome separately for clinical and radiological outcomes[136,137]. High risk of bias for performance, detection bias or both was found in eighteen reviews[59,128,149,150,152,154,155,157–159,131,132,135–137,143,145,146]. This was due to lack of blinding or being unlikely to be blinded. Six reviews only assessed reviews as high risk for performance bias[131,132,146,154,157,158] . Detection bias in these cases were assessed as unclear due to insufficient methodological detail[131,132,146]; low risk due to masking of investigators[154]; or lack or unlikely blinding of investigators felt unlikely to influence outcomes[157]. Anabrees *et al.*[158] only assessed for detection bias in one of three studies as it was felt not to be applicable, and did not assess detection bias in the reviews at high risk of performance bias[158].

Incomplete outcome was assessed as high risk in eight reviews [127,131,135,143,149,150,152,159]. Many studies in Premji *et al.*[133] were noted as not having complete follow up[133]. Low risk for incomplete outcome assessment was generally only assessed where outcomes were there was more than 80 or 90% follow up at least for primary outcomes and some studies with 80% follow up for some outcomes[59,127,150–152,128,131,136,137,143,145,147,148]. Three reviews additionally assessed trials as low risk where exclusions were explained or where data from protocol violations were included in the final analysis[131,136,137]. High risk of bias was assessed where there was high loss to follow up[127]; high post-randomisation exclusions [143]; exclusion due to receiving more than 10% human milk in a trial assessing formulas [150]; 19% exclusion without details of group allocation [127]. In Nasuf *et al.*[152], one study was assessed as high risk as a high proportion of infants were excluded and not analysed with information only determined through correspondence with the study author. Fenton *et al.* 2020[149] assessed more than half of studies as high risk of attrition bias due to >50% of infants withdrawing from the study. Watson *et al.*[135] included two trials at high risk of attrition bias due to very high withdrawal due to reasons including failure to pass the feeding tube, and growth data being reported in one study for only infants who tolerated the procedure, which was 71%

in the nasogastric group and 35% in the transpyloric group. Greene *et al.*[159] assessed a study as low risk when reasons provided despite 11/30 infants excluded post randomisation.

Reporting bias was assessed as high risk in six reviews[127,145,147,148,152,159]. Studies were assessed as high risk where not all pre-specified outcomes were reported[145,159], narrative statement of similarity of baseline characteristics without data as evidence[159]; narrative reporting of outcomes[127,147] and a study which declared it did not include head circumference in the analysis when no relationship was seen in data during the study[148].

Other bias was assessed as high risk in three reviews [127,146,159]. Other bias was assessed as high risk where there were discrepant and unreproducible results reported of two reports of the same trial[146], or a trial was specifically to evaluate the specific formula clinically with the manufacturer giving financial assistance[127]. The review of oral stimulation[159] assessed three studies as high risk due to lack of clarity about who made decisions on increasing feeding volume; an included study suspected (but not declared) to have the same study participants as another included study which was previously published, and potential bias from variability between interventions provided by parents, researchers and nursing staff. Unclear risk was assessed where milk formula companies provided funding [127,145,150], inadequate or unclear baseline data [154]. Brown *et al.*[143] does not assess the outcome but described how trials were affected by funding from formula manufacturers.

10.3.1 Discrepancies noted between reviews

Differences between and within reviews were seen in how reviews assessed risk for bias in studies. These discrepancies most often existed between assessing reviews as unclear or high risk.

These included decisions on risk of randomisation and allocation bias where envelopes were not stated to be sealed[131,137] and if failing to state an envelope is opaque puts a study at unclear[143,148,150] or low risk[127,152].

There were discrepancies found with assessment where randomisation method is not described. Most reviews assessed studies as unclear risk if the method of randomisation was not described, but some reviews assessed studies as low

risk without describing this information[133,134], and Dawson *et al.*[132] assessed selection bias as high risk for a study where it justified this with it not having information provided[132]. Fenton *et al.*[149] has inconsistent assessment of bias from random sequence generation, judging one study as unclear risk with the randomisation sequence not described, but four studies as high risk with no randomisation described. This may be as in the study which was assessed as unclear risk for random sequence generation, the allocation concealment was low risk while in the other four studies this was also unclear, but another study with low risk of random sequence generation was assessed as high risk for allocation, performance and detection bias when methodology was unclear[149]. Two studies were assessed as high risk for random sequence generation, but were reasoned with "assigned randomly"[149]. Similarly, 70% of the studies in Fenton *et al.*[149] were assessed as high risk of selection bias, reasoned with allocation concealment not having been described, or "can't tell", but one study was assessed as "unclear risk" with the same reasoning[149]. Anabrees *et al.*[158] only gave justification for assessment of random sequence generation for one of the included studies[158].

Discrepancies of how blinding, descriptions of blinding or lack of detail affected assessment were seen between studies. Some reviews reasoned a declaration of "double blind" with low risk[149,150], but others assessed these as "unclear risk" if this was declared, but it was not described who was blinded[143,144,148]. Premji *et al.*[133] assessed studies where it was "not feasible" for caregivers and investigators to be blinded as unclear risk. Discrepancies between assessments of studies within the review was again seen in Fenton *et al.*[149], where which assessed some studies at high risk of performance and detection bias if they "couldn't tell", but some as "unclear risk" [149]. It also assessed one trial as unclear risk despite reasoning this with "blinding of intervention: yes", and assessed one trial as "high risk" for performance and detection bias combined, but "unclear risk" when assessed separately despite reasoning this with "no blinding of outcome"[149]. Walsh *et al.*[127] assessed one trial where formulas were colour coded as "low risk" but a trial where formulas were identified by numerical code as "unclear risk".

Interesting to consider is where an assessment of "unclear risk" was also made in studies assessing "what to feed" when formulas were only identifiable by numerical code [127] or bottles were "colour coded by manufacturer[149]; and where the

trial was 'double blind' but the investigators acknowledged different tastes, texture or smell between the two formulas[150], or investigators were "unaware" but it was unclear if carers or parents were aware[150].

The level of attrition which assessed as posing a risk varied but was generally low risk when below 10 or 20% or analysis was stated to be performed by intention-to-treat. Yet, some studies were assessed an unclear or high risk of bias below 10%, including high risk at 5% due to early withdrawal due to respiratory problems or metabolic acidosis[149] and unclear risk despite only 1.6% of infants withdrawn from studies[149]. Conversely, in Amissah *et al.*[148], one study was assessed as unclear risk despite only 22% of infants who met criteria completing the study with no details on attrition[148] and Tan-Dy *et al.*[130] responded "yes" to completeness of follow up despite <80% reaching study day 14. Incomplete outcome was also assessed as unclear risk when assessed in studies where impact of baseline differences were unclear [144], insufficient details were available in the abstract [147] or large group imbalances in 10% infants with adverse outcomes not assessed for growth[145].

There was a discrepancy identified between reviews in assessment of reporting bias regarding the availability of a protocol. Amissah *et al.*[146], and Amissah *et al.*[148] assessed reporting bias as low risk if all outcomes were listed in the methods, even if no protocol was available, while other studies assessed studies as unclear risk if there was no protocol available [127,132,144,145,150,152]. Although one study in Walsh *et al.*[127] was assessed as low risk when all outcomes reported in the methods were reported in the results. Other reasons for an 'unclear' assessment included an unpublished where secondary outcomes had not been stated[152], it was unclear which outcomes were primary outcomes[144] or reporting of an additional non pre-specified outcome[148].

Reviews with more than 2 trials were more likely to have included trials with a high risk of bias, which then affected the GRADE assessment.

10.4 **Methodological quality of included reviews**

We rated the quality of the included reviews using the AMSTAR2 and ROBIS tools

10.4.1 AMSTAR2

Reviews with sufficient studies for meta-analysis were assessed in all 16 criteria, reviews with only a single eligible study were assessed for 14 criteria, omitting criteria requiring meta-analysis (Tables 7 and 8). Reviews with no eligible studies were only assessed for 8 criteria on the methodology, exclusions, plan for publication bias assessment and conflict of interest.

We used a high/moderate/low/critically low assessment of methodological quality, as per the guidance provided by the developers of the AMSTAR2 tool[106]. As per AMSTAR2 we decided on critical and non-critical weaknesses to provide an overall confidence rating in the results of the review.

For this we included the criteria suggested by AMSTAR2 which include protocol registered before commencement of the review; adequacy of the literature search; justification for excluding individual studies; risk of bias from individual studies being included in the review; appropriateness of meta-analytical methods; consideration of risk of bias when interpreting the results of the review; and assessment of presence and likely impact of publication bias. In addition to these criteria we also assessed selection of studies performed in duplicate; discussion and assessment of sources of heterogeneity found and management of declared conflict of interest as critical criteria.

Table 7: Summary table of AMSTAR2 Assessment of "What to feed"

Review	PICO complete	A priori methods with no significant deviation [CRITICAL]	Selection of study designs explained	Comprehensive literature search strategy [CRITICAL]	Study selection in duplicate [CRITICAL]	Data extraction in duplicate [CRITICAL]	Exclusions justified [CRITICAL]	Included studies described	Risk of Bias assessed satisfactorily [CRITICAL]	Funding sources reported	Appropriate meta-analysis method [CRITICAL]	RoB impact assessed	RoB impact discussed [CRITICAL]	Heterogeneity impact discussed [CRITICAL]	Publication bias assessed and discussed [CRITICAL]	Conflict of interest reported & managed [CRITICAL]	Overall rating
		[CRITICAL]		[CRITICAL]	[CRITICAL]	[CRITICAL]	[CRITICAL]		[CRITICAL]		[CRITICAL]		[CRITICAL]	[CRITICAL]	[CRITICAL]	[CRITICAL]	
"What is fed"																	
Thanigainathan 2020[128]	Yes	**Yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Yes	**Yes**	*No*	**Yes**	Yes	**Yes**	**Yes**	**Yes**	**Yes**	High
Brown 2020[143]	Yes	**Yes**	No	**Yes**	**No**	**Yes**	**Yes**	Yes	**Partial Yes**	Yes	**Yes**	No	**Yes**	**Yes**	**Yes**	**No**	Critically low
Amissah 2020a[146]	Yes	**Yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Yes	**Yes**	Yes	**n/a**	n/a	**Yes**	**Yes***	**Yes**	**Yes**	High
Amissah 2020b[148]	Yes	**Yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Yes	**Yes**	Yes	**Yes**	Yes	**Yes**	**Yes**	**Yes**	**Yes**	High
Amissah 2020c[144]	Yes	**Yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Yes	**Yes**	Yes	**n/a**	n/a	**Yes**	**Yes***	**Yes**	**Yes**	High
Fenton 2020[149]	Yes	Yes	No	**Yes**	**Yes**	**Yes**	**Yes**	Yes	**Yes**	No	**Yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Moderate
Quigley 2019[145]	Yes	**Yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Yes	**Yes**	Yes	**Yes**	No	**Yes**	**Yes**	**Yes**	**No**	Low
Ng 2019[150]	Yes	**Yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Yes	**Yes**	Yes	**Yes**	No	**Yes**	**No**	**Yes**	**No**	Critically low
Walsh 2019[127]	Yes	**Yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Yes	**Yes**	Yes	**Yes**	Yes	**Yes**	**Yes**	**Yes**	**No**	Low
Dempsey 2019[142]	Yes	**Yes**	No	**Yes**	**Yes**	**n/a**	**Yes**	n/a	**n/a**	n/a	**n/a**	n/a	**n/a**	**n/a**	**Yes**	**Yes**	High
Brown 2019[126]	Yes	**Yes**	No	**Yes**	**Yes**	**n/a**	**Yes**	n/a	**n/a**	n/a	**n/a**	n/a	**n/a**	**n/a**	**Yes**	**Yes**	High
Basuki 2019[147]	Yes	**Yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Yes	**Yes**	No	**Yes**	Yes	**Yes**	**No**	**Yes**	**Yes**	Low
Premkumar 2019[151]	Yes	**Yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Yes	**Yes**	No	**n/a**	n/a	**Yes**	**Yes***	**Yes**	**Yes**	High
Nasuf 2018[152]	Yes	**Yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Yes	**Yes**	Yes	**Partial yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Critically low
Abiramalatha 2017[59]	Yes	**Yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Yes	**Yes**	No	**n/a**	n/a	**Yes**	**Yes***	**Yes**	**No**	Low
Tan-Dy 2013[130]	Yes	**Partial Yes**	No	**Yes**	**No**	**Yes**	**Yes**	Yes	**Yes**	No	**n/a**	n/a	**Yes**	**Yes***	**Yes**	**Yes**	Critically low
Nehra 2002[129]	Yes	**Partial Yes**	No	**No**	**Yes**	**Yes**	**Yes**	Yes	**Partial Yes**	No	**Yes**	No	**Yes**	**Yes**	**No**	**Yes**	Low

Table 8: Summary table of AMSTAR2 Assessment of "When to feed", "How to feed" and "Adjunctive strategies".

Review	PICO	**A priori methods + no significant deviation**	Selection of study designs	**Compre-hensive literature search strategy**	**Study selection in duplicate**	**Data extraction in duplicate**	**Exclusion justified**	Included studies described	**Risk of Bias assessed satisfactorily**	Funding sources reported	**Appropriate meta-analysis method**	RoB impact assessed	**RoB impact discussed**	**Heterogeneity impact discussed**	**Publication bias assessed and discussed**	**Conflict of interest managed**	Overall rating
		[CRITICAL]		[CRITICAL]	[CRITICAL]	[CRITICAL]	[CRITICAL]		[CRITICAL]		[CRITICAL]		[CRITICAL]	[CRITICAL]	[CRITICAL]	[CRITICAL]	
"When is fed"																	
Yeo 2019[154]	Yes	**Yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Yes	**Partial yes**	No	**n/a**	n/a	**Yes**	**Yes***	**Yes**	**Yes**	Low
Abiramalatha 2019a [155]	Yes	**Yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Yes	**Yes**	No	**Yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Moderate
Oddie 2017 [136]	Yes	**Partial yes**	No	**Yes**	**No**	**Yes**	**Yes**	Yes	**Partial yes**	No	**Yes**	No	**Yes**	**Yes**	**Yes**	**No**	Critically low
Morgan 2014[137]	Yes	**Yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Yes	**Partial yes**	No	**Partial yes**	No	**Yes**	**No**	**No**	**Yes**	Critically low
Morgan 2013[131]	Yes	**Partial yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Yes	**Partial yes**	No	**Yes**	No	**Yes**	**Yes**	**No**	**Yes**	Critically low
"How is fed"																	
Abiramalatha 2019b [156]	Yes	**Yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Paritial Yes	**Yes**	No	**n/a**	n/a	**Yes**	**Yes**	**Yes***	**Yes**	Moderate
Watson 2013 [135]	Yes	**No**	No	**Yes**	**Yes**	**Yes**	**Yes**	Yes	**Partial Yes**	No	**Yes**	Yes	**Yes**	**Yes**	**Yes**	**Yes**	Critically low
Dawson 2012[132]	Yes	**No**	No	**Yes**	**Yes**	**Yes**	**Yes**	Yes	**Yes**	No	**n/a**	n/a	**Yes**	**Yes**	**Yes***	**Yes**	Low
Premji *et al.* 2011[133]	Yes	**Yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Yes	**Partial Yes**	No	**Yes**	No	**Yes**	**Yes**	**No**	**Yes**	Critically low
"Adjunctive strategies"																	
Muelbert 2019[157]	Yes	**Yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Yes	**Yes**	Yes	**Yes**	Yes	**Yes**	**Yes**	**Yes**	**Yes**	High
Greene 2016[159]	Yes	**Yes**	No	**Partial Yes**	**Yes**	**Yes**	**Yes**	Yes	**Yes**	No	**Yes**	No	**Yes**	**No**	**No**	**Yes**	Critically low
Anabrees 2015[158]	Yes	**Partial Yes**	No	**Yes**	**Yes**	**Yes**	**Yes**	Yes	**Yes**	Yes	**Yes**	Yes	**Yes**	**Yes**	**No**	**Yes**	Critically low
Ng 2008[134]	Yes	**Partial Yes**	No	**Yes**	**No**	**Yes**	**Yes**	Partial Yes	**Partial Yes**	No	**No**	No	**Yes**	**Yes**	**No**	**Yes**	Critically low

10.4.1.1 High overall confidence

On consideration of the individual weaknesses in the reviews, eight reviews were assessed as high overall confidence in the results of the review as they only had one non-critical flaw of not explaining their reason for selection of the study designs for inclusion in the review[126,128,142,144,146,148,151,157], of which two reviews had planned good methodology but did not have any eligible studies for inclusion[126,142].

10.4.1.2 Moderate overall confidence

Three reviews were assessed as moderate as they had not satisfied more than one non-critical criterion[149,155,156]. These criteria included no explanation of the chosen study designs eligible for inclusion[149,155,156]; not assessing or planning to assess the impact of risk of bias by sensitivity analysis[149,155], not describing the site of study[156]; and not reporting funding sources of included studies[149,155,156].

10.4.1.3 Low overall confidence

Seven reviews were assessed as low as they included one critical criterion along with non-critical criteria[59,127,129,132,145,147,154]. None included an explanation of the study design included. Two did not assess the impact of risk of bias by sensitivity analysis of risk of bias [129,145], and four had no funding sources reported[59,132,147,154]

The critical criterion was not explaining how a declared conflict of interest was managed which for two reviews included one review author having previous research funding from formula manufacturers[127,145] and for one review was a review author as principle investigator in the only included study in the review [59]. Other critical criterion was not acknowledging or discussing the high heterogeneity found on meta-analysis for head circumference[147], inconsistent heterogeneity assessment[129], not performing risk of bias assessment for selective reporting[132,154] or not declaring changes to the protocol from outcome measures which may impact the results of the review[132].

10.4.1.4 Critically low overall confidence

Twelve reviews were assessed as critically low as they contained violations of multiple critical criteria[130,131,158,159,133–137,143,150,152].

Four reviews had an unresolved conflict of interest, from either an author with previous funding sources from formula manufacturers[143,150,152] or authors were investigators of a large included study with no explanation of how this was managed[136].

Four reviews did not meet the critical criterion of adequate study selection methodology as only one review author assessed all articles[136,143], or it was unclear if it was performed independently with consensus reached[130] or in duplicate[134].

Three reviews did not meet criteria for assessment of sources for heterogeneity due to inconsistent reporting of the presence in the results[150], inconsistent heterogeneity investigation by subgroup analysis [137] or no assessment of the source of heterogeneity[134,159].

Six reviews did not meet the critical criterion for risk of bias assessment as no selective reporting assessment was performed[131,133,135–137,143] or there was only a risk of bias assessment for allocation concealment with only a narrative statement of blinding[134].

Six reviews did not meet criteria due to significant deviations from protocol methodology[129–131,135,136,158]. These included omissions of selective reporting bias assessment as per protocol[136]; no statement of changes to review methodology from previous reviews[131]; additional outcomes reported but not declared[129,130]; changes to the reported outcome measures which may impact the results of the review[135,158] and additional eligibility of cluster-randomised controlled trials[158]. In the latter study, no additional studies were accepted for this reason[158].

Six reviews did not meet the critical criterion of publication bias assessment, as no plan for this was made[129,131,133,137,158,159].

Two reviews were assessed as having an incomprehensive literature search strategy due to limitation by language of publication restricted to English[129] or did not search the bibliographies of included studies[159].

None included an explanation of study design included. Other non-critical violations of criteria in these reviews were no reporting of funding sources of included studies[130,133–136,159], and no assessment of the impact of risk of bias on meta-analysis effect estimate[133,134,136,143,159].

10.4.2 *Risk of Bias in Systematic reviews (ROBIS)*

ROBIS was assessed in three phases. A summary of assessment is seen in table 9-12. Appendix 4 contains complete details of assessment of the initial two phases of assessment.

All reviews were assessed as appropriately considering the relevance of identified studies to the review's research question and avoiding emphasizing the results based on their statistical significance. All reviews summarised findings and discussed results in the context of factors affecting the risk of bias, characteristics of the assessed population and interventions.

Twelve reviews in 'what to feed'[59,126,151,152,127,128,142,144–146,148,149], two reviews in 'when to feed'[154,155], one review in 'how to feed'[156], two reviews in 'adjunctive strategies'[157,159] were assessed as low concern for risk of bias within the review.

Six reviews were assessed as low risk despite unclear concerns in the domains of the second phase[134,144,148,152,155,159]. Two reviews were assessed as low risk as the reason for the unclear risk of bias had been addressed in the reviews[144,148]. One review[152] was assessed as low risk of bias, despite not having addressed the concern in domain 4 within the interpretation of findings, as no subgroup analyses were performed or omission justification, it was not felt to have introduced significant bias. Although one review[155] did not specifically discuss the moderate heterogeneity found, this was assessed as low concern as the quality of evidence was overall acknowledged as low on GRADE assessment. Green *et al.*[159] was unclear about the methods to minimise error in the risk of bias assessment but this was felt to be low risk in the review overall.

Twelve reviews were assessed as unclear concern for the overall judgement of risk of bias within the review[129,130,150,158,131,133–137,143,147].

Seven reviews remained unclear due to concerns about the study selection being performed in duplicate or independently[130,131,134–137,143]. Three reviews[137,147,150] were assessed as unclear as they did not address the heterogeneity observed in discussion of the findings. Eleven reviews[129,130,158,131,133–137,143,147] remained unclear concern as the risk of bias tool had not assessed all areas of bias, most frequently not assessing selective reporting or other bias. In addition to the out of date risk of bias assessment

in Ng *et al.*[134] , although assessed and described, but concerns on independent screening assessment remained. Four reviews remained unclear due to uncertainty of the rigour of the risk of bias assessment[131,133,135,137]. Nehra *et al.*[129] was unclear due to the lack of explanation for language restriction, inconsistent risk of bias assessment. Unclear also remained in two reviews due to the absence of a plan for sensitivity analysis based on bias or to look for publication bias[129,131].

Three reviews continued to have concerns about undeclared changes to the methodology due to changes in outcomes assessed from those predefined[132,135,158] and the reason behind their change[132]. Anabrees *et al.*[158] also continued to have unresolved concerns due to the heterogeneity to length of stay not discussed when summarising the results of the review analyses[158].

The nature of the changes to the outcomes in Dawson *et al.*[132] were felt to be significant, and therefore the overall risk of bias was felt to be high concern.

Table 9: ROBIS summary table (What to feed)

Review	Phase 1: Review question matches target question	Phase 2: ROBIS Domains				Overall Risk of Bias (RoB) Assessment				
		Study eligibility criteria	Identification and selection of studies	Data collection & study appraisal	Synthesis and findings	A: Concerns addressed	B: Relevance considered	C: No emphasis on statistical significance	RoB in review	Rationale for risk:
Thanigainathan *et al.* 2020 [128]	Yes	Low concern	Low concern	Low concern	Low concern	Yes	Yes	Yes	**Low concern**	n/a
Brown *et al.* 2020 [143]	Yes	Low concern	Unclear concern	Unclear concern	Low concern	Yes	Yes	Yes	**Unclear concern**	Selection of reviews not fully duplicated and performed independently. Incomplete RoB assessment.
Amissah *et al.* 2020a[146]	Yes	Low concern	Low concern	Low concern	Low concern	Yes	Yes	Yes	**Low concern**	n/a
Amissah *et al.* 2020b [148]	Yes	Low concern	Low concern	Unclear concern	Low concern	Yes	Yes	Yes	**Low concern**	n/a
Amissah *et al.* 2020c [144]	Yes	Low concern	Low concern	Unclear concern	Low concern	Yes	Yes	Yes	**Low concern**	n/a
Fenton *et al* 2020.[149]	Yes	Low concern	Low concern	Low concern	Low concern	Yes	Yes	Yes	**Low concern**	n/a
Quigley *et al.* 2019 [145]	Yes	Low concern	Low concern	Low concern	Low concern	Yes	Yes	Yes	**Low concern**	n/a
Ng *et al.* 2019[150]	Yes	Low concern	Low concern	Low concern	Unclear concern	No	Yes	Yes	**Unclear concern**	High heterogeneity (I^2=82% for head circumference not acknowledged or discussed)
Walsh *et al.* 2019[127]	Yes	Low concern	Low concern	Low concern	Low concern	Yes	Yes	Yes	**Low concern**	n/a
Dempsey *et al.* 2019 [142]	Yes	Low concern	Low concern	Low concern	Low concern	Yes	Yes	Yes	**Low concern**	n/a

Review	Phase 1: Review question matches target question	Phase 2: ROBIS Domains				Overall Risk of Bias (RoB) Assessment				
		Study eligibility criteria	Identification and selection of studies	Data collection & study appraisal	Synthesis and findings	A: Concerns addressed	B: Relevance considered	C: No emphasis on statistical significance	RoB in review	Rationale for risk:
Brown *et al.* 2019 [126]	Yes	Low concern	Low concern	Low concern	Low concern	Yes	Yes	Yes	**Low concern**	n/a
Basuki *et al.* 2019 [147]	Yes	Low concern	Low concern	Unclear concern	Unclear concern	No	Yes	Yes	**Unclear concern**	Heterogeneity (primary outcome of abdominal distension as a marker of feeding intolerance) not acknowledged or discussed. Incomplete RoB assessment.
Premkumar *et al.* 2019[151]	Yes	Low concern	Low concern	Low concern	Low concern	Yes	Yes	Yes	**Low concern**	n/a
Nasuf *et al.* 2018 [152]	Yes	Low concern	Unclear concern	Low concern	Unclear concern	No	Yes	Yes	**Low concern**	No publication bias or sensitivity analysis performed but justified/rationale recognised. No subgroup analyses performed, no justification but not felt to have introduced significant bias.
Abiramalatha *et al.* 2017[59]	Yes	Low concern	Low concern	Low concern	Low concern	Yes	Yes	Yes	**Low concern**	n/a
Tan-Dy *et al.* 2013 [130]	Yes	Unclear concern	Unclear concern	Unclear concern	Low concern	No	Yes	Yes	**Unclear concern**	Out of date review, unclear if independent study screening, risk of bias tool not assessing all areas of bias.
Nehra *et al.* 2002 [129]	Yes	Unclear concern	Unclear concern	Unclear concern	Unclear concern	No	Yes	Yes	**Unclear concern**	No explanation for language restriction, inconsistent risk of bias assessment, sensitivity analysis not performed, no plan to assess publication bias.

Table 10: ROBIS Summary Table (When to feed)

Review	Phase 1: All PICO Criteria match?	Phase 2: ROBIS Domains				Overall Risk of Bias Assessment				
		Study eligibility criteria	Identification and selection of studies	Data collection and study appraisal	Synthesis and findings	A: Concerns addressed	B: Relevance considered	C: Not emphasised on statistical significance	Risk of bias in the review	Rationale for risk:
Yeo *et al.* 2019 [154]	Low concern	Low concern	Low concern	Low concern	Low concern	Yes	Yes	Yes	**Low concern**	n/a
Abiramalatha *et al.* 2019a [155]	Low concern	Low concern	Low concern	Lowconcern	Unclear concern	No	Yes	Yes	**Low concern**	Moderate heterogeneity not discussed but quality of evidence already acknowledged as low.
Oddie *et al.* 2017 [136]	Low concern	Unclear concern	Unclear concern	Unclear concern	Low concern	No	Yes	Yes	**Unclear concern**	Selection of reviews not fully duplicated and performed independently. Incomplete RoB assessment.
Morgan *et al.* 2014 [137]	Low concern	Low concern	Unclear concern	Unclear concern	Unclear concern	No	Yes	Yes	**Unclear concern**	Unclear if study selection performed independently, unclear if risk of bias assessment performed in duplicate; Substantial heterogeneity for duration of hospital stay, acknowledged but not discussed or explored as planned in protocol. Incomplete RoB assessment.
Morgan *et al.* 2013 [131]	Low concern	Unclear concern	Unclear concern	Unclear concern	Unclear concern	No	Yes	Yes	**Unclear concern**	Unclear if study selection performed independently; unclear if risk of bias assessment performed in duplicate; no plan to check for publication bias/sensitivity analysis based on bias. Incomplete RoB assessment.

Table 11: ROBIS Summary table (How to feed)

Review	Phase 1: All PICO Criteria match?	Phase 2: ROBIS Domains				Overall Risk of Bias Assessment				
		Study eligibility criteria	Identification and selection of studies	Data collection and study appraisal	Synthesis and findings	A: Concerns addressed	B: Relevance considered	C: Not emphasised on statistical significance	Risk of bias in the review	Rationale for risk:
Abiramalatha *et al.* 2019b [156]	Low concern	Low concern	Low concern	Low concern	Low concern	Yes	Yes	Yes	**Low concern**	No concerns
Watson *et al* 2013.[135]	Low concern	Unclear	Unclear	Unclear	High Concern	No	Yes	Yes	**Unclear**	Unclear if study selection and risk of bias assessment were performed in duplicate and independently. Unclear if change to methods were predefined. Incomplete RoB assessment
Dawson *et al.* 2012 [132]	Low concern.	High Concern	Low concern.	Low concern.	High Concern	No	Yes	Yes	**High concern**	Deviations from protocol (changes to outcomes assessed) not explained.
Premji *et al.* 2011 [133]	Low concern	Low concern	Low concern	Unclear	Low concern	No	Yes	Yes	**Unclear concern**	Unclear if risk of bias assessment was performed in duplicate. Incomplete RoB assessment.

Table 12: ROBIS Summary Table (Adjunctive strategies)

Review	Phase 1: All PICO Criteria match?	Phase 2: ROBIS Domains				Overall Risk of Bias Assessment				
		Study eligibility criteria	Identification and selection of studies	Data collection and study appraisal	Synthesis and findings	A: Concerns addressed	B: Relevance considered	C: Not emphasised on statistical significance	Risk of bias in the review	Rationale for risk:
Muelbert *et al.* 2019 [157]	Low concern	Low concern	Low concern	Low concern	Low concern	Yes	Yes	Yes	**Low concern**	Review addressed concerns about primary studies and assessed quality of evidence as very low.
Greene *et al.* 2016 [159]	Low concern	Low concern	Low concern	Unclear concern	Low concern	Yes	Yes	Yes	**Low concern**	Methods to minimised error in risk of bias assessment unclear but felt low risk of bias in review overall.
Anabrees *et al.* 2015 [158]	Low concern	High Concern	Low concern	Unclear concern	High Concern	No	Yes	Yes	**Unclear**	Changes to outcomes assessed not declared Heterogeneity <75% not discussed. Inconsistent RoB assessment.
Ng *et al.* 2008 [134]	Low concern	Low concern	Unclear concern	Unclear concern	Low concern	No	Yes	Yes	**Unclear**	Out of date review, but risk of bias assessed and described, but incomplete. Unclear if study selection performed independently.

10.5 Effects of the intervention: Necrotising enterocolitis

A summary of the effect estimates of the interventions on growth is presented in appendix 5.

10.5.1 What is fed

Thirteen reviews defined the criteria for necrotising enterocolitis[59,127,151,152,128,130,142,143,145,147,149,150]. Four reviews defined necrotising enterocolitis as that confirmed at surgery or autopsy of diagnosed by at least two of the following clinical features - pneumatosis intestinalis, portal venous gas or free air in the abdomen on abdominal x-ray; abdominal distention with gaseous distention and/or frothy appearance of bowel lumen on abdominal x-ray; blood in stool; lethargy hypotonia or apnoea or a combination of these [127,143,145,147]. Nine reviews defined necrotising enterocolitis as that defined by modified Bell's stage 2 or greater[59,128,130,142,143,149–152].

Four reviews did not define criteria for necrotising enterocolitis[129,144,146,148].

The quality of the evidence by GRADE was not assessed in two reviews[129,130], and was assessed by overview authors (WM and VW).

10.5.1.1 Statistically significant difference demonstrated

There was moderate certainty evidence shown of a statistically significant increased risk of necrotising enterocolitis in infants fed formula milk compared with infants fed donor breastmilk[145]. The evidence was downgraded for imprecision. Subgroup analysis was performed for types of formula and fortifier, with only preterm formula versus fortified DBM marginally retaining a statistically significant increase in necrotising enterocolitis in the formula fed group (RR 1.64 [1.03, 2.61], I^2 = 51%). The subgroups of term formula versus unfortified DBM (RR 4.73 [0.52, 43.09], single study) and preterm formula versus unfortified DBM (RR 2.99 [0.90, 9.87], I^2 = 0%) had higher point estimates, but very wide confidence intervals spanning both risks and benefits. Subgroup analysis of the intervention used as a sole diet or supplement to maternal breastmilk were performed. Only sole diet achieved statistical significance (RR 4.62 [1.47, 14.56], I^2 = 0%) with a more imprecise estimate, while as a supplement it had a more precise estimate but whose confidence

interval suggested some possibility of harm as well as benefit of donor breastmilk as a supplement (RR 1.56 [0.98, 2.47] I^2=36%). The subgroup difference was near but did not reach statistical significance.

10.5.1.2 Statistically significant difference not demonstrated

Low certainty evidence for no statistically significant effect of degree of hydrolysis[150], nutrient enriched or standard formula[127], human derived versus bovine milk derived fortifier[151], addition of multi-nutrient fortification[143], early versus late fortification of human milk[128] and treatment of feeds with lactase[130] on necrotising enterocolitis limited by the risk of bias within the trials. Evidence was also downgraded on the GRADE assessment due to imprecise effect size estimate[128,130,143,150,151], post hoc exclusions in two trials[127], inclusion of only a single study[151] and uncertainty about risk of bias from trial methodology[130,143] and blinding[128,143]. Subgroups where able to be assessed in these reviews also did not show statistically significant differences.

Very low certainty evidence showed no statistically significant effects of carbohydrate[146] or protein supplementation[148] and oropharyngeal colostrum[152], high versus standard volume feeds[59], or level of MCT in formula[129] on necrotising enterocolitis. Quality of evidence was limited by concerns of risk of bias from allocation and blinding[59,152], insufficient methodological information[129,146], few participants, events or sample sizes[146,148,152] and imprecision from wide confidence intervals crossing the line of no effect[59,146,148,152] or limited narrative information from withdrawal of infants[129]. The risk of bias for high versus low medium chain triglyceride formula was also limited by inconsistency as one study compared two levels of MCT in formula, while the other study compared four different levels[129]. Subgroup analyses were not possible in these reviews, although no significant difference in weight gain was reported with no standard deviations for a subgroup of small for gestational age infants who were fed high or standard volume feeds in the included study.

Evidence assessing high versus low protein intake[149] on the risk of necrotising enterocolitis was found to be not estimable. The evidence was graded very low certainty due to concerns of risk of bias and imprecision. The authors were uncertain

about what criteria were used to define necrotising enterocolitis within the included studies. No subgroup analyses were possible.

No eligible studies were found assessing necrotising enterocolitis in trials comparing banked preterm versus banked human term milk[142], maternal breastmilk versus formula[126], dilute versus full-strength formula[147], or fat supplementation of human milk[144].

The review assessing early versus late fortification of human milk feeds[128] also assessed for surgical NEC. Low certainty evidence demonstrated no statistically significant difference in surgical NEC between infants who received early versus late fortification of human milk. The review downgraded the evidence due to risk of bias from lack of blinding, and serious imprecision from a small sample size (237 infants). They were unable to perform subgroup analyses due to inadequate data.

10.5.2 When to feed

Three reviews[131,136,137]. defined necrotising enterocolitis as that confirmed at surgery or autopsy of diagnosed by at least two of the following clinical features (pneumatosis intestinalis, portal venous gas or free air in the abdomen on abdominal x-ray; abdominal distention with gaseous distention and/or frothy appearance of bowel lumen on abdominal x-ray; blood in stool; lethargy hypotonia or apnoea or a combination of these) Two reviews[154,155] defined necrotising enterocolitis as that defined by modified Bell's stage 2 or greater.

Yeo *et al.*[154] assessed necrotising enterocolitis both within 48 hours after a blood transfusion and any episode after the blood transfusion. The review assessing routine monitoring of gastric residuals[155] also assessed number of infants with surgical NEC only.

No GRADE assessment was performed in two reviews[131,137]. GRADE was also not assessed for the effect of routine monitoring of quality and quantity of gastric residuals on any NEC, or the review's assessment of either comparison in for surgical NEC only[155]. These were conducted by authors during the overview.

None of the reviews assessing when to feed found a statistically significant difference in risk of necrotising enterocolitis[131,136,137,154,155].

The evidence for risk of necrotising enterocolitis with the rate of feed advancement was of moderate certainty, limited by risk of bias from all trials being unblinded[136]. Subgroup analyses also did not show a statistically significant difference.

Low certainty evidence for risk of necrotising enterocolitis for continuing feeds during transfusion was seen both within 48 hours of a blood transfusion and at any time after the first blood transfusion[154]; and presence or absence of routine monitoring of gastric residuals[155]. Evidence in both reviews was limited by small sample size and low[155] or no[154] event rate. There was also lack of blinding in the included studies, yet this was not included in the downgrading as part of the GRADE assessment. The comparisons of routine monitoring and specific routine monitoring of quality and quantity[155] also showed very wide confidence intervals. No subgroup analyses were possible for these comparisons.

The evidence assessing risk of necrotising enterocolitis with delayed versus early introduction of progressive feeding[137] or early trophic feeding versus enteral fasting[131] were both assessed as very low certainty. The evidence was affected by risk of bias from uncertain methodology and lack of blinding affecting performance bias, and serious imprecision[131,137]. Radiological assessment of necrotising enterocolitis in trials assessing early introduction of progressive feeding was, however, noted as masked in 3/8 studies[137]. Authors reviewing early trophic feeding additionally express concerns that findings may not be applicable to infants at highest risk of necrotising enterocolitis[131]. No statistically significant subgroup difference was demonstrated on subgroup analysis for introduction of progressive feeding in infants with IUGR or abnormal antenatal doppler flow velocities[137].

10.5.3 How is fed

One review defined necrotising enterocolitis as that confirmed at surgery or autopsy of diagnosed by at least two of the following clinical features (pneumatosis intestinalis, portal venous gas or free air in the abdomen on abdominal x-ray; abdominal distention with gaseous distention and/or frothy appearance of bowel lumen on abdominal x-ray; blood in stool; lethargy hypotonia or apnoea or a combination of these)[135]. Three reviews defined necrotising enterocolitis as that defined by modified Bell's stage 2 or greater[132,133,156]. For the review assessing re-

feeding gastric residuals the reported result was for NEC stage 2/3 or spontaneous intestinal perforation (SIP)[156].

No statistically significant differences were found for assessment of NEC in any of the included reviews with data assessing how to feed infants[132,133,135,156]. No eligible studies were found to assess difference in risk of necrotising enterocolitis in push versus gravity bolus tube feeding[132].

The certainty of evidence for the effect of all interventions on how to feed infants was assessed as very low certainty[132,133,135,156]. GRADE was only assessed by review authors for re-feeding versus discarding gastric residuals[156], all other assessments were conducted as part of the overview process.

Evidence assessing re-feeding versus discarding gastric residuals on risk of necrotising enterocolitis was limited due to serious imprecision from very wide confidence intervals and risk of detection bias from study authors including both NEC and SIP[156]. The two subgroup analyses performed for infants fed only human milk and only formula milk both did not show a significant subgroup difference or statistically significant estimate.

The evidence for NEC between transpyloric and gastric tube feeding[135] was significantly affected by high risk of bias across all domains, serious indirectness and very serious imprecision from included data mostly from reasons for study withdrawal.

Evidence comparing NEC in continuous versus intermittent bolus milk feeding[133] was limited by serious imprecision; and very serious risk of bias from uncertain performance and detection bias in most studies and incomplete outcomes from infant withdrawal affecting the results. It assessed proven and probable separately, both of which had serious imprecision from wide confidence intervals with significant benefit and harm, but also narratively described a further study with NEC where one case of each proven and probable NEC occurred in the continuous feeding group, but none in the bolus feeding group. Serious heterogeneity from significant heterogeneity affected the certainty of results on assessment of probable NEC. Subgroup analyses in this review were performed as separate comparisons. None

demonstrated a significant result. Concerns about sample size calculation method were also expressed by review authors.

10.5.4 Adjunctive strategies

Three reviews defined NEC as defined by Bell's stage 2 or more [157–159]. Ng *et al.*[134] defined NEC as the presence of pneumatosis intestinalis and/or portal air on an abdominal X-ray.

No statistically significant difference in risk of necrotising enterocolitis was found for exposure versus no exposure to smell and taste stimulation of milk with tube feeds[157], glycerin prophylaxis[158] or giving erythromycin both for prevention or treatment of feeding intolerance[134]. No eligible studies assessing this outcome for comparisons with oral stimulation[159].

GRADE assessment was only performed by review authors for the comparison with exposure to smell and taste stimulation[157]. This assessed the evidence as low certainty due to imprecision from a single trial with small sample size and wide confidence intervals. It judged that lack of blinding would have been unlikely to influence this outcome.

The evidence for other adjunctive strategies on NEC were assessed by overview authors as very low certainty. The evidence in for glycerine prophylaxis was affected by serious risk of bias, and very serious imprecision[158]. Although there was no statistical heterogeneity, differences in preparations and dosing regimens were seen.

The evidence comparing erythromycin with placebo for preventing feeding intolerance was affected by very serious imprecision from wide confidence interval and serious risk of bias from uncertainty about performance and detection bias[134]. Evidence for erythromycin for treating feeding intolerance was similarly limited by risk of bias, but also very serious imprecision from narrative reporting or data unable to be combined in meta-analysis, and serious inconsistency from heterogeneity between study in design, methodology, population and variability in the dosages and duration. Only one of the studies in this review had defined necrotising enterocolitis and inclusion criteria of feeding intolerance differed widely between the studies.

10.6 Effects of the intervention: Growth

A summary of the effect estimates of the interventions on growth is presented in appendix 6.

10.6.1 What is fed

Short term growth was measured as rate of weight gain (g/kg/day or g/day), time to regain birthweight.

10.6.1.1 Weight gain (g/kg/day)

Nine reviews presented short term growth as rate of weight gain in g/kg/day[59,127,129,143–146,148–150].

There was moderate certainty evidence found of higher weight gain in infants fed formula milk to donor breastmilk[145] (MD 2.51 g/kg/d [1.93, 3.08] I^2 = 90%) with significant subgroup differences (I^2 =83.3%). A higher mean difference was found in infants fed preterm formula versus unfortified DBM (MD 4.16 g/kg/d [3.04, 5.28] I^2 =94%), and lower but still significant mean difference in infants fed term formula versus unfortified DBM (MD 1.74 g/kg/d [0.96, 2.53] I^2 =94) and preterm formula versus fortified DBM (MD 2.37 g/kg/d [1.09, 3.65] I^2 =0%). The evidence was downgraded to moderate due to the high level of heterogeneity.

There was low certainty evidence of an increased weight gain in infants fed fortified breastmilk when compared with unfortified breastmilk (MD 1.76 g/kg/day [1.30, 2.22]), which was higher on subgroup analysis of only very preterm or very low birthweight infants (MD 2.18 g/kg/day [1.54, 2.81])[143]. Evidence was limited by unexplained heterogeneity and uncertainty from risk of bias from selection, performance and detection bias in most trials.

There was low certainty evidence of higher weight gain in infants fed nutrient-enriched formula compared with standard formula[127] (MD 2.43 g/kg/d [1.60, 3.26], I^2=46%), with a subgroup difference present (I^2=76.2%). This difference was higher in infants fed formula as a sole diet (MD 3.87 [2.26, 5.47], I^2=20%) and reduced but still present in infants fed the formula as a supplement (MD 1.90 [0.93, 2.87], single study). This was downgraded due to uncertainty about risk of bias from methodology, and due to the moderate to high heterogeneity.

There was low certainty evidence of a statistically significant higher rate of weight gain in infants fed higher volume versus standard volume of feeds[59] (MD 6.2 g/kg/day [2.71, 9.69]). The quality of evidence was downgraded due to lack of blinding in the single study, and imprecision of the effect estimates. The comparison of the volumes in the only included trial were 300ml/kg/day versus 200ml/kg/day.

There was low certainty evidence of a higher rate of weight gain in infants receiving protein supplemented human milk compared with no supplementation[148] (MD 3.82 [2.94, 4.70], I^2=73%). The evidence was downgraded due to uncertainty about the trial methodology and moderate to high heterogeneity. There was insufficient data to conduct subgroup analyses according to gestational age, birthweight, sex and type of supplement.

There was low certainty evidence of a higher rate of formula fed infants receiving a high protein intake (3≤x≤4 g/kg/day) compared with a low intake (<3 g/kg/day) [149] (MD 2.36 [1.31, 3.4] I^2=57%), downgraded due to concerns about risk of bias and moderate heterogeneity. This outcome was not able to be measured in the very high (≥4 kcal/kg/day) versus low and very high versus high protein intake comparisons.

Low certainty evidence showed infants fed hydrolysed formula gained weight slower (g/kg/day) compared with infants fed non-hydrolysed formula[150] (MD -3.02 [-4.66, -1.38] I^2=19%). There was insufficient data for subgroup analyses. Data is limited by serious imprecision, and a serious risk of bias. The two larger included trials were mostly low risk and the third mostly unclear risk in the assessed categories, but the review notes that the included trials were supported by the formula manufacturers and although they were not involved in trial design or analysis there remains some concern of potential for selective reporting of study findings of their specialist formulas.

There was very low certainty evidence of no statistically significant differences in weight gain with fat supplementation of human milk[144] or different levels of medium chain triglycerides in formula, in each comparison made[129]. Evidence for fat supplementation of human milk was downgraded by review authors for uncertain risk of bias from trial methodology, imprecision from a very small sample size of 14 from a single study, and a wide confidence interval spanning across benefits and harms[144]. Evidence for formula medium chain triglyceride content was assessed by

overview authors to be limited by risk of bias from uncertainty about the details of sequence generation and blinding, serious clinical inconsistency between trials despite no statistical heterogeneity, and serious imprecision[129]. The overview authors note risk of bias introduced from review authors restricting eligible infants by the language restriction in their study search[129].

10.6.1.2 Other short term weight gain assessments

Low certainty evidence showed a statistically significant reduction in time to regain birthweight (MD -3.08 days [-4.38, -1.77]) in the formula group compared with infants fed donor breastmilk[145].A subgroup difference showed higher improvement in infants fed formula was higher in infants fed term formula versus unfortified DBM (MD -4.00 days [-5.81, -2.18]), than infants fed preterm formula versus unfortified DBM (MD -2.10 [-3.97, -0.23]). No eligible studies assessed the subgroup of preterm formula versus fortified DBM. Evidence was limited by very serious risk of bias from unblinded trials and high risk of selection bias in one study using term formula by allocation of every 5th infant was to pooled breast milk. The reduced point estimate for the subgroup estimate of preterm formula versus unfortified DBM is surprising, but the review states in the characteristics of studies that the "Fortification policy was not described" and the author was contacted, but the response is not reported therefore there is a potential that the donor breastmilk may be fortified, therefore reducing the anticipated difference in nutrient content between the two interventions.

Very low certainty evidence demonstrated increased weight at 30 days of life of between 12.4-308.4g in infants fed milk with a carbohydrate (prebiotic) supplement versus no supplement[146]. This evidence was downgraded due to uncertainties about the methodology and lack of blinding, indirectness due to lack of clarity on dose, frequency and duration of administration, imprecision due to small sample sizes, few participants and events and wide confidence intervals.

Low certainty evidence did not demonstrate a statistically significant difference in time to regain birthweight (days), or relative risk of extra-uterine growth restriction at discharge between infants who had fortification of breastmilk introduced earlier than 100ml/kg/day than at 100 ml/kg/day[128]. Evidence for both was downgraded for risk of bias from lack of blinding, and serious imprecision from a small sample size

(237 infants). Subgroup analyses unable to be performed due to inadequate data. The discharge criteria at which extra-uterine growth restriction was measured in the included studies used were not reported in the review.

Low certainty evidence did not demonstrate a statistically significant difference in weight gain at 7 days comparing infants fed dilute versus full strength formula[147]. Evidence was limited by serious imprecision and serious risk of bias from unclear blinding and reporting bias from lack of protocol.

Low certainty evidence did not demonstrate a statistically significant weight gain (g) during the intervention or weight for age z scores in infants fed human derived milk fortifier compared with bovine milk-derived fortifier[151]. The review downgraded the evidence due to imprecision from large confidence intervals and inclusion of a single study. The single trial included was low risk of bias in all areas assessed.

Despite evidence of increased mean weight gain in g/kg/day, low certainty and very low certainty evidence showed no difference in weight at term-equivalent age and weight at the end of the study (at 2200g or when breastfeeding was initiated), respectively, for infants fed human milk supplemented with protein compared with unsupplemented milk[148]. Evidence for term-equivalent age was limited by serious risk of selection and detection bias, and serious imprecision[148]. Evidence for weight at the end of study was limited by very serious risk of bias, with uncertain risk across all domains except high risk for reporting bias as there was evidence of selective reporting when another outcome was not analysed when there was no relation to protein intake.

Low certainty evidence did not demonstrate a difference in weight gain (g) at the end of the study (at 2200g or when breastfeeding was initiated) for infants fed fat supplemented human milk compared with unsupplemented milk[144]. Evidence was limited by serious risk of bias with lack of methodological details, and serious imprecision from wide confidence intervals spanning possible meaningful benefits and harms from a small sample size with few events. Interestingly, the narrative report in the review reports that “there was evidence of a clear difference in weight between the fat-supplemented and the unsupplemented groups”[144].

Very low certainty evidence demonstrated no difference in weight at discharge home in infants given oropharyngeal colostrum compared to control[152]. The evidence was downgraded due to imprecision, unclear selection and reporting bias and inclusion of a single study. This was a secondary outcome reported in the trial that was not pre-specified in the protocol.

Moderate certainty evidence demonstrated no significant difference of increased weight gain to discharge (g/day) in infants with a very high protein intake compared to a high protein intake[149]. The data suggested some possibility of increased weight gain with the evidence limited by serious imprecision as the data breached the line of no effect (MD 3.10 [-0.04, 6.24])[149]. Moderate certainty evidence from the same trial was also not able significant difference between the intervention groups in weight gain (g/d) at term or 12 weeks corrected age.

The review[149] also looked at comparisons of each of the three comparisons but adding studies comparing formulas with differences in other nutrients, resulting in a significantly improved weight gain (g/kg/day) in infants fed high versus low protein intake (MD 2.53 g/kg/d, 95% CI 1.62 to 3.45). It reported a significantly improved weight gain (g/kg/day) and g/day in infants fed a very high protein formula compared with high protein formula, however that based on regression curve calculation, more infants fed very high protein took longer than the calculated time to reach 2200g[149]. No significant difference was found in the rate of weight gain in g/week (between birth and when 2400g gained) between very high protein intake, however these comparisons was assessed in three gestational group categories independently[149].

Although unable to report their primary outcome of weight measured by g/kg/day, the review assessing lactase treated feeds[130] was able to report low certainty evidence on weight gain (g/day) at 7 days, 10 days and 14 days after study entry and on study exit. The results of weight gain were inconsistent, with no statistically significant difference in weight gain at 7 and 14 days and on study exit found for infants fed lactase treated feeds compared with feeds that were treated with placebo, however a marginal increased weight gain was found at 10 days after study entry (MD 4.90 g/day [0.18, 9.62]). Although described in the review[130] that the single study was of high quality, the evidence was limited by a serious risk of bias from unclear selection

and reporting biases as it was uncertain if the time points reported for growth were predetermined. The trial was blinded, but concealed allocation was assumed by the review authors and only 80-81% of infants were followed up by 14 days.

In addition to no clear effect in weight gain in g/kg/day, low certainty evidence demonstrated no clear effect on weight gain (g/day) in infants fed formula with high MCT compared with low MCT[129]. The evidence was limited from serious imprecision and serious risk of bias from uncertain methodology including allocation concealment and the two trials were cross-over in design with no clarity about the wash-out period. There was also serious inconsistency, as the review notes that one trial included reported weight gain as g/kg/day once in the text, but as g/day elsewhere. In the other included trial it was not clear if infants received parenteral nutrition and an assumption was made by the review that percentages of MCT were by weight.

10.6.1.3 Linear growth (mm or cm/week)

Three reviews showed statistically significant differences in linear growth[143,145,148], with one further review showing statistical significance only when the intervention was given as sole diet[127].

Moderate certainty evidence of increased crown-heel length was seen in infants fed formula compared with donor breast milk[145] (MD 1.21 mm/week [0.77, 1.65]). Statistically significant differences were seen in the three subgroups with the largest difference in preterm formula versus unfortified DBM, and smallest difference in term formula versus fortified DBM[145]. The subgroup analysis between preterm formula and fortified breastmilk also remained significant at between 0.33 and 1.87 mm/week. Only one trial within the meta-analysis favoured linear growth in donor breastmilk yet had a wide non-significant confidence interval. The evidence was downgraded due to high heterogeneity (I^2=68%). Some included trials had unclear methodological quality and several trials received funding from the formula manufacturer. The review additionally showed very low certainty evidence of both a statistically significant increased crown-rump (MD 0.59 mm/week [0.08, 1.10]) and femoral length gain (MD 0.34 mm/week [0.13, 0.55]). It gave a narrative report of one trial that was unable to detect a difference in average length at 15 days of life or 36 weeks' postmenstrual age, but the data for this was not reported.

There was low certainty evidence of a clinically marginal increased length gain (MD 0.12 cm/week [0.07, 0.17], I^2=89%) in infants fed protein supplementation compared with no supplementation[148]. The certainty of the evidence was reduced due to largely uncertain methodology, and moderate-to-high heterogeneity. Low certainty evidence of no difference was found at term-equivalent age, assessed by a single study blinded to caregivers.

Low certainty evidence showed an increased length gain (MD 0.11 cm/week [0.08, 0.15], I^2=69%) in infants fed fortified compared with unfortified breastmilk[143]. A significant difference was also seen in all three subgroup analyses, with a higher difference seen in trials only assessing very preterm or VLBW infants and trials conducted in low or middle-income countries, and a lower, only marginally significant result when preterm formula was used as a fortifier. The certainty of the evidence was assessed as low due to unexplained heterogeneity (which increased when subgroup analyses were performed), and uncertain risk of bias from methodology of randomisation, allocation and blinding in most trials.

Low certainty evidence showed no clear difference in length gain in the overall comparison of nutrient-enriched versus standard formula, however a significant improvement in length gain in infants fed the nutrient-enriched formula sole diet (MD 1.72 mm/week [0.23, 3.20], I^2=47%)[127]. The evidence was downgraded due to concerns of uncertain risk of bias from random sequence generation and allocation concealment, each in two trials, concerns from lack of blinding, and moderate to high heterogeneity (I^2=67%). The heterogeneity was only reduced to moderate by subgroup analysis of sole diet.

Low certainty evidence showed no statistically significant difference in linear growth in infants who had breastmilk fortification introduced earlier than at 100 ml/kg/day[128]. The review downgraded the evidence due to risk of bias from lack of blinding, and serious imprecision from a small sample size (237 infants). They were unable to perform subgroup analyses due to inadequate data.

Very low certainty evidence showed no clear difference in length gain in infants fed fat supplemented human milk compared with milk not supplemented[144]. The certainty of the evidence was affected by an uncertain risk of bias from insufficient

methodological detail and imprecision from a very small sample size of 14 infants, resulting in wide confidence intervals.

Very low certainty evidence of no significant difference in length gain was seen in the comparison of hydrolysed versus non-hydrolysed formula [150]. Certainty was affected by serious risk of bias and serious imprecision from small trials funded by formula manufacturers with otherwise mostly low risk of bias.

Very low certainty evidence of no clear difference in length gain was seen in the comparison of high MCT formula versus low MCT formula[129]. The certainty of the evidence was affected by serious risk of bias, imprecision and inconsistency from moderate heterogeneity (I^2=50%), although this was stated as "no heterogeneity" by the review authors. There was lack of clarity with the methodology of the included trials, with two crossover trials with an uncertain length of washout period between the interventions.

The review comparing lactase treated feeds compared to placebo/no intervention was only able to assess length gain (cm/week) on study day 14 or study exit if this occurred earlier[130]. Low certainty evidence was not able to demonstrate a statistically significant difference in length gain. Review authors described the single included study as "high quality", however on overview GRADE assessment the certainty was low due to serious imprecision and serious risk of bias with uncertainty regarding methodology and predetermined growth outcome time points.

Low certainty evidence assessed by review authors showed no significant difference in linear growth in the comparison of high versus low protein intake in formula[149]. Certainty of the evidence was affected by imprecision and risk of bias from selection, performance, detection, attrition bias. Although not statistically significant the difference had a confidence interval was wide and was more in favour of high protein intake (MD 0.16 cm/week [-0.02, 0.34]). The review demonstrated moderate certainty evidence of no difference in linear growth measured at discharge in infants receiving a very high versus high protein intake (MD 0.00 cm/week [-0.14, 0.14]) but a confidence interval showing little improvement with infants with a high protein intake compared to high protein intake to no difference when measured at term corrected age (MD 0.10 cm/week [0.00, 0.20]) or 12 weeks corrected gestational age (MD 0.00 [-0.49, 0.49])[149]. The three outcomes were all taken from the same single

trial and while the review authors downgraded the certainty of the evidence at discharge due to imprecision, the overview assessments of the evidence of measurements at term and 12 weeks corrected were assessed as high certainty as we did not judge imprecision to be serious. There was no evidence to assess this outcome in the comparison of very high versus low protein intake[149].

The review comparing human milk-derived fortifier to bovine milk-derived fortifier[151] did not find any eligible studies to assess rate of length gain, however found low certainty evidence of no difference between groups on assessment of change in length during the intervention and length-for-age z-score. Certainty of the evidence was limited by imprecision from wide 95% confidence intervals and inclusion of only a single eligible study. This study[151] was low risk of bias for all domains.

There were no eligible trials to assessed linear growth in the comparisons of banked preterm versus banked term human milk[142], formula versus maternal breastmilk[126], dilute versus full-strength formula[147], carbohydrate supplementation of human milk[146], exposure to oropharyngeal colostrum[152], or high-volume versus standard volume feeds[59].

10.6.1.4 Head growth

Five reviews showed statistically significant differences to head growth [127,143,145,148,149].

Moderate certainty evidence showed an increase in head growth (MD 0.85 mm/week [0.47, 1.23], I^2=74%) in infants fed formula milk compared with donor breast milk[145]. This increase in head growth in infants fed formula milk remained in the subgroups assessing term formula versus unfortified formula (MD 0.81 [0.15, 1.47], I^2=0%), preterm formula versus unfortified DBM (MD .01 [1.21, 2.81], I^2=84%), but not in the comparison of preterm formula versus fortified DBM (MD 0.30 [-0.27, 0.86], I^2=0%). Statistical significance also remained in the subgroup analysis of when the intervention was used as the sole diet (MD 1.36 [0.85, 1.88], I^2=77%), but not when only used as a supplement to human milk (MD 0.24 [-0.32, 0.80], I^2=0%). The quality of the evidence for this was assessed as moderate due to the high heterogeneity (I^2=74%). Some included trials had unclear methodological quality and several trials received funding from the formula manufacturer.

Low certainty evidence showed increased rate of head growth (MD 1.04 mm/week [0.18, 1.89], I^2=57%) in infants fed nutrient-enriched formula compared with standard formula [127]. The quality of evidence was downgraded due to uncertainty about randomisation and masking methods used in the trials, and moderate to high heterogeneity. One of the included trials did not report standard deviations, so they were imputed from the trial of the nearest size. The significant increase in head circumference with nutrient-enriched formula compared to standard formula was higher in the subgroup analysis of trials using the formula as sole diet (MD 2.26 mm/week [1.00, 3.52], I^2=85%) but this comparison showed even higher heterogeneity. No difference was seen in subgroup analysis when formulas compared were used as a supplement to human milk (MD 0.00 mm/week [-1.6, 1.16]).

Low certainty evidence was reported of a marginal increase in rate of head circumference growth (MD 0.06 mm/week [0.01, 0.12], I^2=84%) in infants fed protein supplemented compared to unsupplemented human milk[148]. Certainty was downgraded due to uncertain methodology in the included trials, and moderate to high heterogeneity.

Low certainty evidence showed a marginal increase in head circumference (MD 0.06 cm/week [0.03, 0.08], I^2=42%) of infants fed multi-nutrient fortified breastmilk compared with infants fed unfortified milk[143]. This marginal increase remained on subgroup analysis of trials recruiting only very preterm or VLBW infants (MD 0.07 [0.03, 0.11], I^2=69%), and trials conducted in low or middle-income countries (MD 0.04 [0.01, 0.08], I^2=53%), however reached the line of no effect for trials using preterm formula powder as fortifier (MD 0.05 [0.00, 0.11], I^2=8%). No significant difference found on the test for subgroup differences. Certainty of the evidence was downgraded due to high risk of bias from randomisation and allocation methodology and lack of blinding in most trials. Unlike some other growth parameters in the review, there was low heterogeneity. The review authors describe the funnel plot to not be asymmetrical.

Very low certainty evidence showed a marginal increase in rate of head growth in infants fed of high protein intake in formula versus low protein intake[149] (MD 0.37 cm/week, [0.16, 0.58], single study). Two further studies in the review did not

provide data but reported "no significant difference" in head growth. The certainty of the evidence was reduced due to risk of bias (uncertainty in randomisation and selection bias, more than 20% loss to follow up), and serious imprecision from inclusion of a single very small trial (18 infants).

Low certainty evidence showed no statistically significant difference in head growth in infants who had fortification of human milk introduced earlier than at 100 ml/kg/day[128]. The review downgraded the evidence due to risk of bias from lack of blinding, and serious imprecision from a small sample size (237 infants). They were unable to perform subgroup analyses due to inadequate data.

Very low certainty evidence showed no significant difference in head circumference (MD 0.2 cm/week [-0.07, 0.4], single study) between infants fed fat supplemented compared with control milk[144]. Certainty of the evidence was limited by an uncertain risk of bias from insufficient methodological detail, imprecision from a very small sample size (14 infants), and wide confidence intervals spanning across benefits and harms.

Very low certainty evidence of no statistically significant difference in rate of head growth was shown between infants fed hydrolysed protein and non-hydrolysed formula[150]. Certainty of the evidence was limited by serious imprecision, serious inconsistency from high heterogeneity (I^2=82%) between the two studies and serious risk of bias. The studies were mostly low risk of bias apart from lack of abstract availability, funding by the formula manufacturer and in one study a lack of clarity about the method of randomisation. A further single study in the review with some unclear methodology was unable to demonstrate a significant difference in head circumference at term-equivalent age.

The review comparing human milk-derived versus bovine-milk derived fortifier[151] did not find any eligible trials for rate of head growth, however found low certainty evidence of no statistically significant difference demonstrated by a single eligible trial for change in head circumference during the trial period or head circumference for age z-score. The certainty of the evidence was reduced due to imprecision of the estimate with confidence intervals spanning both a potential benefit or harm of the intervention, and due to evidence from a single study (118 infants). The single included study was assessed as low risk in all domains.

Very low certainty evidence showed no significant difference found in head circumference gain between infants fed high MCT formula and low MCT formula[129]. Certainty was limited by serious imprecision, serious inconsistency and serious risk of bias from uncertain methodology, lack of consistent blinding in most trials and about the period of time for the wash-out period between three cross-over trials of the five eligible trials for this outcome, and how these may affect the outcome. Although there was no statistical heterogeneity, inconsistency was from lack of clarity between some trials on milk used and if parenteral nutrition was given.

Low certainty evidence showed no difference found for head circumference gain at 14 days after study entry or at study exit if occurred earlier, between infants fed lactase treated feeds versus placebo[130]. The study is described as high quality and there was blinding of researchers and caregivers but certainty was assessed as low by overview authors or serious risk of bias from uncertainty about whether the time points used to report growth were predetermined, and serious imprecision from the estimate of a single small study (130 infants).

There were no eligible studies to assess rate of head growth in the reviews assessing banked preterm versus banked term human milk[142], formula versus maternal breastmilk[126], dilute versus full-strength formula[147], carbohydrate supplementation of human milk[146], oropharyngeal colostrum[152], high versus standard-volume feeds[59].

10.6.1.5 Other short term growth measures

Other short term growth measures reported by reviews addressing what to feed were skinfold thickness[127,129,148].

Very low certainty evidence showed statistically significant but clinically marginal increases in rate of skinfold thickness gain in infants fed nutrient-enriched formula versus standard formula[127], both at the triceps (MD 0.12 mm/week [0.07, 0.17], I^2=0%) and subscapular (MD 0.10mm/week [0.04, 0.16], I^2=25%) areas. A significant difference remained on subgroup analysis of triceps skinfold for both when the formula was used as sole diet (MD 0.16 mm/week [0.05, 0.27], I^2=0%) and as supplement to human milk (MD 0.12 mm/week [0.07, 0.17]). It only remained for infants fed the formula for sole diet on subgroup analysis of subscapular skinfold thickness (MD 0.15mm/week [0.17, 0.24], I^2=0%), but not reaching significance when used as a supplement to human milk, with a moderate to significant subgroup

difference seen (I^2=58.1%). Certainty was limited by serious imprecision and a very serious risk of bias. Risk of bias affecting the outcome stemmed from all trials were sponsored by pharmaceutical companies, with three trials sponsored by the formula manufacturer in the trial (and one trial where this is uncertain) and blinding was not present or uncertain in all but one trial. There was low heterogeneity, but there were too few trials to assess publication bias.

Low certainty evidence showed no statistically significant differences in skinfold thickness in the triceps and subscapular areas between infants fed protein supplemented compared with unsupplemented human milk[148]. Subgroup analyses were unable to be assessed. Certainty was limited by serious risk of bias and serious imprecision. There was general uncertainty about the risk of bias in the trial assessing this outcome due to insufficient methodological details, and the very few participants (20 infants).

Low certainty evidence showed no statistically significant difference in skinfold thickness gain was found between infants fed high MCT formula versus low MCT formula[129]. Certainty was limited by serious risk of bias and serious imprecision. The evidence was provided by a very small trial (14 infants) with unequal numbers assessing this outcome in each group (4 vs 10 infants). There was uncertainty in several of the bias assessments and there was lack of clarity regarding taurine supplementation and parenteral nutrition.

No further short term growth outcomes were assessed or reported[59,126,148,149,151,152,128,130,142–147]

10.6.1.6 Long term growth measures

Three reviews reported data for long term growth measures, with evidence assessed by overview authors as moderate certainty due to risk of bias in all three reviews[127,143,145].

No statistically significant differences were found between infants fed formula compared with donor breastmilk[145] in weight, length or head circumferences when measured at 9 months, 18 months post term and 7.5-8 years of age. The studies providing data compared preterm formula with unfortified donor breastmilk, no studies assessing other combinations of comparisons of formula type and added

fortification to the donor breastmilk provided evidence for this outcome. Subgroup analyses for formula and milk provided as sole diet or as a supplement to breastmilk all did not show a statistically significant difference. All outcomes were provided by two trials which were funded by the trial formula manufacturer and there was uncertainty in both trials if blinding was present and reporting bias due to no protocol available. Long term outcomes were reported to be present in more than 80% of participants. Although publication bias was not assessed the two trials were two parts of the same trial design (differentiated by whether the formula or donor milk was used as sole diet or supplement), therefore low heterogeneity is not surprising.

No statistically significant differences were found for weight, height, head circumference, triceps skinfold thickness subscapular skinfold thickness or BMI at 18 months post term or 7.5-8 years post term when nutrient-enriched formula was compared with standard formula[127]. There was no statistically significant difference for waist-to-hip ratio measured at 7.5-8 years post term. No statistically significant was found for the subgroups of the formula used as sole diet and supplement to human milk were performed. Evidence for all outcomes were provided by two trials run in parallel which received financial assistance and supply of the trial diets by the trial formula manufacturer and there was uncertainty in both trials if blinding was present (despite the description that formulas were identified by numerical code so that neonatal staff, parents and follow-up staff were blinded) and reporting bias due to no protocol having been available. The follow up for long term growth outcomes was for more than 80% of surviving infants. Heterogeneity was low for all outcomes, which as the two trials were of the same trial design apart from the intervention and comparators being used as sole diet or supplement to human milk, is not surprising. There were insufficient trials to perform a funnel plot to identify possible publication or reporting bias.

No statistically significant differences were found for weight, length or head circumference at 12 to 18 months in the comparison of breastmilk with multi-nutrient fortifier compared with unfortified breastmilk[143]. Subgroup analyses were not performed for these outcomes. Evidence was provided by two unblinded trials. The larger trial was mostly low risk apart from blinding, while the smaller trial did not perform intention-to-treat analyses of growth outcome data and infants were

noted to have significantly lower weight at birth and study entry, and shorter length at study entry. There were insufficient trials to assess publication or reporting biases through a funnel plot.

There were no eligible studies providing evidence for long term growth outcomes in the reviews assessing banked preterm versus banked term human milk[142], formula versus maternal breastmilk[126], dilute versus full-strength formula[147], human milk-derived versus bovine milk derived fortifier[151], addition of carbohydrate, protein or fat supplements to human milk[144,146,148], high-volume versus standard-volume feeds[59], higher versus lower protein in formula-fed VLBW infants[149] or early versus later fortification of human milk feeds[128].

No long term outcomes after hospital discharge were assessed in the comparison of hydrolysed versus non-hydrolysed formula[150], providing oropharyngeal colostrum[152], lactase-treated feeds[130], or high versus low MCT formula[129].

10.6.2 When is fed

10.6.2.1 Weight gain (g/kg/day)

Only the comparison of delayed versus early introduction of progressive feeding[137] had eligible studies to assess weight gain (g/kg/day). There was no quantitative data available, however the review narratively described that the two trials assessed did not detect a statistically significant difference. Certainty of the evidence was assessed as very low certainty due to the very serious imprecision and serious risk of bias. The risk of bias in the larger study was mostly low risk with uncertainty only lying with the how the random sequence was generated and high risk due to lack of blinding of carers and investigators. The smaller study had very few participants (12) with global uncertainty about methodological details due to only the abstract being available.

There were no eligible studies to assess this outcome for the comparison of slow versus faster rates of feed advancement[136], stopping versus continuing feeds during transfusion[154], presence or absence of routine monitoring of gastric residuals[155], and early trophic feeding versus enteral fasting[131].

10.6.2.2 Other short term weight gain assessments

The review comparing rate of feed advancements[136] reported time to regain birthweight and weight z-score at hospital discharge. Evidence for time to regain birthweight was given narratively that seven trials showed a longer time to regain birthweight in infants with a slow rate of advancement, with two trials reporting a mean difference of 2 [1 to 3] and 3.8 (no CI) days, trials giving median differences of 2, 5 and 6 days, and 2 trials with no data available. The review reported no difference in weight z-score at discharge between the two intervention groups. This data was reported from a single large multicentre trial. There was no specific GRADE assessment for growth, as this was a secondary outcome of the review, however the evidence for the primary outcomes was assessed as moderate due to concerns about the lack of blinding and how it may affect overestimation of feeding intolerance and NEC in faster fed infants. Yet, although the majority of the reported data showed increased growth in infants who had a faster rate of feeding advancement, the data for weight gain was from several small studies with varied methods of measurement that could not be combined for analysis.

Yeo *et al.*[154] did not report any other weight outcomes in addition to no eligible studies in weight gain g/kg/day.

Time to regain birthweight was assessed in the review comparing delayed versus early introduction of progressive feeding[137], with limited data provided from two trials showing no clear differences between the two groups (one trial had no difference, the other a difference of 1 day longer in the early introduction group, no ranges reported). The review additionally reported narrative statements from two trials of no significant differences in the rate of weight gain. The main concern for both outcomes was the lack of blinding, with one trial in the comparison of weight gain having very little methodological detail to make a judgement about other forms of bias.

Low quality evidence showed a marginally increased time to regain birthweight (MD 1.70 [0.01,3.39]) in infants who received routine monitoring of gastric residuals compared to no routine monitoring of the gastric residuals[155], but no difference in the number of infants with extrauterine growth restriction at discharge. The quality of the evidence for both outcomes was downgraded due to imprecision from a small

sample size (80 participants from 1 trial). In addition to concerns of lack of blinding in this trial, there was also a concern of inadequate allocation concealment from a predictable allocation from a fixed block size in an open level trial. There was no significant difference in these outcomes when specifically, presence or absence of routine monitoring of gastric residual quality and quantity was assessed. The evidence for this comparison was similarly affected by small sample size from one trial (87 participants) and lack of blinding.

The comparison of early trophic versus enteral fasting[131] showed no statistically significant difference in days to regain birthweight on meta-analysis. It described two further trials that were not able to be included that agreed with this conclusion. This review also reported no significant differences in weight gain during the trial period g/week, weight gain by day 21 and "growth until 60 days of life", all of which were evidence provided by single trials. The review did report that infants receiving trophic feeds had a higher increase in weight from birthweight to day 30 (223g (SD 125) vs 95 (SD 161) in one small trial involving only 29 participants, however no significant difference (264g (SD 126) vs 213g (SD 142)) from another small trial involving 47 participants and mostly unclear risk of bias. GRADE was not performed, however there was uncertainty in the majority of trials in a variety of domains with several trials unblinded to caregivers and clinical assessors, the data provided either had imprecise estimates with very wide confidence intervals or provided limited or only narrative descriptions of outcomes.

10.6.2.3 Linear growth

There were no eligible studies to assess rate of linear growth in the reviews assessing rates of feed advancement[136], delayed versus early introduction of progressive feeding[137], routine monitoring of gastric residuals[155].,

The rate of linear growth was not assessed in the review comparing stopping versus continuing feeds during transfusion[154].

The only review which provided evidence, albeit limited, on rate of linear growth was the review comparing early trophic feeding versus enteral fasting[131]. This review was only able to report evidence from one possibly unblinded small study (47 infants) which had mostly unclear risk of bias, that narratively stated that increases in

length were similar for both groups. This was assessed as very low certainty evidence due to serious risk of bias and very serious imprecision.

10.6.2.4 Head growth

Only one review reported evidence for rate of head growth[131], with one further review reporting head circumference z-score at hospital discharge[136].

A "borderline significantly higher" head circumference growth (MD 0.7 cm/week [0.1, 1.3], single study) was reported in infants receiving trophic feeds compared with enteral fasting[131]. The review narratively reported that another study reported increases in head circumference to be "similar for both groups". Certainty of the evidence was very low due to serious risk of bias and inconsistency and very serious imprecision. The review providing data[131] was at low risk of bias, but unblinded to caregivers and investigators except laboratory staff (radiological assessment unclear). Parenteral nutrition was provided to both groups during the trial. There was imprecision from a wide 95% confidence interval and evidence from a single study with number of infants involved in the analysis not reported in the review. The trial providing narrative description was a smaller trial with unclear risk of bias in several domains due to unclear methodology and lack of blinding to caregivers or investigators.

The review comparing slow versus faster rates of feed advancement[136] did not assess rate of head growth but reported no difference in head circumference z-score at hospital discharge. The evidence was assessed as moderate certainty due to the risk of bias from a single large (1400 participants) trial unblinded for clinical assessments.

The review comparing stopping feeds to continuing feeds during transfusion[154] did not assess this outcome.

The reviews assessing delayed versus early introduction of progressive feeding[137] and the two comparisons within routine monitoring of gastric residuals versus no routine monitoring[155] planned to assess head growth within the secondary outcomes, but did not report these outcomes in their results. It was assumed that none of the eligible studies reported this outcome.

10.6.2.5 Other short term growth measures

Only one review[131] provided any evidence of other short-term growth measures.

The review comparing early trophic feeding versus enteral fasting[131] only provided very low certainty evidence from a narrative statement from a single study which stated that increase in mid-arm circumference was "similar for both groups". This was a small trial (47 infants) with unclear risk of bias in most domains due to unclear methodology, therefore certainty of the evidence was reduced due to serious risk of bias and very serious imprecision.

No further short-term outcomes were reported in the reviews comparing slow versus faster rates of feed advancement[136] and delayed versus early introduction of progressive feeding[137].

Short term outcomes other than weight, length and head circumference were not planned to be assessed in the reviews assessing stopping feeds versus continuing feeds during transfusion[154] or routine monitoring of gastric residuals[155].

10.6.2.6 Long term growth measures

None of the reviews were able to assess long-term growth from the studies eligible for inclusion[131,136,137,154,155].

10.6.3 How is fed

10.6.3.1 Weight gain (g/kg/day)

Only the review comparing continuous with intermittent bolus milk feeding[133] had eligible studies to assess weight gain (g/kg/day). Low certainty evidence showed no statistically significant effect, with evidence limited by very serious risk of bias, very serious inconsistency from high heterogeneity (83%) and serious imprecision. Although the point estimate suggested a potential reduction in weight gain with continuous milk feeding, this did not reach statistical significance (MD -1.1 [-2.28, 0.03], I^2=83%), whereas the second comparison of continuous versus nasogastric intermittent bolus feeding only did not suggest a difference with confidence intervals spanning well across the line of no effect.

There were no eligible studies to assess weight gain (g/kg/day) for the comparisons of transpyloric versus gastric tube feeding[135] or re-feeding versus discarding gastric residuals[156].

The comparison of push versus gravity bolus tube feeding[132] did not assess this outcome.

10.6.3.2 Other short term weight gain assessments

Very low certainty evidence showed no statistically significant differences in time to regain birthweight and number of infants growth restricted (<10^{th} centile) at discharge, respectively, between infants where gastric residuals were re-fed or discarded[156]. Certainty was reduced due to high risk of bias from blinding and imprecision. Both outcomes were assessed from the same single small trial (59 infants), and the evidence for time to regain birthweight was downgraded due to concerns about attrition and risk of bias from lack of blinding (high risk) and serious imprecision for a wide confidence interval. Infants who developed NEC/SIP (12 of 72 randomised infants) and one further infant (unclear, but presumably due to death) were excluded for all other analyses. The risk of bias was assessed as low risk apart from the performance and detection bias from lack of blinding. There was no significant difference found in the subgroups of only human milk fed infants, and formula fed infants, with no statistical differences between subgroups when tested.

The comparison of continuous versus intermittent bolus milk feeding[133] reported days to regain birthweight, weight gain (g/week), weight gain (g/day) and days to 2040g, with an overall conclusion of no difference in growth. These analyses were reported for all infants fed by either nasogastric or orogastric tube, and comparisons where only nasogastric tube feeding or specific weight groups were assessed. There was no statistically significant difference in days to regain birthweight in any comparisons, or weight gain (g/week) in the comparison of all infants between continuous and intermittent bolus feeding by nasogastric or orogastric feeding tube (not assessed for the stratified weights). There was no statistically significant finding of weight gain in g/day. Although the confidence interval was more in favour of continuous milk feeding, the two included trials had findings favouring opposite intervention groups and the weighting of the effect estimate was heavily influenced by one trial in which data on the complete study sample was not by intention to treat (while analysis by individual weight groups was). When analysis was performed in the post priori weight groups, findings were in favour of continuous NG milk feeding for infants <1000g (MD 2.0 [0.54, 3.46]) and infants >1000g and <1249g (MD 2.0 [0.16, 3.84]), but not infants >1250g and <1499g (MD 0.0 [-1.77, 1.77]). The number of infants included in each of these comparisons were very few (30-32). There were no differences in days to the discharge weight of 2040g found between groups in comparisons including all infants or specified weight categories of <1000g and >1000g and <1249g. This was not assessed in infants >1250g and <1499g. These comparisons were also from a single study (80 participants, with 40 participants in each analysis by weight). The review did not assess GRADE but commented on the lack of blinding of caregivers potentially introducing bias on feeding management decisions, and potential effect of incomplete outcomes from infants dropped from feeding protocols in analyses of these studies. The review stated in the conclusion in 2011 that they had requested subgroup data from the four reviews involved in the assessment of growth to further assess this difference in the weight gain between the subgroups. We assessed the evidence for days to regain birthweight and weight gain (g/week and g/day) as low certainty due to serious risk of bias and imprecision, while days to 2040g was very low certainty due to very serious imprecision.

Low certainty evidence showed no clear difference in change in weight (g/week) in the comparison of transpyloric versus gastric tube feeding[135]. Evidence was reduced due to very serious risk of bias from trials assessed as high risk of bias in at least 2 domains. Five reviews were at high risk of bias in randomisation, allocation and blinding. There was a statistically significant, but clinically potentially marginal reduced weight in infants fed by the transpyloric route compared with gastric tube feeding measured at the expected date of delivery (MD -0.3 kg [-0.6, -0.03]). Evidence was low certainty from a quasi-randomised unblinded small (44 infants, assumed 36 infants assessed) single trial at high risk of bias.

There were no eligible studies assess the outcome of days to regain birthweight in the comparison of push versus gravity bolus tube feeding[132].

10.6.3.3 Linear growth

No statistically significant difference between continuous and intermittent bolus milk feeding groups [133] was found in the overall comparison, or in any of the further comparisons exploring NG tube only feeding or different weight categories. Certainty of the evidence for these outcomes was low, due to very serious risk of bias from unblinded trials, attrition and an included trial with alternate assignment. It additionally reported very low certainty evidence from a further study assessing growth rate of the lower leg from birth to 32 weeks and birth to 36 weeks, limited by serious risk of bias and very serious imprecision. The narrative reports that a significantly faster growth to both 32 weeks and 36 weeks PMA was found in the continuous nasogastric group when compared to the two control groups of intermittent nasogastric and intermittent orogastric tube combined but was only found to 32 weeks PMA for infants <1000g (control groups NG and OG combined), and only found to 36 weeks PMA when compared to infants fed intermittently by NG tube only. The review comments how outcomes are influenced by differences in management of feeding intolerance and how some studies reporting higher gastric residuals in continuous feeding groups therefore there may be differing number of feeding interruptions between studies and groups. There was also variation between studies in timing of initiation of the feeds and the type of feeds used.

The review assessing transpyloric versus gastric tube feeding[135] did not find a significant difference in rate of length gain (mm/week) in either the meta-analysis or

the narrative report of two further studies. Certainty of the evidence was low due to very serious risk of bias. All three trials in the meta-analysis were unblinded, and one trial was high risk in all domains with alternate allocation. The review reports that in many of the trials growth data from infants who had developed complications or there were protocol violations such as tube placement were not reported, therefore incomplete outcome data with no true intention-to-treat analysis may influence the reported effect to differ from the true effect.

There were no eligible studies found to assess length in the comparison of re-feeding versus discarding gastric residuals[156] and there was no assessment of length gain in the review assessing push versus gravity for intermittent bolus tube feeds[132].

10.6.3.4 Head growth

No statistically significant differences were found in rate of head circumference (cm/week) found between infants fed by continuous or intermittent bolus milk feeding in the overall comparison, or in any of the further comparisons exploring NG tube only feeding or different weight categories[133]. One study was excluded from the two analyses which included infants of all weight categories, due to review author concern of a typographical error for which clarification was sought, as presented data appeared significant but was reported as insignificant with no differences shown between subgroups of weight within the study. Certainty was assessed as low by overview authors. One trial in the analyses of all weight categories was quasi-experimental and all trials were not blinded to caregivers. There was uncertainty about attrition bias in all three trials included in the overall analysis, with one trial removing infants from the treatment protocol if unable tolerate more than 1 week, and in two trials not including post-randomisation excluded infants in their analysis. In this outcome there is also the uncertainty about what effect the excluded trial would have if the data published was a significant difference rather than the presumed typographical error which led to its exclusion. As with other growth outcomes reported by this review, outcomes may be influenced by differences in management of feeding intolerance, timing of initiation of the feeds and the type of feeds used.

No statistically significant difference in rate of head growth was found between infants fed by transpyloric feeding compared with gastric tube feeding[135]. Although

five trials stated no statistically significant difference, only two of the trials could contribute data to the meta-analysis. Certainty of the evidence was low, with very serious risk of bias. Both trials contributing to meta-analysis were unblinded and have risk of bias as one trial was quasi-randomised with alternate allocation and in the other method of random sequence generation is unclear. Both trials also have significant concerns about attrition bias with growth data only reported in infants who survived to the end of the study period (36 weeks) in one trial (75% of transpyloric group, 86% of the nasogastric group), and only reported in infants successfully tolerating the allocated feeding route (36% of nasoduodenal group. 71% of the nasogastric group) in the other trial. The conclusion is consistent with the other three unblinded trials unable to be included in the meta-analysis. In the larger trial in the meta-analysis and two of the trials not included in the meta-analysis, the comparison was between continuous nasojejunal and intermittent nasogastric, therefore site of feeding tube tip was not the only differentiating factor. Yet, this review also found low certainty evidence of a statistically significant reduction in head circumference (MD -1.0 cm [-1.7, -0.3]) in infants fed by the transpyloric route compared with gastric tube when measured at the expected date of delivery, in a single small (44 infants) trial that was at high risk of bias from alternate monthly allocation and lack of blinding. As with the previous comparison site of feeding tube tip was not the only differentiating factor, with groups also differentiating by continuous versus intermittent feeding.

The review assessing push versus gravity for intermittent bolus tube feeds did not assess this outcome[132], and there were no eligible trials to assess the outcome in the review comparing re-feeding versus discarding gastric residuals[156].

10.6.3.5 Other short term growth measures

Two reviews additionally assessed short term skinfold thickness[133,135].

There were no statistically significant differences found in change in triceps skinfold thickness (mm/week) between infants fed by continuous feeding versus intermittent bolus milk feeding[133]. Certainty of the evidence was low due to very serious risk of bias. One study providing evidence had incomplete follow-up with infants unable to feed as per protocol excluded from analysis and was blinded only to outcome

assessors, the other study was a small quasi-experimental unblinded trial (16 infants) with alternate assignment.

Low certainty evidence showed no statistically significant difference in change in skinfold thickness (site unspecified) between infants fed by the transpyloric route compared with the gastric route[135]. Evidence was provided by a small (21 infants) unblinded trial with unclear randomisation and allocation methodology. The risk of bias assessment was low risk for attrition bias, but three of twelve infants from the nasojejunal group were withdrawn from growth comparisons due to persistent displacement of the tube back into the stomach, extensive abdominal distension and duodenal perforation.

No other short-term growth outcomes were assessed in the comparisons of push versus gravity bolus feeding[132] or re-feeding versus discarding gastric residuals[156].

10.6.3.6 Long term growth measures

Only one study reported evidence for long term growth measures[135].

The review comparing transpyloric versus gastric tube feeding[135] only provided a narrative statement of "no statistically significant differences" in weight and head circumference measured at 3 and 6 months between infants fed via a feeding tube with the tip sited in the jejunum compared with in the stomach. The evidence was very low certainty due to very serious risk of bias and very serious imprecision. Evidence was narrative evidence of "no statistically significant differences" provided by a small (44 infants) single trial with high risk of bias due to alternate monthly allocation and being unblinded. It was noted to have considerable loss to follow up for these outcomes which was unbalanced as it was mainly reduced in the continuous nasojejunal group (only 18/28 at 3 months, 16/28 infants assessed at 6 months) compared with the intermittent nasogastric group (15/16 assessed at 3 and 6 months).

There were no eligible trials to assess growth following discharge in the comparison of re-feeding versus discarding gastric residuals[156].

Long term growth was not planned to be assessed in the comparison of push versus gravity tube feeding[132], and not reported in the comparison of continuous versus intermittent tube feeding[133].

10.6.4 Adjunctive strategies

10.6.4.1 *Weight gain (g/kg/day)*

The review assessing exposure to smell and taste stimulation of milk with tube feeds versus no exposure[157] was the only review that assessed this outcome, however was unable to combine the data of the two studies for meta-analysis due to differences in assessment. The review estimated mean growth rate for two studies using exponential model estimates, demonstrating a faster mean growth rate in the intervention group compared with the control, however, did not provide confidence intervals for these findings. Certainty of the evidence was assessed by review authors as very low due to very serious risk of bias and imprecision, and serious inconsistency. Evidence included one quasi-randomised trial and one unblinded trial. There was indirectness from differences in participants and interventions, and imprecision from small and few trials which had no confidence intervals and differences in assessment.

There were no eligible studies to assess this outcome in the comparisons of erythromycin versus placebo for prevention of feeding intolerance[134] or oral stimulation for promoting oral feeding[159].

The comparison of glycerine prophylaxis versus placebo/no intervention[158] did not assess this outcome.

10.6.4.2 Other short term weight gain assessments

There was low certainty evidence of no significant difference in weight at discharge home (g) between infants receiving glycerin prophylaxis versus placebo/no intervention[158]. Certainty of the evidence was limited by serious imprecision and serious risk of bias with concerns about blinding and unexplained protocol violations (15/36 in intervention group did not receive enema, 8/42 in control group did receive enema) in the single included trial.

For the comparison of erythromycin versus placebo for prevention of feeding intolerance[134] there was low certainty evidence with no significant differences found in time to regain birthweight or weight gain from birth to discharge, limited by very serious imprecision.

In the comparison of erythromycin compared to placebo for treatment of feeding intolerance[134], the very low certainty evidence for high dose did not support a significant difference in time to regain birthweight in two studies, one reporting mean and another median in time to regain birthweight, and a single study reporting median weight at discharge. Evidence was limited by very serious inconsistency and very serious imprecision[134].

Low dose erythromycin versus placebo for treatment of feeding intolerance did not show a significant difference in weight at study day 8, and weight gain since enrolment (unclear at which time point this was measured)[134]. Evidence was assessed as very low certainty and low certainty, respectively. The included studies were assessed at low risk of bias apart from two studies providing evidence for time to regain birthweight where there was no clarity on blinding of outcome measures. Yet, for weight gain at study day 8 overview authors assessed a serious risk of bias from discrepancy between feed volume at study entry 36 (39) ml/kg/d vs 28 (21) ml/kg/d. The studies were all small or very small studies where the data was not able to be combined and presented individually, therefore there was very serious imprecision. Widespread heterogeneity was discussed by the review authors among study design, methodology, population, intervention, and definitions and reporting of outcome measures.

The review investigating oral stimulation for promoting oral feeding in preterm infants[159] showed low certainty evidence for no significant effect of oral stimulation compared with standard care on weight gain (time period not reported). It did not include the outcomes of one study reporting a percentage weight gain, and one study describing weight change from four/four to eight/eight oral feeds a day until discharge. The certainty of evidence was downgraded by review authors for moderate heterogeneity and high risk of selection, performance, attrition and reporting bias. No evidence was found for assessing weight gain on oral stimulation compared with other non-oral interventions.

There were no other short term weight gain assessments in the review assessing exposure to smell and taste stimulation with milk tube feeds[157].

10.6.4.3 Linear growth

There were no eligible trials to assess linear growth in the reviews assessing exposure to smell and taste stimulation of milk with tube feeds[157] and oral stimulation for promoting oral feeding[159].

The outcome of linear growth was not assessed by the reviews assessing glycerin prophylaxis[158] and erythromycin[134] for prevention or treatment of feeding intolerance.

10.6.4.4 Head growth

There was no evidence assessing head growth in any of the reviews assessing adjunctive strategies, with no eligible studies in the reviews assessing exposure to smell and taste stimulation[157] and oral stimulation[159], and no plan for assessment in the reviews assessing glycerin[158] and erythromycin[134] for prevention and treatment of feeding intolerance[92,158].

10.6.4.5 Other short term growth measures

None of the reviews provided evidence for other short term growth measures[92,157–159].

The reviews assessing glycerin[158] and erythromycin[134] for prevention of feeding intolerance, and oral stimulation[159] to promote oral feeding did not plan any other growth outcome assessments. The reviews assessing exposure to smell and taste stimulation of milk with tube feeding did not report findings of any other short-term growth measures[157].

10.6.4.6 Long term growth measures

None of the reviews provided evidence for long term growth outcomes[92,157–159].

Three reviews did not plan to assess long term growth[92,157,158], while the review assessing oral stimulation for promoting oral feeding in preterm infants[159] planned to assess all outcomes for immediate change, three to six months and beyond six months but did not report longer term growth data.

10.7 Summary of effectiveness of the interventions

Tables 13 to 16 summarise the effect of the interventions.

10.7.1 What to feed

No conclusions were possible for short and long term growth or necrotising enterocolitis in twelve reviews[59,126,150,152,128–130,142,144,146,147,149]. These assessed interventions involving breastmilk including giving early oropharyngeal colostrum[152], time of introduction of breastmilk fortification[128], breastmilk supplementation of carbohydrates[146] or fat[144], and preterm versus term donor human milk[142]. These also assessed differences in formulas including dilution of formula[147], hydrolysis of protein in formula[150], MCT content of formula[129], and between high and low protein intake in formula-fed infants[149]. Within this category was also included reviews comparing formula and maternal breastmilk[126], high versus standard volume feeds[59] and lactase treatment of feeds[130].

Generally no conclusions were possible for protein supplementation of breastmilk[148] on long term growth or necrotising enterocolitis, or the majority of short term growth, except head circumference which was probably ineffective. This moderate certainty evidence of no effect at term-equivalent age although this contradicts with the low certainty evidence of between 0.01-0.12cm/week increased rate in head circumference[148].

Similarly, no conclusions are possible for the effect of human milk-derived versus bovine-milk derived fortifier[151] on long term growth and necrotising enterocolitis, and the majority of short term growth, except that it is probably ineffective for affecting length-for-age z-score.

Formula when compared with donor breast milk[145] was shown to be a promising intervention for improving most short term weight, length and head growth, but probably ineffective for long term growth, but that donor breast milk was a promising intervention for reduction of *harm* from NEC.

Addition of a multi-nutrient fortifier[143] was shown to be probably ineffective for long term growth, with no conclusions possible on necrotising enterocolitis and generally no conclusions possible for short term growth, except moderate certainty evidence of

an increased rate of head circumference gain of 0.03-0.08cm/week when infants are given multi-nutrient fortified milk.

Similarly, the certainty of the evidence showed that no conclusions are possible on the effect of nutrient-enriched formula on short term growth and necrotising enterocolitis, and is probably ineffective for long term growth[127].

Although no conclusions were possible for growth and necrotising enterocolitis on the comparison of high versus low protein intake, the comparison of very high versus high protein intake was seen to be probably ineffective for weight and length gain with no conclusions possible for other short term growth outcomes, long term growth or necrotising enterocolitis[149].

10.7.2 When to feed

No conclusions were possible for short and long term growth and necrotising enterocolitis in three reviews[137,154,155].

Limited conclusions were possible for short term growth and no conclusions were possible for long term growth for the other two reviews in this category [131,136]. Only one of these reviews was able to draw probable conclusions about necrotising enterocolitis[136].

Evidence from slow versus faster rates of feed advancement[136] showed evidence that slower feeding advancement was probably ineffective for reducing necrotising enterocolitis and change in rate of feed advancement was probably ineffective for affecting weight and head circumference z-score at discharge, but no conclusions were possible for overall rate of short term growth.

Evidence from early trophic feeding versus enteral fasting[131] showed that difference in intervention taken was probably ineffective for changing time to regain birthweight, but no conclusions were possible for other short term growth outcomes, long term growth or necrotising enterocolitis.

10.7.3 How to feed

No conclusions were possible for short and long term growth and necrotising enterocolitis in any reviews addressing how to feed[132,133,135,156].

10.7.4 Adjunctive interventions

No conclusions were possible for short and long term growth and necrotising enterocolitis in any reviews addressing adjunctive feeding strategies[157].

Table 13: Summary of the effects of interventions on growth and necrotising enterocolitis, and quality of the systematic review: "What is fed"

Review	Short term growth				Long term growth	Necrotising enterocolitis	AMSTAR2	ROBIS
	Weight gain	**Head circumference gain**	**Length gain**	**Other short term growth**				
Thanigainathan *et al.* 2020[128] Early versus late fortification of breastmilk	No conclusions possible				No conclusions possible	No conclusions possible	**High confidence**	**Low concern**
	Low certainty evidence of no effect on time to regain birthweight and extra-uterine growth restriction at discharge.	*Low certainty evidence of no effect.*	*Low certainty evidence of no effect.*	*Nil evidence.*	*Nil evidence.*	*Low certainty. No effect.* *Surgical NEC: Low certainty. No effect*		
Brown *et al.* 2020[143] Multi-nutrient fortified vs unfortified breast milk	No conclusions possible, except promising intervention for marginal increase in short-term head circumference gain.				Probably ineffective.	No conclusions possible.	**Critically low confidence**	**Unclear concern**
	Low certainty evidence of 1.76 (1.3-2.22) g/kg/day increased weight gain with fortification of breastmilk.	*Moderate certainty evidence of 0.06 (0.03-0.080 cm/week increased head circumference gain with fortification of breastmilk.*	*Low certainty evidence of 0.11 (0.08-0.15) cm/week increased weight gain with fortification of breastmilk.*	*Nil further assessed.*	*Moderate certainty evidence of no effect on weight, length or head circumference at 12-18 months.*	*Low certainty. No effect.*		
Amissah *et al.* 2020a[146] Carbohydrate (prebiotic) supplement vs no supplement	No conclusions possible				No conclusions possible	No conclusions possible	**High confidence**	**Low concern**
	Very low certainty of increased weight (MD 160.4g [12.4-308.4g]) at 30 days with carbohydrate supplementation	No eligible trials	No eligible trials	None reported	None reported	Very low certainty. No effect.		

Review	Short term growth				Long term growth	Necrotising enterocolitis	AMSTAR2	ROBIS
	Weight gain	**Head circumference gain**	**Length gain**	**Other short term growth**				
Amissah *et al.* 2020b[148] Protein supplementation versus no supplementation	No conclusions possible, except probably ineffective for head circumference at term equivalent age.				No conclusions possible	No conclusions possible	**High confidence**	**Low concern**
	Low certainty. Higher weight gain 3.8 (2.94-4.70) g/kg/day with protein supplementation. No effect on weight at term (Low certainty) & end of study (Very low certainty).	*Low certainty evidence of marginal increase in head circumference gain (0.06 [0.01-0.12] cm/week. Moderate certainty evidence of no effect at term-equivalent age.*	*Low certainty. Marginal increase length gain (MD 0.12 [0.07-0.17] cm/week) with protein supplementation. Low certainty evidence of no effect on length at term equivalent age.*	*Low certainty evidence of no effect on triceps and subscapular skinfold thickness growth.*	*None reported*	*Very low certainty. No effect.*		
Amissah *et al.* 2020c[144] Fat supplemented human milk versus control	No conclusions possible				No conclusions possible	No conclusions possible	**High confidence**	**Low concern**
	Very low certainty evidence of no effect on weight gain (g/kg/day) & Low certainty evidence of no effect on weight at end of study	*Very low certainty evidence of no effect.*	*Very low certainty. No effect.*	*Nil eligible studies*	*None reported*	*No eligible trials*		

Review		Short term growth				Long term growth	Necrotising enterocolitis	AMSTAR2	ROBIS
		Weight gain	**Head circumference gain**	**Length gain**	**Other short term growth**				
Fenton *et al* 2020[149]	High versus low protein intake in formula fed infants	No conclusions possible				No conclusions possible	No conclusions possible	**Moderate confidence**	**Low concern**
		Low certainty. 2.36 (1.31-3.4) g/kg/day higher weight gain in infants with high protein intake.	*Very low certainty evidence of 0.37 (0.16-0.58) cm/week increased head growth in high protein intake.*	*Low certainty. No effect (but trend towards higher growth in high protein intake).*	*No evidence*	*No evidence*	Very low certainty. Not estimable.		
	Very high versus high protein intake	Probably ineffective for weight gain and length gain. No conclusions possible on head circumference and other short term growth.				No conclusions possible.	No conclusions possible		
		Moderate certainty evidence of no effect on weight gain (g/day) to discharge (but trend) or term or at 12 weeks corrected age.	*No evidence.*	*Moderate certainty evidence of no difference at discharge, High certainty at term or 12 weeks corrected age (but trend towards marginal increase at term).*	*No evidence*	*No evidence*	*Not assessed*		
Quigley *et al.* 2019[145] Formula milk vs Donor breast milk		Formula is a promising intervention for marginally increased short term weight, head and longitudinal growth. No conclusions possible on crown-rump, femoral and other short-term growth outcomes.				Probably ineffective.	Donor breast milk is a promising intervention for reduction in harm by NEC.	**Low confidence**	**Low concern**

Review	Short term growth				Long term growth	Necrotising enterocolitis	AMSTAR2	ROBIS
	Weight gain	**Head circumference gain**	**Length gain**	**Other short term growth**				
Quigley *et al.* 2019[145] Formula milk vs Donor breast milk (continued)	*Moderate certainty evidence of higher weight gain 2.51 [1.93, 3.08] g/kg/day, and low certainty evidence of shorter time to regain birthweight in formula fed infants (higher difference when DBM unfortified).*	*Moderate certainty evidence of increased head growth (MD 0.85 [0.47, 1.23] mm.week) in infants fed formula (no effect when DBM fortified or formula as supplement).*	*Moderate certainty evidence of marginally increased crown-heel length (1.21 [0.77, 1.65] mm/week in infants fed formula. Highest difference when DBM unfortified. Very low certainty evidence of increased crown-rump MD 0.59 [0.08, 1.10] mm/week and femoral (MD 0.34 [0.13, 0.55] mm/week) lengths.*	*Nil reported*	*Moderate certainty evidence of no effect on weight, length/height or head circumference at 9 months, 18 months and 7.5-8 years of age.*	*Moderate certainty increased risk in infants fed formula vs DBM*	**Critically low confidence**	**Unclear concern**
Ng *et al.* 2019[150] [150]Hydrolysed versus non-hydrolysed formula,	No conclusions possible.				No conclusions possible	No conclusions possible		
	Low certainty evidence of lower weight gain (-3.02 [-4.66, -1.38] g/kg/day) with hydrolysed formul a	*Very low certainty evidence of no effect.*	*Very low certainty evidence of no effect.*	*Nil reported*	*Not assessed*	*Low certainty. No effect*		

Review	Short term growth				Long term growth	Necrotising enterocolitis	AMSTAR2	ROBIS
	Weight gain	**Head circumference gain**	**Length gain**	**Other short term growth**				
Walsh *et al.* 2019[127] Nutrient-enriched formula versus standard formula	No conclusions possible.				Probably ineffective.	No conclusions possible.	**Low confidence**	**Low concern**
	Low certainty evidence of higher weight gain (2.43 [1.60, 3.26] g/kg/day) in infants fed nutrient enriched formula. Higher difference when sole diet.	*Low certainty. Increased head growth in infants fed nutrient enriched formula (MD 1.04 [0.18, 1.89] mm.week), higher when fed as sole diet MD 2.26 [1.00, 3.52] But no difference when supplement only.*	*Low certainty evidence of no effect overall but increased length gain in nutrient enriched formula as sole diet only (MD 1.72 [0.23, 3.20] mm/week)*	*Very low certainty of increased rate of skinfold thickness gain (MD 0.12 [0.07, 0.17] mm/week) present in both subgroups. Increase rate of subscapular skinfold thickness gain 0.10 [0.04, 0.16] mm/week, only present overall and in sole diet subgroup.*	*Moderate certainty evidence of no effect on weight gain, head growth or length/height, skinfold thickness or BMI at 18 months and 7.5-8 years post term. No effect on weight-hip ratio at 7.5-8 years post term.*	*Low certainty. No effect.*		
Dempsey *et al.* 2019[142] Banked preterm versus banked term human milk	No conclusions possible: No eligible trials						**High confidence**	**Low concern**
Brown et al. 2019[126] Formula versus maternal breastmilk	No conclusions possible: No eligible trials						**High confidence**	**Low concern**
Basuki *et al.* 2019[147] Dilute versus full-strength formula	No conclusions possible				No conclusions possible	No conclusions possible	**Low confidence**	**Unclear concern**
	Low certainty of no effect on weight gain at 7 days (no evidence for weight gain g/kg/day)	*No eligible trials*	*No eligible trials*	*Nil reported*	*Nil studies eligible*	*No eligible trials*		

Review	Short term growth				Long term growth	Necrotising enterocolitis	AMSTAR2	ROBIS
	Weight gain	**Head circumference gain**	**Length gain**	**Other short term growth**				
Premkumar 2019[151] Human milk-derived fortifier versus bovine milk-derived fortifier	No conclusions possible except probably ineffective for length-for-age z-score.				No conclusions possible	No conclusions possible	**High confidence**	**Low concern**
	Low certainty of no effect on weight gain during intervention or weight-for-age z scores.	*Low certainty evidence of no effect on change in absolute change in head circumference or head circumference-for-age z-score.*	*Low certainty evidence for no effect on change in length during intervention & moderate certainty evidence of no effect on length-for-age z score.*	*Nil further assessed*	*Nil studies eligible.*	*Low certainty. No effect.*		
Nasuf *et al.* 2018[152] Oropharyngeal colostrum (OPC) compared to control (water, saline or no intervention) in preterm infants	No conclusions possible				No conclusions possible	No conclusions possible	**Critically low confidence**	**Low concern**
	Very low certainty evidence of no effect on weight at discharge home.	*Not assessed*	*Not assessed*	*Nil further assessed*	*Not assessed*	*Very low certainty. No effect.*		
Abiramalatha *et al.* 2017[59]	No conclusions possible.				No conclusions possible	No conclusions possible	**Low confidence**	**Low concern**

Review	Short term growth				Long term growth	Necrotising enterocolitis	AMSTAR2	ROBIS
	Weight gain	**Head circumference gain**	**Length gain**	**Other short term growth**				
High-volume vs standard-volume feeds	*Low certainty evidence of higher weight gain (6.2 [2.71, 9.69] g/kg/day) in infants fed higher volume feeds.*	*No eligible trials*	*No eligible trials*	*Nil further assessed*	*Nil eligible studies*	*Very low certainty. No effect.*		
Tan-Dy *et al.* 2013[130] Lactase treated feeds vs placebo	No conclusions possible				No conclusions possible	No conclusions possible	**Critically low confidence**	**Unclear concern**
	Low certainty evidence of marginal increase at 10 days after study entry (4.9 [0.18, 9.62] g/day) but no effect of weight gain (g/day at 7 and 14 days after study entry.	*Low certainty evidence of no effect on head growth at study day 14/study exit.*	*Low certainty evidence of no effect on length gain at study day 14/study exit.*	*Nil further assessed*	*No long term growth data*	*Low certainty evidence of no effect*		
Nehra *et al.* 2002[129] High MCT formula versus low MCT formula	No conclusions possible				No conclusions possible	No conclusions possible	**Low confidence**	**Unclear concern**
	Very low certainty evidence of no effect on weight gain g/kg/d or g/day	*Very low certainty evidence of no effect on head growth.*	*Very low certainty evidence of no effect on length gain*	*Low certainty evidence of no effect on skinfold thickness gain*	*No eligible studies for longterm growth data*	*Very low certainty evidence of no effect.*		

Table 14: Summary of effect of interventions on growth and necrotising enterocolitis, and quality of the systematic review: "When is fed"

Review		Short term growth				Long term growth	Necrotising enterocolitis	AMSTAR2	ROBIS
		Weight gain	**Head circumference gain**	**Length gain**	**Other short term growth**				
Yeo *et al.* 2019[154] Stopping feeds vs continuing feeds during transfusion		No conclusions possible				No conclusions possible	No conclusions possible	**Low confidence**	**Low concern**
		Nil eligible studies	*Not assessed*	*Not assessed*	*Not assessed*	*Not assessed*	*Low certainty. No effect.*		
Abiramalatha *et al.* 2019a[155]	Routine monitoring of gastric residuals vs no routine monitoring of gastric residuals	No conclusions possible				No conclusions possible	No conclusions possible	**Moderate confidence**	**Low concern**
		Nil eligible studies	*Nil eligible studies*	*Nil eligible studies*	*Not assessed*	*Nil eligible studies*	*Low certainty of no effect on NEC and very low certainty evidence of no effect on surgical NEC*		
	Routine monitoring of gastric residuals quality vs routine monitoring of gastric residuals quality and quantity	No conclusions possible				No conclusions possible	No conclusions possible		
		Nil eligible studies	*Nil eligible studies*	*Nil eligible studies*	*Not assessed*	*Nil eligible studies*	*Very low certainty evidence of no effect on NEC or surgical NEC.*		

Review	Short term growth				Long term growth	Necrotising enterocolitis	AMSTAR2	ROBIS
	Weight gain	**Head circumference gain**	**Length gain**	**Other short term growth**				
Oddie *et al.* 2017[136] Slow versus faster rates of feed advancement	Probably ineffective intervention for weight and head circumference z-score at discharge, but no conclusions possible for overall weight, length and head circumference gain.				No conclusions possible	Probably ineffective.	**Critically low confidence**	**Unclear concern**
	Moderate certainty evidence of no effect on weight z-score at hospital discharge. Very low certainty evidence of increased time to regain birthweight in the slow-rate-of-advancement. Weight gain (g/kg/day) not assessed	*Moderate certainty evidence of no effect on head circumference z-score at hospital discharge. Nil studies eligible for head growth (mm/week).*	*Nil eligible studies*	*Nil further reported*	*Nil eligible studies*	*Moderate certainty. No effect.*		
Morgan *et al.* 2014[137] early introduction of progressive feeding	No conclusions possible				No conclusions possible	No conclusions possible	**Critically low confidence**	**Unclear concern**
	Very low certainty evidence of no effect on weight gain (g/kg/day) and time to regain birthweight.	*Nil eligible studies*	*Nil eligible studies*	*Nil further outcomes reported*	*Nil eligible studies*	*Very low certainty evidence of no effect*		

Review	Short term growth				Long term growth	Necrotising enterocolitis	AMSTAR2	ROBIS
	Weight gain	**Head circumference gain**	**Length gain**	**Other short term growth**				
Morgan *et al.* 2013[131] early trophic feeding versus enteral fasting	Probably ineffective intervention for time to regain birthweight, no conclusions possible for other weight gain, head growth, length gain or other growth outcomes.				No conclusions possible.	No conclusions possible	**Critically low confidence**	**Unclear concern**
	Moderate certainty evidence of no effect on days to regain birthweight. Very low certainty evidence of inconsistent differences in weight gain measured at g/week, at day 21, day 30 and growth until 60 days of life.	*Very low certainty conflicting evidence of marginal increase in head circumference (0.7 [0.1, 1.3] cm/week) and "no differences".*	*Very low certainty evidence of no effect.*	*Very low certainty evidence of no effect.*	*Nil eligible studies.*	*Very low certainty evidence of no effect*		

Table 15: Summary of effect of interventions on growth and necrotising enterocolitis, and quality of the systematic review: "How is fed"

Review	Short term growth				Long term growth	Necrotising enterocolitis	AMSTAR2	ROBIS
	Weight gain	**Head circumference gain**	**Length gain**	**Other short term growth**				
Abiramalatha *et al.* 2019b [156] Re-feeding vs discard gastric residuals	No conclusions possible				No conclusions possible	No conclusions possible	**Moderate confidence**	**Low concern**
	Very low certainty evidence of no effect on time to regain birthweight or number of infants with weight <10th centile at discharge.	*Nil eligible studies*	*Nil eligible studies*	*None additionally assessed*	*Nil eligible studies*	*Very low certainty. No effect*		
Watson *et al.* 2013[135] Transpyloric versus gastric tube feeding	No conclusions possible				No conclusions possible	No conclusions possible	**Critically low confidence**	**Unclear concern**
	Low certainty evidence of lower weight at EDD with transpyloric feeding (MD -0.3 [-0.6, -0.03] kg) but no effect on change in weight (g/week)	*Low certainty evidence of reduce head circumference at EDD in the transpyloric feeding group, but low certainty evidence of no effect on growth in mm/week.*	*Low certainty evidence of no effect*	*Low certainty evidence of no effect on change in subscapular skinfold thickness.*	*Very low certainty evidence of no effect on weight, head circumference at 3 and 6 months.*	*Very low certainty evidence of no effect*		

Review	Short term growth				Long term growth	Necrotising enterocolitis	AMSTAR2	ROBIS
	Weight gain	**Head circumference gain**	**Length gain**	**Other short term growth**				
Dawson *et al.* 2012 [132] Push versus gravity bolus tube feeding	No conclusions possible				No conclusions possible	No conclusions possible	**Low confidence**	**High concern**
	Not assessed. No eligible trials to assess days to regain birthweight.	*Not assessed*	*Not assessed*	*Nil additional assessed.*	*Not assessed*	*No eligible studies*		
Premji *et al.*2011 [133] Continuous vs intermittent bolus feeding	No conclusions possible				No conclusions possible	No conclusions possible	**Critically low confidence**	**Unclear concern**
	Low certainty evidence of no effect on weight gain and days to regain birthweight (but trend towards lower weight gain with continuous feeding).	*Low certainty evidence on no effect.*	*Low certainty evidence of no effect on length gain (but trend for increased linear growth with continuous feeding). Very low certainty narrative evidence of "significantly faster" lower leg growth at 32 and 36 weeks in continuous milk feeding group.*	*Low certainty evidence on no effect on change in triceps skinfold thickness.*	*None assessed*	*Very low certainty evidence of no effect on proven NEC (Stage ≥II) or probable NEC*		

Table 16: Summary of effect of interventions on growth and necrotising enterocolitis, and quality of the systematic review: "Adjunctive strategies"

Review		Short term growth				Long term growth	Necrotising enterocolitis	AMSTAR2	ROBIS
		Weight gain	**Head circumference gain**	**Length gain**	**Other short term growth**				
Muelbert *et al.* 2019[157] Exposure to smell and taste stimulation of milk with tube feeds versus no exposure		No conclusions possible				No conclusions possible	No conclusions possible	**High confidence**	**Low concern**
		Very low certainty evidence of faster mean growth rate with exposure to smell and taste stimulation of milk.	*Nil eligible studies*	*Nil eligible studies*	*Nil further reported*	*Not assessed*	*Low certainty evidence of no effect*		
Greene *et al.* 2016 [159]	Oral stimulation vs standard care for promoting oral feeding	No conclusions possible				No conclusions possible	No conclusions possible	**Critically low confidence**	**Low concern**
		Low certainty evidence of no effect on weight gain (g).	*Nil eligible*	*Nil eligible*	*Nil further assessed*	*Not reported*	*No eligible studies*		
	Oral stimulation vs non-oral intervention for promoting oral feeding	No conclusions possible				No conclusions possible	No conclusions possible		
		Nil eligible	*Nil eligible*	*Nil eligible*	*Nil further assessed*	*Not reported*	*No eligible studies*		
Anabrees *et al.*2015[158] Glycerin prophylaxis versus placebo/no intervention		No conclusions possible				No conclusions possible	No conclusions possible	**Critically low confidence**	**Unclear concern**
		Low certainty evidence of no effect	*Not assessed*	*Not assessed*	*Nil further assessed*	*Not assessed*	*Very low certainty*		

Review		Short term growth				Long term growth	Necrotising enterocolitis	AMSTAR2	ROBIS
		Weight gain	Head circumference gain	Length gain	Other short term growth				
		on weight at discharge home (g)					*evidence of no effect.*		
Ng *et al.* 2008[134]	Erythromycin vs placebo for prevention of feeding intolerance	No conclusions possible				No conclusions possible	No conclusions possible	**Critically low confidence**	**Unclear concern**
		Low certainty evidence of no effect on time to regain birthweight, weight gain from birth to discharge	*Not assessed*	*Not assessed*	*Nil further assessed*	*Not assessed*	*Very low certainty evidence of no effect.*		
	Erythromycin vs placebo for treatment of feeding intolerance	No conclusions possible				No conclusions possible	No conclusions possible		
		Very low certainty evidence of no effect on time to regain birthweight (all) and weight at study day 8 (as low dose). Low certainty evidence of no effect on weight at discharge (as high dose).	*Not assessed*	*Not assessed*	*Nil further assessed*	*Not assessed*	*Very low certainty evidence of no effect.*		

11 Discussion

11.1 Summary of main results

We identified 30 completed Cochrane reviews that evaluated the effects of early enteral feeding strategies on necrotising enterocolitis and growth in very preterm or very low birth weight infants. Broadly, these included strategies for 'what was fed' (17 reviews), 'when to feed' (5), 'how to feed' (4), and 'adjunctive strategies' (4).

No interventions were concluded to provide high certainty or probable evidence that they improve short term growth while reducing or having high or probable certainty of no effect on necrotising enterocolitis. Most reviews concluded only low or very low certainty evidence for short term growth, long term growth and necrotising enterocolitis therefore assessing as no conclusions possible.

Two reviews provided more (moderate) certainty to their conclusions on both growth and necrotising enterocolitis[136,145]. The comparison between donor breast milk and formula showed that formula was a promising intervention for a small increase in short term weight, head and most longitudinal growth but probably ineffective for long term growth[145]. Yet importantly, in contrast, donor breastmilk was a promising intervention for reducing necrotising enterocolitis (with formula milk probably increasing harm)[145]. Different rates of feed advancement were probably ineffective for weight and head growth when measured at discharge and probably ineffective for changes in necrotising enterocolitis[136]. Only a few other reviews were able to conclude isolated growth parameters as probably ineffective[131,144,149,151] or promising for growth[143].

Evidence for long term growth was very sparse, only able to be assessed by four reviews[127,135,143,145]. Three were able to conclude that they are probably ineffective for difference in long term growth[127,143,145]. This evidence was from trials performed more than 20-30 years ago and funded by the study formulas.

11.2 Quality of the evidence

In general, the quality of these Cochrane reviews was high, as expected as the editorial process includes a published peer-reviewed protocol and a requirement to list all study characteristics and assessments. Some procedural and methodological uncertainties and weaknesses were identified using the AMSTAR2 and ROBIS assessments, which highlight areas for improvement in future Cochrane reviews.

11.2.1 Review content

The total number of included trials and infants in each review varied considerably. Only four reviews included more than 1000 infant participants[136,137,143,145]. Consistent with this sample size, these reviews provided more precise estimates of effects on primary outcomes than do reviews with fewer trails or participants. For example, moderate-certainty evidence suggests that feeding with formula milk versus donor human milk increases the risk of necrotising enterocolitis[145], while the Cochrane reviews of the rate of feed volume advancement suggest that faster (versus slower) enteral feeding regimens may not be associated with an increased risk of necrotising enterocolitis[136]. Similarly, the review of multi-nutrient fortification of human milk for feeding preterm infants has sufficient trial data to generate a precise estimate of the effect on in hospital growth rates, although the certainty of the evidence was only moderate for a marginally increased head growth with fortification of breastmilk and low for weight gain and longitudinal growth[143]. This body of evidence on key enteral feeding practices for very preterm or very low birth weight infants has informed national and international guidelines, influenced policy and practice, and underpinned initiatives to improve care and outcomes globally, including three sets of World Health Organisation guidelines[160–162].

Despite these strengths, however, it is notable that none of these reviews, nor other reviews in general, provide high-certainty evidence about the long term effects of enteral feeding strategies, particularly on growth and development beyond infancy.

In contrast, six reviews included fewer than 100 infants[59,132,146,148,154,155] and two reviews contained no eligible studies/infants[126,142]. These reviews, therefore, are unable to provide precise or reliable estimates of effect on any important outcomes

and are of limited value for informing practice or policy. Reviews with few trials or data, however, can inform research priorities and agendas. For example, the review of routine monitoring of gastric residual for prevention of necrotising enterocolitis in preterm infants has contributed to the UK National Institute of Health Research Health Technology Assessment decision to commission a large pragmatic trial of this intervention to provide evidence of sufficient validity and applicability to inform practice[155]. In other cases, it is unlikely that any further randomised controlled trials will be undertaken. For example, trials of formula versus maternal milk are not likely to be ethical, feasible, or necessary, and this review could not be considered "dormant"[126,163].

As expected, the quality of the trials included in the reviews was highly variable. The most common domains to be assessed as high risk of bias were random sequence generation and allocation concealment (selection bias), most commonly due to lack of allocation concealment. Six reviews included at least one quasi-randomised trial with predictable allocation. Lack of blinding was the second most common reason for high risk of bias in a domain and affected 15 reviews. In the majority of trials, the interventions were unable to be masked from caregivers due to the appearance or preparation of the milk being given, the timing of introduction or type of equipment used and investigations required due to the placement of feeding tubes.

The Cochrane risk of bias tool now allows acknowledgement that unmasked trials can be at low risk of bias if there were no deviations from the intervention due to the trial context and also how passive detection of outcomes in a masked trial may lead to bias in measurement of the outcome[164]. The new tool acknowledges the influence of baseline differences signalling a problem with randomisation, non-adherence to an intervention and how appropriate analyses are with a high level of missing outcome data. Selection bias was not a priority for Cochrane in previous reviews and was not reported in six reviews. The new risk of bias tool puts a greater emphasis on this outcome with specific signalling questions to probe the effect of this on bias[164]. It is more specific about applying the criterion to the outcome of interest, rather than the quality of the evidence being influenced by non-reporting of another irrelevant outcomes. The tool also provides guidance for the judgement of risk of bias for each

domain using an algorithm which may provide a more consistent approach to risk of bias between reviews.

Some reviews, especially those assessing how to feed, were limited by attrition bias and exclusion of infants who did not adhere to the trial protocol. This is especially significant as when selecting the route of feeding, or tip placement, the implication of the additional imaging, disruptions with replacing tubes, feeding intolerance are all major factors in feeding infants.

The potential contribution of methodological weaknesses to bias in trials and systematic reviews is well-described[165]. In particular, quasi-randomised trials and randomised trials with inadequate concealment of allocation tend to over-estimate effect size estimates compared with randomised trials with adequately concealed allocation[166]. For example, the Cochrane review of transpyloric feeding versus intra-gastric feeding for preterm infants suggested that the intervention conferred important harms, including a higher risk of death[135]. These effects were no longer statistically significant when the trial with inadequate sequence generation and concealment of allocation was excluded.

11.2.2 Quality of the evidence assessment by the included reviews

We used both AMSTAR2 and ROBIS to assess the quality of the included reviews. Confidence rating in the results of the review as assessed by AMSTAR2 varied, with 11 rated as moderate or high confidence, seven as low confidence and 12 as critically low. A strength of Cochrane reviews is that protocols are mandatory, and all reviews first published after the establishment of Prospero, the international prospective register of systematic reviews, in 2011, had a protocol available. Most reviews kept to protocol or justified their changes. Generally, changes did not affect the validity, but some large undeclared and unjustified changes were noted where the methods were largely revised[132]. All reviews did not explain their decisions of study design inclusion, which although standard Cochrane protocol, there was a variation noticed in a priori declared inclusion of crossover studies. Other quality limitations related to sensitivity analysis of effects of intervention by risk of bias, and complete reporting of study site, funding sources or absence of this information, and inconsistent or incomplete risk of bias assessment (most commonly not reporting selective reporting

or other bias). Other concerns in some reviews were lack of clarity about independent assessment by study authors, how they managed conflict of interest, and undeclared changes from their published protocol or previous review. Although AMSTAR2 has categorised these studies into low and critically low confidence, we would regard these as significant flaws within the conduct of the review that should be addressed in updates of the review. Furthermore, several of these identified problems with methodology are due to continuous improvements in Cochrane methodology where reviews are either out of date or have not fully updated their methodology to that of best practice on updates.

Similar features for improvement were found on assessment by ROBIS, although this assessment had a further step for overview authors to consider if the methodological weakness was likely to introduce a risk of bias in the review. Therefore 57% of the reviews were assessed as low concern, and only one review was assessed as remaining at high concern due to the significance of the undeclared changes in outcomes reported from the protocol.

11.3 Overall completeness and applicability of evidence

As no reviews concluded high certainty evidence and only one review concluded an intervention to be promising for growth but possibly harmful for necrotising enterocolitis, the overall completeness of evidence is poor for the effect of early feeding interventions on growth and necrotising enterocolitis as assessed by randomised controlled trials.

The evidence for necrotising enterocolitis requires large improvements throughout neonatal feeding trials. Most reviews were unable to demonstrate a difference with wide confidence intervals for the estimate suggesting both potential benefits and harms. The quality of the evidence was most frequently limited by risk of bias, imprecision and assessments reasoning their GRADE assessments with small sample sizes or single included trials. Reviews assessing when to feed also noted limitations in low or no event rate, where studies with zero events cannot contribute to meta-analysis conclusions[154]. To address this, future studies will need to be large enough to be able to include enough cases to detect a meaningful difference between interventions if there is one.

Similarly, the evidence provided for short term growth was mostly concluded as no conclusions possible due to the certainty of the evidence being low, very low, or no evidence reported at all. In addition to improvements in risk of bias and imprecision of the evidence, these need improved consistency in measurement between trials.

11.3.1 Absence of evidence

There is currently no evidence from randomised controlled trials to assess necrotising enterocolitis or growth for type of donor breastmilk[142] or, unsurprisingly, maternal breastmilk versus formula[126]. Although there is at least some evidence for growth, there is currently no evidence reported by randomised controlled trials assessing necrotising enterocolitis in comparisons of dilution of formula milk[147], fat supplementation of breastmilk[144], push versus gravity feeding[132] and oral stimulation for promoting oral feeding[159]. Short term growth evidence is currently absent in the reviews assessing monitoring of gastric residuals[155] and feeding during transfusion[154].

11.3.2 Interventions where there is no systematic review currently

Early feeding strategies in very preterm and very low birthweight infants are broadly covered by Cochrane systematic reviews, yet as new interventions emerge the coverage by systematic reviews will need to adapt. Some decisions made in clinical care, such as suggested improved tolerability of partially hydrolysed cow's milk protein fortifier or donkey milk sourced fortifier will need to be included in systematic analysis. A systematic review looking at cup feeding in preterm infants exists, which includes a subgroup analysis for preterm infants, yet was excluded from this review due to no a priori outcome measure of necrotising enterocolitis, and the only studies introducing this within the first week did not include very preterm infants[123].

One review included two studies published in 2010 and 2013 which excluded infants with presence of umbilical vessel catheters[136]. Enteral feeding in clinical practice while umbilical venous catheters (UVC) and umbilical arterial catheter (UAC) in place is variable but thought to be more widely practiced than literature suggests[167].

If continued variation exists regarding this practice, a systematic review and further research into this topic may need to be considered.

Five reviews included studies excluding infants with necrotising enterocolitis[133,134,143,147,154]. One review included studies excluding infants who received surgery[129]. Evidence has shown how development of necrotising enterocolitis has severe consequences on short term infant growth[9]. By excluding infants who have developed gastrointestinal morbidity, a more homogenous patient pool is present. Yet no Cochrane systematic review has explored re-introduction of enteral feeding after necrotising enterocolitis. This highlights a gap where evidence summaries are required on the subsequent early feeding strategies in these infants on development of further episodes of necrotising enterocolitis and in improving their subsequent opportunity for growth and development. At least one systematic review (and one meta-analysis outside of Cochrane) has explored the evidence base yet were only able to conclude no increase in negative outcomes with earlier re-feeding based on observational studies[168,169]. Both explored infection, NEC recurrence and strictures, but neither explored subsequent growth[168,169]. A recent literature review on NEC explored nutrition following medical and surgical necrotising enterocolitis, highlighting the requirement for further research on the topic[170].

11.3.3 Subgroup analyses for infants at risk of necrotising enterocolitis

Almost all reviews planned subgroup analyses, most often based on gestational age groups, birthweight, as well as differences specific to the interventions tested. Only eight reviews, however, were able to conduct at least one subgroup analysis[127,133,136,137,143,145,150,156]. Similarly, ten reviews planned specific subgroup analyses for infants who were small for gestational age[59,127,128,152,155], growth restricted at birth[131,136,137,151,157], or with absent or reversed end diastolic flow[131,136,137]. Yet only three reviews were able to perform this, of which one only reported this for NEC[137]. In addition, several reviews included studies which excluded infants with absent or reversed end diastolic flow (AREDF)[136,155] or evidence of intrauterine growth restriction[137,154]. The terms small for gestational age (SGA) and intrauterine growth restriction (IUGR) are often used interchangeably, but while SGA is defined numerically (<10th centile), intrauterine growth restriction

describes the clinical features. Yet evidence of intrauterine growth restriction and AREDF is an important consideration for neonatologists, with a significant association between the two, especially for infants who were not born due to spontaneous rupture of membranes where between 53 and 65% of infants between 25 and 33 weeks were born <10th centile (16-20% in spontaneous rupture of membranes)[171]. Research into enteral feeding in AREDF is said to be of even greater importance in less developed countries such as India due to higher incidence[172].

11.3.4 GRADEing the evidence

GRADE assessment informs the certainty of the review conclusions of the outcomes. Although this is a transparent and systematic approach performed according to guidance and criteria, this this assessment is still found to include subjectivity. Although inter-rater reliability has been shown to improve with training, but differences still exist[173]. The main reasons for reduced certainty on GRADE were risk of bias, inconsistency from statistical heterogeneity, and imprecision. Indirectness was rarely considered, only seen in this overview by our own GRADE assessments where necrotising enterocolitis was not reported as pre-determines outcomes but only reported as reasons for withdrawal. Imprecision was found to be widely subjective. What do review authors decide as a threshold between when a confidence interval is “wide” or “narrow”? In growth assessments, a “wide” confidence interval was variably defined as a range of a few grams/kg/day of weight gain, and mm/week of length and head growth. These imprecisions may be below the errors found for measurements between observers in clinical practice[174,175]. To improved comparisons between trials, clarity and consensus will need to be established both on what constitutes a clinically significant increase in growth and what margin constitutes a precise confidence interval. Finally, how much bias, what kind and what proportion of the evidence with high or uncertain risk of bias influences the certainty? As GRADE considers all the evidence, inclusion of a quasi-randomised study at high risk of bias may immediately reduce the certainty even when other studies can be of higher quality. How to balance the risk of excluding important evidence and including potential bias will need to be explored. Several reviews used the GRADE approach to define the certainty of the evidence with respect to risk of bias, directness of evidence, heterogeneity, and precision of effect

estimates and risk of publication bias[176]. For many reviews, however, it was unclear how the scientific quality of the included trials had informed the conclusions.

Where evidence was reported to contribute to meta-analysis, evidence was often limited by unexplained heterogeneity and risk of bias from selection, performance and detection bias. For example, reviews assessing specialised formulas, such as those nutrient-enriched, treated with lactase, hydrolysed formula, and breastmilk fortification, were affected by risk of bias from funding by formula manufacturers[127,130,143,150]. Reviews, especially those which did not demonstrate significant differences, were additionally frequently limited by imprecision.

Narrative evidence, reporting of median times, or data not able to be combined on meta-analysis were frequently the only evidence on which to base conclusions for certain outcomes[131,134,136,137]. In most cases, narrative evidence which was reported in addition to meta-analysis results were similar to the conclusion already drawn or stated no significant differences.

11.3.5 Publication bias

Trials that report a statistically significant effect are more likely to be submitted and accepted for publication than studies that do not[177]. Few of the reviews, however, assessed the potential for publication bias, although most did not have enough trials for meaningful funnel plot asymmetry or regression testing. Only three reviews included meta-analyses with data from at least 10 studies to assess publication bias by visual assessment of a funnel plot[135,136,143]. Absence of a risk of bias assessment, however, does not mean that there is an absence of risk, and the important consideration of publication bias from the impact of small study effects, where small studies can predispose to a more beneficial intervention effect estimate must still be considered[178,179].

11.3.6 Industry involvement in trials

The pharmaceutical industry and formula manufacturers are perceived as the dark side of preterm feeding and baby friendly initiatives advocate for promotion of breastfeeding and ensuring that advice on breastmilk substitutes is factual and

unbiased by commercial interests[180]. Formula, or breastmilk substitutes, are nutritional products and do need evidence to show their safety and efficacy. If these products are part of nutritional trials, negative and positive results should be published and therefore randomised controlled trials assessing these will continue to be included in Cochrane reviews. To allow these trials to be done in the most rigorous process, criteria for conducting trials of breastmilk substitutes have been developed which protects breastfeeding and removes the manufacturer from the scientific conduct and conclusions of the trial[181]. Conflict of interest, which is not only restricted to funding from pharmaceutical companies, can have profound effects on trial outcomes, and it is important to ensure transparency and demonstrating how this was managed[182].

11.3.7 Age of the evidence

Currently, Cochrane neonatal systematic reviews include evidence from trials regardless of time since the trials. New evidence, therefore, may be clinically heterogenous to previous trials in terms of confounding factors related to the outcomes of infants such as improvements in routine respiratory care. No eligible trials provided new evidence in the past two decades for five intervention reviews, questioning how applicable the evidence is in our current practice[127,129,130,144,147]. As new evidence emerges, it may be worth exploring differences in evidence found by subgroup analysis by time the trial was performed, or decisions to exclude evidence from prior to a certain date. Yet, as discussed in the Cochrane handbook[183], authors need to strike a balance between loss of valuable data if studies are excluded, and impact on applicability of the effect estimate if studies are included.

11.3.8 Consideration of effect on NEC and Growth in the context of other outcomes

This overview has concentrated on summarising the effect of interventions on necrotising enterocolitis and growth, however feeding interventions have implications on a much wider range of important outcomes such as neurodevelopment and sepsis. Most conclusions remained uncertain with evidence of lower certainty or no significant differences found, although moderate certainty

evidence suggested no significant effect on mortality for donor breastmilk compared with formula[145].

The only review showing evidence of effect on neurodevelopment was that assessing nutrient-enriched formula, which showed an isolated finding with low certainty evidence of an improved psychomotor subsection of the Bayley I Development Index at 18 months post term, but no other differences found in the other categories assessed[127].

The only review to show a difference in invasive infection, demonstrated a "borderline increased risk" (RR 1.15, 95% CI 1.00 to 1.32) when feeds were advanced more slowly[136]. This may be due to duration of central access due to a longer requirement of parenteral nutrition.

Feeding intolerance is a large barrier to growth, as when clinicians believe this to have occurred then changes such as cessation, reduction or delay in further advancement of feeds occurs, therefore limiting nutrient intake during each episode. Feeding intolerance definitions were highly variable between trials both within and between reviews and therefore was frequently assessed individually by many different categories or producing high heterogeneity in results.

11.4 Agreements and disagreements with other studies and/or reviews

The overall low certainty of evidence around interventions in very preterm and very low birthweight infants despite a high number of included trials is not restricted to early enteral feeding strategies. Throughout neonatal medicine, there is little high quality or certainty evidence to conclude on decisions for decision and policy makers, leading to variation in clinical practice[12]. A systematic literature review of Cochrane reviews in neonatology found that almost half of systematic reviews were inconclusive with no specific recommendations, and these are increasing in proportion with time[95]. The main reasons for inconclusive evidence are small sample sizes, insufficient data, insufficient methodological quality and heterogeneity of studies[12,95]. Heterogeneity between outcomes reported between trials is considered an important contribution towards this problem[12]. The problem of inconsistency in outcomes measured between trials and the effect this has on meta-analysis has been

discussed and has led the recent development of a core outcome set for neonatal research[12]. This looked at priorities from healthcare professionals, researchers and former parents and patients. Necrotising enterocolitis was one of the 12 important outcomes identified as a research priority. Growth, although discussed as a core outcome measure in earlier rounds of establishing consensus, was not felt to be a priority by former patients and parents, the definition was felt to be unclear by healthcare professionals and was felt by researchers to be a general measure that did not influence other outcomes[12].

While growth may not be a priority in overall outcomes of neonatal trials, in neonatal nutrition it remains an important marker. The evidence for short term growth included in this overview were assessed by a variety of measurements, including weight gain, head growth, linear growth, and skinfold thickness, and extrauterine growth restriction. Yet, each measurement had further variation in the units reported (e.g., g/day, g/kg/day or absolute weights), timepoint and length of assessment. A similar core outcome set for growth measurements would be helpful for consistent reporting in future trials on early enteral feeding strategies to reduce heterogeneity in measurements to reduce research waste by measurements unable to be combined. For this we need to reach a consensus on units of measurement as well as timeframe of measurement. If measuring z-scores or measurements at discharge, we also need to consider variation in where infants start on the trajectory to understand success.

Although systematic reviews of interventions focus on clear evidence of effect of single interventions, if each individual intervention shows potential benefit albeit clinically minimal, it may be advantageous for future trials and reviews to approach packages of interventions which may together be shown to provide clinically significant benefit. A statistically significant increase in growth, may not necessarily be clinically significant enough to outweigh cost, harm from additional interventions for monitoring, or harm in other clinically significant outcomes, especially considering the accuracy of growth measurements which are susceptible to user error[174,175].

11.5 Potential biases in the overview process

As seen in this review, authors who embark on an overview are likely to have an interest and skills in systematic review methodology and the topic of interest, in this case preterm feeding strategies. This means that overview authors are likely to be aware of the current evidence base included by the Cochrane reviews, which has the potential to introduce bias. The major potential source of bias in the process is that the author and co-investigators are also authors of some of the included Cochrane reviews[112]. Where only one overview author was a co-author of an included review, bias was minimised through assessment by a second assessor not involved in the review. Two reviews were identified where both main overview authors (VW and WM) were co-authors[126,127], where to minimise bias, a third assessor not affiliated with the review was involved in the assessment (SO), extracting and assessing data independently before comparison and discussion to reach consensus among all three review authors.

Bias introduced by prior knowledge of the systematic review conclusions and dual authorship were mitigated by planning methodology as a protocol according to guidance from the Cochrane handbook. This included prior decision made about inclusion criteria, a data extraction proforma, and use of recommended assessment frameworks including GRADE, AMSTAR2 and ROBIS. Although review eligibility and selection were performed in duplicate and independently, data extraction was only performed by one author (VW) and checked for accuracy by the second author (WM). As this data extraction is objective, rather than subjective, this is unlikely to have introduced bias. Assessment of the quality of the included reviews by AMSTAR2 and ROBIS, as well as the GRADE assessments of outcomes not performed by the original review authors, was also performed by the first review author (VW), before duplicate assessment by a second review author. Any disagreement to the initial assessment was marked for discussion until consensus was reached. As assessment was more subjective and the second review author was not blind to the initial assessment, this may have introduced bias.

A separate potential concern is that many had not been updated within the past two years, as per Cochrane guidelines. Of the 30 reviews, 8 (7 at the time of the search)

had not been updated within the past five years. Cochrane Neonatal recognizes the challenges in keeping reviews up-to-date and determines priorities for updating based on expert opinion and focused searches. Furthermore, some reviews may be considered as "complete" or "dormant", and no longer be updated as new or modified interventions become established.

12 Conclusions

1. There are few trial data for many common interventions

In particular, the evidence for effects on necrotising enterocolitis requires large improvements throughout the neonatal feeding trials. Most reviews were unable to demonstrate a difference with wide confidence intervals for the estimate suggesting both potential benefits and harms. Evidence assessing necrotising enterocolitis was generally of low or very low quality with limitations including the risk of bias from masking of interventions from caregivers and assessors.

2. Many included trials had high or uncertain risk of bias.

Several outcome assessments included trials at high risk of selection bias from the method of randomisation or allocation concealment, and many were downgraded due to lack of blinding. Despite reviews planning a priori to perform sensitivity analyses with trials only at low risk of bias, only one review conducted an additional analysis excluding one trial due to clinically significant baseline differences but included other higher risk trials. Where meta-analysis included high risk and uncertain risk in these studies, this may lead to conclusions of reduced certainty on GRADE assessment, despite inclusion of some higher quality trials.

3. Several interventions are not addressed in Cochrane reviews

The area of what to feed in early feeding strategies in very preterm and very low birthweight infants is broadly covered, yet as new interventions emerge the coverage by systematic reviews will need to adapt. Some decisions made in clinical care, such as suggested improved tolerability of partially hydrolysed cow's milk protein

fortifier or donkey milk sourced fortifier will need to be included in systematic analysis.

4. Paucity of subgroup data

In many cases, subgroup data from included trials were not reported or available, a major limitation to external validity.

5. *Continuous improvement in Cochrane methodology*

Several areas for improvement have been identified in Cochrane methodology used in the included reviews, when assessed by AMSTAR2 and ROBIS. These are predominantly where reviews are either out of date or have not fully updated their methodology to that of best practice. Review authors should consider changes in best practice and consider implementing emerging tools, such as the Risk of Bias 2, as they update their reviews.

13 References

1. WHO. ICD-11 for Mortality and Morbidity Statistics (Version : 05/2021) [Internet]. Icd.who.int. 2021 [cited 2021 Sep 9]. Available from: https://icd.who.int/browse11/l-m/en

2. Agostoni C, Buonocore G, Carnielli VP, De Curtis M, Darmaun D, Decsi T, et al. Enteral Nutrient Supply for Preterm Infants: Commentary From the European Society of Paediatric Gastroenterology, Hepatology, and Nutrition Committee on Nutrition. J Pediatr Gastroenterol Nutr. 2010 Jan;50(1):85–91.

3. Hay, Jr. WW. Strategies for Feeding the Preterm Infant. Neonatology. 2008 Oct 2;94(4):245–54.

4. Thureen PJ. Early Aggressive Nutrition in the Neonate. Pediatr Rev. 1999 Sep 1;20(9):e45–55.

5. Keunen K, van Elburg RM, van Bel F, Benders MJNL. Impact of nutrition on brain development and its neuroprotective implications following preterm birth. Pediatr Res. 2015 Jan;77(1):148–55.

6. Glass HC, Costarino AT, Stayer SA, Brett CM, Cladis F, Davis PJ. Outcomes for Extremely Premature Infants. Anesth Analg. 2015 Jun;120(6):1337–51.

7. Ehrenkranz RA. Growth in the Neonatal Intensive Care Unit Influences Neurodevelopmental and Growth Outcomes of Extremely Low Birth Weight Infants. Pediatrics. 2006 Apr 1;117(4):1253–61.

8. McGuire W, Henderson G, Fowlie PW. Feeding the preterm infant. BMJ. 2004 Nov 20;329(7476):1227–30.

9. Pike K, Brocklehurst P, Jones D, Kenyon S, Salt A, Taylor D, et al. Outcomes at 7 years for babies who developed neonatal necrotising enterocolitis: the ORACLE Children Study. Arch Dis Child - Fetal Neonatal Ed. 2012 Aug 29;97(5):F318–22.

10. Tanner SM, Berryhill TF, Ellenburg JL, Jilling T, Cleveland DS, Lorenz RG, et al. Pathogenesis of Necrotizing Enterocolitis. Am J Pathol. 2015 Jan 1;185(1):4–16.

11. Janvier A, Farlow B, Baardsnes J, Pearce R, Barrington KJ. Measuring and communicating meaningful outcomes in neonatology: A family perspective. Semin Perinatol. 2016 Dec 1;40(8):571–7.

12. Webbe JWH, Duffy JMN, Afonso E, Al-Muzaffar I, Brunton G, Greenough A, et al. Core outcomes in neonatology: development of a core outcome set for neonatal research. Arch Dis Child Fetal Neonatal Ed. 2020 Jul 1;105(4):425–31.

13. Janvier A, Bourque CJ, Dahan S, Robson K, Barrington KJ, on behalf of the Partenariat Famille (PAF) team. Integrating Parents in Neonatal and Pediatric Research. Neonatology. 2019 Jun;115(4):283–91.

14. Webbe JWH, Duffy JMN, Afonso E, Al-Muzaffar I, Brunton G, Greenough A, et al. Core outcomes in neonatology: Development of a core outcome set for neonatal research. Arch Dis Child Fetal Neonatal Ed. 2020;105(4):Supplementary material.

15. Pereira-Da-Silva L, Virella D, Fusch C. Nutritional assessment in preterm infants: A Practical Approach in the NICU. Nutrients. 2019 Aug 23;11(9):1999.

16. Schünemann HJ, Gunn EV, Higgins JP, Sansetto N, Deeks JJ, Glasziou P, et al. Chapter 15: Interpreting results and drawing conclusions. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Page MJ, Welch VA, editors. Cochrane Handbook for Systematic Reviews of Interventions version 60 (updated July 2019). Cochrane; 2019.

17. Klingenberg C, Embleton ND, Jacobs SE, O'Connell LAF, Kuschel CA. Enteral feeding practices in very preterm infants: an international survey. Arch Dis Child - Fetal Neonatal Ed. 2012 Aug 18;97:F56–61.

18. Cooke R, Embleton N. Feeding issues in preterm infants. Arch Dis Child - Fetal Neonatal Ed. 2000 Nov 1;83(3):F215–8.

19. Maly J, Burianova I, Vitkova V, Ticha E, Navratilova M, Cermakova E. Preterm human milk macronutrient concentration is independent of gestational age at birth. Arch Dis Child - Fetal Neonatal Ed. 2019 Jan;104(1):F50–6.

20. Castellote C, Casillas R, Ramírez-Santana C, Pérez-Cano FJ, Castell M, Moretones MG, et al. Premature Delivery Influences the Immunological Composition of Colostrum and Transitional and Mature Human Milk. J Nutr. 2011 Jun 1;141(6):1181–7.

21. Ballard O, Morrow AL. Human Milk Composition. Pediatr Clin North Am. 2013 Feb;60(1):49–74.

22. Gidrewicz DA, Fenton TR. A systematic review and meta-analysis of the nutrient content of preterm and term breast milk. BMC Pediatr. 2014 Dec 30;14:216.

23. Macedo I, Pereira-da-Silva L, Cardoso M. The fortification method relying on assumed human milk composition overestimates the actual energy and macronutrient intakes in very preterm infants. Matern Heal Neonatol Perinatol. 2018 Sep 17;4:22.

24. Hill PD, Aldag JC, Chatterton RT, Zinaman M. Comparison of Milk Output Between Mothers of Preterm and Term Infants: The First 6 Weeks After Birth. J Hum Lact. 2005 Feb 25;21(1):22–30.

25. Montjaux-Régis N, Cristini C, Arnaud C, Glorieux I, Vanpee M, Casper C. Improved growth of preterm infants receiving mother's own raw milk compared with pasteurized donor milk. Acta Paediatr. 2011 Jun 24;100(12):1548–54.

26. Madore LS, Bora S, Erdei C, Jumani T, Dengos AR, Sen S. Effects of Donor Breastmilk Feeding on Growth and Early Neurodevelopmental Outcomes in Preterm Infants: An Observational Study. Clin Ther. 2017 Jun 1;39(6):1210–20.

27. Arslanoglu S, Corpeleijn W, Moro G, Braegger C, Campoy C, Colomb V, et al. Donor Human Milk for Preterm Infants. J Pediatr Gastroenterol Nutr. 2013 Oct;57(4):535–42.

28. Weaver G, Bertino E, Gebauer C, Grovslien A, Mileusnic-Milenovic R, Arslanoglu S, et al. Recommendations for the Establishment and Operation of Human Milk Banks in Europe: A Consensus Statement From the European

Milk Bank Association (EMBA). Front Pediatr. 2019 Mar 4;7.

29. National Institute for Health and Care Excellence. Donor milk banks: service operation [CG93]. London; 2010 Feb.

30. Singhal A, Cole TJ, Lucas A. Early nutrition in preterm infants and later blood pressure: two cohorts after randomised trials. Lancet. 2001 Feb 10;357(9254):413–9.

31. Singhal A, Cole TJ, Fewtrell M, Lucas A. Breastmilk feeding and lipoprotein profile in adolescents born preterm: follow-up of a prospective randomised study. Lancet. 2004 May 15;363(9421):1571–8.

32. Hay WW, Hendrickson KC. Preterm formula use in the preterm very low birth weight infant. Semin Fetal Neonatal Med. 2017 Feb 1;22(1):15–22.

33. Boirie Y, Dangin M, Gachon P, Vasson MP, Maubois JL, Beaufrère B. Slow and fast dietary proteins differently modulate postprandial protein accretion. Proc Natl Acad Sci. 1997 Dec 23;94(26):14930–5.

34. Traves D. Understanding infant formula. Paediatr Child Health (Oxford). 2019 Sep 1;29(9):384–8.

35. Kunz C, Lönnerdal B. Re-evaluation of the whey protein/casein ratio of human milk. Acta Paediatr. 1992 Feb;81(2):107–12.

36. Mihatsch WA, Högel J, Pohlandt F. Hydrolysed protein accelerates the gastrointestinal transport of formula in preterm infants. Acta Paediatr. 2001 Feb 2;90(2):196–8.

37. Zuppa AA, Visintini F, Cota F, Maggio L, Romagnoli C, Tortorolo G. Hydrolysed milk in preterm infants: An open problem. Acta Paediatr. 2007 Jan 2;94(s449):84–6.

38. Heyman MB. Lactose Intolerance in Infants, Children, and Adolescents. Pediatrics. 2006 Sep 1;118(3):1279–86.

39. Shulman RJ, Wong WW, Smith EO. Influence of changes in lactase activity and small-intestinal mucosal growth on lactose digestion and absorption in preterm infants. Am J Clin Nutr. 2005 Feb;81(2):472–9.

40. Erasmus HD, Ludwig-Auser HM, Paterson PG, Sun D, Sankaran K. Enhanced weight gain in preterm infants receiving lactase-treated feeds: a randomized, double-blind, controlled trial. J Pediatr. 2002 Oct 1;141(4):532–7.

41. Kien CL. Digestion, absorption, and fermentation of carbohydrates in the newborn. Vol. 23, Clinics in Perinatology. Elsevier; 1996. p. 211–28.

42. Shulman RJ, Schanler RJ, Lau C, Heitkemper M, Ou C-N, Smith EO. Early feeding, feeding tolerance, and lactase activity in preterm infants. J Pediatr. 1998 Nov 1;133(5):645–9.

43. Thymann T, Møller HK, Stoll B, Støy ACF, Buddington RK, Bering SB, et al. Carbohydrate maldigestion induces necrotizing enterocolitis in preterm pigs. Am J Physiol Gastrointest Liver Physiol. 2009 Dec 1;297(6):G1115–25.

44. Buddington RK, Davis SL, Buddington KK. The Risk of Necrotizing Enterocolitis Differs Among Preterm Pigs Fed Formulas With Either Lactose or Maltodextrin. J Pediatr Gastroenterol Nutr. 2018 Mar;66(3):e61–6.

45. Martin CR, Cheesman A, Brown J, Makda M, Kutner AJ, DaSilva D, et al. Factors Determining Optimal Fatty Acid Absorption in Preterm Infants. J Pediatr Gastroenterol Nutr. 2016 Jan;62(1):130–6.

46. Jensen GL, Jensen RG. Specialty Lipids for Infant Nutrition. II. Concerns, New Developments, and Future Applications. J Pediatr Gastroenterol Nutr. 1992 Nov;15(4):382–94.

47. Mazzocchi A, D'Oria V, Cosmi V, Bettocchi S, Milani G, Silano M, et al. The Role of Lipids in Human Milk and Infant Formulae. Nutrients. 2018 May 4;10(5):567.

48. Arslanoglu S, Boquien C-Y, King C, Lamireau D, Tonetto P, Barnett D, et al. Fortification of Human Milk for Preterm Infants: Update and Recommendations of the European Milk Bank Association (EMBA) Working Group on Human Milk Fortification. Front Pediatr. 2019 Mar 22;7:76.

49. Arslanoglu S, Moro GE, Ziegler EE, The Wapm Working Group On Nutrition. Optimization of human milk fortification for preterm infants: new concepts and recommendations. J Perinat Med. 2010 Feb 25;38(3):233–8.

50. Radmacher PG, Adamkin DH. Fortification of human milk for preterm infants. Semin Fetal Neonatal Med. 2017 Feb 1;22(1):30–5.

51. de Halleux V, Pieltain C, Senterre T, Rigo J. Use of donor milk in the neonatal intensive care unit. Semin Fetal Neonatal Med. 2017 Feb 1;22(1):23–9.

52. de Halleux V, Rigo J. Variability in human milk composition: benefit of individualized fortification in very-low-birth-weight infants. Am J Clin Nutr. 2013 Aug 1;98(2):529S-535S.

53. Rochow N, Fusch G, Choi A, Chessell L, Elliott L, McDonald K, et al. Target Fortification of Breast Milk with Fat, Protein, and Carbohydrates for Preterm Infants. J Pediatr. 2013 Oct 1;163(4):1001–7.

54. Arslanoglu S, Moro GE, Ziegler EE. Adjustable fortification of human milk fed to preterm infants: Does it make a difference? J Perinatol. 2006 Aug 3;26:614–21.

55. Arslanoglu S, Moro GE, Ziegler EE. Preterm infants fed fortified human milk receive less protein than they need. J Perinatol. 2009 May 14;29:489–92.

56. Stanger J, Zwicker K, Albersheim S, Murphy JJ. Human milk fortifier: An occult cause of bowel obstruction in extremely premature neonates. J Pediatr Surg. 2014 May 1;49(5):724–6.

57. Pearson F, Johnson MJ, Leaf AA. Milk osmolality: does it matter? Arch Dis Child - Fetal Neonatal Ed. 2013 Mar;98(2):F166–9.

58. Ellis Z-M, Tan HSG, Embleton ND, Sangild PT, van Elburg RM. Milk feed osmolality and adverse events in newborn infants and animals: a systematic review. Arch Dis Child Fetal Neonatal Ed. 2019 May 1;104(3):F333–40.

59. Abiramalatha T, Thomas N, Gupta V, Viswanathan A, McGuire W. High versus standard volume enteral feeds to promote growth in preterm or low birth weight infants. Cochrane Database Syst Rev. 2017 Sep 12;(9):CD012413.

60. Dorling J, Kempley S, Leaf A. Feeding growth restricted preterm infants with abnormal antenatal Doppler results. Arch Dis Child - Fetal Neonatal Ed. 2005

Aug 19;90(5):F359–63.

61. McClure R. Trophic feeding of the preterm infant. Acta Paediatr. 2007 Jan 2;90(s436):19–21.

62. McClure R, Newell S. Randomized controlled study of digestive enzyme activity following trophic feeding. Acta Paediatr. 2007 Jan 2;91(3):292–6.

63. Lucas A, Bloom SR, Aynsley-Green A. Gut hormones and "minimal enteral feeding". Acta Paediatr Scand. 1986 Sep 1;75(5):719–23.

64. McKeown RE, Marsh TD, Amarnath U, Garrison CZ, Addy CL, Thompson SJ, et al. Role of delayed feeding and of feeding increments in necrotizing enterocolitis. J Pediatr. 1992 Nov 1;121(5):764–70.

65. Fanaro S. Feeding intolerance in the preterm infant. Early Hum Dev. 2013 Oct;89(Supplement 2):S13–20.

66. Parker LA, Weaver M, Murgas Torrazza RJ, Shuster J, Li N, Krueger C, et al. Effect of Gastric Residual Evaluation on Enteral Intake in Extremely Preterm Infants. JAMA Pediatr. 2019 Apr 29;173(6):534–43.

67. Mizumoto H, Kawai M, Yamashita S, Hata D. Intraday glucose fluctuation is common in preterm infants receiving intermittent tube feeding. Pediatr Int. 2016 May 1;58(5):359–62.

68. Aynsley-Green A, Adrian TE, Bloom SR. Feeding and the development of enteroinsular hormone secretion in the preterm infant: effects of continuous gastric infusions of human milk compared with intermittent boluses. Acta Paediatr. 1982 May;71(3):379–83.

69. Aynsley-Green A. New Insights into the nutritional management of newborn infants derived from studies of metabolic and endocrine inter-relations during the adaptation to post-natal life. Proc Nutr Soc. 1989 Jul;48(2):283–92.

70. Shulman RJ, Redel CA, Stathos TH. Bolus versus continuous feedings stimulate small-intestinal growth and development in the newborn pig. J Pediatr Gastroenterol Nutr. 1994 Apr;18(3):350–4.

71. El-Kadi SW, Suryawan A, Gazzaneo MC, Srivastava N, Orellana RA,

Nguyen H V., et al. Anabolic signaling and protein deposition are enhanced by intermittent compared with continuous feeding in skeletal muscle of neonates. Am J Physiol Endocrinol Metab. 2012 Mar 15;302(6):E674-886.

72. Greer FR, McCormick A, Loker J. Changes in fat concentration of human milk during delivery by intermittent bolus and continuous mechanical pump infusion. J Pediatr. 1984 Nov 1;105(5):745–9.

73. Jawaheer G, Shaw NJ, Pierro A. Continuous enteral feeding impairs gallbladder emptying in infants. J Pediatr. 2001 Jun 1;138(6):822–5.

74. de Ville K, Knapp E, Al-Tawil Y, Berseth CL. Slow infusion feedings enhance duodenal motor responses and gastric emptying in preterm infants. Am J Clin Nutr. 1998 Jul 1;68(1):103–8.

75. Eichenwald EC, AAP COMMITTEE ON FETUS AND NEWBORN. Diagnosis and Management of Gastroesophageal Reflux in Preterm Infants. Pediatrics. 2018 Jul 1;142(1):e20181061.

76. Cresi F, Martinelli D, Maggiora E, Locatelli E, Liguori SA, Baldassarre ME, et al. Cardiorespiratory events in infants with gastroesophageal reflux symptoms: Is there any association? Neurogastroenterol Motil. 2017 Dec 21;30(5):e13278.

77. Abbasi S, Sivieri EM, McGowan M, Bhutani VK. Effect of intermittent gavage feeds on the pulmonary status of preterm infants. 1118. Pediatr Res. 1996 Apr;39:189.

78. Corvaglia L, Martini S, Aceti A, Capretti MG, Galletti S, Faldella G. Cardiorespiratory events with bolus versus continuous enteral feeding in healthy preterm infants. J Pediatr. 2014 Dec;165(6):1255–7.

79. Fang S, Kempley ST, Gamsu HR. Prediction of early tolerance to enteral feeding in preterm infants by measurement of superior mesenteric artery blood flow velocity. Arch Dis Child Fetal Neonatal Ed. 2001 Jul 1;85(1):F42–5.

80. Pezzati M, Dani C, Tronchin M, Filippi L, Rossi S, Rubaltelli F. Prediction of early tolerance to enteral feeding by measurement of superior mesenteric artery blood flow velocity: appropriate- versus small-for-gestational-age

preterm infants. Acta Paediatr. 2004 Jun 1;93:797–802.

81. Dani C, Pratesi S, Barp J. Continuous milk feeding versus intermittent bolus feeding in preterm infants. Early Hum Dev. 2013 Oct;89:S11–2.

82. Lane AJP, Coombs RC, Evans DH, Levin RJ. Effect of feed interval and feed type on splanchnic haemodynamics. Arch Dis Child Fetal Neonatal Ed. 1998 Jul;79:F49-53.

83. Maggio L, Costa S, Zecca C, Giordano L. Methods of enteral feeding in preterm infants. Early Hum Dev. 2012 May;88 Suppl 2:S31-3.

84. Malcolm WF, Smith PB, Mears S, Goldberg RN, Cotten CM. Transpyloric tube feeding in very low birthweight infants with suspected gastroesophageal reflux: Impact on apnea and bradycardia. J Perinatol. 2009;29:372–5.

85. Macdonald PD, Skeoch CH, Carse H, Dryburgh F, Alroomi LG, Galea P, et al. Randomised trial of continuous nasogastric, bolus nasogastric, and transpyloric feeding in infants of birth weight under 1400 g. Arch Dis Child. 1992 Apr 1;67:429–31.

86. Lang S, Lawrence CJ, Orme RL. Cup feeding: an alternative method of infant feeding. Arch Dis Child. 1994 Oct;71:365–9.

87. Hafström M, Kjellmer I. Non-nutritive sucking in the healthy pre-term infant. Early Hum Dev. 2000 Nov;60:13–24.

88. Pineda R, Dewey K, Jacobsen A, Smith J. Non-Nutritive Sucking in the Preterm Infant. Am J Perinatol. 2019 Feb 6;36(3):268–76.

89. Harding C, Cockerill H, Cane C, Law J. Using non-nutritive sucking to support feeding development for premature infants: A commentary on approaches and current practice. J Pediatr Rehabil Med. 2018;11(3):147–52.

90. Marchini G, Lagercrantz H, Feuerberg Y, Winberg J, Uvnäs-Moberg K. The effect of non-nutritive sucking on plasma insulin, gastrin, and somatostatin levels in infants. Acta Paediatr Scand. 1987 Jul;76(4):573–8.

91. Widström A-M, Marchini G, Matthiesen AS, Werner S, Winberg J, Uvnäs-Moberg K. Nonnutritive Sucking in Tube-Fed Preterm Infants. J Pediatr

Gastroenterol Nutr. 1988 Jul;7(4):517–23.

92. Ng PC. Use of oral erythromycin for the treatment of gastrointestinal dysmotility in preterm infants. Neonatology. 2009 Oct 2;95(2):97–104.

93. Lund M, Pasternak B, Davidsen RB, Feenstra B, Krogh C, Diaz LJ, et al. Use of macrolides in mother and child and risk of infantile hypertrophic pyloric stenosis: nationwide cohort study. BMJ. 2014 Mar 11;348:g1908.

94. Al Faleh K, Al-Omran M. Reporting and methodologic quality of Cochrane Neonatal review group systematic reviews. BMC Pediatr. 2009 Jun 17;9:38.

95. Willhelm C, Girisch W, Gottschling S, Gräber S, Wahl H, Meyer S. Systematic Cochrane reviews in neonatology: a critical appraisal. Pediatr Neonatol. 2013 Aug 1;54(4):261–6.

96. Brok J, Greisen G, Madsen LP, Tilma K, Faerk J, Borch K, et al. Agreement between Cochrane Neonatal reviews and clinical practice guidelines for newborns in Denmark: a cross-sectional study. Arch Dis Child - Fetal Neonatal Ed. 2008 May 1;93:F225–9.

97. Pollock M, Fernandes RM, Becker LA, Pieper D, Hartling L. Chapter V: Overviews of Reviews. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. Cochrane Handbook for Systematic Reviews of Interventions version 61 (updated September 2020). 2020.

98. Methods C. Editorial Decision Tree for Overviews [Internet]. Cochrane Methods; [cited 2021 Mar 27]. Available from: https://methods.cochrane.org/sites/methods.cochrane.org.cmi/files/public/uploads/DecisionChart.pdf

99. Chandler J, Cumpston M, Thomas J, Higgins JP, Deeks JJ, Clarke MJ. Chapter I: Introduction. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. Cochrane Handbook for Systematic Reviews of Interventions version 62 (updated February 2021). Cochrane; 2021.

100. Pollock M, Fernandes RM, Hartling L. Evaluation of AMSTAR to assess the methodological quality of systematic reviews in overviews of reviews of healthcare interventions. BMC Med Res Methodol. 2017 Mar 23;17(1):48.

101. Useem J, Brennan A, LaValley M, Vickery M, Ameli O, Reinen N, et al. Systematic Differences between Cochrane and Non-Cochrane Meta-Analyses on the Same Topic: A Matched Pair Analysis. Smalheiser NR, editor. PLoS One. 2015 Dec 15;10(12):e0144980.

102. Gates M, Elliott SA, Johnson C, Thomson D, Williams K, Fernandes RM, et al. A descriptive analysis of non-Cochrane child-relevant systematic reviews published in 2014. BMC Med Res Methodol. 2018 Oct 1;18(1):99.

103. Crick K, Thomson D, Fernandes RM, Nuspl M, Eurich DT, Rowe BH, et al. Descriptive analysis of cochrane child-relevant systematic reviews: an update and comparison between 2009 and 2013. BMC Pediatr. 2017 Jul 11;17(1):155.

104. Shea B, Moher D, Graham I, Pham B, Tugwell P. A comparison of the quality of Cochrane reviews and systematic reviews published in paper-based journals. Eval Heal Prof. 2002 Mar;25(1):116–29.

105. Pollock M, Fernandes RM, Becker LA, Featherstone R, Hartling L. What guidance is available for researchers conducting overviews of reviews of healthcare interventions? A scoping review and qualitative metasummary. Syst Rev. 2016 Dec 14;5(1):190.

106. Shea BJ, Reeves BC, Wells G, Thuku M, Hamel C, Moran J, et al. AMSTAR 2: a critical appraisal tool for systematic reviews that include randomised or non-randomised studies of healthcare interventions, or both. BMJ. 2017 Sep 21;358:j4008.

107. Reeves BC, Deeks JJ, Higgins JP, Shea B, Tugwell P, Wells G. Chapter 24: Including non-randomized studies on intervention effects. In: Higgins JP, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al., editors. Cochrane Handbook for Systematic Reviews of Interventions version 62 (updated February 2021). Cochrane; 2021.

108. Pieper D, Buechter R, Jerinic P, Eikermann M. Overviews of reviews often have limited rigor: a systematic review. J Clin Epidemiol. 2012 Dec;65(12):1267–73.

109. Brown JVE, Walsh V, McGuire W. Birth room transition support for term and near-term infants: a Cochrane overview. Cochrane Database Syst Rev. 2019 Sep 5;(9):CD013411.

110. Brown JVE, Walsh V, McGuire W. Birth room transition support for preterm infants: a Cochrane overview. Cochrane Database Syst Rev. 2019 Sep 12;(9):CD013428.

111. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al., editors. Cochrane Handbook for systematic reviews of interventions version 6.0 (updated July 2019). In Cochrane; 2019.

112. Büchter RB, Pieper D. Most overviews of Cochrane reviews neglected potential biases from dual authorship. J Clin Epidemiol. 2016 Sep 1;77:91–4.

113. Whiting P, Savović J, Higgins JPT, Caldwell DM, Reeves BC, Shea B, et al. ROBIS: A new tool to assess risk of bias in systematic reviews was developed. J Clin Epidemiol. 2016 Jan;69:225–34.

114. Pieper D, Puljak L, González-Lorenzo M, Minozzi S. Minor differences were found between AMSTAR 2 and ROBIS in the assessment of systematic reviews including both randomized and nonrandomized studies. J Clin Epidemiol. 2019 Apr 1;108:26–33.

115. Banzi R, Cinquini M, Gonzalez-Lorenzo M, Pecoraro V, Capobussi M, Minozzi S. Quality assessment versus risk of bias in systematic reviews: AMSTAR and ROBIS had similar reliability but differed in their construct and applicability. J Clin Epidemiol. 2018 Jul 1;99:P24-32.

116. Bühn S, Mathes T, Prengel P, Wegewitz U, Ostermann T, Robens S, et al. The risk of bias in systematic reviews tool showed fair reliability and good construct validity. J Clin Epidemiol. 2017 Nov 1;91:121–8.

117. Lorenz RC, Matthias K, Pieper D, Wegewitz U, Morche J, Nocon M, et al. A psychometric study found AMSTAR 2 to be a valid and moderately reliable appraisal tool. J Clin Epidemiol. 2019 Oct 1;114:133–40.

118. Shepherd E, Salam RA, Middleton P, Han S, Makrides M, McIntyre S, et al. Neonatal interventions for preventing cerebral palsy: an overview of Cochrane

Systematic Reviews. Cochrane database Syst Rev. 2018 Jun 20;(6):CD012409.

119. Medley N, Vogel JP, Care A, Alfirevic Z. Interventions during pregnancy to prevent preterm birth: an overview of Cochrane systematic reviews. Cochrane database Syst Rev. 2018 Nov 14;(11):CD012505.

120. Amari S, Shahrook S, Namba F, Ota E, Mori R. Branched-chain amino acid supplementation for improving growth and development in term and preterm neonates. Cochrane database Syst Rev. 2020 Oct 2;(10):CD012273.

121. Moon K, Rao SC, Schulzke SM, Patole SK, Simmer K. Longchain polyunsaturated fatty acid supplementation in preterm infants. Cochrane database Syst Rev. 2016 Dec 20;(12):CD000375.

122. Watson J, McGuire W. Responsive versus scheduled feeding for preterm infants. Cochrane Database Syst Rev. 2016 Aug 31;(8):CD005255.

123. Flint A, New K, Davies MW. Cup feeding versus other forms of supplemental enteral feeding for newborn infants unable to fully breastfeed. Cochrane Database Syst Rev. 2016 Aug 31;(8):CD005092.

124. Foster JP, Psaila K, Patterson T. Non-nutritive sucking for increasing physiologic stability and nutrition in preterm infants. Cochrane database Syst Rev. 2016 Oct 4;(10):CD001071.

125. Watson J, McGuire W. Nasal versus oral route for placing feeding tubes in preterm or low birth weight infants. Cochrane database Syst Rev. 2013 Feb 28;(2):CD003952.

126. Brown JVE, Walsh V, McGuire W. Formula versus maternal breast milk for feeding preterm or low birth weight infants. Cochrane database Syst Rev. 2019 Aug 12;(8):CD002972.

127. Walsh V, Brown JVE, Askie LM, Embleton ND, McGuire W. Nutrient-enriched formula versus standard formula for preterm infants. Cochrane database Syst Rev. 2019 Jul 17;(7):CD004204.

128. Thanigainathan S, Abiramalatha T. Early fortification of human milk versus late fortification to promote growth in preterm infants. Cochrane database

Syst Rev. 2020 Jul 29;(7):CD013392.

129. Nehra V, Genen LH, Brumberg HL. High versus low medium chain triglyceride content of formula for promoting short-term growth of preterm infants. Cochrane Database Syst Rev. 2002 Jul 22;(1):CD002777.

130. Tan-Dy CRY, Ohlsson A. Lactase treated feeds to promote growth and feeding tolerance in preterm infants. Cochrane database Syst Rev. 2013 Mar 28;(3):CD004591.

131. Morgan J, Bombell S, McGuire W. Early trophic feeding versus enteral fasting for very preterm or very low birth weight infants. Cochrane database Syst Rev. 2013 Mar 28;(3):CD000504.

132. Dawson JA, Summan R, Badawi N, Foster JP. Push versus gravity for intermittent bolus gavage tube feeding of premature and low birth weight infants. Cochrane database Syst Rev. 2012 Nov 14;(11):CD005249.

133. Premji SS, Chessell L. Continuous nasogastric milk feeding versus intermittent bolus milk feeding for premature infants less than 1500 grams. Cochrane Database Syst Rev. 2011 Nov 9;(11):CD001819.

134. Ng E, Shah VS. Erythromycin for the prevention and treatment of feeding intolerance in preterm infants. Cochrane database Syst Rev. 2008 Jul 16;(3):CD001815.

135. Watson J, McGuire W. Transpyloric versus gastric tube feeding for preterm infants. Cochrane database Syst Rev. 2013 Feb 28;(2):CD003487.

136. Oddie SJ, Young L, McGuire W. Slow advancement of enteral feed volumes to prevent necrotising enterocolitis in very low birth weight infants. Cochrane Database Syst Rev. 2017 Aug 30;(8):CD001241.

137. Morgan J, Young L, McGuire W. Delayed introduction of progressive enteral feeds to prevent necrotising enterocolitis in very low birth weight infants. Cochrane Database Syst Rev. 2014 Dec 1;(3):CD001970.

138. Walsh V, Brown JVE, Copperthwaite BR, Oddie SJ, McGuire W. Early full enteral feeding for preterm or low birth weight infants. Cochrane Database Syst Rev. 2020 Mar 12;2020(3).

139. Fabrizio V, Trzaski JM, Brownell EA, Esposito P, Lainwala S, Lussier MM, et al. Targeted or adjustable versus standard diet fortification for growth and development in very low birth weight infants receiving human milk. Cochrane Database Syst Rev. 2019 Nov 4;(11):CD013465.

140. Ibrahim NR, Van Rostenberghe H, Ho JJ. Short versus long feeding interval for bolus feedings in very preterm infants. Cochrane Database Syst Rev. 2016 Aug 17;(8):CD012322.

141. Gao C, Miller J, Collins CT, Rumbold A. Comparison of different protein concentrations of human milk fortifier for promoting growth and neurological development in preterm infants. Cochrane Database Syst Rev. 2008 Apr 16;(2).

142. Dempsey E, Miletin J. Banked preterm versus banked term human milk to promote growth and development in very low birth weight infants. Cochrane database Syst Rev. 2019 Jun 28;(6):CD007644.

143. Brown JVE, Lin L, Embleton ND, Harding JE, McGuire W. Multi-nutrient fortification of human milk for preterm infants. Cochrane Database Syst Rev. 2020 Jun 3;(5):CD000343.

144. Amissah EA, Brown J, Harding JE. Fat supplementation of human milk for promoting growth in preterm infants. Cochrane database Syst Rev. 2020 Aug 25;(8):CD000341.

145. Quigley M, Embleton ND, McGuire W. Formula versus donor breast milk for feeding preterm or low birth weight infants. Cochrane Database Syst Rev. 2019 Jul 19;(4):CD002971.

146. Amissah EA, Brown J, Harding JE. Carbohydrate supplementation of human milk to promote growth in preterm infants. Cochrane database Syst Rev. 2020 Sep 8;(9):CD000280.

147. Basuki F, Hadiati DR, Turner T, McDonald S, Hakimi M. Dilute versus full-strength formula in exclusively formula-fed preterm or low birth weight infants. Cochrane database Syst Rev. 2019 Jun 27;(6):CD007263.

148. Amissah EA, Brown J, Harding JE. Protein supplementation of human milk

for promoting growth in preterm infants. Cochrane database Syst Rev. 2020 Sep 23;(9):CD000433.

149. Fenton TR, Al-Wassia H, Premji SS, Sauve RS. Higher versus lower protein intake in formula-fed low birth weight infants. Cochrane database Syst Rev. 2020 Jun 23;(6):CD003959.

150. Ng DHC, Klassen JRL, Embleton ND, McGuire W. Protein hydrolysate versus standard formula for preterm infants. Cochrane database Syst Rev. 2019 Jul 24;(7):CD012412.

151. Premkumar MH, Pammi M, Suresh G. Human milk-derived fortifier versus bovine milk-derived fortifier for prevention of mortality and morbidity in preterm neonates. Cochrane database Syst Rev. 2019 Nov 7;2019(11):CD013145.

152. Nasuf AWA, Ojha S, Dorling J. Oropharyngeal colostrum in preventing mortality and morbidity in preterm infants. Cochrane database Syst Rev. 2018 Sep 7;(9):CD011921.

153. Ng DHC, Klassen J, Embleton ND, McGuire W. Protein hydrolysate versus standard formula for preterm infants. Cochrane Database Syst Rev. 2017 Oct 2;2016(10).

154. Yeo KT, Kong JY, Sasi A, Tan K, Lai NM, Schindler T. Stopping enteral feeds for prevention of transfusion-associated necrotising enterocolitis in preterm infants. Cochrane database Syst Rev. 2019 Oct 28;(10):CD012888.

155. Abiramalatha T, Thanigainathan S, Ninan B. Routine monitoring of gastric residual for prevention of necrotising enterocolitis in preterm infants. Cochrane database Syst Rev. 2019 Jul 9;(7):CD012937.

156. Abiramalatha T, Thanigainathan S, Balakrishnan U. Re-feeding versus discarding gastric residuals to improve growth in preterm infants. Cochrane Database Syst Rev. 2019 Jul 8;(7):CD012940.

157. Muelbert M, Lin L, Bloomfield FH, Harding JE. Exposure to the smell and taste of milk to accelerate feeding in preterm infants. Cochrane database Syst Rev. 2019 Jul 16;(7):CD013038.

158. Anabrees J, Shah VS, AlOsaimi A, AlFaleh K. Glycerin laxatives for prevention or treatment of feeding intolerance in very low birth weight infants. Cochrane database Syst Rev. 2015 Sep 30;(9):CD010464.

159. Greene Z, O'Donnell CPF, Walshe M. Oral stimulation for promoting oral feeding in preterm infants. Cochrane database Syst Rev. 2016 Sep 20;(9):CD009720.

160. World Health Organization (WHO). Feeding of very-low-birth-weight infants [Internet]. 2019 [cited 2021 Mar 24]. Available from: https://www.who.int/elena/titles/feeding_vlbw_infants/en/

161. World Health Organization (WHO). Mother's milk for low-birth-weight infants [Internet]. 2019 [cited 2021 Mar 24]. Available from: https://www.who.int/elena/titles/mothersmilk_infants/en/

162. World Health Organization (WHO). Donor human milk for low-birth-weight infants [Internet]. 2019 [cited 2021 Mar 24]. Available from: https://www.who.int/elena/titles/donormilk_infants/en/

163. Ryan AS, Hay WW. Challenges of infant nutrition research: a commentary. Nutr J. 2016 Dec 22;15(1):42.

164. Sterne JAC, Savović J, Page MJ, Elbers RG, Blencowe NS, Boutron I, et al. RoB 2: a revised tool for assessing risk of bias in randomised trials. BMJ. 2019 Aug 28;366:l4898.

165. Moher D, Pham B, Jones A, Cook DJ, Jadad AR, Moher M, et al. Does quality of reports of randomised trials affect estimates of intervention efficacy reported in meta-analyses? Lancet. 1998 Aug 22;352(9128):609–13.

166. Odgaard-Jensen J, Vist GE, Timmer A, Kunz R, Akl EA, Schünemann H, et al. Randomisation to protect against selection bias in healthcare trials. Cochrane Database Syst Rev. 2011 Apr 13;(4):MR000012.

167. Tiffany KF, Burke BL, Collins-Odoms C, Oelberg DG. Current practice regarding the enteral feeding of high-risk newborns with umbilical catheters in situ. Pediatrics. 2003 Jul 1;112(1):20–3.

168. Patel EU, Wilson DA, Brennan EA, Lesher AP, Ryan RM. Earlier re-

initiation of enteral feeding after necrotizing enterocolitis decreases recurrence or stricture: a systematic review and meta-analysis. J Perinatol. 2020 Nov 18;40:1679–87.

169. Hock A, Chen Y, Miyake H, Koike Y, Seo S, Pierro A. Initiation of Enteral Feeding After Necrotizing Enterocolitis. Eur J Pediatr Surg. 2018 Feb 24;28(01):44–50.

170. Ou J, Courtney CM, Steinberger AE, Tecos ME, Warner BW. Nutrition in Necrotizing Enterocolitis and Following Intestinal Resection. Nutrients. 2020 Feb 18;12(2):520.

171. Zeitlin J, Ancel PY, Saurel-Cubizolles MJ, Papiernik E. The relationship between intrauterine growth restriction and preterm delivery: an empirical approach using data from a European case-control study. BJOG An Int J Obstet Gynaecol. 2000 Jun;107(6):750–8.

172. Subramanian S, Murki S. The feeding conundrum. Transl Pediatr. 2017 Apr 1;6(2):86–7.

173. Mustafa RA, Santesso N, Brozek J, Akl EA, Walter SD, Norman G, et al. The GRADE approach is reproducible in assessing the quality of evidence of quantitative evidence syntheses. J Clin Epidemiol. 2013 Jul 1;66(7):736-742.e5.

174. West J, Manchester B, Wright J, Lawlor DA, Waiblinger D. Reliability of routine clinical measurements of neonatal circumferences and research measurements of neonatal skinfold thicknesses: findings from the Born in Bradford study. Paediatr Perinat Epidemiol. 2011 Mar;25(2):164–71.

175. Wood AJ, Raynes-Greenow CH, Carberry AE, Jeffery HE. Neonatal length inaccuracies in clinical practice and related percentile discrepancies detected by a simple length-board. J Paediatr Child Health. 2013 Mar 1;49(3):199–203.

176. Guyatt GH, Oxman AD, Vist GE, Kunz R, Falck-Ytter Y, Alonso-Coello P, et al. GRADE: an emerging consensus on rating quality of evidence and strength of recommendations. BMJ. 2008 Apr 26;336(7650):924–6.

177. Easterbrook P., Gopalan R, Berlin J., Matthews D. Publication bias in clinical

research. Lancet. 1991 Apr 13;337(8746):867–72.

178. Sterne JAC, Gavaghan D, Egger M. Publication and related bias in meta-analysis: power of statistical tests and prevalence in the literature. J Clin Epidemiol. 2000 Nov 1;53(11):1119–29.

179. Boutron I, Page MJ, Higgins JP, Altman DG, Lundh A, Hróbjartsson A. Chapter 7: Considering bias and conflicts of interest among the included studies. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, et al., editors. Cochrane Handbook for Systematic Reviews of Interventions version 62 (updated February 2021). Cochrane; 2021.

180. UNICEF. Guide to the UNICEF UK Baby Friendly Initiative Standards (Second edition). 2017.

181. Jarrold K, Helfer B, Eskander M, Crawley H, Trabulsi J, Caulfield LE, et al. Guidance for the conduct and reporting of clinical trials of breast milk substitutes. JAMA Pediatr. 2020 Sep 1;174(9):874–81.

182. Østengaard L, Lundh A, Tjørnhøj-Thomsen T, Abdi S, Gelle MHA, Stewart LA, et al. Influence and management of conflicts of interest in randomised clinical trials: qualitative interview study. BMJ. 2020 Oct 27;371:m3764.

183. Higgins J, Thomas J, Chandler J, Cumpston M, Li T, Page M, et al., editors. Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021). Cochrane; 2021.

184. Quigley MA, Poulsen G, Boyle E, Wolke D, Field D, Alfirevic Z, et al. Early term and late preterm birth are associated with poorer school performance at age 5 years: a cohort study. Arch Dis Child Fetal Neonatal Ed. 2012 May;97:F167-73.

14 Definitions

NEC: Necrotising enterocolitis

VLBW: Very Low Birthweight

NNTB: Number Needed to Treat for additional Benefit

NNTH: Number Needed to Treat for additional Harm

MID: Minimal Important Difference

IgA: Immunoglobulin A

DBM: Donor breast milk

LCT: Long Chain Triglyceride

MCT: Medium Chain Triglyceride

PNALD: Parenteral Nutrition Associated Liver Disease

GOR: Gastro-oesophageal Reflux

NG: Nasogastric

OG: Orogastric

CDSR: Cochrane Database of Systematic Reviews

VW: Verena Walsh

WM: William McGuire

SO: Sam Oddie

GRADE: Grading of Recommendations Assessment, Development and Evaluation

AMSTAR2: A Measurement Tool to Assess Systematic Reviews 2

ROBIS: Risk of Bias in Included Studies

IRR: Inter-rater Reliability

NICU: Neonatal Intensive Care Unit

SIP: Spontaneous Intestinal Perforation

RoB: Risk of Bias

PICO: Population, Intervention, Comparison, Outcome

RR: Relative Risk or Risk Ratio

IUGR: Intrauterine Growth Restriction

PDA: Patent Ductus Arteriosus

MD: Mean Difference

CI: Confidence Interval

SD: Standard Deviation

PMA: Postmenstrual age

RS: Randomsisation sequence

SB: Selection Bias

PB: Performance Bias

DB: Detection Bias

IO: Incomplete Outcome/Attrition bias

RB: Reporting Bias/Selective Reporting

OB: Other bias

AREDF: Absent or Reversed End Diastolic Flow

15 Appendix

15.1 Appendix 1: Data collection form

Title			
Author			
Publication date			
Date of most recent search/update			
Study types considered			
Population (as per protocol)	Gestational age		
	Birthweight		
	Setting		
	Other		
Intervention			
Comparator			
Outcomes	Reported (eligible studies for these outcomes)		
	Not reported (no data)		
Search strategy	Databases searched	Cochrane Central Register of Controlled Trials (CENTRAL)	
		MEDLINE	
		Embase	
		Cumulative Index to Nursing and Allied Health Literature (CINAHL)	

		Clinicaltrials.gov	
		World Health Organization International Trials Registry and Platform	
		International Standard Randomized Controlled Trials Number (ISRCTN) Registry	
		Reference lists of articles included in the review	
		Additional conference proceedings	
		Ovid Maternity & Infant Care Database	
		Other	
	Language restriction		
	Search terms and limits		
Number of studies included			
Study designs			
Years studies conducted			
Size of smallest and largest studies included	Smallest		
	Largest		
Number of participants included			
Population (of participants in included studies)	Gestational age		
	Birthweight		
	Setting		
	Other		

Methodological quality of included studies:			
Cochrane Risk of Bias Tool/Jadad	Assessed in review?	Decision (state how many studies low risk, unclear risk, high risk)	Reasoning (Reasons given by review authors for assessing studies as low risk, unclear risk, high risk)
Random sequence generation (selection bias)			
Allocation concealment (selection bias)			
Blinding of participants and personnel (performance bias)			
Blinding of outcome assessment (detection bias)			
Incomplete outcome data (attrition bias)			
Selective reporting (reporting bias)			
Other bias			

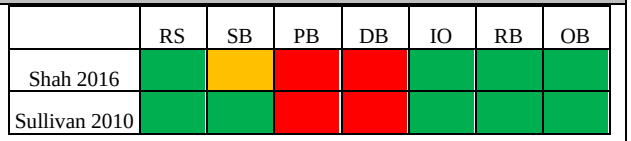
Effects of interventions							
Outcome reported	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	GRADE reason	Analysed with random/fixed effect model?

15.2 Appendix 2: Review author assessment of Risk of bias in the included studies

15.2.1 Appendix 2a: What to feed

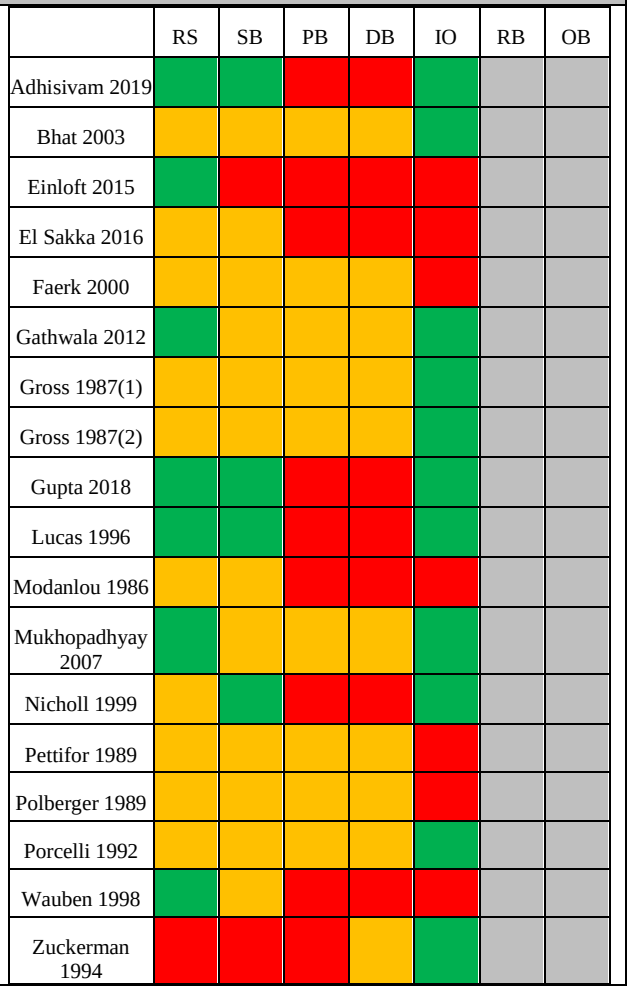
RS = Randomsisation sequence; SB = Selection Bias; PB = Performance Bias; DB = Detection Bias; IO = Incomplete Outcome; RB= Reporting Bias; OB = Other bias

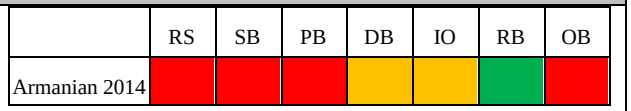
Table 17: Appendix 2a: Review author assessment of quality of the included studies (What to feed)

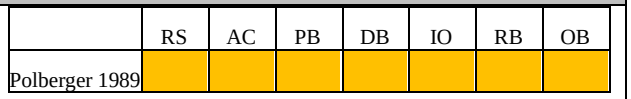
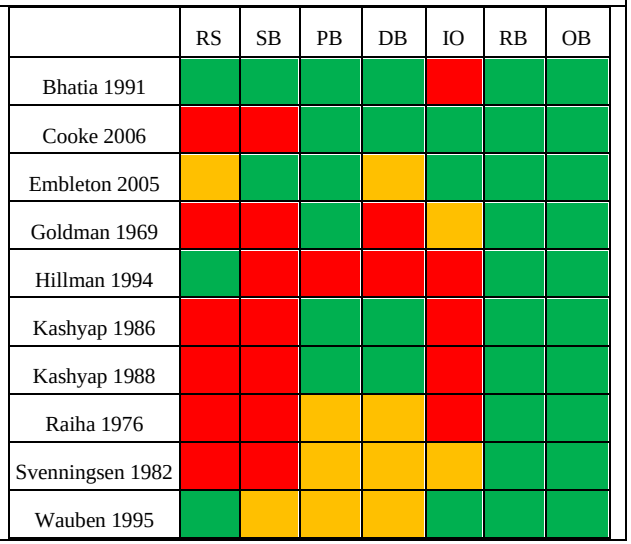
Study	Assessment		RoB table
Thanigainathan *et al.* 2020[128]: Early vs later fortification of human milk	Sequence generation	2 low risk	
	Allocation concealment	1 low risk, 1 unclear risk	
	Performance bias	2 high risk	
	Detection bias	2 high risk	
	Incomplete outcome	2 low risk	
	Reporting bias	Low risk	
	Other bias	Low risk	

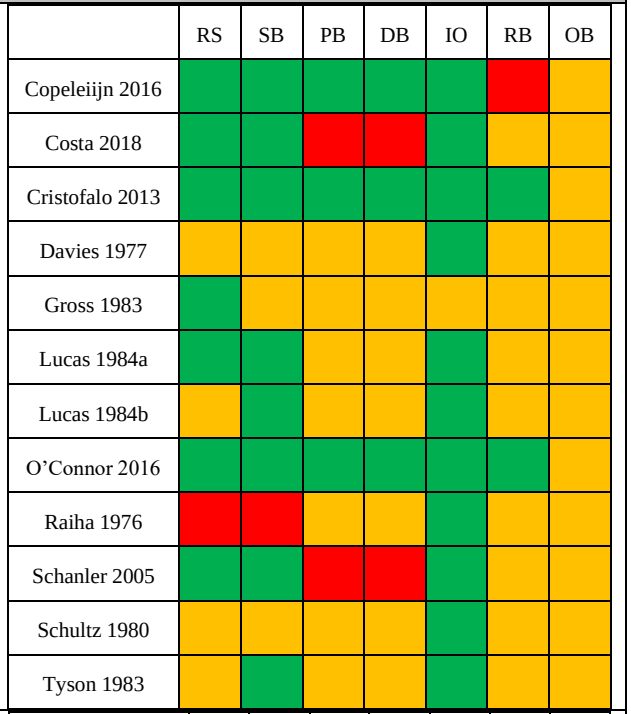
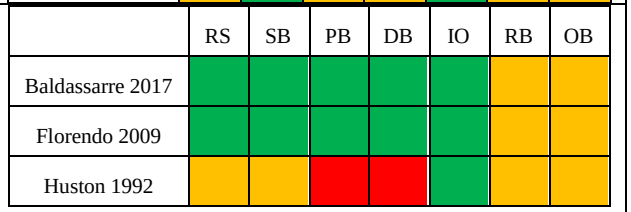
Study	Assessment	
Brown *et al.* 2020[143]: Multi-nutrient fortified vs unfortified breast milk	Sequence generation	7 Low risk, 10 unclear, 1 high risk
	Allocation concealment	4 Low risk, 12 unclear, 2 high risk
	Performance bias	9 unclear, 9 high risk
	Detection bias	10 unclear, 8 high risk
	Incomplete outcome	11 Low risk, 7 high risk
	Reporting bias	Not assessed
	Other bias	Not assessed

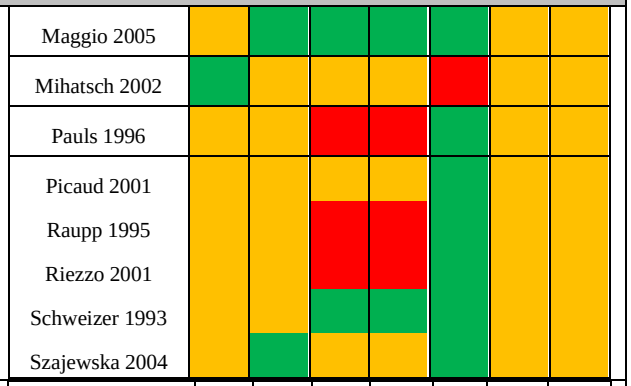
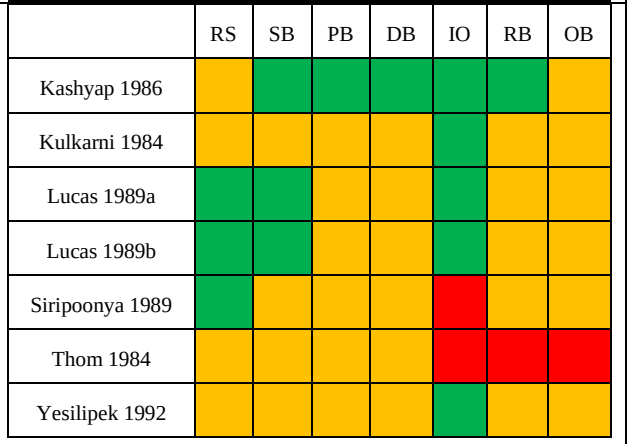
RoB table

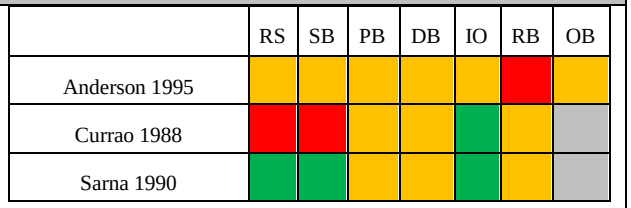
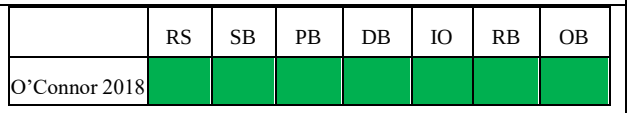


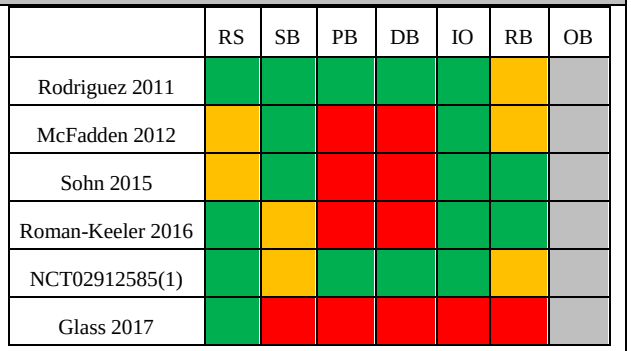
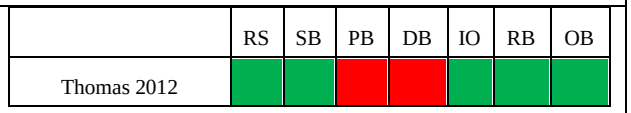
Study	Assessment		RoB table
Amissah *et al.* 2020a[146]: "Carbohydrate (prebiotic) supplementation vs no supplementation of human milk"	Sequence generation	High risk	
	Allocation concealment	High risk	
	Performance bias	High risk	
	Detection bias	Unclear risk	
	Incomplete outcome	Unclear risk	
	Reporting bias	Low risk	
	Other bias	High risk	
Amissah *et al.* 2020b[148]: "Protein supplementation versus no supplementation of human milk"	Sequence generation	6 Unclear risk	
	Allocation concealment	6 unclear risk	
	Performance bias	1 low risk, 5 unclear risk	
	Detection bias	6 unclear risk	
	Incomplete outcome	2 low risk, 4 unclear risk	
	Reporting bias	4 low risk 1 unclear risk, 1 high risk	
	Other bias	2 low risk, 4 unclear risk	

Study	Assessment		RoB table
Amissah *et al.* 2020c[144]: "Fat supplemented human milk versus control" Amissah *et al.* 2020c[144]: "Fat supplemented human milk versus control" (continued)	Sequence generation	1 unclear risk	
	Allocation concealment	1 unclear risk	
	Performance bias	1 unclear risk	
	Detection bias	1 unclear risk	
	Incomplete outcome	1 unclear risk	
	Reporting bias	1 unclear risk	
	Other bias	1 unclear risk	
Fenton *et al.* 2020[149]: High versus low / Very high versus high protein intake in formula fed infants	Sequence generation	3 low risk, 2 unclear risk, 3 not assessed	
	Allocation concealment	2 low risk, 4 unclear risk, 2 not assessed	
	Performance bias	3 low risk, 3 unclear risk, 2 high risk	
	Detection bias	3 low risk, 3 unclear risk, 2 high risk	
	Incomplete outcome	3 low risk, 5 high risk	
	Reporting bias	2 low risk, 6 not assessed	
	Other bias	2 unclear risk, 6 not assessed	

Study	Assessment		RoB table
Quigley *et al.* 2019[145] : "Formula milk vs Donor breast milk"	Sequence generation	7 Low risk, 4 unclear, 1 high risk	
	Allocation concealment	6 Low risk, 5 unclear, 1 high risk	
	Performance bias	4 Low risk, 7 unclear, 2 high risk	
	Detection bias	4 Low risk, 7 unclear, 2 high risk	
	Incomplete outcome	11 Low risk, 1 unclear, 0 high risk	
	Reporting bias	2 Low risk, 10 unclear, 1 high risk	
	Other bias	0 Low risk, 12 unclear, 0 high risk	
Ng *et al.* 2019[150]: "Hydrolysed versus non-hydrolysed formula"	Sequence generation	3 Low risk, 9 unclear risk	
	Allocation concealment	4 Low risk, 8 unclear risk	
	Performance bias	4 Low risk, 4 unclear, 4 high risk	
	Detection bias	4 Low risk, 4 unclear, 4 high risk	

Study	Assessment		RoB table
Ng *et al.* 2019[150] (continued)	Incomplete outcome	11 Low risk, 1 high risk	
	Reporting bias	12 unclear risk	
	Other bias	12 unclear risk	
Walsh *et al.* 2019[127]: "Nutrient-enriched formula versus standard formula"	Sequence generation	3 Low risk, 4 unclear	
	Allocation concealment	3 Low risk, 4 unclear	
	Performance bias	1 Low risk, 6 unclear	
	Detection bias	1 Low risk, 6 unclear	
	Incomplete outcome	5 Low risk, 2 high risk	
	Reporting bias	1 Low risk, 5 unclear, 1 high risk	
	Other bias	6 unclear, 1 high risk	
Dempsey *et al.* 2019[142]: "Banked preterm versus banked term human milk"	No eligible studies found		

Study	Assessment		RoB table
Brown *et al.* 2019[126]: "Formula versus maternal breastmilk"	No eligible studies found		
Basuki *et al.* 2019[147]: "Dilute versus full-strength formula"	Sequence generation	1 Low risk, 1 unclear, 1 high risk	
	Allocation concealment	1 Low risk, 1 unclear, 1 high risk	
	Performance bias	3 unclear risk	
	Detection bias	3 unclear risk	
	Incomplete outcome	2 Low risk, 1 unclear	
	Reporting bias	2 unclear risk, 1 high risk	
	Other bias	3 Low risk	
Premkumar *et al.* 2019[151] : "Human milk-derived fortifier versus bovine milk-derived fortifier"	Sequence generation	1 Low risk	
	Allocation concealment	1 Low risk	
	Performance bias	1 Low risk	
	Detection bias	1 Low risk	
	Incomplete outcome	1 Low risk	
	Reporting bias	1 Low risk	
	Other bias	1 Low risk	

Study	Assessment		RoB table
Nasuf *et al.* 2018[152] : "Oropharyngeal colostrum (OPC) compared to control (water, saline or no intervention) in preterm infants"	Sequence generation	4 Low risk, 2 unclear risk	
	Allocation concealment	3 Low risk, 2 unclear, 1 high risk	
	Performance bias	2 Low risk, 4 high risk	
	Detection bias	2 Low risk, 4 high risk	
	Incomplete outcome	5 Low risk,1 high risk	
	Reporting bias	2 Low risk, 3 unclear, 1 high risk	
	Other bias	Not assessed	
Abiramalatha *et al.* 2017[59]: "High-volume vs standard-volume feeds"	Sequence generation	Low risk	
	Allocation concealment	Low risk	
	Performance bias	High risk	
	Detection bias	High risk	
	Incomplete outcome	Low risk	
	Reporting bias	Low risk	
	Other bias	Low risk	

<table>
<tr><th>Study</th><th colspan="2">Assessment</th><th>RoB table</th></tr>
<tr><td rowspan="7">Tan-Dy et al. 2013[130]:
“Lactase treated feeds vs placebo”</td><td>Sequence generation</td><td>Unclear risk</td><td rowspan="7"><table><tr><th></th><th>RS</th><th>SB</th><th>PB</th><th>DB</th><th>IO</th><th>RB</th><th>OB</th></tr><tr><td>Erasmus 2002</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table></td></tr>
<tr><td>Allocation concealment</td><td>Unclear risk</td></tr>
<tr><td>Performance bias</td><td>Low risk</td></tr>
<tr><td>Detection bias</td><td>Low risk</td></tr>
<tr><td>Incomplete outcome</td><td>Low risk</td></tr>
<tr><td>Reporting bias</td><td>Unclear risk</td></tr>
<tr><td>Other bias</td><td>Not assessed</td></tr>
<tr><td rowspan="7">Nehra et al. 2002[129]:
“High MCT formula versus low MCT formula”</td><td>Sequence generation</td><td>Risk not assessed: 5 “Can’t tell”</td><td rowspan="7"><table><tr><th></th><th>RS</th><th>SB</th><th>PB</th><th>DB</th><th>IO</th><th>RB</th><th>OB</th></tr><tr><td>Huston 1983</td><td>Can’t tell</td><td>Yes</td><td>Can’t tell</td><td>Can’t tell</td><td>Complete follow up</td><td></td><td></td></tr><tr><td>Okamtoto 1982</td><td>Can’t tell</td><td>Can’t tell</td><td>CC yes, RT no.</td><td>Can’t tell</td><td>Complete follow up</td><td></td><td></td></tr><tr><td>Sulkers 1992</td><td>Can’t tell</td><td>Can’t tell</td><td>Can’t tell</td><td>Can’t tell</td><td>Complete follow up</td><td></td><td></td></tr><tr><td>Hamosh 1989a</td><td>Can’t tell</td><td>Can’t tell</td><td>CC yes. RT can’t tell</td><td>Can’t tell</td><td>Complete follow up</td><td></td><td></td></tr><tr><td>Hamosh 1991b</td><td>Can’t tell</td><td>Can’t tell</td><td>CC yes. RT can’t tell</td><td>Can’t tell</td><td>Complete follow up</td><td></td><td></td></tr><tr><td>Whyte 1966</td><td>Can’t tell</td><td>Yes</td><td>Yes</td><td>Can’t tell</td><td>Complete follow up</td><td></td><td></td></tr></table>CC = clinical caretakers, RT = research team</td></tr>
<tr><td>Allocation concealment</td><td>2 Low risk, 4 unclear risk</td></tr>
<tr><td>Performance bias</td><td>Risk not assessed: 2 “Can’t tell”; 1 “clinical caretakers (CC) yes, research team (RT) no”; 2 “clinical caretakers yes, research team can’t tell”</td></tr>
<tr><td>Detection bias</td><td>Risk not assessed: 5 “Can’t tell”</td></tr>
<tr><td>Incomplete outcome</td><td>Risk not assessed: 5 “Complete follow up”</td></tr>
<tr><td>Reporting bias</td><td>Not assessed</td></tr>
<tr><td>Other bias</td><td>Not assessed</td></tr>
</table>

Randomisation sequence generation (RS) and Allocation concealment (SB)

The majority of systematic reviews included evidence with a combination of reviews mostly at low risk or unclear risk of bias in randomisation sequence generation and allocation concealment.

Two reviews included a single study that was assessed as low risk of bias for both randomisation sequence generation and allocation concealment [59,151].

Three reviews[130,144,148] assessed all of their studies as unclear risk of randomisation bias, of which two studies only included a single study [130,144]. The studies assessed as unclear risk provided no or insufficient description of how the randomisation sequence was generated.

One review assessed all of their studies as low risk of randomisation bias as both randomised by block randomisation, with one study also low risk at allocation concealment but one study at unclear risk due to not describing how this was done, further adding that a fixed block size of 4 in an unmasked single centre trial would make allocation of each 4th infant predictable [128].

Three reviews[143,145,147] identified one of their studies as high risk for random sequence generation with a further study [146] assessing the only included study as high risk. These were assessed as high risk as the randomisation sequence was at least in part predictable [145], had allocation odd/even numbers [143,146,147].

One out of date review only assessed risk of selection bias, but all included studies were stated as "cannot tell" for sequence generation in the characteristics of studies[129]. Allocation concealment was present for 2/6 reviews, otherwise “can't tell”. They described that the outcome describing that random allocation was claimed in all included studies, but the technique used for sequence generation was not described in any of the studies [129].

Three reviews assessed all included studies as unclear risk for both random sequence generation and allocation concealment [130,144,148], reasoning these with lack of methodological details. Some trials assessed reviews as unclear risk when trials described using sealed envelopes, but not specifying if these were opaque [143,148,150], while other studies marked these studies as low risk [127,152].

The recently updated review by Fenton *et al.*[149] has inconsistent assessment of bias from random sequence generation, judging one study as unclear risk with the randomisation sequence not described, but four studies as high risk with no randomisation described. This may be as in the study which was assessed as unclear risk, the allocation concealment was low risk while in the other four studies this was also unclear. Two studies were assessed as high risk for random sequence generation but were reasoned with "assigned randomly".

On assessment of allocation concealment (selection bias), Fenton *et al.* [149] assessed 70% of the included studies at high risk of selection bias, reasoning this with that the allocation concealment was not described or "can't tell". Only one study was assessed as unclear risk with this reasoning. Six reviews assessed some of their included studies as high risk [143,145–147,149,152]. Two of these included only a single study [146,152]. Amissah *et al.*[146] assessed the single included study as high risk as the caregivers were not blinded to the carbohydrate supplementation, therefore the sequence was easy to predict. Quigley *et al.*[145] assessed a study as high risk as every 5th infant was allocated to the donor breastmilk arm. Two reviews assessed studies at high risk as allocation by odd/even numbers were easy to predict [143,147]. Two reviews assessed studies at high risk as they stated that allocation concealment was not used or "not applicable" [143,152].

Performance and detection bias

The majority of reviews had a combination of outcomes for the risk of performance and detection bias. Some trials assessed performance and detection bias as one outcome [145]. All except three reviews assessed the two categories as the same risk of bias [143,148,149].Only two studies had all included studies assessed as low risk [130,151]. The single study in Premkumar *et al.*[151] achieved masking through amber coloured tubing and coloured paper wrapped bottles. In Tan-Dy *et al.* [130] the masking was achieved by placebo solution in an identical carrier agent and bottles, with randomisation information only known to the research nurse and central food production staff.

Low risk of bias for masking was also assessed in some of the included studies [127,145,148–150]. In some reviews trials were assessed at low risk if they were described as "double blind" [149,150] or described families, clinicians, caregivers and in some

cases specified investigators as "blinded" or "masked" [127,145,146,149]. In some trials intervention and control formulas were described as "identical in colour and smell" [150] or colour coded [127]. Nasuf *et al.*[152] assessed studies as low risk when clinical and research staff were blinded with preparation performed by an independent research member, or syringes were covered by opaque tape.

No included studies were found to be low risk for performance and detection bias in six revires [59,128,143,144,146,147].

High risk of bias for performance and detection bias was found in eight reviews [59,128,143,145,146,149,150,152]. Trials were assessed as high risk if they were "unblinded" [143,150,152,155], families and clinicians were not masked [145], "unlikely to be blinded" with only abstract available [143,150] or not blinded to allow "proper handling of mother's own milk" [128]. Thanigainathan *et al.* [128] assessed one included study as high risk for detection bias despite it being "not mentioned" but had assessed the performance bias at high risk due to lack of true blinding of infants' caregivers. Nasuf *et al.* [152] assessed one trial as high risk where the allocation was explicitly labelled at the bedside. Fenton *et al.* [149] assessed one trial as high risk for detection bias as a physician aware of allocation codes performed assessment of infants.

Different levels of bias for the two categories were assessed in three studies. In Amissah *et al.*[146], performance bias was assessed at high risk as the intervention was unblinded, but assessed the detection bias at unclear risk due to insufficient methodological detail. Brown *et al.*[143] assessed one trial as high risk for performance bias as it was unblinded, but unclear risk for detection bias as radiographers were reported to have been blinded.

In the review by Nasuf *et al.* [152], when detection bias for outcomes were assessed at high risk, the additional outcome for detection bias for death before discharge home was assessed as low risk, as it was felt unlikely to be affected.

All studies were assessed as unclear risk in two studies [144,147]. The trials in Basuki *et al.* were assessed as this when not described but unlikely given nature of the trial, while three trials [143,144,148] assessed this when trials stated they were double blinded but did not specify who was blinded. Other trials assessed studies as unclear risk when no information was given [127,148,184]. Walsh *et al.*[127] assessed one study at unclear risk of performance bias despite formulas only identifiable by numerical

code. Ng *et al.*[150] assessed one trial as unclear risk despite being described as "double blind" for both performance and detection bias, as the investigators acknowledged different tastes, texture or smell between the two formulas. It also assessed one trial unclear risk for both despite investigators being unaware of the formula, but unclear if carers or parents aware. Fenton *et al.*[149] assessed a trial as unclear risk for detection bias if "bottle colour coded by manufacturer" but outcome assessment blinding was not described.

Inconsistent assessment of this outcome was seen in Fenton *et al.*[149], which assessed some studies at high risk of performance and detection bias if they "couldn't tell", but some as "unclear risk". It also assessed one trial as unclear risk despite reasoning this with "blinding of intervention: yes". Assessed one trial as "high risk for performance and detection bias combined, but "unclear risk" when assessed separately despite reasoning this with "no blinding of outcome".

The out of date review by Nehra *et al.*[129] stated in the characteristics of included studies that blinding of the intervention was present in one study, "can't tell" for two studies, present for clinical caretakers but not research team for 1 study and present for clinical caretakers but unknown for the research team for two studies. Blinding of outcome was stated to be "can't tell" for all included studies.

Incomplete outcome assessment

In fi reviews, the majority of studies were assessed as low risk for incomplete outcome assessment, with most with >90% follow up at least for primary outcomes and some studies with 80% follow up for some outcomes [127,128,145,147,150]. Two reviews assessed their only included studies as low risk for bias from incomplete outcome assessment with the 3% excluded only due to death before the intervention [151] or <5% loss to follow up [59].

One review assessed the majority of outcomes as low risk however split this outcome between all outcomes and length of stay[152]. For all outcomes 5/6 studies were assessed as low risk with complete follow up, or where missing information could be obtained from study authors and included in intention to treat analysis.

Length of stay was only reported in 4 of the 6 studies in Nasuf *et al.*[152]. Of these two studies were assessed as low risk with complete follow up or where information

could be obtained from study authors and included in intention to treat analyses. One study was assessed as high risk as it did not include the data for length of stay for three infants who remained inpatients at the end of the study period.

The studies assessed as low risk in Brown *et al.*[143] had >80% or complete follow up reported or assumed, all infants included in intention-to-treat analyses. A third of studies were assessed as low risk in Amissah *et al.* with >80% follow up post discharge for most outcomes. Only 30% of studies were assessed as low risk in Fenton[149] supported with "outcomes objective" and low loss to follow up of short and long term outcomes.

The studies assessed as high risk of bias for incomplete outcome assessment were judged high risk due to high loss to follow up Walsh 2019[127]; high post-randomisation exclusions[143]; exclusion due to receiving >10% human milk[150]; 19% exclusion without details of group allocation Walsh *et al.*[127]. In Nasuf *et al.*[152], one study was assessed as high risk as a high proportion of infants were excluded and not analysed with information only determined through correspondence with the study author.

Fenton *et al.* [149] assessed more than half of studies as high risk of bias for incomplete outcome assessment due to >50% of infants withdrawing from the study. Two studies were assessed as high risk despite <20% loss to follow up, with one study with 12% infant withdrawn due to NEC/human milk provision; and another study with 3% withdrawn in first 3 days due to respiratory problems, 2% further withdrawn due to metabolic acidosis and nitrogen retention.

Unclear risk was assessed in the only included study in Amissah *et al.*[146]and Amissah *et al.*[144]. These were assessed due to transfer of infants[146] or <20% missing data[144] but without clear comparison of baseline differences or if an intention to treat approach was used. In Amissah *et al.*[148], one study was assessed as unclear risk despite only 22% of infants who met criteria completing the study with no details on attrition. Other were assessed as unclear risk due to no clarity if difference between excluded infants or no details. Other reviews assessed studies as unclear risk due to only the abstract with insufficient details[147] or a large group imbalance of group allocation in 10% infants with adverse outcomes not assessed for growth (23% formula vs 2.4% donor)[145]. Fenton *et al.* [149] assessed one study as unclear risk despite only 1.6% of

infants withdrawn from studies. Reasons for withdrawal included death and diarrhoea.

The two out of date reviews by Tan-Dy *et al.*[130] and Nehra *et al.* [129] did not assess risk of bias, but judged completeness of follow up. Complete follow up was present in all included studies in Nehra *et al.* [129]. Tan-Dy *et al.*[130] responded "yes" to completeness of follow up despite <80% reaching study day 14.

Reporting bias

All studies were assessed as low risk in four reviews due to all outcomes reported [59,128,146,151], "no concerns"[149] or no deviation from the protocol[150]. Studies were assessed as low risk if there was no deviation from the protocol [145,148,152]. Amissah *et al.* [146], and Amissah *et al.*[148] assessed reporting bias as low risk if all outcomes were listed in the methods, even if no protocol was available, while other studies assessed studies as unclear risk if there was no protocol available [127,144,145,150,152], although one study in Walsh *et al.*[127] was assessed as low risk when all outcomes reported in the methods were reported in the results.

All studies in Ng *et al.*[150], the majority of studies in three reviews [127,145,147] and two studies in Nasuf *et al.*[152] were assessed as unclear risk due to no protocol available. An additional study in Nasuf *et al.*[152] was assessed as unclear risk as the study was unpublished and secondary outcomes had not been stated. One study in Amissah *et al.*[148] was assessed as unclear risk when reporting an additional non pre-specified outcome. The single study in Amissah *et al.*[144] was assessed as unclear risk when there was no protocol and no details which outcomes were primary.

Five reviews each had one study assessed as high risk [127,145,147,148,152]. These were due to long term growth and neurodevelopment outcomes specified in protocol not reported[145], growth outcomes only reported as "not statistically significant" and not reported numerically[127] some outcomes only narratively reported in the abstract[147] , head circumference not included in analysis when no relationship seen in data during study[148]. In Nasuf *et al.*[152] one study was assessed as high risk as outcomes were reported as per study protocol, but no explanation why only half of sample size was included.

Two reviews did not assess this outcome [129,130].

Other bias

Two reviews assessed one study as high risk [127,146]. In Amissah *et al.*[146] this was assessed when two reports of same trial had discrepant methods and numbers of included infants, and were unable to reproduce analyses from published data. One review assessed this when a trial was undertaken explicitly to evaluate the specific formula clinically with the manufacturer giving financial assistance[127].

Four reviews assessed other bias as low risk for all included studies [59,128,149,151], with a third of studies assessed as low risk in Amissah *et al.*[148].

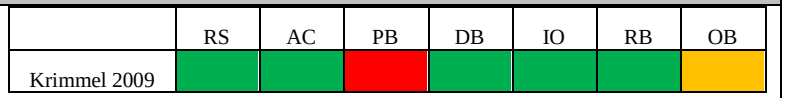
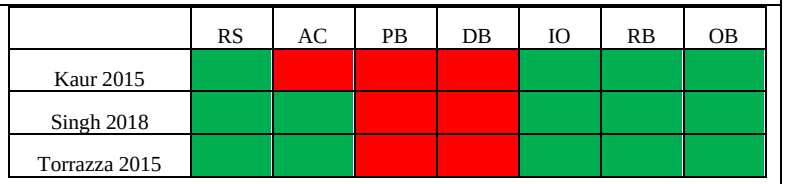
Unclear risk was assessed in all studies in Quigley *et al.*[145] and Ng *et al.*[150], and the majority of studies in Walsh *et al.*[127]. 7 of the studies in Quigley *et al.*[145] were at least part funded in a pharmaceutical company. 7 studies in Ng *et al.*[150] and 5 studies in Walsh *et al.*[127] were funded by milk formula companies. At least 5 studies in Ng *et al.*[150] and 4 in studies Walsh *et al.*[127] were funded by the manufacturer of the trial formula. One study in Walsh *et al.*[127] was unclear about whether the pharmaceutical company providing funding was the trial formula manufacturer, and in one trial the employee of the pharmaceutical company helped in statistical analysis. 3 studies in each Quigley *et al.*[145] and Ng *et al.*[150], and one study in Walsh *et al.*[127] review did not state their funding.

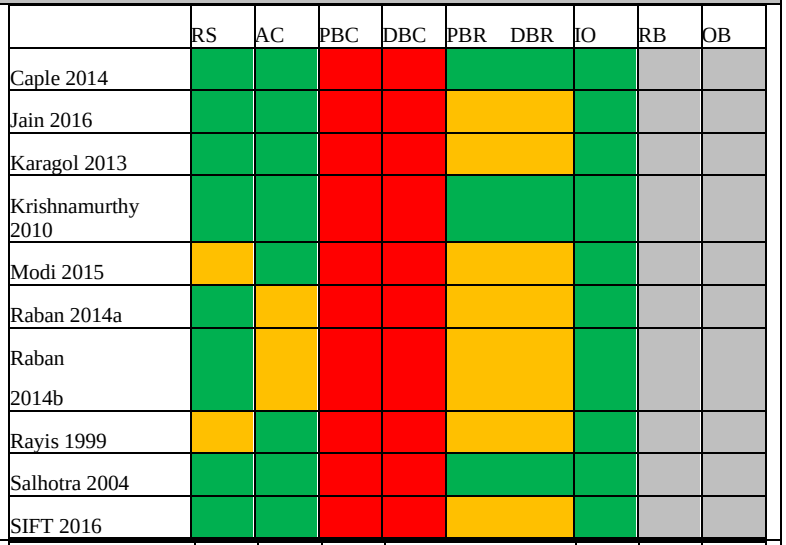
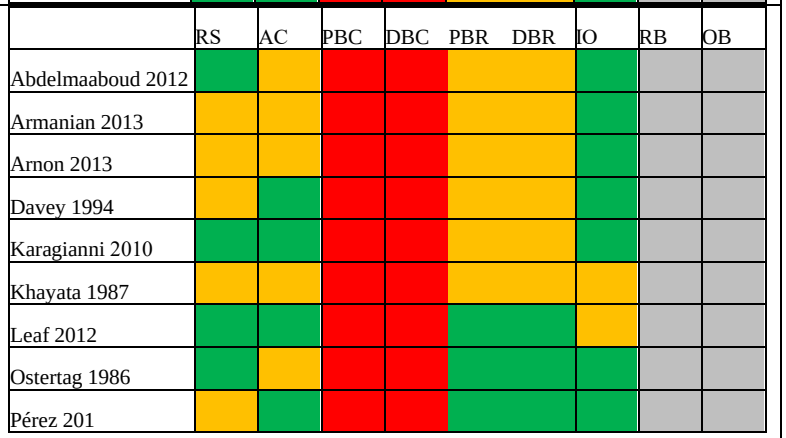
Four trials in in Amissah *et al.*[148] and one trial in Amissah *et al.*[144] were assessed as unclear risk, due to no baseline demographic details to make assessment, of which one in each review stated that there was a difference sex distribution for which analyses confirmed “no implication on results”.

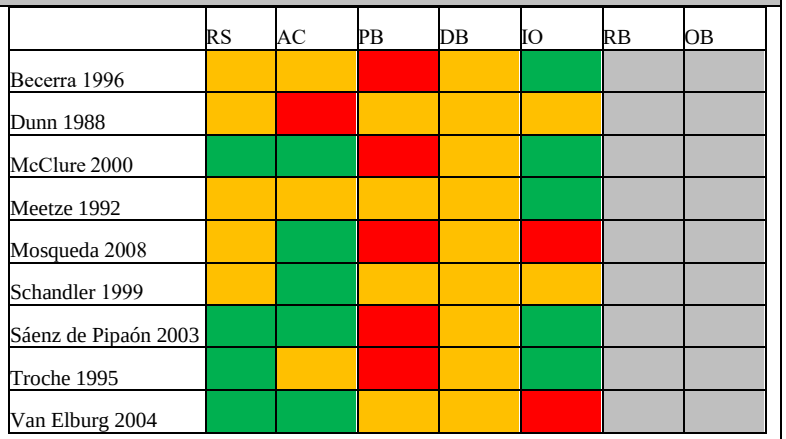
Basuki *et al.*[147] only assessed this outcome for one of the three studies, stating it to be unclear to assess due to being an abstract only. The outcome was not formally assessed in Brown *et al.*[143], but narrative description states that three report authors were employees of fortifier manufacturer, and three trials were funded by the fortifier manufacturer. It was not assessed in three reviews [129,130,152].

15.2.2 Appendix 2b: When to feed

Table 18:Appendix 2b: Review author assessment of quality of the included studies (When to feed)

Study	Assessment		RoB table
Yeo *et al.* 2019[154]: "Stopping enteral feeds for prevention of transfusion associated NEC"	Sequence generation	Low risk	
	Allocation concealment	Low risk	
	Performance bias	High risk	
	Detection bias	Low risk	
	Incomplete outcome	Low risk	
	Reporting bias	Low risk	
	Other bias	Unclear risk	
Abiramalatha *et al.* 2019*a*[155] "Routine monitoring of gastric residuals"	Sequence generation	Low risk	
	Allocation concealment	2 low risk, 1 high risk	
	Performance bias	High risk	
	Detection bias	High risk	
	Incomplete outcome	Low risk	
	Reporting bias	Low risk	
	Other bias	Low risk	

Study	Assessment		RoB table
Oddie *et al.* 2017[136] "Slow versus faster enteral feed advancement to prevent NEC"	Sequence generation	8 low risk, 2 unclear risk	
	Allocation concealment	8 low risk, 2 unclear risk	
	Performance bias	10 high risk	
	Detection bias	10 high risk (all high risk for clinical outcomes; 3 low risk and 7 unclear risk for radiological outcomes)	
	Incomplete outcome	10 low risk	
	Reporting bias	Not assessed	
	Other bias	Not assessed	
Morgan *et al.* 2014[137] "Early versus delayed introduction of progressive enteral feeds"	Overall	4 low risk, 5 unclear risk	
	Allocation concealment	4 low risk, 5 unclear risk	
	Performance bias	High risk	
	Detection bias	3 low risk, 6 unclear risk	
	Incomplete outcome	7 low risk, 2 unclear risk	
	Reporting bias	Not assessed	
	Other bias	Not assessed	

Study	Assessment		RoB table
Morgan *et al.* 2013[131] "Early trophic feeding versus enteral fasting"	Sequence generation	4 low risk, 5 unclear risk	
	Allocation concealment	5 low risk, 3 unclear risk 1 high risk	
	Performance bias	4 unclear risk, 5 high risk	
	Detection bias	9 unclear risk	
	Incomplete outcome	5 low risk, 2 unclear risk, 2 high risk	
	Reporting bias	Not assessed	
	Other bias	Not assessed	

Randomisation sequence generation (RB) and Allocation concealment (AC)

All reviews included in this section assessed the included studies as low risk or unclear risk of bias from randomised sequence generation and assessed bias from allocation concealment mostly as low or unclear risk of bias.

One review assessed all included studies as low risk of bias from random sequence generation and allocation concealment[154].

Abiramalatha *et al.*[155] assessed one study as high risk as although computer-generated block randomisation was performed and the sequence was kept in sequentially numbered sealed opaque envelopes, the authors felt that the fixed block size of 4 may allow allocation prediction in the study unblinded to participants, personnel and outcome assessors. The other studies included in this review used variable block sizes in the computer-generated randomisation.

Morgan *et al.*[131] assessed one of their nine studies as high risk for allocation concealment as it was unclear if the envelopes were sealed, it described random sequence generation as unclear risk as groups were stratified by birthweight and randomised by cards in paired envelopes.

The reviews assessed their studies as a combination of low risk and unclear risk. Trials were assessed as low risk for random sequence generation when this was done by computer generated randomisation[131,136,137], random number table [131,137], random number sequence and sealed opaque envelopes [131,136], selection of cards from sealed envelopes[131].

Studies were assessed as unclear risk for randomisation if the method was not stated [131,136,137], they had "stratified block randomisation" with no further details[136] or randomisation using cards in paired envelopes but unclear if sealed Morgan *et al.*[137].

Allocation concealment was assessed as unclear risk when allocation concealment was not described [131,136,137], and low risk when allocation concealment performed using sealed opaque envelopes[131,136,137], a blinded draw from envelopes by caregivers not involved in the study[136], computer based random allocation[136], central telephone randomisation was used[137] or investigators stated to be blinded at the time of randomisation[131]. Morgan *et al.* [137] assessed one trial as low risk despite stating that allocation concealment was not described.

Performance and detection bias

The majority of studies in the included reviews were assessed as high risk for performance bias due to lack of blinding of clinical staff and caregivers [131,136,137,154,155], with only Morgan *et al.*[131] assessing some studies as unclear risk due to lack of reporting.

Oddie *et al.*[136] and Morgan *et al.*[137] assessed performance bias and detection bias together, but assessed separate outcomes for clinical and radiological outcomes, therefore where detection bias was assessed as unclear or low risk, performance bias for the radiological outcome was assessed at the same risk.

Yeo *et al.*[154]: assessed detection bias of the only included study as low risk, as investigators remained masked to the feeding assignment of the infant. The performance bias for this, however, was at high risk.

Abiramalatha *et al.*[59] assessed detection bias for included studies as high risk due to lack of blinding, similarly to all assessments for publication and detection bias of clinical outcomes in Oddie *et al.*[136] and Morgan *et al.*[137]. These were assessed as high risk when caregivers and clinical caregivers not blinded once allocation had been performed, or unclear if investigators were blinded after allocation stage. For performance and detection bias for radiological outcomes studies were assessed as low risk when radiologists were blinded to the intervention group[136] or an independent review of NEC were performed[137]. They were assessed as unclear risk when masking of the radiological assessors was not specifically stated[136].

Morgan *et al.*[131] all as unclear risk of detection bias of outcome as this was either not reported or no reference whether radiograph interpretations was blind even when laboratory staff blind.

Incomplete outcome assessment

All studies in Oddie *et al.*[136] and the majority of studies in Morgan *et al.*[137] and Morgan *et al.*[131] were assessed as low risk of bias from incomplete outcome assessments as there was near-complete or complete follow up for primary outcomes, there was <20% loss to follow up, exclusions were explained, or protocol violations occurred but data was included in the final analysis.

Morgan *et al.*[137] assessed 2 studies as unclear risk as they were either not described or there was post-randomisation exclusion due to recruitment error or consent withdrawal. Morgan *et al.*[131] assessed 2 studies as unclear risk as although infant withdrawal explained and included in intention to treat analysis, there is uncertainty whether these infants developed necrotising enterocolitis and one only stated "intention-to-treat analysis". Morgan *et al.* [131] assessed 2 studies as high risk as these studies had >20% loss to follow up with one study unbalanced in numbers lost to follow up.

Reporting bias

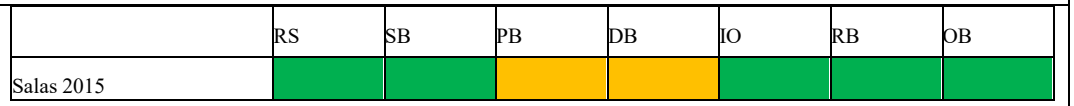
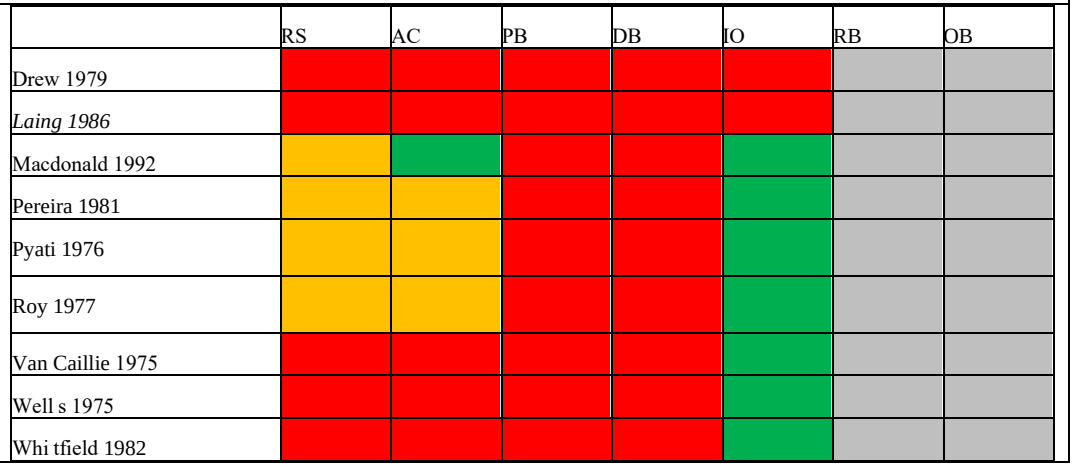
Only Yeo *et al.*[154]: and Abiramalatha *et al.*[155] assessed this outcome. All studies were assessed as low risk as they found that all prespecified outcome measures were reported, with Abiramalatha *et al.*[155] additionally commenting on that a protocol had been published.

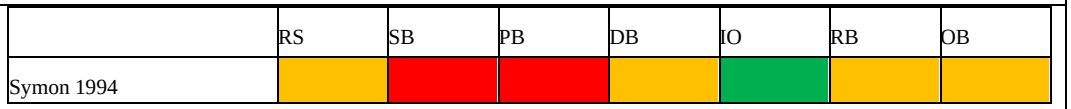
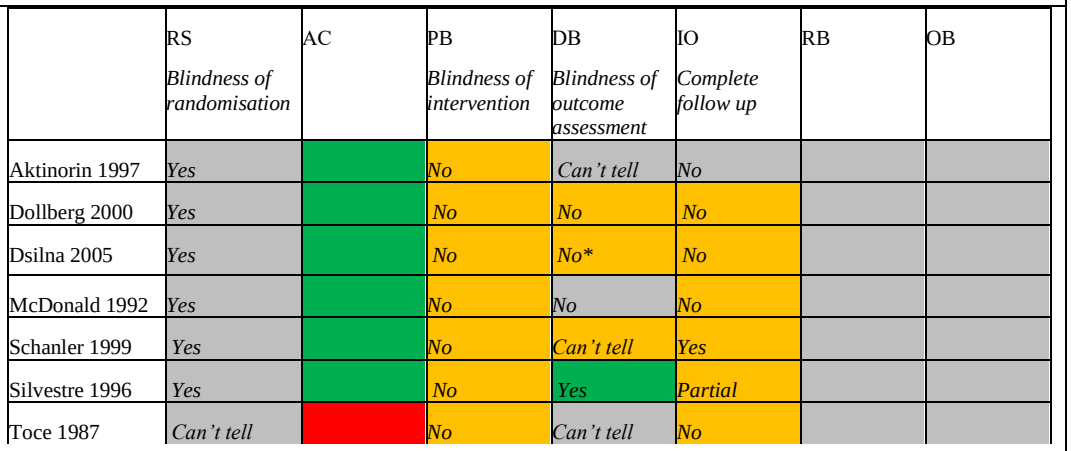
Other bias

Only Yeo *et al.*[154] and Abiramalatha *et al.*[155] assessed this outcome. The single study in Yeo *et al.*[154] was assessed as unclear risk due to inadequate baseline data available and lack of clarity whether infants required more than one transfusion and if so, whether allocated to same intervention for subsequent transfusion episodes. Abiramalatha *et al.*[155] assessed all included studies as low risk with no other bias detected.

15.2.3 Appendix 2c: How to feed

Table 19: Appendix 2c: Review author assessment of quality of the included studies (How to feed)

Study	**Assessment**		**RoB table**
Abiramalatha *et al.*2019[156] "Re-feeding versus discarding gastric residuals"	Sequence generation	Low risk	
	Allocation concealment	Low risk	
	Performance bias	High risk	
	Detection bias	High risk	
	Incomplete outcome	Low risk	
	Reporting bias	Low risk	
	Other bias	Low risk	
Watson *et al.* 2013[135] "Transpyloric versus gastric tube feeding"	Sequence generation	5 high risk,4 unclear risk	
	Allocation concealment	1 low risk, 3 unclear risk, 5 high risk	
	Performance bias	Assessed together	
	Detection bias	All high risk	
	Incomplete outcome	2 high risk, 7 low risk	
	Reporting bias	Not assessed	
	Other bias	Not assessed	

Study	**Assessment**		**RoB table**
Dawson *et al.*2012[132] "Push versus gravity for intermittent bolus gavage tube feeding"	Sequence generation	Unclear risk	
	Allocation concealment	High risk	
	Performance bias	Unclear risk	
	Detection bias	High risk	
	Incomplete outcome	Unclear risk	
	Reporting bias	Low risk	
	Other bias	Unclear risk	
Premji *et al.* 2011[133] "Continuous nasogastric milk feeding versus intermittent bolus milk feeding"	Sequence generation	6 "Yes", 1 "Can't tell"	
	Allocation concealment	6 low risk, 1 high risk	
	Performance bias	7 "No"	
	Detection bias	1 "Yes", 3 "Can't tell", 3 "No"	
	Incomplete outcome	1 "Yes", 1 "Partial", 5 "No"	
	Reporting bias	Not assessed	
	Other bias	Not assessed	

*Radiographic assessors for NEC outcome blinded to group assignment

Randomisation sequence generation (RB) and Allocation concealment (AC)

Abiramalatha *et al.*[156] assessed the only included study as low risk for both randomisation sequence generation and selection bias, with use of simple randomisation procedures and sequentially numbered opaque, sealed envelopes. All but one study in Premji *et al.*[133] was assessed as low risk for allocation concealment, described as using random number tables, opaque sealed envelopes, some stratifying for birthweight, gestational age or diet and one using an uneven blocked designed. Of not is that this review also assessed one study which only stated "randomly assigned" as low risk.

One study in Premji *et al.*[133] was assessed as high risk due to using an alternate assignment method. This was also used in the studies assessed as high risk of allocation concealment bias in Watson *et al.*[135], with one using alternate monthly allocation. The remaining studies were assessed as unclear risk due to insufficient information regarding randomisation and allocation, with only one reported as low risk of allocation concealment only due to using sealed envelopes using a random sequence.

Dawson *et al.*[132] assessed the only included study as unclear risk for randomisation as it had only stated "allocated randomly", and high risk for selection bias as no information was provided.

Performance and detection bias

All studies in the reviews by Abiramalatha *et al.* [156] and Watson *et al.*[135] were assessed this as high risk due to lack of blinded, it was said to be unfeasible. Performance and detection bias were assessed as a single outcome in Watson *et al.*[135].

Premji *et al.*[133] assessed studies where it was "not feasible" for caregivers and investigators to be blinded as unclear risk. It assessed one study as low risk for outcome assessment as although investigators were not masked, the outcome assessors were blinded. In a further trial outcome assessment was assessed as unclear risk as only the radiologists assessing NEC were blinded. One risk assessed as unclear but used a designed serial assessment method to ensure objective assessment

of the major outcome variable. Three studies did not have detection bias within the risk of bias assessment, but two reported "can't tell" and one "no".

Dawson *et al.*[132] assessed the only included study as high risk of performance bias as all feeds were given by one unmasked researcher and unclear risk of detection bias as no information was provided.

Incomplete outcome assessment (attrition bias)

Abiramalatha *et al.*[156] and Dawson *et al.*[132] assessed their only included studies as low risk. Abiramalatha *et al.*[155] reported complete follow up for mortality and necrotising enterocolitis, while the crossover trial by Dawson *et al.*[132] only assessed short term outcomes.

Watson *et al.*[135] assessed the majority of trials as low risk for attrition bias with near-complete follow up. In one of these studies, however short term weight gain data was only presented in infants below 1.4kg. Two studies were assessed as high risk. One study had a third of participants withdrawn after allocation with explanations including failure to pass the feeding tube, assisted ventilation and "insufficient data to compute". The second study was assessed as high risk as growth data was only reported in the 71% (nasogastric)-35% (nasoduodenal) of infants who tolerated allocated feeding route.

Premji *et al.*[133] only assessed this outcome in six of the seven included studies, which were mostly assessed as unclear risk. Follow up was variable in these studies from a study with complete follow up yet unbalanced infants removed from feeding protocol due to being unable to adhere, only 3% not analysed but unclear if analysed as intention to treat, 12% of infants removed from treatment protocols and excluded from overall analysis however receiving complete follow up for analysis within stratified groups, 14% exclusion due to hospital transfer during or death prior to intervention, and 32% exclusion from analysis. One study in this review was assessed high risk due to 36% exclusion from analysis of an unblinded quasi-experimental study.

Reporting bias

This was assessed as low risk in Abiramalatha *et al.*[156] as all outcomes in the protocol were reported, and unclear risk in Dawson[132] as they were unable to obtain the study protocol. This was not assessed in the remaining reviews [133,135].

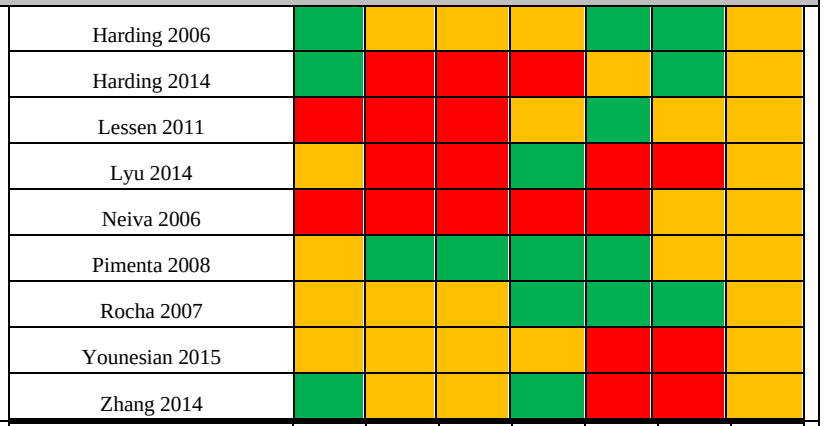
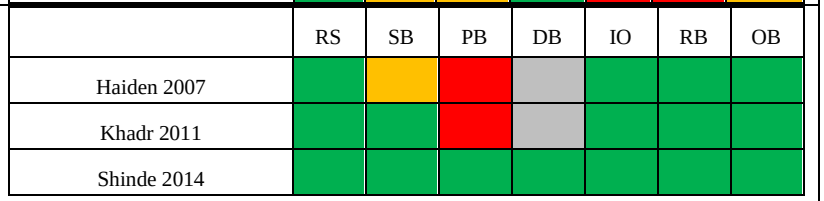
Other bias

Abiramalatha *et al.*[156] assessed the this as low risk with no further concerns of bias, while Dawson *et al.*[132] assessed this as unclear risk due to no information about the period between the crossover trial's changeover, therefore potentially being affected by carryover effect. This outcome was not assessed in the remaining reviews [133,135].

15.2.4 Appendix 2d: Adjunctive studies

Table 20: Appendix 2d: Review author assessment of quality of the included studies (Adjunctive strategies)

Study	Assessment		RoB table
Muelbert *et al.* 2019[157] "Exposure to smell and taste stimulation of milk with tube feeds versus no exposure"	Sequence generation	1 low risk, 1 unclear risk, 1 high risk	
	Allocation concealment	1 low risk, 1 unclear risk, 1 high risk	
	Performance bias	1 trial high risk, 1 unclear risk, 1 low risk.	
	Detection bias	Unclear risk	
	Incomplete outcome	2 unclear risk, 1 low risk	
	Reporting bias	2 low risk, 1 unclear risk	
	Other bias	2 low risk, 1 unclear risk	
Greene *et al.* 2016[159] Oral stimulation for promoting oral feeding	Sequence generation	7 low risk 7 unclear risk, 2 high risk	
	Allocation concealment	1 low risk, 9 unclear risk, 6 high risk	
	Performance bias	3 low risk, 5 unclear risk, 8 high risk	
	Detection bias	6 low risk, 4 unclear risk, 6 high risk	
	Incomplete outcome	8 low risk, 1 unclear risk, 7 high risk	
	Reporting bias	7 low risk, 3 unclear risk, 6 high risk	

Study	Assessment		RoB table
Greene *et al.* (2016)[159] (continued)	Other bias	13 unclear risk, 3 high risk	
Anabrees 2015[158] Glycerin prophylaxis versus placebo/no intervention	Sequence generation	All low risk	
	Allocation concealment	1 unclear risk. 2 low risk	
	Performance bias	2 high risk, 1 low risk	
	Detection bias	2 not assessed, 1 low risk	
	Incomplete outcome	3 low risk	
	Reporting bias	3 low risk	
	Other bias	3 low risk	

Study	Assessment		RoB table
Ng *2008* [134] Erythromycin vs placebo for prevention of feeding intolerance	Sequence generation	Not assessed	
	Allocation concealment	9 had blinding of randomisation, 1 "Can't tell"	
	Performance bias	8 had blinding of intervention, 1 "Can't tell", 1 "No"	
	Detection bias	4 had blinding of "outcome measures", 6 "Can't tell"	
	Incomplete outcome	All "complete follow up"	
	Reporting bias	Not assessed	
	Other bias	Not assessed	

	RS "Blinding of randomisation"	SB "Allocation concealment"	PB "Blinding of intervention"	DB "Blinding of outcome measures"	IO "Complete follow up	RB	OB
Aly 2007	*Yes*		*Yes*	*Yes*	*Yes*		
Cairns 2002	*Yes*		*Yes*	*Can't tell*	*Yes*		
El Hanawy 2003	*Yes*		*Yes*	*Yes*	*Yes*		
Madani 2007	*Can't tell*		*Can't tell*	*Can't tell*	*Yes*		
Ng PC 2001	*Yes*		*Yes*	*Can't tell*	*Yes*		
Ng PC 2003	*Yes*		*Yes*	*Can't tell*	*Yes*		
Nuntnaran 2006	*Yes*		*Yes*	*Yes*	*Yes*		
Oei 2001	*Yes*		*Yes*	*Can't tell*	*Yes*		
Patole 2000	*Yes*		*Yes*	*Yes*	*Yes*		
Stenson 1996	*Yes*		*No*	*Can't tell*	*Yes*		

Randomisation sequence generation (RB) and Allocation concealment (AC)

Muelbert *et al.*[157] assessed random sequence generation as low risk when a computer generated random number table was used, unclear risk when not stated and high risk due to sequential allocation based on date of admission. They assessed studies at low risk when sequentially numbered opaque sealed envelopes were used, unclear risk when not stated, and high risk when no allocation concealment was performed.

Anabrees *et al.*[158] assessed all studies as low risk for random sequence generation with only one described method in characteristics of studies (random block assignment). It only assessed 2 of the three studies for allocation concealment, with one unclear risk and one low risk. No explanations were given.

Ng *et al.*[134] only assessed allocation concealment within risk of bias, which it assessed as low risk in all but one study. Randomisation sequence generation was described narratively as present in all but one study, with studies using block randomisation, sealed envelopes, or computer-generated random numbers. One study was assessed as low risk despite only stating "randomisation was stratified by study centre and postmenstrual age" with no additional methodological information, while one was assessed as unclear risk as no information regarding method of randomisation. Explanation of allocation concealment was inconsistently described narratively, with the method of some studies describing concealed randomisation codes known only to pharmacists off-site, and sealed envelopes.

Greene *et al.*[159] assessed seven studies as low risk, seven studies as unclear risk and two studies at high risk of bias from sequence generation. It assessed studies at high risk when infants randomly assigned in blocks of 2 or described as "distributed in a random manner" but where there was no attempt to conceal group allocation. It assessed studies as unclear risk when a "simple randomisation method, "convenience sampling" or "randomly assigned" was used but not clearly described. One study was assessed as unclear risk when infants assumed to be randomised as per a previous study by same author, but not explicitly stated. Another study was assessed as unclear risk when selection bias with random number generator in excel was suspected. It assessed studies as low risk when block randomisation, computer generated randomisation was used or computer generated matched paired design was used.

Allocation concealment was assessed by Greene *et al.*[159] as unclear risk in nine, high risk in six and low risk in one study. It assessed studies as high risk when the study was described as unblinded, allocation would be seen by the researcher opening the envelope with order of allocation sequence, speech and language therapist delivering all interventions or assignment in blocks of two (with subsequent participant enrolled to the group where the previous infant had been withdrawn from, or all interventions and assessments are carried out by the researcher. It assessed studies as unclear risk when there was insufficient information about the method of allocation, or in one study where infants are assumed to be the same infants randomised in a previous study by same author, but this is not explicitly stated. Low risk assessed studies described computer generated random numbers, sequential and/or sealed opaque envelopes.

Performance and detection bias

The majority of studies in Muelbert *et al.*.[157], Ng *et al.*[134] and Greene *et al.*[159] were assessed as either high risk or unclear risk of bias.

Studies were assessed as low risk of performance bias when study subjects and neonatologists were said to be blinded despite the method not being reported [157], caregivers were masked to intervention using a sham procedure (Anabrees 2015, Greene 2016), and the speech therapist assessing suck capacity blinded and independent from the speech therapist delivering the intervention[159].

Performance bias was assessed as unclear risk when not stated[157], when unclear whether parents, nurses or physician performing the intervention were aware of group allocation[159] or breaking of blinding possible due to knowledge of the intervention methods or when seen to be delivering the intervention to other infants in the unit not involved in the study[159].

Performance bias was assessed as high risk when participants and personel were not blinded [157,158], investigators are not blinded[159], when an unblinded researcher administered interventions and assessed outcomes[159], if blinding is disrupted by one speech and language therapist delivering all interventions [159], when protocols when protocols posted on the isolettes making caregivers and therapists aware of group allocation [159], when presence of blinding of caregivers and medical staff is not reported[159].

Detection bias was assessed as a combination of high risk, unclear risk and low risk. It was only assessed for one of three studies in Anabrees *et al.*[158] as it was felt not to be applicable, it assessed this one study as low risk with no further explanation. Muelbert *et al.*[157] assessed all studies as unclear risk of detection bias, with outcome assessors not blinded or unlikely to be blinded but unlikely to influence outcomes. Greene *et al.*[159] with the outcome assessor as one of the researchers.

Greene *et al.*[159] assessed an equal number of included studies as low risk and high risk, with four studies assessed as unclear risk. Studies were assessed as low risk when staff conducting treatments or aware of allocation were not involved with outcome assessment or making decisions on ability to initiate oral feeding, when a single nurse masked to allocation recorded all duration and volume of feed outcomes, or when feeding decisions were performed by physician who is unclear if they are blinded but feeding variables monitored by researcher blinded to group allocation. The review assessed detection bias as unclear when there was uncertainty who performed measurements or decided on progression of oral feeds, and uncertainty whether outcome assessors were masked. Studies were assessed as high risk when the investigators or outcome assessors were aware of group assignment, especially when the unblinded researcher administered interventions and assessed outcomes. Some assessed as high risk were not blinded "due to its nature". One had the protocols posted on the isolettes making caregivers and therapists aware of group allocation, with the outcome assessor as one of the researchers.

Ng *et al.*[134] assessed if blinding of intervention was present using a yes/can't tell/no scale, with blinding of intervention mostly present, except "can't tell" in one study where no methodological information was available and "no" in one study. Decisions were inconsistently explained describing in some studies that the placebo was made to look identical to the intervention drug and not present when a placebo was not given. Blinding of outcome measures present in four studies and "can't tell" in 6 studies. Only one of the studies marked that outcome assessment was present described who was blinded, with three studies only stating that it was achieved. In studies assessed as "can't tell", it was stated to not be clear or it was stated that no other methodological information was available, or in one study the treatment drug had a distinct odour and therefore masking was incomplete.

Incomplete outcome assessment

All studies in Anabrees *et al.*[158] were low risk for attrition bias, and Ng *et al.*[134] states that complete follow up is present in all studies. One study in Muelbert *et al.*[157] and half the studies in Greene *et al.*[159] were assessed as low risk of attrition bias. These were assessed as low risk when all data for outcomes were reported, or due to using intention to treat analysis[157,159] even if all reasons for withdrawal were not provided[159]. One study was assessed as low risk when reasons provided despite 11/30 infants excluded post randomisation[159].

Muelbert *et al.*[157] assessed two studies as unclear risk due to limited information on excluded participants or data reported. Greene *et al.*[159] assessed one study as unclear risk when no intention to treat analysis completed but groups were equally balanced, acceptable reasons provided for missing data.

Greene *et al.*[159] assessed studies as high risk when baseline characteristics were only narratively reported as similar without data provided to confirm, data was missing or outcomes were not reported. In one included study summary statistics were provided but no information on number of infants they described. In another, there was lack of data to confirm narrative statement. Greene *et al.*[159] also assessed studies as high risk when parental visits not reported. A predefined outcome found not to be reported in several studies was that of behavioural state at the start and/or end of feeding.

Reporting bias

Only three of the reviews assessed reporting bias. The review by Anabrees *et al.*[158] assessed all studies as low risk, but no justification given. Muelbert *et al.*[157] described that two studies as low risk all outcomes reported, and one as unclear risk when protocol was not available.

Greene *et al.*[159] assessed seven studies as low risk when all prespecified and expected outcomes of interest were reported, and three at unclear risk when individual data for outcomes not available and abbreviations in tables not explained. Five studies were assessed as high risk. These studies provided no data to confirm baseline characteristics that were narratively reported as similar, did not report several co-variates prespecified in the protocol, did not report all prespecified outcomes or had

lack of data to confirm narrative statement of an outcome. Behavioural state around feeding was the outcome not reported in several of these studies.

Other bias

Three reviews reported their assessment of other factors affecting bias. Muelbert *et al.*[157] assessed two studies as low risk as they had no significant baseline characteristic differences and no loss to follow up, and one study with limited information as unclear risk. Anabrees *et al.*[158] assessed all studies as low risk, but no explanation was provided.

Greene *et al.*[159] assessed thirteen studies as unclear risk due to no adverse events having been reported, the report was difficult to interpret or "unclear", or there was insufficient information to assess the outcome. Three studies were assessed as high risk. One study had lack of clarity about who made decisions on increasing feeding volume. Another study was suspected to have the same study participants as another included review, previously published, but this was not stated. Potential bias in the third study was due to parents of infants in the intervention group having been given instructions on how to hold infants during feeding, and as interventions were carried out by parents, researchers and nursing staff there is the potential for variability between interventions.

15.3 Appendix 3: Methodological quality of included reviews: AMSTAR2 Full description

We rated the quality of the included reviews using the AMSTAR and ROBIS tools

1) Did the research questions and inclusion criteria for the review include the components of PICO (Population, Intervention, Comparator Group, Outcome)?

   All eligible reviews in all four categories met all criteria including the optional timeframe for follow up.

2. Did the report of the review contain an explicit statement that the review methods were established prior to the conduct of the review and did the report justify any significant deviations from the protocol?

   Fifteen of seventeen reviews in "what to feed", three of five reviews in "when to feed", two reviews in "how to feed" and two of four reviews in "adjunctive strategies" met all criteria.

   For six of these reviews the original published protocol was not available, therefore the categories assessing protocol content were assessed as present if found in the methods section of the review and there was no statement of deviation from the protocol [126,143–146,148].

   Nine reviews were only assessed as a partial yes [128–132,134–136,158]. Two reviews had no plan to investigate causes of heterogeneity [129,134]. Two reviews differed from protocol as in trial design inclusion criteria [128,158]. Greene *et al.*[159] stated inclusion of cluster-randomised controlled trials in the methods while the protocol states that they will only consider parallel studies for the review[158]. Thanigainathan *et al.*[128] did not declare or explain why the protocol planned inclusion of quasi-randomised and cluster-randomised trials were omitted from the review methods[128]. No studies were listed to be included or excluded for this reason. Thanigainathan *et al.*[128] also did not state whether conference proceedings were searched as stated in the protocol[128]. Oddie *et al.*[136] did not declare of justify why the inclusion of selection bias assessment in the methods was not performed or reported in the

review[136]. Two review updates did not state and justify changes in methods from the previously published review [131,135]. Morgan *et al.*[131] had stated the update to the methods from the initial protocol in 1997, however did not state and justify the further changes since 2009[131].

Three reviews did not declare or justify significant deviations from the protocol by inclusion of additional outcomes [129,130,158], although Tan-Dy *et al.*[130] did declare one of two deviations [6].

Three reviews did not declare changes or removal of prespecified outcomes assessed. Watson *et al.*[135] removed neurodevelopment and time to establish full oral feeds as outcomes and time to establish full enteral tube feeds was made a primary outcome[135]. Two reviews did not declare the changed of the prespecified outcome of weight gain to other measures of time to regain birthweight [132] and weight at discharge home [158]. Dawson *et al.*[132] also did not declare that they combined the protocol outcomes of apnoea, bradycardia and oxygenation during gavage into one outcome "severe apnoea", did not assess four prespecified outcomes, and performed additional subgroup analyses in methods[132].

3. Did the review authors explain their selection of the study designs for inclusion in the review?

   None of the reviews in any of the four categories explained their selection of study designs for inclusion in the reviews, stating only their intention to include the trial designs of randomised controlled trials with some reviews also including quasi-randomised controlled trials, or cross-over trials. This, however, is standard Cochrane Neonatal protocol.

4. Did the review authors use a comprehensive literature search strategy?

   All but two reviews in all four categories were found to use a comprehensive literature search strategy, with no language restrictions or additional restrictions made. One study assessing “what to feed” did not meet criteria as it restricted the search to only include studies published in English, with no justification made[129]. One study assessing “adjunctive feeding strategies” met

criteria for a partial yes as it did not report that it searched the reference lists/bibliographies of included studies[159].

5. Did the review authors perform study selection in duplicate?

   Almost all reviews in the four categories ("what to feed", "when to feed", "how to feed", "adjunctive strategies") performed study selection in duplicate, with at least two reviewers had independently agreeing on selection of eligible studies and achieving consensus.

   Three reviews did not meet criteria as it was unclear if there was independent assessment as only one author is reported to have assessed all titles and abstracts [136,143], it is described that screening was performed in duplicate but only articles selected by the principle review author were assessed [152] or it was unclear how many authors were involved in the screening and selection [130]. Where screening was performed by two reviewers with decision for inclusion in the review made by consensus, but it was unclear if screening was performed independently we accepted this as meeting criteria [131,134,135,137].

6. Did the review authors perform data extraction in duplicate?

   All reviews were reported to perform data extraction in duplicate with at least two reviewers achieving consensus on which data to extract from included studies.

7. Did the review authors provide a list of excluded studies and justify the exclusions?

   All reviews provided a list of excluded studies and justified their exclusions. Two reviews did not exclude any studies [130,132].

8. Did the review authors describe the included studies in adequate detail?

   Almost all reviews described the included studies in adequate detail. Two reviews were downgraded to a partial yes as the study settings were not described[134,156].

9. Did the review authors use a satisfactory technique for assessing the risk of bias (RoB) in individual studies that were included in the review?

Most reviews with eligible studies performed a satisfactory assessment of the risk of bias in the individual studies.

Eight reviews did not report on the selection of the reported result among multiple measurements or analyses of a specified outcome [129,131,133,135–137,143,154]. One of these reviews stated the intention for selection assessment in the protocol but did not report an assessment in the review [136]. One review had inconsistent reporting with allocation concealment and blinding of participants and personel reported for all studies, but blinding of outcome assessment and method of randomisation only assessed for two thirds of studies [133]. One described areas of bias but only assessed allocation concealment and narratively stated if blinding was achieved[129].

10. Did the review authors report on the sources of funding for the studies included in the review?

Half of reviews in "adjunctive strategies", almost half of reviews in "what to feed", and all reviews in "when to feed" and "how to feed" did not report on funding of the included studies or absence of this information in the included studies.

11. If meta-analysis was performed did the review authors use appropriate methods for statistical combination of results?

Ten of seventeen reviews assessing "what to feed", 4/5 reviews assessing "when to feed", 2/5 reviews assessing "how to feed" and all reviews assessing adjunctive strategies had sufficient studies to perform meta-analysis. All reviews which performed meta-analysis used a weighted technique, but no adjustment was made for heterogeneity. We did not downgrade the overall rating for this category.

One review in "what to feed" had no heterogeneity to adjust for [129]. Six reviews planned to investigate heterogeneity but were unable to as studies did not vary [128,147] or there was insufficient information for the subgroup analysis [129,131,143,149]. These reviews therefore met criteria.

Two reviews were assessed as a partial yes due to inconsistency [137,152]. Nasuf *et al.*[152] investigated causes for heterogeneity present, excluding a single

study in two outcomes with 91% and 49% heterogeneity, but gave a narrative statement of analysis only [152]. One review only performed assessment of heterogeneity for primary outcomes, despite a secondary outcome with high heterogeneity [137]. One review did not meet criteria as it had no plan to investigate the heterogeneity found [134].

12. If meta-analysis was performed, did the review authors assess the potential impact of RoB in individual studies on the results of the meta-analysis or other evidence synthesis?

    Four of ten reviews in "what to feed", two of four reviews in "adjunctive strategies" and none of the reviews in "when to feed" and "how to feed" met this criteria.

    Six studies had planned the analysis but were unable to perform due to insufficient trials or all trials with a similar risk of bias [127,128,144,147,157,158]. One review achieved criteria as it performed sensitivity analysis excluding a high risk study which was also found to have infants with different baseline characteristics[135].

    Six studies did not meet criteria as they had planned analyses to investigate the potential impact of risk of bias on the summary estimates of effect, did not conduct this or provide a statement of why this was not performed [133,136,137,150,152,159]. One review included studies at a similar risk of bias, but did not state this as a reason for not conducting the planned sensitivity analysis [155].

    Two reviews did not plan a sensitivity analysis to assess risk of bias [131,134].

13. Did the review authors account for RoB in individual studies when interpreting/discussing the results of the review?

    In all four categories, all reviews with eligible studies met these criteria.

14. Did the review authors provide a satisfactory explanation for, and discussion of, any heterogeneity observed in the results of the review?

Thirteen reviews in "what to feed", three reviews in "when to feed", all four reviews in "how to feed", and three of four reviews in "adjunctive strategies" met criteria.

Eight reviews had no heterogeneity due to inclusion of a single study [59,130,132,144,146,151,154,156]. Six studies the heterogeneity was not felt to be significant with the highest I^2 being 50% or below [128,129,131,135,155,157].

In eight reviews the authors acknowledged the heterogeneity, attempted to investigate source of heterogeneity through subgroup analyses [127,133,143,145,148,149,157] or exclusion of a study [152] and discussed the impact of this on the results of the review, usually through the GRADE assessment in the summary of outcomes [127,143,148,149,157,184]. One review stated that they were unable to identify reasons for the heterogeneity due insufficient data[148].

Two reviews did not directly acknowledge heterogeneity, however described and discussed the impact of differences between studies on the results of the review [134,158]. In contrast to most reviews, Anabrees *et al.*[158] planned to only investigate heterogeneity over 75%[158], and although Ng *et al.*[134] provides a narrative description of differences between studies, Chi^2 is only presented for one outcome[134].

Three reviews did not meet the criteria as they did not discuss the high or substantial heterogeneity they observed on analysis [137,147,150]. One review did not meet criteria as although heterogeneity was stated in GRADE, there was no investigation or discussion of potential sources of this[159].

15. If they performed quantitative synthesis did the review authors carry out an adequate investigation of publication bias (small study bias) and discuss its likely impact on the results of the review?

Sixteen reviews in "what to feed", three reviews in "when to feed", three studies in "how to feed" and one study in "adjunctive strategies" were assessed as having carried out an adequate investigation of publication bias.

All except three of these reviews discussed that they had planned to perform a funnel plot to assess publication bias, but had insufficient studies to

perform. The threshold to perform a funnel plot was 10 studies in the majority of studies.

Three reviews investigated publication bias using a funnel plot [135,136,143]. One review performed analysis for publication bias by funnel plot, but for growth outcome data only[143]. One conducted a funnel plot analysis for outcomes with included studies above a threshold of five trials[135].

Seven reviews did not meet criteria. Six of these reviews neither planned nor assessed for publication bias [129,131,133,134,137,158]. One only mentioned publication bias as part of the GRADE assessment and risk of bias, but neither planned nor reported any investigation[159].

16. Did the review authors report any potential sources of conflict of interest, including any funding they received for conducting the review?

Twelve of seventeen studies in "what to feed", four of five reviews in "when to feed", and all reviews in "how to feed" and "adjunctive strategies" met criteria.

Seventeen studies reported no potential sources of conflict of interest [128,131,148,152,154–156,158,159,132–135,137,144,146,147]. A further four authors described funding sources and how they managed potential conflicts of interest [126,142,151,157]. One reported that three study authors were involved in an ongoing trial investigating this topic (where two authors involved in design and are on the steering committee and one author is part of the research team) and one author is a co-author of an included trial[157]. The review states that the author who is a co-author of an included trial had no role in the assessment, data extraction, or analysis of data from this trial within the review.

Six reviews described how they managed the conflict of interest of core editorial and administrative support for this review being provided by a grant, by having an editor outside of the Cochrane Neonatal core editorial team receiving any financial remuneration from the grant as the Sign-off Editor for this review and by a senior editor from the Cochrane Children and Families Network assessing and signing off on the Cochrane Review [126,127,142,150,151,184].

Three of these reviews did not meet the criteria of this assessment point as there was a further unresolved conflict of interest [127,145,150].

Four reviews did not meet criteria as they declared that one study author had a conflict of interest as they had received multiple previous episodes of funding from formula manufacturers for research, however it was not explained how conflict of interest was managed [127,143,145,150].

Two reviews did not meet criteria due to author involvement as investigators in included trials either as an author was a principle investigator in the only included study[59] or authors were investigators of a very large included study[136]. Both of these reviews do not discuss how this conflict of interest was managed.

15.4 Appendix 4: Methodological quality of included reviews by ROBIS (Full description)

Phase one

The first phase of ROBIS assessed the relevance of the question being addressed in the assessed review to the overview target question.

All reviews were assessed as matching the overview target question. The intervention, comparator and outcomes matched our predefined criteria.

Two reviews fully met the review population criteria of very preterm infants and/or very low birthweight infants [142,144].

The other reviews had a target population or assessed population which had a larger scope than that targeted by our overview. This was either by including infants with a higher birthweight or a more advanced gestation up to 37 weeks' and a higher birthweight of up to 1750-2500g. Some reviews reported studies where either the gestational age at birth or birthweights were mostly unspecified.

Necrotising enterocolitis was assessed only as a secondary outcome in twelve reviews [59,127,148,149,129,130,132,135,142–144,146].

Growth was only assessed as a secondary outcome in fifteen reviews, including all reviews in "when to feed" and "adjunctive strategies".

Phase two: Identifying concerns with the review process

a) Domain 1: Study eligibility criteria

15/17 studies in 'what to feed' were assessed as low concern in assessment of the first domain of study eligibility criteria[59,126,148–152,127,128,142–147]. Three reviews in 'when to feed' were assessed as low concern for this domain[137,154,155]. Two reviews in 'how to feed' were assessed as low concern [133,156]. Three reviews in 'adjunctive strategies' were assessed as low concern [134,157,159]. Ng *et al.*[134] had a post hoc decision to include studies with erythromycin doses >12 mg/kg/day yet justified this decision.

Studies were assessed as unclear concern as they had undeclared differences to the protocol or previous review methods [131,136], they had significant deviations from the protocol where not all additional outcomes were declared and no justification was given [129,130], or significant changes to the pre-defined outcomes of the reviews [132,135].

Anabrees *et al.*[158] and Dawson *et al.* [132] were assessed as high concern due to the unexplained changes and additions to the secondary outcomes assessed.

b) Domain 2: Identification and selection of studies

13/17 reviews in 'what to feed'[59,126,149–151,127,128,142,144–148], 2/5 reviews in "when to feed"[154,155], 3/4 reviews in 'how to feed'[132,133,156], and all 3/4 reviews in "adjunctive strategies[157–159] were assessed as low concern. Of these it should be noted that seven studies [126–128,145,150,155,156] stated that although all citations were screened, only declared independent screening of the full reviews, therefore were marked as "probably yes" for this criteria. Dempsey *et al.*[142] applied machine learning using the Cochrane Classifier tool to assess for and remove reports with 0-2% probability of being randomised and having infants in the population[142].

Eight reviews were assessed as unclear concern[129–131,134–137,143,152]. Four reviews[131,134,135,137] were assessed as unclear concern as it was unclear if screening was performed independently. In Oddie *et al.* [136] it was unclear if study selection was conducted by two reviewers, in Tan-Dy *et al.*[130] it was not clear on how many authors assessed the abstracts and full reports. Brown *et al.*[143] was assessed as unclear concern as it stated that only one review author screened all titles and abstracts, a second review author only assessed all records coded as "order" and

made the final decision. Eligibility was assessed by the one author and checked (but not independently) by a second author[143]. Nasuf *et al.*[152] conducted screening of all citations by two reviewers, but only assessed the full text of all articles selected by the principle review author[152]. Nehra *et al.* [129] have no explanation for the language restriction in the search[129].

c) Domain three: Data collection and study appraisal

Thirteen reviews in 'what to feed'[59,126,151–153,127,128,142,143,145–147,149], two reviews in 'when to feed'[136,154,155]; two reviews in 'how to feed'[132,156], two reviews in 'adjunctive strategies'[157,158] were assessed as low concern. Within these reviews there were a few studies with insufficient methodological details from the studies available for authors and readers to be able to interpret the results. Muelbert *et al.*[157] does not say that data extraction was in duplicate, but it was felt to be implied by reporting that disagreements were resolved by discussion with a third assessor[157].

Two reviews were assessed as unclear concern as the majority [148] or all [144] risk of bias assessments had insufficient methodological details from the studies available for authors and readers to be able to interpret the results. These were addressed in the reviews therefore in the overall ROBIS assessment was not a risk for bias in the review process.

Eleven reviews[129,130,158,131,133–137,143,147] remained unclear concern as the risk of bias tool had not assessed all areas of bias, most frequently not assessing selective reporting or other bias. Tan-Dy *et al.*[130] was assessed as unclear concern as although a risk of bias assessment was conducted, the quality of included trials were assessed using a more basic tool (Assessed as blinding of randomisation, blinding of intervention, blinding of outcome measure assessment, completeness of follow up), with some methodological details insufficient, and other bias not assessed[130]. Nehra *et al.*[129] gave only a narrative inconsistent description of areas of bias of included studies with mostly insufficient methodological details for assessment[129]. The risk of bias assessment used in Premji *et al.* 2011[133] is also inconsistent in reporting and incomplete. Similarly, Ng *et al.*[134] had an incomplete risk of bias assessment, but described bias narratively and assessed using defined criteria.

Six reviews were assessed as unclear concern as there was insufficient methodological information to assess if error of risk of bias assessment was

minimised, including if the risk of bias assessment was performed in duplicate [131,133–135,137,159].

d) Domain four: Synthesis and findings

Thirteen reviews in "what to feed"[59,126,148,149,151,127,128,130,142–146], two reviews in 'when to feed'[136,155], two reviews in "how to feed"[133,156] and three reviews in "adjunctive strategies"[134,157,159] were assessed as low concern in this domain.

Four reviews performed subgroup analysis [127,143,145,149]. Brown *et al.*[143] assessed publication bias in outcomes with sufficient studies to perform, and sensitivity analyses were performed where there was >50% heterogeneity[143].

Most reviews were unable to conduct their planned sensitivity analyses for risk of bias due to no studies being at low or similar risk of bias and had insufficient studies to perform their planned subgroup analyses and assess for publication bias.

Ng *et al.*[134] was assessed as low risk as it acknowledged and justified departure from methods and recognised and described how large heterogeneity made it impossible to combine outcomes.

Seven reviews were assessed as unclear concern[129,131,137,147,150,152,155]. Four reviews neither acknowledged nor discussed their findings of moderate, high or substantial heterogeneity [137,147,150,155], two reviews had no plan for assessment for publication bias [129,131], one review did not report a planned subgroup analysis and did not give a justification for its absence [152], and one review did not plan or perform a sensitivity analysis for assessment of low risk of bias [129].

Three reviews were assessed as high concern in this domain. Watson *et al.*[135] was assessed as high concern as the funnel plot for publication bias suggested publication bias due to assymmetry for GI intolerance[135]. Dawson *et al.*[132] was assessed as high concern due to the unexplained changes to the outcome measures in addition to the insufficient information to assess if the findings were robust due to insufficient studies for funnel plot or sensitivity analyses[132]. Anabrees *et al.*[158] was assessed as high concern as there were changes and additions to the secondary outcomes assessed and moderate heterogeneity in duration of hospital stay (I^2=66%), which it did not acknowledge when discussing that secondary outcomes were not influenced in the summary of main results[158].

Phase three: Judging Risk of bias in the review

All reviews were assessed as appropriately considering the relevance of identified studies to the review's research question and avoiding emphasizing the results based on their statistical significance. All reviews summarised findings and discussed results in the context of factors affecting the risk of bias, characteristics of the assessed population and interventions.

Twelve reviews in 'what to feed'[59,126,151,152,127,128,142,144–146,148,149], two reviews in 'when to feed'[154,155], one study in 'how to feed'[156], two reviews in 'adjunctive strategies'[157,159] were assessed as low concern for risk of bias within the review.

Six reviews were assessed as low risk despite unclear concerns in the domains of the second phase[134,144,148,152,155,159]. Two reviews were assessed as low risk as the reason for the unclear risk of bias had been addressed in the reviews[144,148]. One review[152] was assessed as low risk of bias, despite not having addressed the concern in domain 4 within the interpretation of findings, as no subgroup analyses were performed or omission justification, it was not felt to have introduced significant bias. Although one review[155] did not specifically discuss the moderate heterogeneity found, this was assessed as low concern as the quality of evidence was overall acknowledged as low on GRADE assessment. Green *et al.*[159] was unclear about the methods to minimise error in the risk of bias assessment but this was felt to be low risk in the review overall. The risk of bias assessment in Ng *et al.*[134] was out of date but assessed and described, but concerns on independent screening assessment remained.

Twelve reviews were assessed as unclear concern for the overall judgement of risk of bias within the review[129,130,150,158,131,133–137,143,147].

Seven reviews remained unclear due to concerns about the study selection being performed in duplicate or independently[130,131,134–137,143]. Three reviews[137,147,150] were assessed as unclear as they did not address the heterogeneity observed in discussion of the findings. Eleven reviews[129,130,158,131,133–137,143,147] remained unclear concern as the risk of bias tool had not assessed all areas of bias, most frequently not assessing selective reporting or other bias. Four reviews remained unclear due to uncertainty of the rigour of the risk of bias assessment[131,133,135,137]. Nehra *et al.*[129]was unclear due to the lack of explanation for language restriction, inconsistent risk of bias assessment.

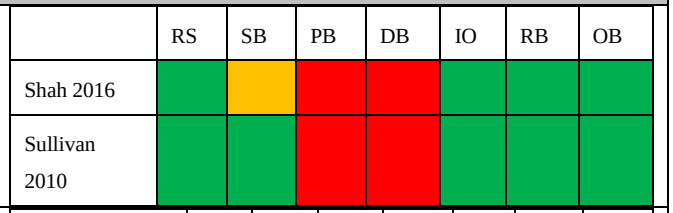
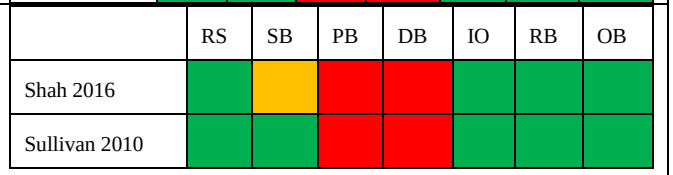
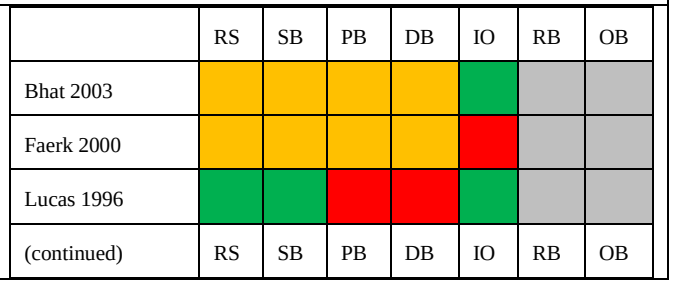
Unclear also remained in two reviews due to the absence of a plan for sensitivity analysis based on bias or to look for publication bias[129,131].

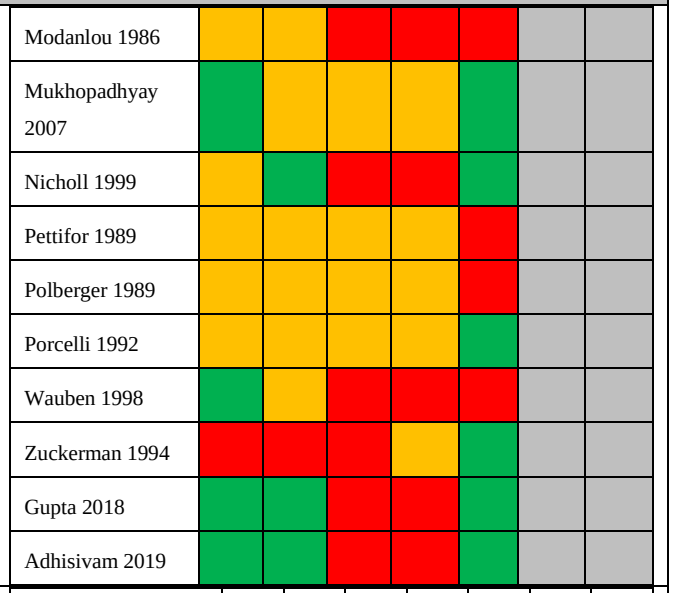
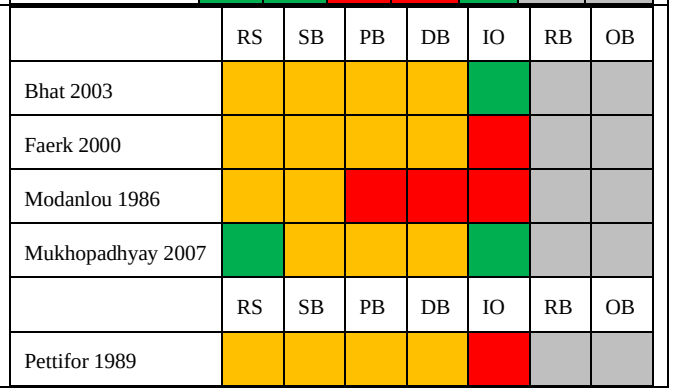
Three reviews continued to have concerns about undeclared changes to the methodology due to changes in outcomes assessed from those predefined [132,135,158] and the reason behind their change[132]. Anabrees *et al.*[158] also continued to have unresolved concerns due to the heterogeneity to length of stay not discussed when summarising the results of the review analyses[158].

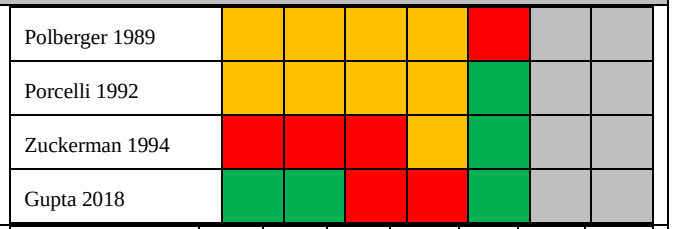
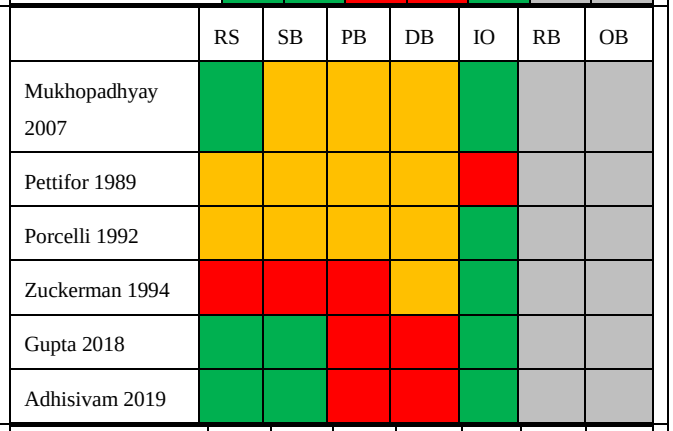
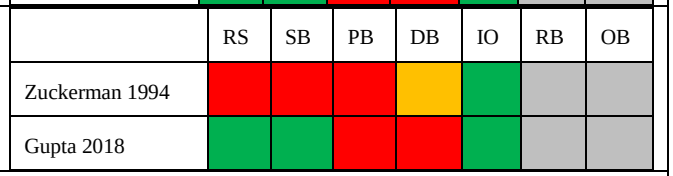
The nature of the changes to the outcomes in Dawson *et al.*[132] were felt to be significant, and therefore the overall risk of bias was felt to be high concern.

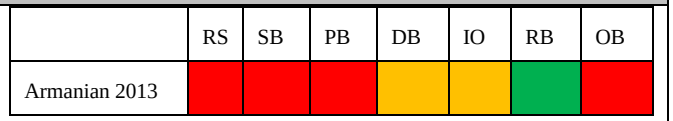
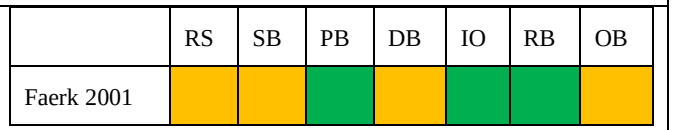
15.5 Appendix 5: Summary tables of effects of the interventions on Necrotising Enterocolitis

Table 21: Appendix 5a: Summary table of effects of the interventions on Necrotising Enterocolitis: 'What to feed'

Outcome assessed: Necrotising Enterocolitis	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Thanigainathan *et al.* 2020[128] NEC stage 2/3	Early versus late fortification	237 (2 studies) 120 vs 117	RR 1.36 [0.44, 4.16]	0%	Low certainty: • Lack of blinding • Small sample size	
Thanigainathan *et al.* 2020[128] Surgical NEC	Early versus late fortification	237 (2 studies) 120 vs 117	RR 0.98 [0.14, 6.85]	0%	Not assessed in review *Low certainty [Overview]* • *Risk of bias: lack of blinding* • *Small sample size*	
Brown *et al.* 2020[143]	Multi-nutrient fortified vs unfortified breast milk	1110 (13 studies) 565 vs 546	RR 1.37 [0.72, 2.63]	0%	Low certainty • Uncertainty about methods used to generate random sequence, conceal allocation and	

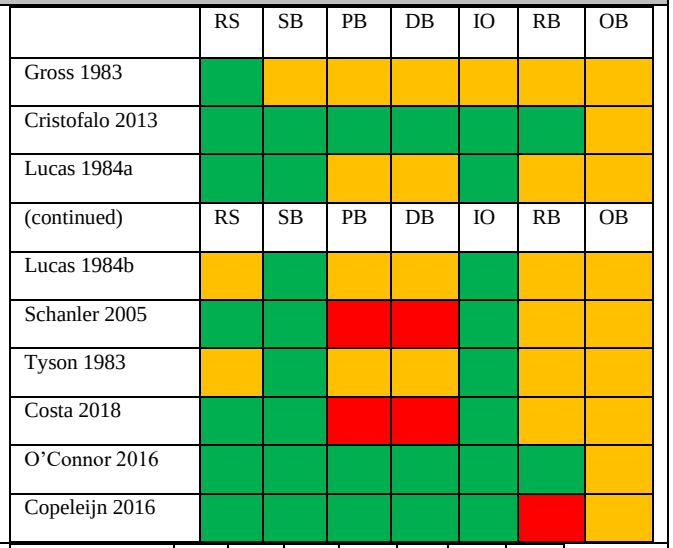
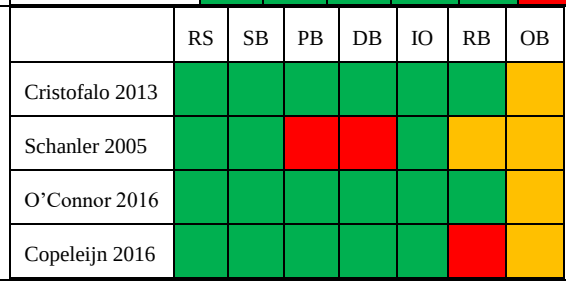
Outcome assessed: Necrotising Enterocolitis	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Brown *et al.* 2020[143] (continued)					blind assessments in most trials • Imprecision of estimated of effect (95% CI of RR consistent with substantial harm of benefit)	
	Subgroup: Trials recruiting only very preterm or VLBW infants	(9 studies) 359 vs 342	RR 1.28 [0.55, 2.99]	0%		

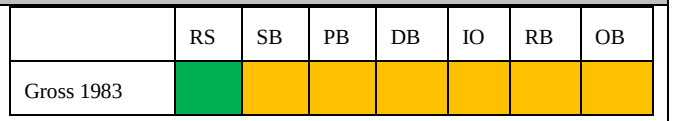
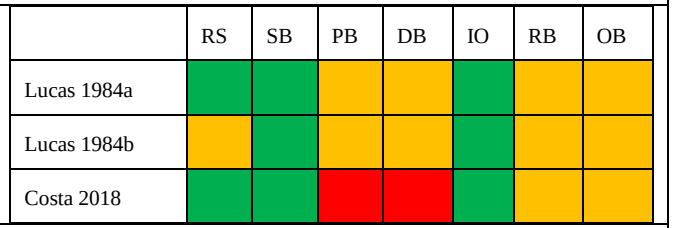
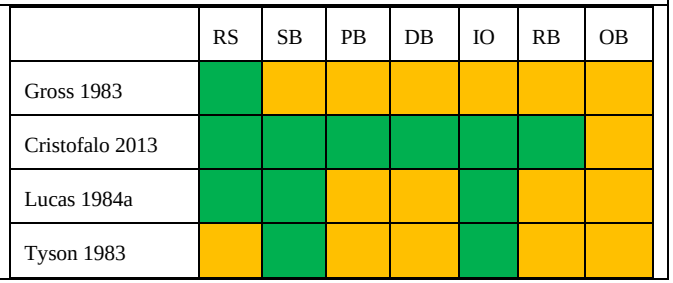
Outcome assessed: Necrotising Enterocolitis	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Brown *et al.* 2020[143] (continued)						
	Subgroup: Trials conducted in low- or middle-income countries	(6 studies) 286 vs 266	RR 1.10 [0.36, 3.38]	0%		
	Subgroup: Trials using preterm formula powder as fortifier	301 (2 studies) 104 vs 97	RR 1.49 [0.21, 10.76]	0%		
	Test for subgroup differences: $Ch^2 = 0.13$, df = 3 (P=0.990, I^2=0%)					

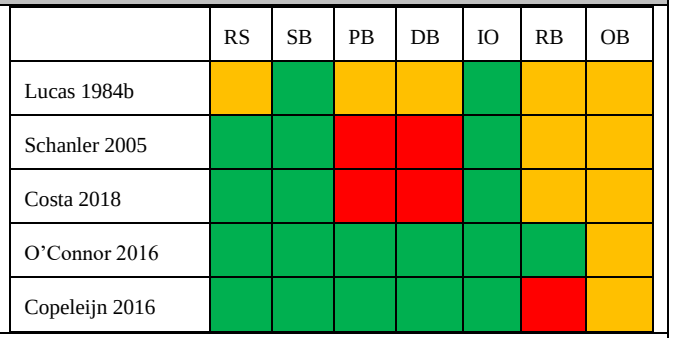
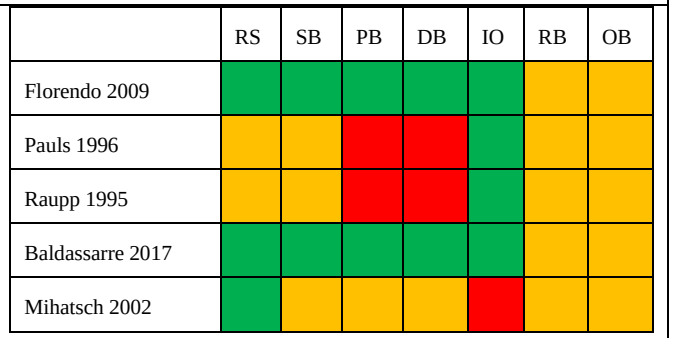
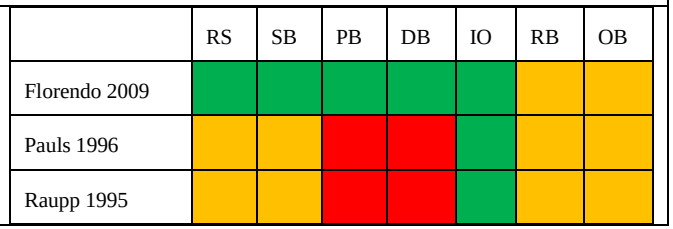
Outcome assessed: Necrotising Enterocolitis	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Amissah *et al.* 2020a[146]	Carbohydrate (prebiotic) supplement vs no supplement *No subgroup analyses*	75 (1 study) 25 vs 50	RR 0.18 [0.02 to 1.33]	n/a single study	Very low-quality evidence: • Methodological information was insufficient for judgement of risk of bias, participants • Events were few, and confidence intervals were wide	
Amissah *et al.* 2020b[148]	Protein supplementation versus no supplementation *No subgroup analyses*	76 (1 study) 36 vs 40	RR 1.11 [0.07 to 17.12]	n/a	Very low quality evidence • Few patients, few events and very wide confidence intervals. Downgraded two levels.	
Amissah *et al* 2020c[144]	Fat supplemented human milk versus control	No eligible studies reported this outcome.	n/a	n/a	n/a	n/a

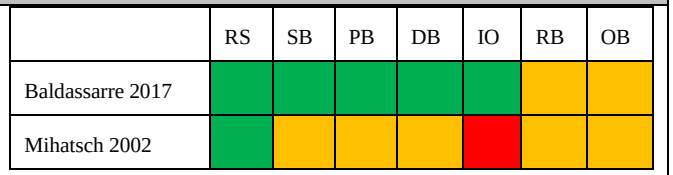
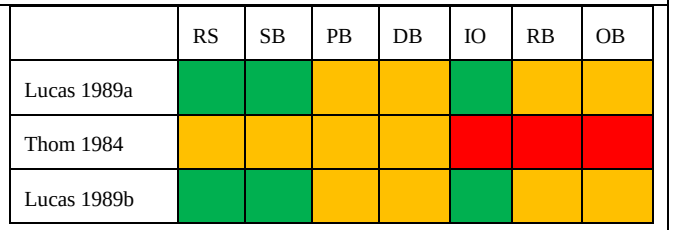
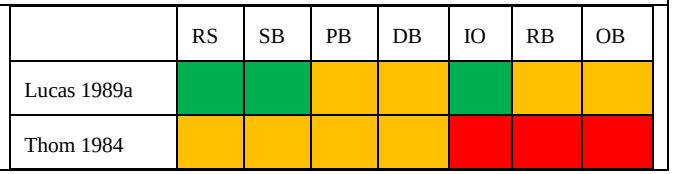
<table>
<tr><th>Outcome assessed: Necrotising Enterocolitis</th><th>Comparison</th><th>Number of subjects (studies)</th><th>Measure of effect (95% CI)</th><th>I^2</th><th>GRADE assessment</th><th>Risk of Bias</th></tr>
<tr><td rowspan="2">Fenton et al 2020[149]</td><td>High versus low protein intake

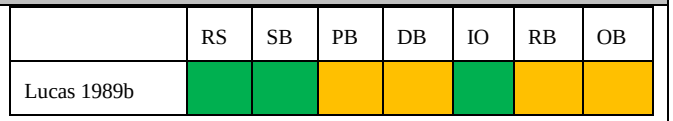
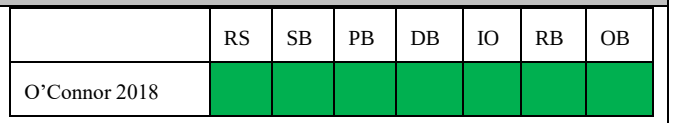
No subgroup analyses</td><td>46 (2 studies) 24 vs 22</td><td>Not estimable.
RD 0.00, 95% CI -0.12 to 0.12</td><td>0%</td><td>Very low certainty:
Risk of bias (selection, performance, detection, attrition) and Imprecision (Uncertain what criteria were used to define necrotizing enterocolitis in these studies).</td><td><table><tr><th></th><th>RS</th><th>SB</th><th>PB</th><th>DB</th><th>IO</th><th>RB</th><th>OB</th></tr><tr><td>Svenningsen 1982</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr><tr><td>Wauben 1995</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table>Uncertain what criteria were used to define necrotizing enterocolitis in these studies.</td></tr>
<tr><td colspan="6">n.b. not assessed in comparisons of high vs very high protein content, or in comparisons including studies with differences in other nutrients.</td></tr>
</table>

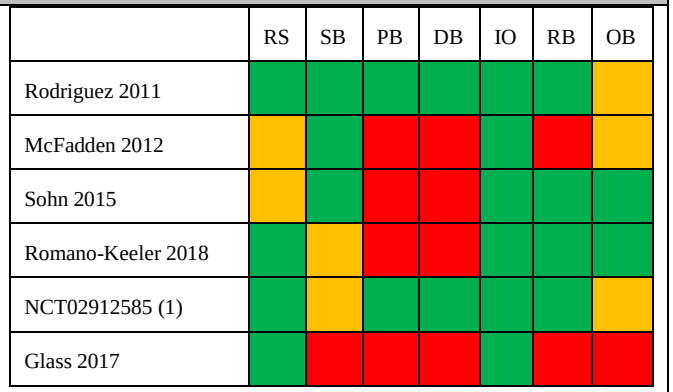
Outcome assessed: Necrotising Enterocolitis	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Quigley *et al* 2019[145]	Formula milk vs Donor breast milk Formula milk vs Donor breast milk (continued)	869 (6 studies) 431 vs 438	RR 1.87 [1.23, 2.85]	14%	moderate-certainty evidence (downgraded for imprecision)	
	Subgroup comparison 1: preterm formula versus fortified DBM	955 (2 studies) 484 vs 471	RR 1.64 [1.03, 2.61]	51%		

Outcome assessed: Necrotising Enterocolitis	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Quigley *et al* 2019[145] (Continued)	Subgroup comparison 1: term formula versus unfortified DBM	67 (1 study) 26 vs 41	RR 4.73 [0.52, 43.09]	n/a		
	Subgroup comparison 1: preterm formula versus unfortified DBM	653 (4 studies) 328 vs 325	RR 2.99 [0.90, 9.87]	0%		
	Test for subgroup differences: Chi^2=1.57, df = 2 (P=0.46), I^2=0.0%					
	Subgroup comparison 2: sole diet	360 (4 studies) 170 vs 190	RR 4.62 [1.47, 14.56]	0%		

Outcome assessed: Necrotising Enterocolitis	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Quigley *et al* 2019[145] (Continued)	Subgroup comparison 2: supplement to maternal breastmilk	1315 (5 studies) 668 vs 647	RR 1.56 [0.98, 2.47]	36%		
	Test for subgroup differences: $Chi^2 = 2.97$, df = 1 (P=0.08), I^2=66%					
Ng *et al* 2019[150]	Hydrolysed versus non-hydrolysed formula,	385 (5 studies) 199 vs 186	RR 1.10 (0.36 to 3.34)	0%	Low certainty: Methodological limitations of included trials (including uncertainty about allocation concealment and blinding) and imprecise effect size estimate	
	Subgroup: Partially hydrolysed	(3 studies) 123 vs 115	RD 0.01 [-0.03, 0.06]	0%		

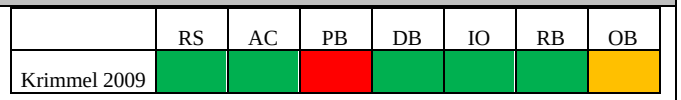
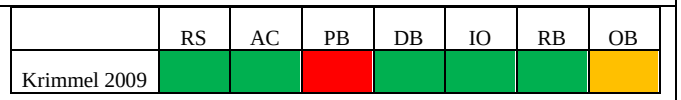
Outcome assessed: Necrotising Enterocolitis	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Ng *et al* 2019[150]	Subgroup: extensively hydrolysed	(2 studies) 76 vs 71	RD -0.02 [-0.07, 0.04]	0%		
	Test for subgroup differences: Chi^2=0.71, df = 1 (P=0.40) I^2= 0%					
Walsh *et al* 2019[127]	Nutrient-enriched formula versus standard formula	489 (3 studies) 248 vs 241	RR 0.72 [0.41, 1.25]	18%	Low risk: • Uncertainty about methods used to generate random sequence, conceal allocation, and mask assessments in trials • Post hoc exclusions in two trials.	
	Subgroup: Sole diet	225 (2 studies) 116 vs 109	RR 0.67 [0.27, 1.65]	58%		

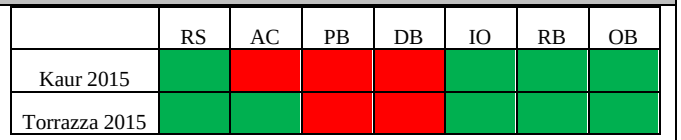
Outcome assessed: Necrotising Enterocolitis	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Walsh *et al* 2019[127] (continued)	Subgroup: Supplemental to human milk	264 (1 study) 132 vs 132	RR 0.75 [0.37, 1.52]	n/a		
	Test for subgroup differences: $Chi^2 = 0.04$, df=1 (P=0.85), I^2=0%					
Dempsey *et al* 2019[142]	Banked preterm versus banked term human milk	No eligible studies found	n/a	n/a	n/a	n/a
Brown *et al* 2019[126]	Formula versus maternal breastmilk	No eligible studies found	n/a	n/a	n/a	n/a
Basuki *et al* 2019[147]	Dilute versus full-strength formula	No studies reported this outcome	n/a	n/a	n/a	n/a
Premkumar *et al* 2019[151]	Human milk-derived fortifier versus bovine milk-derived fortifier *No subgroup analyses*	125 (1 study) 64 vs 61	RR 0.95 (95% CI 0.20 to 4.54	n/a	Low certainty evidence downgraded due to imprecision and inclusion of a single study	

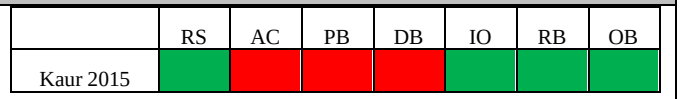
Outcome assessed: Necrotising Enterocolitis	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Nasuf *et al* 2018[152]	Oropharyngeal colostrum (OPC) compared to control (water, saline or no intervention) in preterm infants *No subgroup analyses*	335 (6 studies) 172 vs 163 n.b. 2 studies had no NEC cases therefore risk estimate is based on 4 studies with 290 participants.	RR 1.42 (0.50 to 4.02)	0%	Very low certainty • Studies with the highest weight involved concern about allocation concealment and blinding. • Small sample size. The confidence interval was wide and crossed the line of no effect.	
Abiramalatha *et al* 2017[59]	High-volume vs standard-volume feeds *No subgroup analyses*	61 (1 study) 30 vs 31	RR 1.03 [0.07, 15.78]	n/a	Very low certainty evidence Risk of bias (lack of blinding) and serious imprecision.	

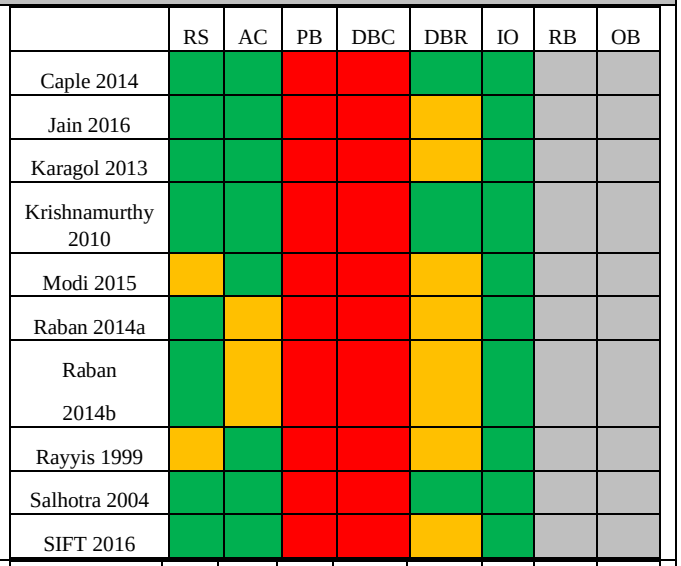
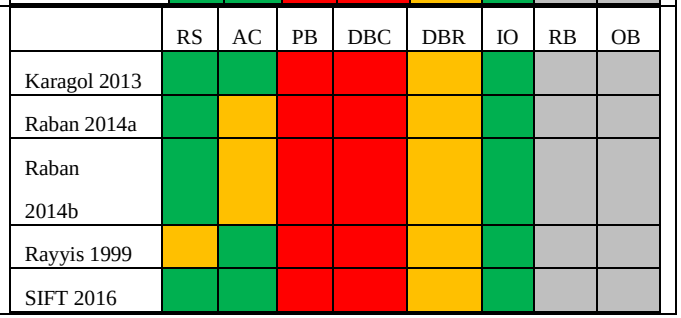
<table>
<tr><th>Outcome assessed: Necrotising Enterocolitis</th><th>Comparison</th><th>Number of subjects (studies)</th><th>Measure of effect (95% CI)</th><th>I^2</th><th>GRADE assessment</th><th>Risk of Bias</th></tr>
<tr><td>Tan-Dy et al 2013[130]</td><td>Lactase treated feeds vs placebo
No subgroup analyses</td><td>130 (1 study) 66 vs 64</td><td>RR 0.32 [0.01, 7.79]</td><td>n/a</td><td>Low certainty evidence [Overview]:
• Serious risk of bias (unclear selection bias insufficient details)
• Serious imprecision</td><td>
<table>
<tr><th></th><th>RS</th><th>SB</th><th>PB</th><th>DB</th><th>IO</th><th>RB</th><th>OB</th></tr>
<tr><td>Erasmus 2002</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
</table>
</td></tr>
<tr><td>Nehra et al 2002[129]</td><td>High MCT formula versus low MCT formula</td><td colspan="3">Narrative only:
Limited information from 2 studies
Small numbers from withdrawal of infants. No apparent evidence of difference in NEC incidence</td><td>Very low certainty evidence [Overview]:
• Serious risk of bias (selection bias unclear methodology)
• Serious inconsistency (One study compared 2 levels of MCT, other study compared 4 different levels)
• Very serious imprecision (limited narrative)</td><td>
<table>
<tr><th></th><th>RS</th><th>SB</th><th>PB</th><th>DB</th><th>IO</th><th>RB</th><th>OB</th></tr>
<tr><td>Wu 1993</td><td>Can’t tell</td><td>Can’t tell</td><td>Can’t tell</td><td>Can’t tell</td><td>Complete follow up</td><td></td><td></td></tr>
<tr><td>Whyte 1966</td><td>Can’t tell</td><td>Yes</td><td>Yes</td><td>Can’t tell</td><td>Complete follow up</td><td></td><td></td></tr>
</table>
</td></tr>
</table>

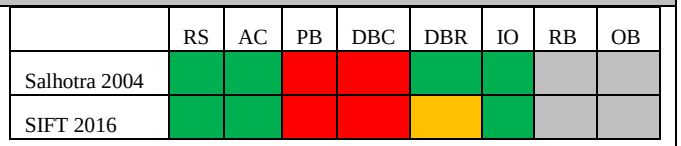
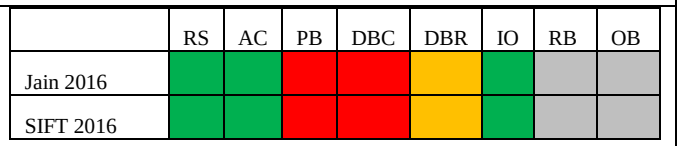
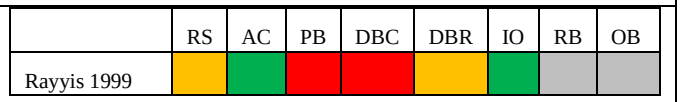
Table 22: Appendix 5b: Summary table of effects of the interventions on Necrotising Enterocolitis: 'When to feed'

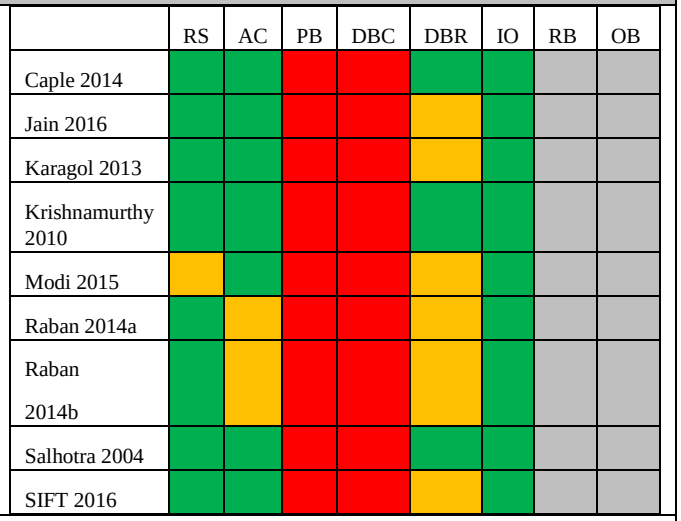
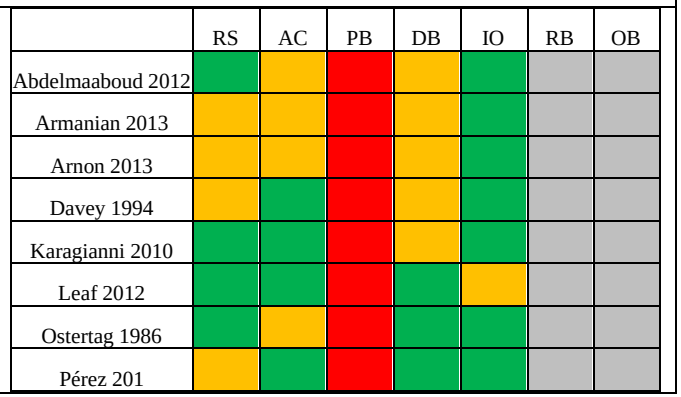
Outcome assessed: Necrotising Enterocolitis	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Yeo *et al* 2019[154]: Necrotising enterocolitis within 48 hours of transfusion	Stopping feeds vs continuing feeds during transfusion *No subgroup analysis*	22 (1 study) 11 vs 11	RR 0.0 [0.0, 0.0]	n/a	Low quality: Single study, no reported events, very few participants	
Yeo *et al* 2019[154]: Incidence of NEC any time after first transfusion	Stopping feeds vs continuing feeds during transfusion *No subgroup analysis*	22 (1 study) 11 vs 11	RR 0.0 [0.0, 0.0]	n/a	Low quality: Single study, no reported events, very few participants	

Outcome assessed: Necrotising Enterocolitis	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Abiramalatha *et al.* 2019a[155]: Number of infants with NEC stage 2 or 3	Routine monitoring of gastric residuals vs no routine monitoring of gastric residuals	141 (2 studies) 70 vs 71	RR 3.07 [0.50, 18.77]	0%	Low quality: Imprecise due to small sample size and low event rate.	
	Routine monitoring of gastric residuals quality vs routine monitoring of gastric residuals quality and quantity	87 (1 study) 42 vs 45	RR 5.35 [0.26, 108.27]	n/a	*Very low certainty [Overview]: - Serious risk of bias* *-Very serious imprecision*	
	No subgroup analyses possible					

Outcome assessed: Necrotising Enterocolitis	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Abiramalatha *et al* 2019a[155]: Number of infants with surgical NEC	Routine monitoring of gastric residuals vs no routine monitoring of gastric residuals	80 (1 study) 40 vs 40	RR 1.25 [0.36, 4.32]	n/a	*Very low certainty [Overview]: -Very serious risk of bias* *-Very serious imprecision*	
	Routine monitoring of gastric residuals quality vs routine monitoring of gastric residuals quality and quantity	87 (1 study) 42 vs 45	RR 5.35 [0.26, 108.27]	n/a	*Very low certainty [Overview]: -Serious risk of bias* *-Very serious imprecision*	
	No subgroup analyses possible					

Outcome assessed: Necrotising Enterocolitis	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Oddie *et al* 2017[136]	Slow versus faster rates of feed advancement	3742 (10 studies) 1886 vs 1856	RR 1.07 [0.83, 1.39]	21%	Moderate Risk of bias – all trials unblinded	
	Subgroup: Extremely LBW (<1kg) or extremely preterm (<28 weeks)	1299 (5 studies) 658 vs 641	RR 1.01 [0.74, 1.38]	59%		

Outcome assessed: Necrotising Enterocolitis	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Oddie *et al.* 2017[136] (continued)	Subgroup: SGA/Growth restricted infants	639 (2 studies) 317 vs 322	RR 1.26 [0.67, 2.37]	36%		
	Subgroup: Absent/Reversed EDFV	465 (2 studies) 241 vs 224	RR 1.59 [0.74, 3.40]	10%		
	Test for subgroup differences: Chi^2 =1.38 df =3 (P=0.71), I^2 = 0%					
	Subgroup: Trials where most infants were exclusively formula fed	185 (1 study) 98 vs 87	RR 1.44 [0.63, 3.32]	n/a		

Outcome assessed: Necrotising Enterocolitis	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Oddie *et al.* 2017[136] (continued)	Subgroup: Trials where most infants were at least partially fed with human milk	3557 (9 studies) 1788 vs 1769	RR 1.04 [0.79 to 1.37]	26%		
	Test for subgroup differences not reported					
Morgan *et al.* 2014[137]	Delayed versus early introduction of progressive feeding	1092 (8 studies) 527 vs 565	RR 0.93 [0.64 to 1.34]	0%	*Very low certainty [Overview]:* *Serious risk of bias* *Very serious imprecision*	

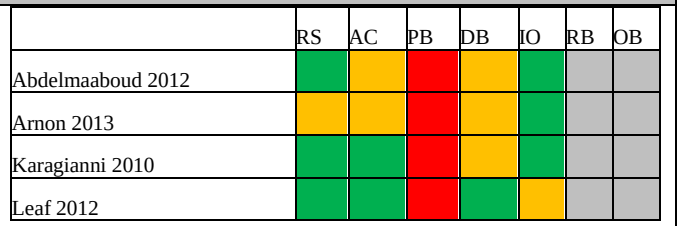
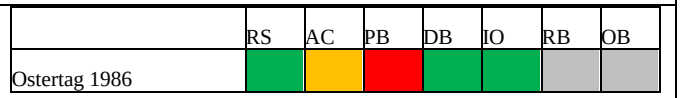
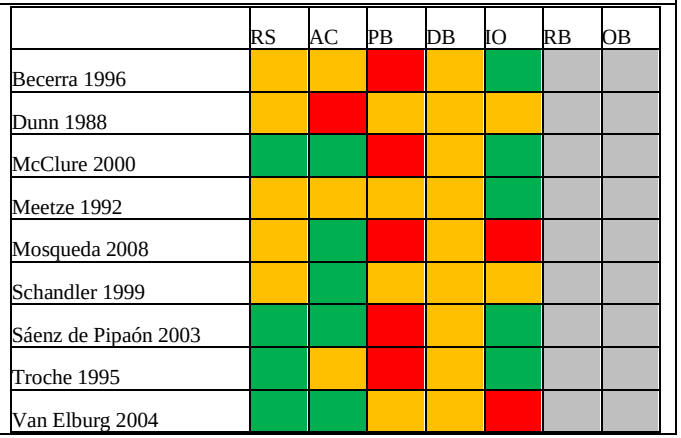
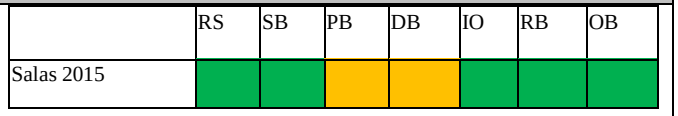
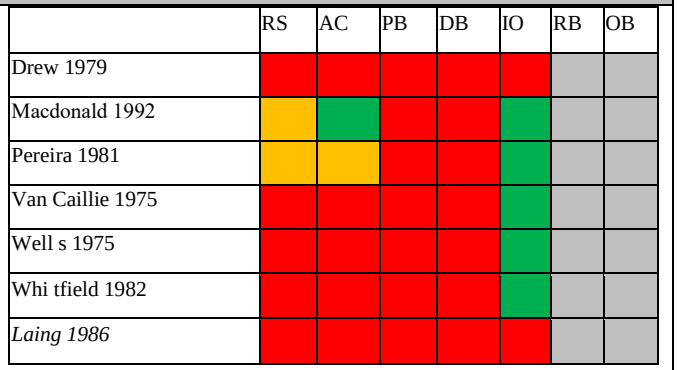
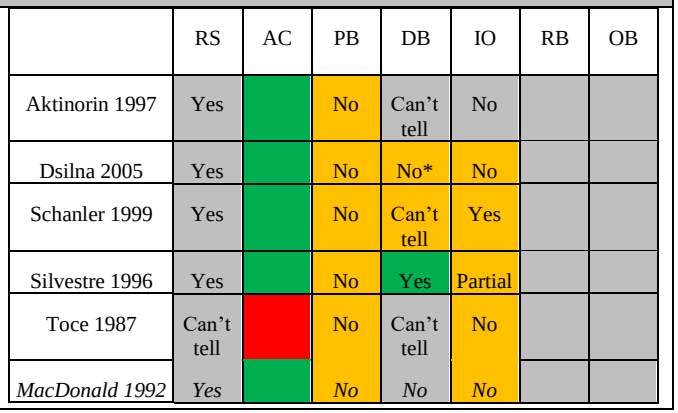
Outcome assessed: Necrotising Enterocolitis	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Morgan *et al.* 2014[137] (continued)	Subgroup: IUGR/abnormal antenatal doppler flow velocities	674 (4 studies) 337 vs 336	RR 0.87 [0.54, 1.41]	0%		
	Test for subgroup differences: $Chi^2 = 0.05$, df = 1 (P=0.830, I^2=0%					
	Subgroup: Exclusive formula fed infants	38 (1 study)	RR 1.08 [0.40, 2.94]	n/a		
	No test for subgroup differences reported					
Morgan *et al.* 2013[131]	Early trophic feeding vs enteral fasting *No subgroup analyses possible*	748 (9 studies) 374 vs 374	RR 1.07 [0.67, 1.70]	0%	*Very low certainty [Overview]: -Very serious risk of bias* *-Serious imprecision*	

Table 23: Appendix 5c: Summary table of effects of the interventions on Necrotising Enterocolitis 'How to feed'

Outcome assessed: Necrotising Enterocolitis	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Abiramalatha *et al.* 2019b[156]	Re-feeding vs discard gastric residuals	72 (1 study) 36 vs 36	RR 0.71 [0.25, 2.04]	n/a	Very Low: - Serious Imprecision due to wide confidence interval. Detection bias as study authors included both NEC and SIP.	
	Subgroup: Infants fed only human milk	49 (1 study) 25 vs 24	RR 0.96 [0.06, 14.50]	n/a		
	Subgroup: Infants fed only formula milk	22 (1 study) 11 vs 11	RR 0.80 [0.29, 2.21]	n/a		
	Test for subgroup differences: Chi^2=0.05, df=2 (P=0.98), I^2=0%					

Outcome assessed: Necrotising Enterocolitis	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Watson *et al.* 2013[135]	Transpyloric versus gastric tube feeding *Subgroup analyses not performed*	298 (7 studies) 153 vs 145 *Excluding Laing 1986: 218 (6 studies) 108 vs 110*	RR 0.63 [0.26, 1.53] *Excluding Laing 1986 – due to differences in baseline characteristics:* *RR 0.91 [0.32, 2.58]*	14% *(0%)*	*Very low certainty [Overview]:* • *Very serious risk of bias* • *Serious indirectness* • *Very serious imprecision*	
Dawson *et al.* 2012[132]	Push versus gravity bolus tube feeding	Nil eligible studies	n/a	n/a	n/a	n/a
Premji *et al.* 2011[133]: Proven NEC (Stage ≥II)	Continuous vs intermittent bolus milk feeding (Comparison 1)	465 (5 studies) 219 vs 246	RR 1.09 [0.58, 2.07] *n.b. further study reported cases of proven & suspected NEC solely on radiographic finding (not described):*	0%	*Very low certainty [Overview]:* • *Very serious risk of bias* • *-Serious imprecision*	

<table>
<tr><th>Outcome assessed: Necrotising Enterocolitis</th><th>Comparison</th><th>Number of subjects (studies)</th><th>Measure of effect (95% CI)</th><th>I²</th><th>GRADE assessment</th><th>Risk of Bias</th></tr>
<tr><td rowspan="4">Premji et al. 2011[133]:

Proven NEC (Stage ≥II)

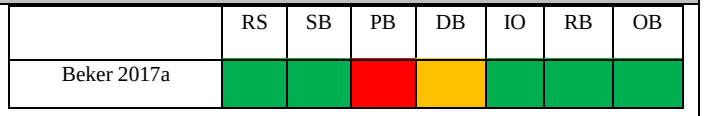
(continued)</td><td>Continuous vs intermittent bolus milk feeding (Comparison 1)

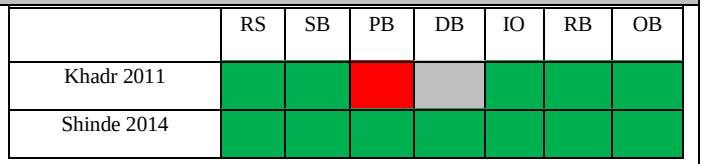
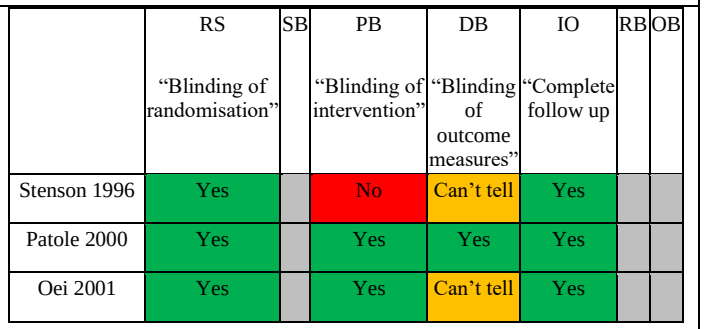
(continued)</td><td></td><td>- Continuous: 1 proven and 1 probable NEC.
• Bolus group: No NEC</td><td></td><td></td><td></td></tr>
<tr><td>Continuous vs NG intermittent bolus (Comparison 2)</td><td>270 (4 studies) 136 vs 134</td><td>2.23 [0.58, 8.57]</td><td>0%</td><td></td><td><table>
<tr><th></th><th>RS</th><th>AC</th><th>PB</th><th>DB</th><th>IO</th><th>RB</th><th>OB</th></tr>
<tr><td>Aktinorin 1997</td><td>Yes</td><td></td><td>No</td><td>Can't tell</td><td>No</td><td></td><td></td></tr>
<tr><td>Dsilna 2005</td><td>Yes</td><td></td><td>No</td><td>No*</td><td>No</td><td></td><td></td></tr>
<tr><td>Silvestre 1996</td><td>Yes</td><td></td><td>No</td><td>Yes</td><td>Partial</td><td></td><td></td></tr>
<tr><td>Toce 1987</td><td>Can't tell</td><td></td><td>No</td><td>Can't tell</td><td>No</td><td></td><td></td></tr>
</table></td></tr>
<tr><td>Continuous vs intermittent bolus (NG or OG) milk feeding in infants <1000g (Comparison 3)</td><td>44 (1 study) 22 vs 22</td><td>MD 5.0 [0.25, 98.52]</td><td>n/a</td><td></td><td><table>
<tr><th></th><th>RS</th><th>AC</th><th>PB</th><th>DB</th><th>IO</th><th>RB</th><th>OB</th></tr>
<tr><td>Dsilna 2005</td><td>Yes</td><td></td><td>No</td><td>No*</td><td>No</td><td>Yes</td><td></td></tr>
</table></td></tr>
<tr><td colspan="6">Continuous vs intermittent bolus (NG) milk feeding in infants <1000g (Comparison 4), Continuous vs intermittent bolus (NG) milk feeding in infants >1000g and <1249g (Comparison 5), Continuous versus intermittent bolus (nasogastric tube) milk feeding in infants > 1250g and < 1499g (Comparison 6) – nil trials eligible.</td></tr>
</table>

<table>
<tr><th>Outcome assessed: Necrotising Enterocolitis</th><th>Comparison</th><th>Number of subjects (studies)</th><th>Measure of effect (95% CI)</th><th>I²</th><th>GRADE assessment</th><th>Risk of Bias</th></tr>
<tr><td rowspan="3">Premji et al. 2011[133]:

Probable NEC</td><td>Continuous vs intermittent bolus milk feeding</td><td>133 (2 studies) 69 vs 64</td><td>RR 1.53 [0.40, 5.89]</td><td>55%</td><td>Very low certainty [Overview]:
Very serious risk of bias
Serious inconsistency
Very serious imprecision</td><td><table><tr><th></th><th>RS</th><th>AC</th><th>PB</th><th>DB</th><th>IO</th><th>RB</th><th>OB</th></tr><tr><td>Aktinorin 1997</td><td>Yes</td><td></td><td>No</td><td>Can’t tell</td><td>No</td><td></td><td></td></tr><tr><td>Toce 1987</td><td>Can’t tell</td><td></td><td>No</td><td>Can’t tell</td><td>No</td><td></td><td></td></tr></table></td></tr>
<tr><td>Continuous vs NG intermittent bolus (Comparison 2)</td><td>133 (2 studies) 69 vs 64</td><td>RR 1.53 [0.40, 5.89]</td><td>55%</td><td></td><td><table><tr><th></th><th>RS</th><th>AC</th><th>PB</th><th>DB</th><th>IO</th><th>RB</th><th>OB</th></tr><tr><td>Aktinorin 1997</td><td>Yes</td><td></td><td>No</td><td>Can’t tell</td><td>No</td><td></td><td></td></tr><tr><td>Toce 1987</td><td>Can’t tell</td><td></td><td>No</td><td>Can’t tell</td><td>No</td><td></td><td></td></tr></table></td></tr>
<tr><td colspan="6">Continuous vs intermittent bolus (NG or OG) milk feeding in infants <1000g (Comparison 3), Continuous vs intermittent bolus (NG) milk feeding in infants <1000g (Comparison 4), Continuous vs intermittent bolus (NG) milk feeding in infants >1000g and <1249g (Comparison 5), Continuous versus intermittent bolus (nasogastric tube) milk feeding in infants > 1250g and < 1499g (Comparison 6) – nil trials eligible</td></tr>
</table>

Table 24: Appendix 5d: Summary table of effects of the interventions on Necrotising Enterocolitis 'Adjunctive strategies'

Outcome assessed: Necrotising Enterocolitis	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Muelbert *et al.* 2019[157]	Exposure to smell and taste stimulation of milk with tube feeds versus no exposure *No subgroup analyses possible*	51 (1 study) 28 vs 23	RR 0.62 [0.15, 2.48]	n/a	Low quality: • imprecision (single trial with small sample size, and wide confidence intervals). • Lack of blinding judge to have unlikely to have influenced the assessment of this outcome.	
Greene 2016[159]	Comparison 1: Oral stimulation vs standard care	Nil eligible studies	n/a	n/a	n/a	n/a
	Comparison 2: Oral stimulation versus other non-oral intervention	Nil eligible studies	n/a	n/a	n/a	n/a

Outcome assessed: Necrotising Enterocolitis	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Anabrees *et al.* 2015[158]	Glycerin prophylaxis versus placebo/no intervention *No subgroup analyses possible*	96 (2 studies) 50 vs 46	RR 2.75 [0.58, 13.08]	0%	*Very low certainty [Overview]:* • *Serious risk of bias* • *Very serious imprecision*	
Ng *et al.* 2008[134]	Erythromycin vs placebo for prevention of feeding intolerance	3 studies. 2 included in meta-analysis. 149 (2 studies) 71 vs 78.	RR 0.59 [0.11, 3.01] Oei 2001 (excluded one infant in each group from analysis due to development of NEC) No events in one group.	n/a.	*Very low certainty [Overview]:* • *Serious risk of bias* • *Very serious imprecision*	

<table>
<tr><th>Outcome assessed: Necrotising Enterocolitis</th><th>Comparison</th><th>Number of subjects (studies)</th><th>Measure of effect (95% CI)</th><th>I^2</th><th>GRADE assessment</th><th>Risk of Bias</th></tr>
<tr>
<td>Ng et al. 2008[134]</td>
<td>Erythromycin vs placebo for treatment of feeding intolerance</td>
<td>5 studies.
n.b. Only Nuntnarumit 2006 clearly defined NEC (as > Bell stage I).</td>
<td>Nil reported significant difference.

Aly 2007: 3 (10%) vs. 4 (13.3%) (p = NS

Madani 2004: 2 (7%) vs. 3 (11%), (p = NS)

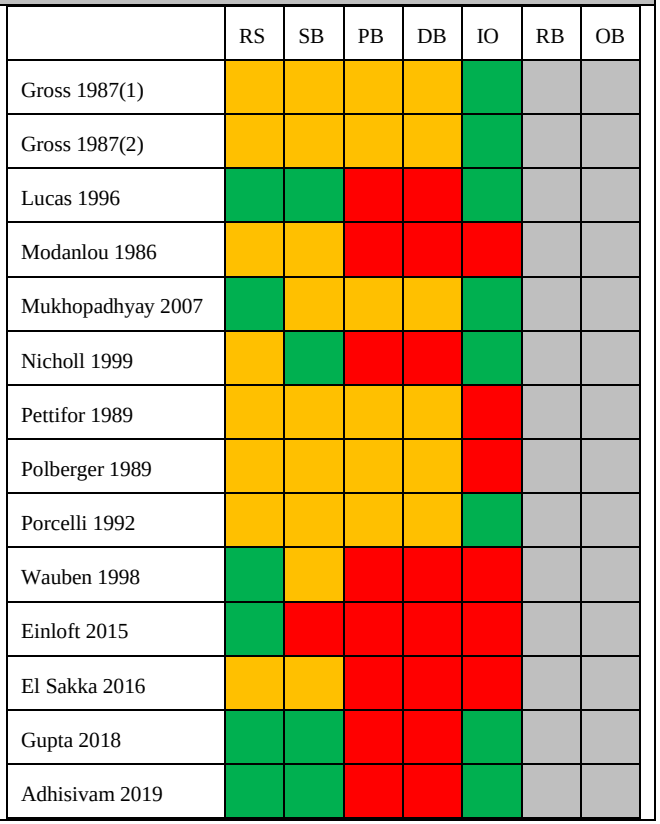
Nuntnarumit 2006: 1 (4%) vs. 4 (13%), (p = 0.61

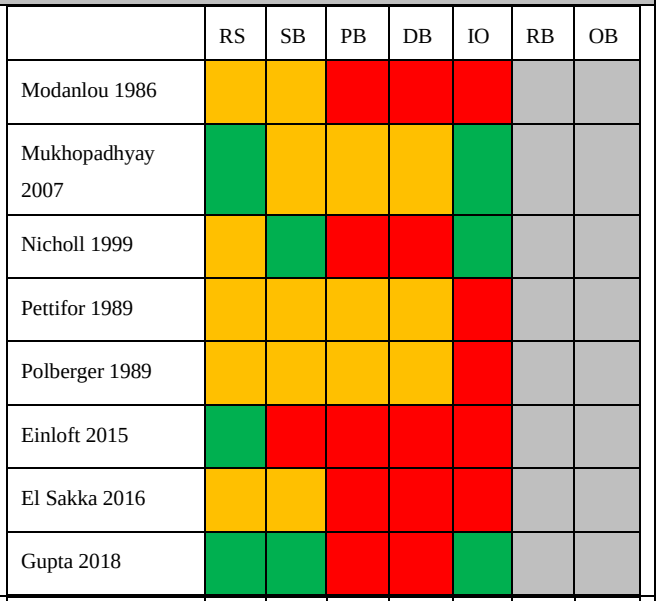
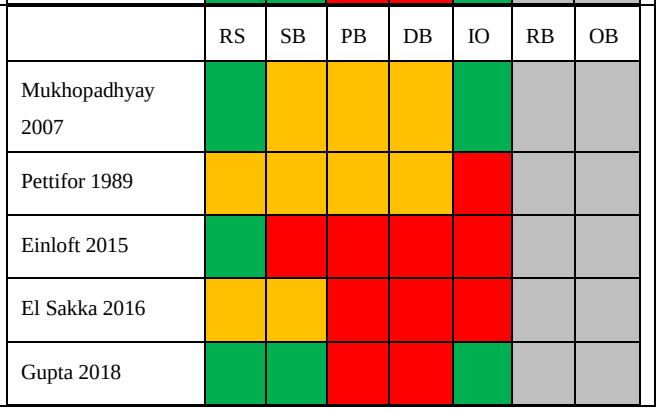
Ng PC 2001: no NEC events

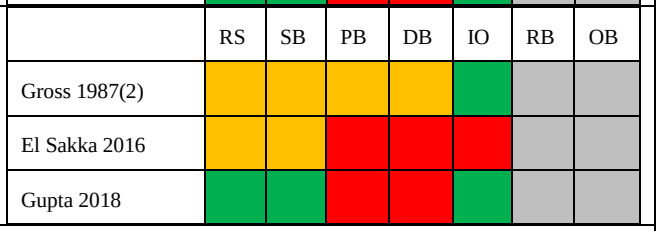
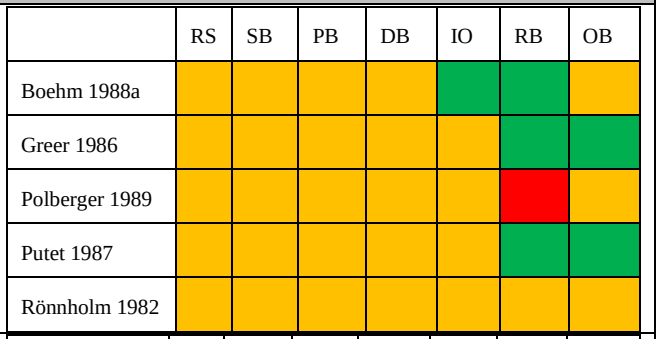
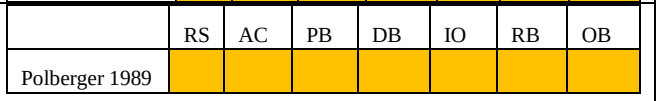
Ng SC 2003: 1 infant in placebo group (1 month after feeds attained).</td>
<td></td>
<td>Very low certainty [Overview]:
• Serious risk of bias
• Serious inconsistency
• Very serious imprecision</td>
<td>
<table>
<tr><th></th><th>RS
“Blinding of randomisation”</th><th>SB</th><th>PB
“Blinding of intervention”</th><th>DB
“Blinding of outcome measures”</th><th>IO
“Complete follow up</th><th>RB</th><th>OB</th></tr>
<tr><td>Nuntarumit 2006</td><td>Yes</td><td></td><td>Yes</td><td>Yes</td><td>Yes</td><td></td><td></td></tr>
<tr><td>Aly 2007</td><td>Yes</td><td></td><td>Yes</td><td>Yes</td><td>Yes</td><td></td><td></td></tr>
<tr><td>Madani 2004</td><td>Can’t tell</td><td></td><td>Can’t tell</td><td>Can’t tell</td><td>Yes</td><td></td><td></td></tr>
<tr><td>Ng PC 2001</td><td>Yes</td><td></td><td>Yes</td><td>Can’t tell</td><td>Yes</td><td></td><td></td></tr>
<tr><td>Ng PC 2003</td><td>Yes</td><td></td><td>Yes</td><td>Can’t tell</td><td>Yes</td><td></td><td></td></tr>
</table>
</td>
</tr>
</table>

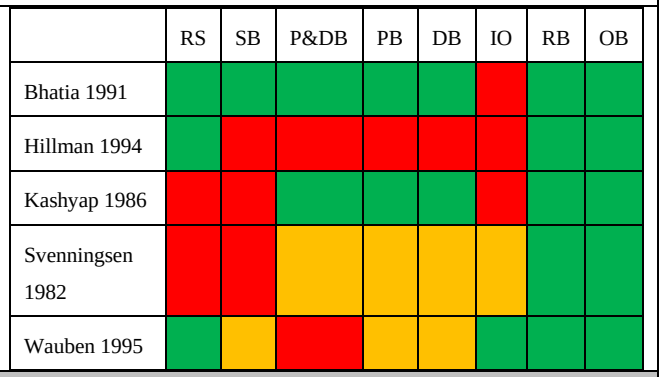
15.6 Appendix 6: Summary tables of effects of the interventions on Growth

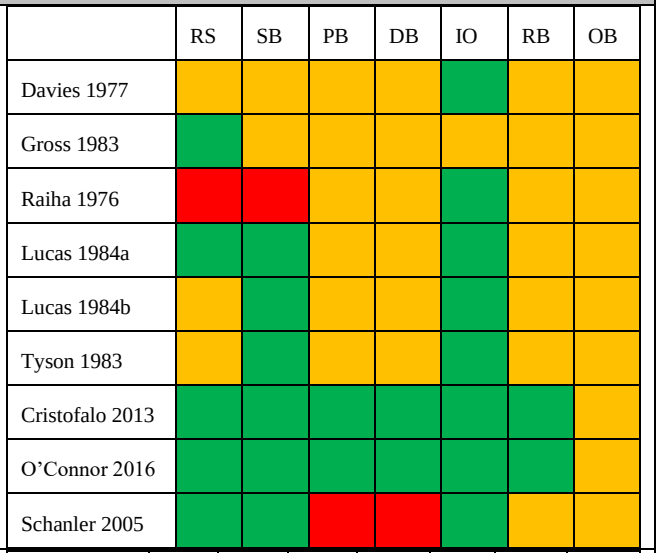
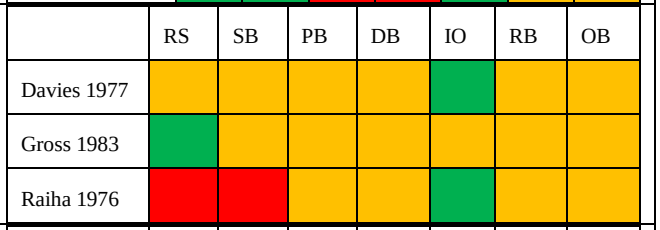
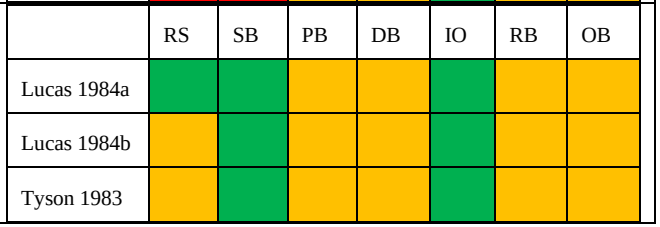
Table 25: Appendix 6: Summary table of effects of the interventions on weight gain (g/kg/day): 'What to feed'

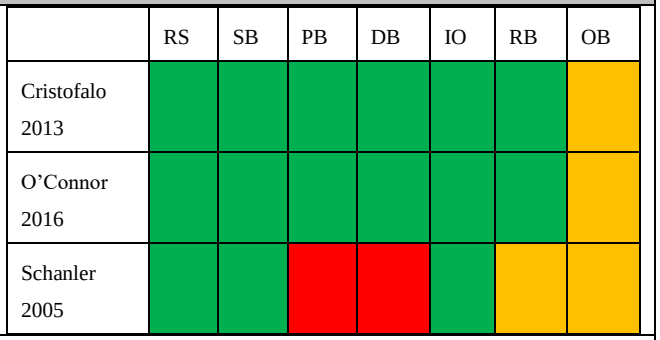
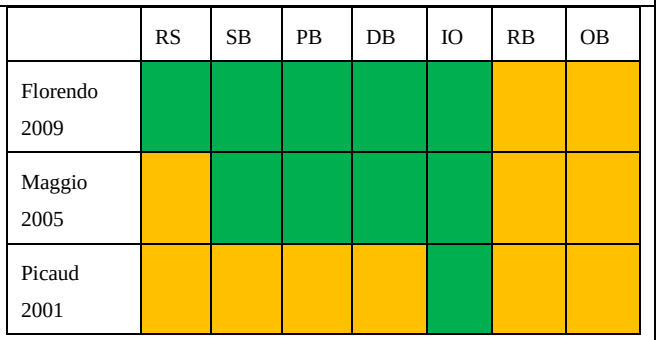
Weight gain (g/kg/day)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Thanigainathan *et al.* 2020[128]	Early versus late fortification of breastmilk	Nil studies eligible	n/a	n/a	n/a	n/a
Brown *et al.* 2020[143]	Multi-nutrient fortified vs unfortified breast milk	951 (14 studies) 484 vs 467	MD 1.76 [1.30, 2.22]	72%	Low certainty • Unexplained heterogeneity • Uncertainty about methods used to generate random sequence, conceal allocation and blind assessments in most trials	

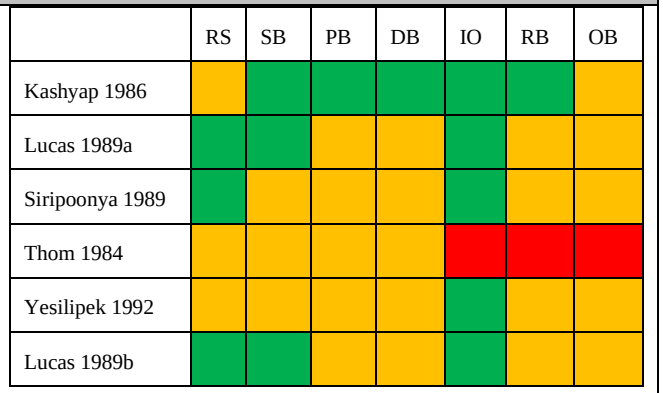
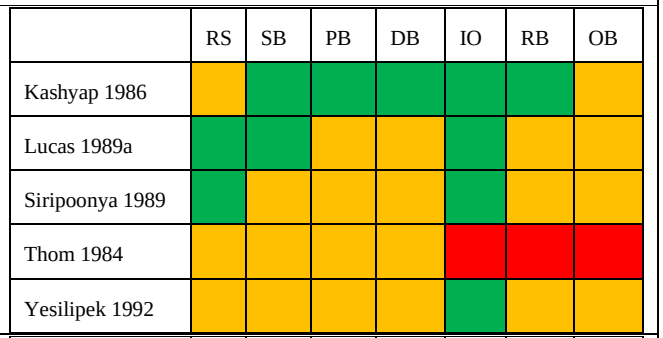
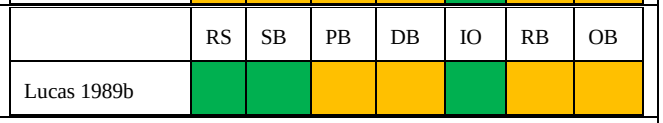
Weight gain (g/kg/day)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Brown *et al.* 2020[143] (continued)	Very preterm/VLBW infants	505 (8 trials) 258 vs 247	MD 2.18 [1.54, 2.81]	70%		
	Trials conducted in low- or middle-income countries	530 (6 trials) 270 vs 260	MD 1.73 [1.10, 2.35]	22%		

Weight gain (g/kg/day)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Brown *et al.* 2020[143] (continued)						
	Trials using preterm formula as fortifier	224 (3 trials) 117 vs 107	MD 2.20 [1.36, 3.04]	16%		
	Test for subgroup differences: Chi^2 = 1.85, df = 3 (P=0.60), I^2 = 0%					
Amissah *et al.* 2020a[146]	Carbohydrate (prebiotic) supplement vs no supplement	Nil eligible (only weight at day 30)	n/a	n/a	n/a	n/a
Amissah *et al.* 2020b[148]	Protein supplementation versus no supplementation	101 (5 studies) 52 vs 49	MD 3.82 [2.94, 4.70]	73%	Low • Uncertain methodology • Moderate-to-high heterogeneity	
Amissah *et al.* 2020c[144]	Fat supplemented human milk versus control	14 (1 study) 7 vs 7	MD 0.6 [−2.4 to 3.6]	n/a	Very low quality • Uncertain risk of bias	

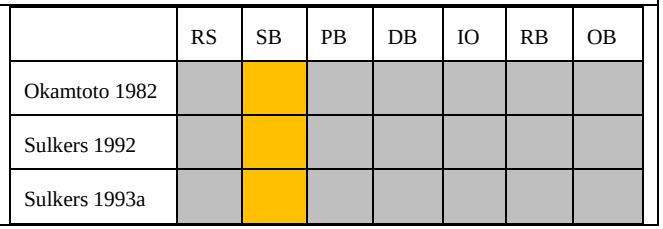
Weight gain (g/kg/day)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Amissah *et al.* 2020c[144] (continued)					(insufficient methodological detail). • Imprecision (very small sample size) Wide confidence intervals spanning across benefits and harms	
Fenton *et al.* 2020[149]	High versus low protein intake in formula fed infants	114 (5 studies) 67 vs 47	MD 2.36 [1.31, 3.40]	57%	Low quality evidence • Risk of bias • Heterogeneity	
	Very high versus high protein intake	Not assessed	n/a	n/a	n/a	n/a

Weight gain (g/kg/day)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Quigley *et al.* 2019[145]	Formula milk vs Donor breast milk	1028 (9 studies) 540vs488	MD 2.51 [1.93, 3.08]	90%	Moderate-certainty evidence • high level of heterogeneity	
	Subgroup: Term formula versus unfortified DBM	234 (3 studies) 138v96	MD 1.74 [0.96, 2.53]	94%		
	Subgroup: preterm formula versus unfortified DBM	249 (3 studies) 128 v 121	MD 4.16 [3.04, 5.28]	94%		

Weight gain (g/kg/day)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Quigley *et al.* 2019[145] (continued)	Subgroup: Preterm formula versus fortified DBM	545 (3 studies) 274 v 271	MD 2.37 [1.09, 3.65]	0%		
	Test for subgroup difference: Chi^2=11.95, df=2 (P=0.00) I^2=83%					
Ng *et al.* 2019[150]	Hydrolysed versus non-hydrolysed formula,	113 (3 studies) 59 vs 54	MD -3.02 [-4.66, -1.38]	19%	*Low certainty [Overview]* • *Serious risk of bias (unclear allocation, funding from formula manufacturers)* • *Serious imprecision*	

Weight gain (g/kg/day)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Walsh *et al.*2019[127]	Nutrient-enriched formula versus standard formula	440 (6 studies) 220 vs 220	MD 2.43 [1.60, 3.26] g/kg/day	46%	Low certainty: • Uncertainty about methods used to generate random sequence, conceal allocation, and mask assessments in trials. • Moderate to high heterogeneity	
	Subgroup: formula as a sole diet	225 (5 studies) 115 v 110	MD 3.87 [2.26, 5.47 g/kg/day	20%		
	Subgroup: formula as a supplement	215 (1 study) 105 v 110	MD 1.90 [0.93, 2.87] g/kg/day	n/a		
	Test for subgroup difference: Chi^2=4.20, df=1 (P=0.04) (I^2=76.2%)					

Weight gain (g/kg/day)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Dempsey *et al.* 2019[142]	Banked preterm versus banked term human milk	Nil studies eligible	n/a	n/a	n/a	n/a
Brown *et al.* 2019[126]	Formula versus maternal breastmilk	Nil studies eligible	n/a	n/a	n/a	n/a
Basuki *et al.*2019[147]	Dilute versus full-strength formula	Nil eligible studies (only weight at 7 days)	n/a	n/a	n/a	n/a
Premkumar *et al.* 2019[151]	Human milk-derived fortifier versus bovine milk-derived fortifier	Nil eligible (only absolute change in weight during intervention)	n/a	n/a	n/a	n/a
Nasuf 2018[152]	Oropharyngeal colostrum (OPC) compared to control (water, saline or no intervention) in preterm infants	Nil studies eligible (data for weight at discharge only)	n/a	n/a	n/a	n/a
Abiramalatha *et al.* 2017[59]	High-volume vs standard-volume feeds	61 (1 study) 30 vs 31	MD 6.20 g/kg/d [2.71, 9.69]	n/a	Low • Unblinded • Imprecision.	

Weight gain (g/kg/day)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Tan-Dy *et al.* 2013[130]	Lactase treated feeds vs placebo	Nil eligible studies	n/a	n/a	n/a	n/a
Nehra *et al.* 2002[129]	High MCT formula versus low MCT formula	109 (5 studies) 53 vs 56	MD -0.35 [-1.44, 0.74]	0%	*Very Low certainty [Overview]:* • *Serious risk of bias (unclear selection, performance and detection bias)* • *Serious unclear clinical inconsistency between trials* • *Serious imprecision*	
	(31-40% MCT vs low MCT)	62 (3 studies) 31 vs 31	MD 0.44 [-1.01, 1.89]	0%	Not assessed	

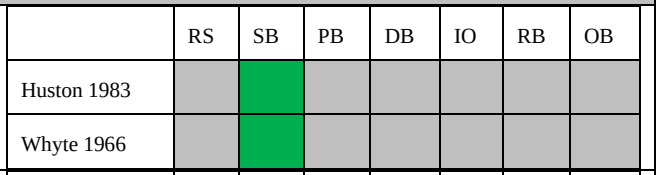
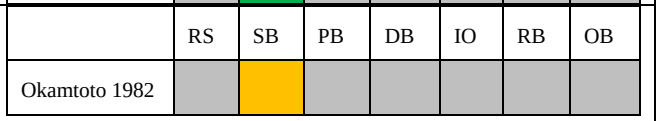
Weight gain (g/kg/day)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Nehra *et al.* 2002[129] (continue)	(41-50% MCT vs low MCT)	50 (2 studies) 25 vs 25	MD -0.86 [-2.45, 0.73]	0%	Not assessed	
	(71-80% MCT vs low MCT)	14 (1 study) 4 vs 10	MD -0.40 [-3.06, 2.26]	n/a	Not assessed	

Table 26: Appendix 6: Summary table of effects of the interventions on weight gain (g/kg/day): 'When to feed'

Outcome assessed: Weight gain (g/kg/day)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Yeo *et al.* 2019[154]	Stopping feeds vs continuing feeds during transfusion	Nil eligible studies	n/a	n/a	n/a	n/a
Abiramalatha *et al.* 2019a[155]	Routine monitoring of gastric residuals vs no routine monitoring of gastric residuals	Nil eligible studies	n/a	n/a	n/a	n/a
	Routine monitoring of gastric residuals quality vs routine monitoring of gastric residuals quality and quantity	Nil eligible studies	n/a	n/a	n/a	n/a
Oddie *et al.* 2017[136]	Slow versus faster rates of feed advancement	Not assessed	n/a	n/a	n/a	n/a

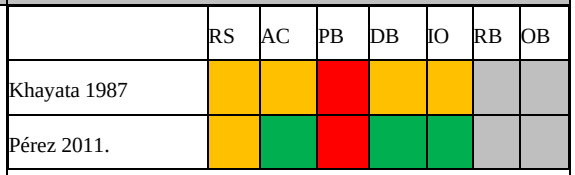
Outcome assessed: Weight gain (g/kg/day)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Morgan *et al.* 2014[137]	Delayed versus early introduction of progressive feeding *Subgroup analysis not possible.*	251 (2 trials) 111 vs 140	"did not detect statistically significant differences in the rate of weight gain". Quantitative data not available	n/a	Not assessed *Overview assessment: Very Low certainty* • *Serious risk of bias (unblinded trials, some unclear methodology)* • *Very serious imprecision (narrative evidence only)*	
Morgan *et al.* 2013[131]	Early trophic feeding vs enteral fasting	Nil eligible studies	n/a	n/a	n/a	n/a

Table 27: Appendix 6: Summary table of effects of the interventions on weight gain (g/kg/day): 'How to feed'

Outcome assessed: Weight gain (g/kg/day)	**Comparison**	**Number of subjects (studies)**	**Measure of effect (95% CI)**	**I^2**	**GRADE assessment**	**Risk of Bias**
Abiramalatha *et al.* 2019b[156]	Re-feeding vs discard gastric residuals	No eligible studies	n/a	n/a	n/a	n/a
Watson *et al.* 2013[135]	Transpyloric versus gastric tube feeding	No eligible studies	n/a	n/a	n/a	n/a
Dawson *et al.* 2012[132]	Push versus gravity bolus tube feeding	Not assessed	n/a	n/a	n/a	n/a

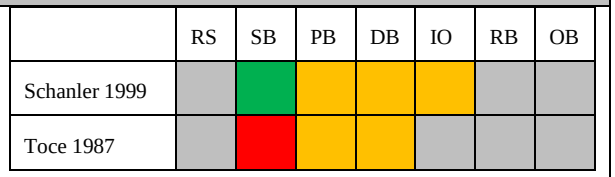
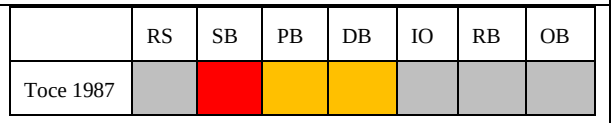
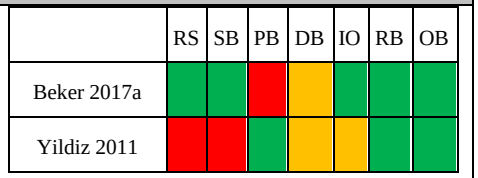
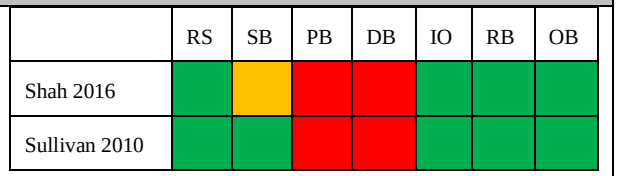
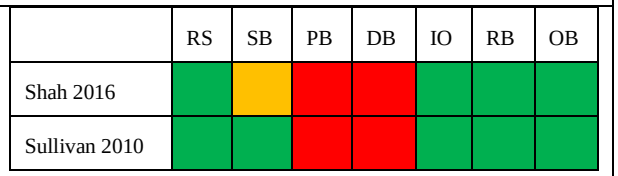
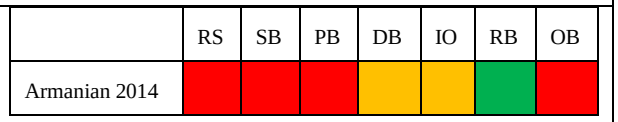
Outcome assessed: Weight gain (g/kg/day)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Premji *et al.* 2011[133]	Continuous vs intermittent bolus milk feeding (Comparison 1)	224 (2 studies) 113 vs 111	MD -1.13 [-2.28, 0.03]	83%	Not assessed *Overview assessment:* *Low certainty* • *Very serious risk from high risk allocation bias, unblinded trials & unclear attrition bias* • *Very serious inconsistency (high heterogeneity)* • *Serious imprecision*	
	Continuous vs NG intermittent bolus (Comparison 2)	53 (1 study) 30 vs 23	MD 1.20 [-1.01, 3.41]	n/a	Not assessed	

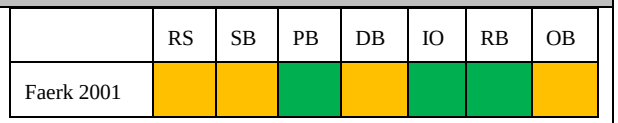
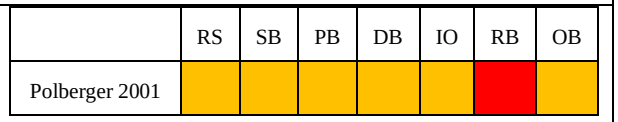
Table 28: Appendix 6: Summary table of effects of the interventions on weight gain (g/kg/day): 'Adjunctive strategies'

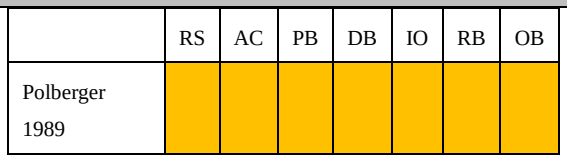
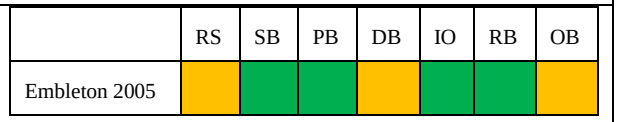
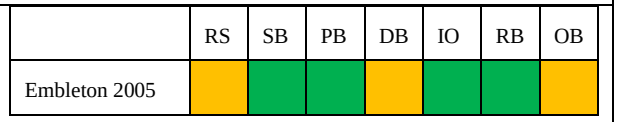
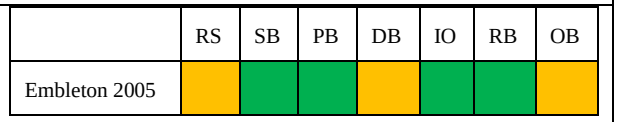
Outcome assessed: Weight gain (g/kg/day)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Muelbert *et al.* 2019[157]	Exposure to smell and taste stimulation of milk with tube feeds versus no exposure	Unable to combine to perform meta-analysis. Mean growth rates estimates using exponential model estimates	Faster mean growth rates in intervention group compared with control: 14.2 g/kg/day versus 12.8 g/kg/day in Beker 2017a and 14.0 g/kg/day versus 7.9 g/kg/day in the study of Yildiz 2011	n/a	*Very low certainty [Overview]:* • *Very serious risk of bias* • *Serious inconsistency* • *Very serious imprecision*	
Anabrees *et al.* 2015[158]	Glycerin prophylaxis versus placebo/no intervention	Not assessed	n/a	n/a	n/a	n/a
Ng *et al.* 2008[134]	Erythromycin vs placebo for prevention of feeding intolerance	Secondary outcome, but nil studies eligible.	n/a	n/a	n/a	n/a

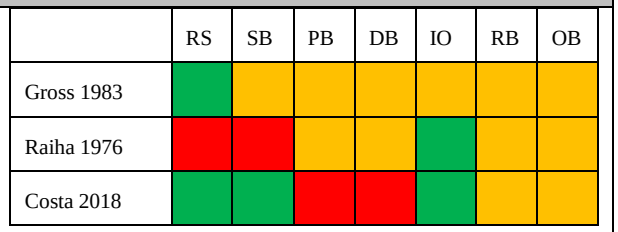
Outcome assessed: Weight gain (g/kg/day)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Greene *et al*. 2016[159]	Oral stimulation for promoting oral feeding	Nil studies eligible	n/a	n/a	n/a	n/a

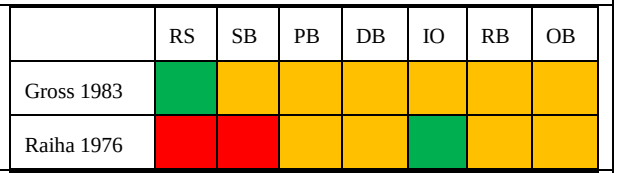
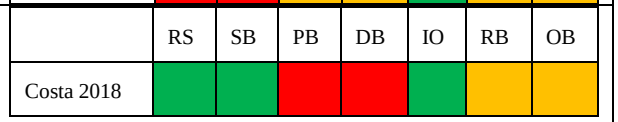
Table 29: Appendix 6: Summary table of effects of the interventions on weight gain (other measures): 'What to feed'

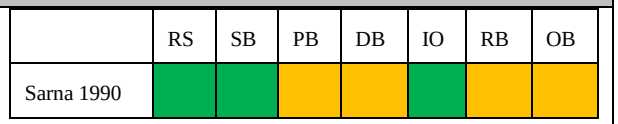
Review Weight gain (other measure)	Outcome assessed:	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Thanigainathan *et al.* 2020[128]	Time to regain birth weight (days)	Early versus late fortification	237 (2 studies) 120 vs 117	MD -0.06 [-1.32, 1.20]	0%	Low certainty: • Lack of blinding • Small sample size	
	Extra-uterine growth restriction at discharge		237 (2 studies) 120 vs 117	RR 1.06 [0.81, 1.39]	0%	Low certainty: • lack of blinding • small sample size	
Amissah *et al.* 2020a[146]	Weight at day 30	Carbohydrate (prebiotic) supplement vs no supplement	75 (1 study)	MD 160.4 [12.4 to 308.4] g	n/a	Very low certainty • Uncertain methodology • Few participants and events • Wide confidence intervals	

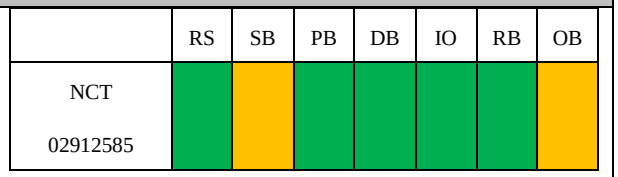
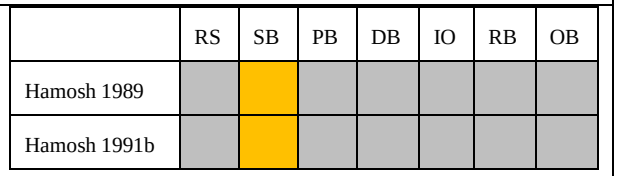
Review Weight gain (other measure)	Outcome assessed:	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Amissah *et al*. 2020b[148]	Weight at term-equivalent age (g)	Protein supplementation versus no supplementation	76 (1 study) 36 vs 40	MD 61.0 [-160.23, 282.23] g	n/a	*Low certainty [Overview]* • *Serious risk of selection and detection bias* • *Serious imprecision*	
	Weight at end of study (grams)	Protein supplementation versus no supplementation	14 (1 study) 7 vs 7	MD 250.00 [-41.56, 541.56]	n/a	*Very Low certainty [Overview]* • *Very serious risk of selection, performance, detection and attrition bias. High risk of reporting bias (but not for this outcome)* • *Serious imprecision*	

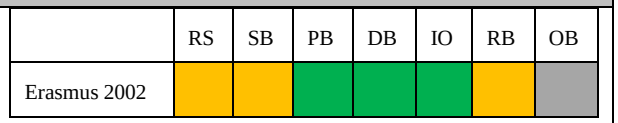
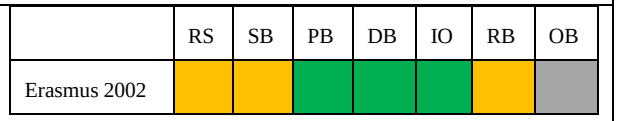
Review Weight gain (other measure)	Outcome assessed:	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Amissah *et al.* 2020c[144]	Weight gain at end of study (g)	Fat supplemented human milk versus control	14 (1 study) 7 vs 7	MD 40.0 [–258.6, 338.6]	n/a	*Low certainty [Overview]* • *Serious risk of bias due to uncertain methodology* • *Serious imprecision*	
Fenton *et al.* 2020[149]	Weight gain at discharge (g/d)	Very high versus high protein intake	77 (1 study) 25 vs 52	MD 3.10 [-0.04, 6.24]	n/a	Moderate certainty. Imprecision	
	Weight gain at term (g/d)	Very high versus high protein intake	74 (1 study) 24 vs 50	MD 2.20 [-1.15, 5.55]	n/a	*Moderate certainty [Overview]* *Serious imprecision*	
	Weight gain at 12 weeks corrected age (g/d)	Very high versus high protein intake	73 (1 study) 24 vs 49	-0.04 [-0.53, 0.45]	n/a	*Moderate certainty [Overview]* *Serious imprecision*	

Review Weight gain (other measure)	Outcome assessed:	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Quigley *et al.* 2019[145]	Time to regain birthweight (days)	Formula milk vs Donor breast milk	236 (3 studies) 139v97	MD -3.08 [-4.38, -1.77] Schultz 1980 "no statistically significant difference" but no SD reported. Lucas 1984a: median time to regain birthweight 10v16, no SDs reported).	37%	*Low certainty [Overview]* • *Very serious risk of bias (including unclear risk of bias from funding sources in 2 studies)*	

Review Weight gain (other measure)	Outcome assessed:	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Quigley *et al.* 2019[145]				Lucas 1984b: median time to regain birthweight 13v15, no SDs reported).			
		Subgroup: Term formula versus unfortified DBM	166 (2 studies) 35v35	MD -4.00 [-5.81, -2.18]	11%		
		Subgroup: preterm formula versus unfortified DBM	70 (1 study) 35 v 35	MD -2.10 [-3.97, -0.23]	n/a		
		Subgroup: Preterm formula versus fortified DBM	Not reported				
	Test for subgroup difference: Chi^2=11.95, df=2 (P=0.00) I^2=83%						
	Subgroup analysis for formula versus donor breast milk as sole diet or supplement to maternal expressed breast milk not reported.						

Review Weight gain (other measure)	Outcome assessed:	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Basuki *et al.* 2019[147]:	Weight gain at 7 days.	Dilute formula vs full strength formula	50 (22 vs 28) 1 study	MD 0.05 kg, 95% CI −0.06 to 0.15	n/a	Not done *Low certainty [Overview]* • *Serious risk of bias from unclear blinding and reporting bias* • *Serious imprecision*	
Premkumar *et al.* 2019[151]	Weight gain (g) during intervention.	Human milk derived fortifier vs bovine milk-derived fortifier	118 (1 study) 61 vs 57	MD -179 [-386.38, 28.38] g	n/a		
Premkumar *et al.* 2019[151]	Weight for age z scores	Human milk derived fortifier vs bovine milk-derived fortifier	118 (1 study) 61 vs 57	MD -0.2 [-0.73 to 0.33]	n/a	Low certainty - Imprecision	

Review Weight gain (other measure)	Outcome assessed:	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Nasuf *et al.* 2018[152]	Weight at discharge to home (?units ?g)	Oropharyngeal colostrum (OPC) versus control (water, saline, or no intervention)	149 (1 study) 81 vs 68	MD 24.5 [-69.66, 118.66] Discrepancy between text and analysis (text reads MD -15.00, 95% CI -50.83 to 20.83)	n/a	Very low quality • Imprecision • Unclear selection • Unclear reporting bias • Single study.	
Nehra *et al.* 2002[129]	Weight gain, g/d	High MCT formula versus low MCT formula	42 (2 studies) 21 vs 21	MD 2.09 [-1.46, 5.64]	0%	*Very Low certainty [Overview]* • *Serious risk of bias from uncertain methodology* • *Serious inconsistency due to unclear measurement unit* • *Serious imprecision*	

Review Weight gain (other measure)	Outcome assessed:	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Tan-Dy *et al.* 2013[130]	Weight gain (g/day) at 7 days after study entry	Lactase treated feeds vs placebo	130 (1 study) 66 vs 64	MD 4.50 [-0.76, 9.76]	n/a	GRADE not performed *Low certainty [Overview]* - *Serious risk of bias from unclear selection and reporting bias* - *Serious imprecision*	
	Weight gain (g/day) at 10 days after study entry	Lactase treated feeds vs placebo	130 (1 study) 66 vs 64	MD 4.90 [0.18, 9.62]	n/a	*Low certainty [Overview]* - *Serious risk of bias from unclear selection and reporting bias* - *Serious imprecision*	

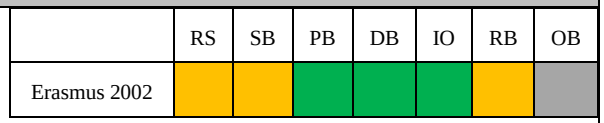
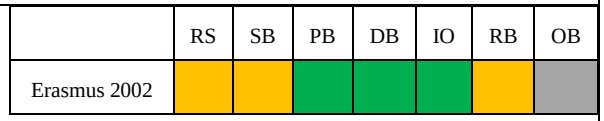
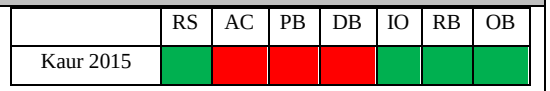
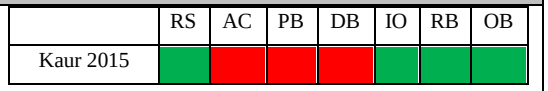
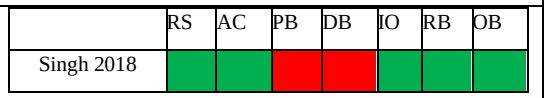
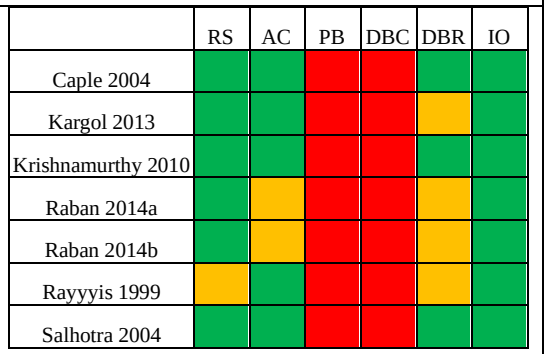
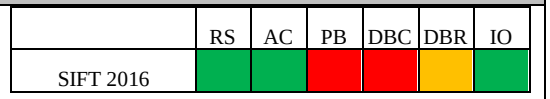
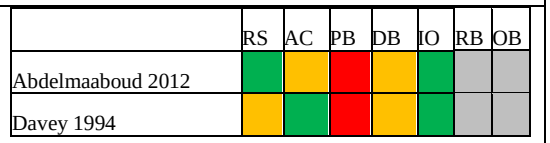
Review Weight gain (other measure)	Outcome assessed:	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Tan-Dy *et al.* 2013[130]	Weight gain (g/day) at 14 days after study entry/study exit if occurred earlier	Lactase treated feeds vs placebo	130 (1 study) 66 vs 64	MD 2.70 [-1.47, 6.87]		*Low certainty [Overview]* - *Serious risk of bias from unclear selection and reporting bias* - *Serious imprecision*	
	Weight gain (g/day) on study exit	Lactase treated feeds vs placebo	130 (1 study) 66 vs 64	MD 2.20 [-0.98, 5.38]		*Low certainty [Overview]* - *Serious risk of bias from unclear selection and reporting bias* - *Serious imprecision*	

Table 30: Appendix 6: Summary table of effects of the interventions on Weight gain (other measures): 'When to feed''

Review	Outcome assessed:	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Yeo *et al.* 2019[154]	Nil additional assessed	n/a	n/a	n/a	n/a	n/a	n/a
Abiramalatha *et al.* 2019a[155]:	Time to regain birth weight (days)	Routine monitoring of gastric residuals vs no routine monitoring of gastric residuals	80 (1 study) 40 vs 40	MD 1.70 [0.01, 3.39]	n/a	Low: Imprecise due to small sample size.	
		Routine monitoring of gastric residuals quality vs routine monitoring of gastric residuals quality and quantity	87 (1 study) 42 vs 45	MD 1.0 [-0.37, 2.37]	n/a	Not assessed *Overview assessment: Very low certainty* • *Very serious risk of selection bias* • *Serious imprecision*	

Review	Outcome assessed:	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Abiramalatha *et al.* 2019a[155]:	Number of infants with extrauterine growth restriction at discharge	Routine monitoring of gastric residuals vs no routine monitoring of gastric residuals	80 (1 study) 40 vs 40	RR 0.89 [0.75, 1.05]	n/a	Low: Imprecise due to small sample size.	
	Number of infants with extrauterine growth restriction at discharge	Routine monitoring of gastric residuals quality vs routine monitoring of gastric residuals quality and quantity	87 (1 study) 42 vs 45	RR 0.54 [0.14, 2.01]	n/a	*Low certainty [Overview]* • *Very serious risk of selection bias*	
Oddie *et al.* 2017[136]	Time to regain birthweight	Slow versus faster rates of feed advancement	7 trials	Infants in the slow-rate-of-advancement group took a longer time to regain birth weight: Median differences 2,5 and 6 days, Mean difference 2 [1,3] to 3.8 [no CI] days. Data from 2 trials not available.	n/a	*Very low certainty [Overview]* • *Serious risk of bias as unblinded* • *Very serious imprecision*	

Review	Outcome assessed:	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Oddie *et al.* 2017[136]	Weight z-score at hospital discharge	Slow versus faster rates of feed advancement	2602 (1 trial) 1295 vs 1307	MD 0.0 [-0.08, 0.08]	n/a	*Moderate certainty [Overview]:* • *Serious risk of bias as unblinded*	
Morgan *et al.*2014[137]:	Time to regain birthweight	Delayed versus early introduction of progressive feeding	2 trials (62 & 125 infants)	1 trial (62 infants): median time 13 days vs 13 days (range not reported)	n/a	*Very low certainty [Overview]:* • *Serious risk of bias as unblinded. Unclear methodology and selection bias.* • *Very serious imprecision*	
Morgan *et al.* 2013[131]	Days to regain birth weight	Early trophic feeding vs enteral fasting	51 (5 studies) 257 vs 261 2 further trials reported	MD -0.01 [-0.96, 0.95]	23%	Not assessed *Overview assessment: Moderate certainty* • *Serious risk of bias*	

Review	Outcome assessed:	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Morgan *et al.* 2013[131]	Days to regain birth weight (continued)		median and range with no statistically significant difference (13v12, 11v10d)				
	weight gain (g/week)	Early trophic feeding vs enteral fasting	2 studies	McClure: MD 130 (95% CI 1 to 250) grams/week. Mosqueda MD -7.3 (95% CI -19.2 to 4.6) grams/week	n/a	*Very Low certainty [Overview]:* • *Very serious risk of bias* • *Serious inconsistency* • *Very serious imprecision*	

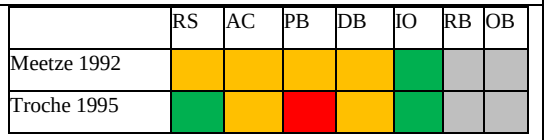
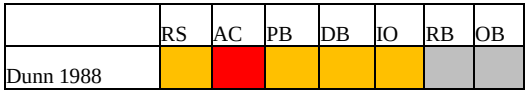
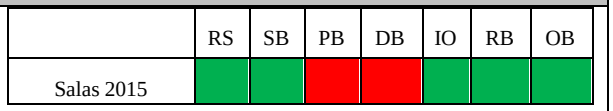
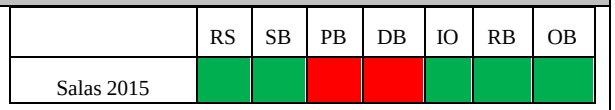
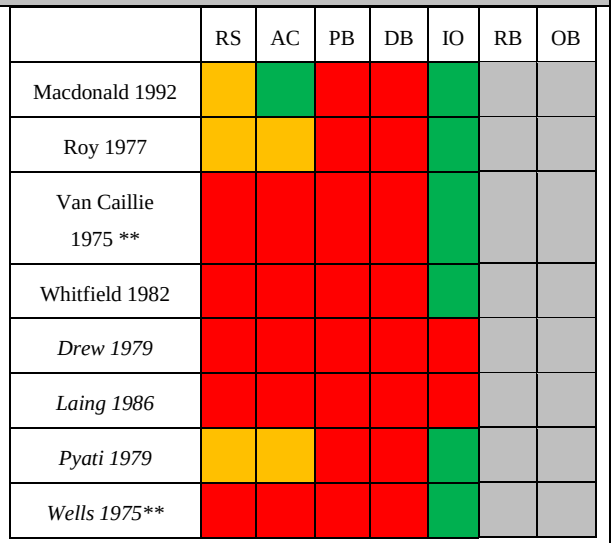
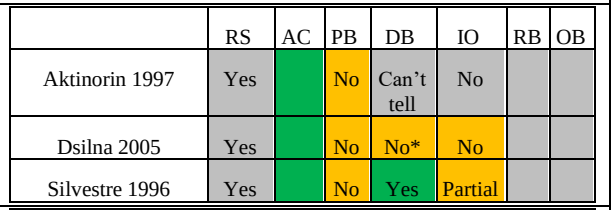
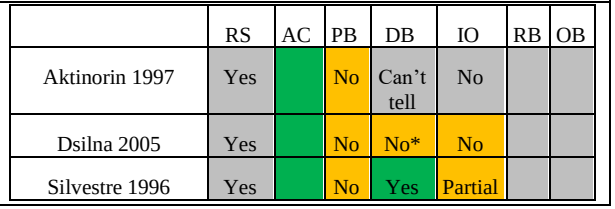
Review	Outcome assessed:	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Morgan *et al.* 2013[131]	Weight gain by day 21	Early trophic feeding vs enteral fasting	1 study	Sáenz de Pipaón 2003: 188 vs190g	n/a	*Very Low certainty [Overview]:* • *Serious risk of bias* • *Very serious imprecision*	
	Weight gain by day 30	Early trophic feeding vs enteral fasting	2 studies	Troche 1995 : weight gain by day 30: (223 (SD 125) versus 95 (SD 161) grams) Meetze 1992: weight gain at day 30: 264 (SD 126) grams versus 213 (SD 142) grams		*Very Low certainty [Overview]:* • *Serious risk of bias* • *Very serious inconsistency* • *Very serious imprecision*	
	Growth until 60 days of life	Early trophic feeding vs enterafasting	1 study	Dun 1988 "did not detect any significant differences between the two groups"	n/a	*Very Low certainty [Overview]:* • *Very serious risk of bias* • *Very serious imprecision (narrative)*	

Table 31: Appendix 6: Summary table of effects of the interventions on Weight gain (other measures): 'How to feed

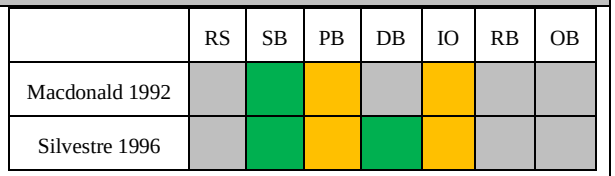
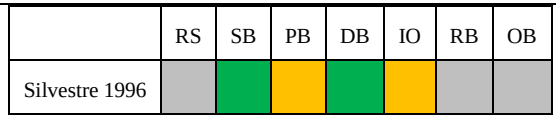
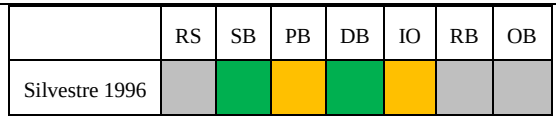
Outcome assessed: Weight gain (other measure)	Outcome measured	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Abiramalatha *et al.* 2019b[156]:	Time to regain birthweight	Re-feeding vs discard gastric residuals	59 (1 study) 29 vs 30	MD 0.40 [-2.89, 3.69] Subgroup analyses not statistically significant.	n/a	Very Low • Concern due to attrition • Serious Imprecision due to wide confidence interval.	
		Subgroup: Infants fed only human milk	41 (1 study) 20 vs 21	MD -1.80 [-5.51, 1.91]	n/a	Not assessed	
		Subgroup: Infants fed formula milk	18 (1 study) 9 vs 9	MD 3.0 [-4.03, 10.03]	n/a	Not assessed	
		Test for subgroup differences: Chi^2=1.64 df=1 (0.44) I^2=0% p-0.44					

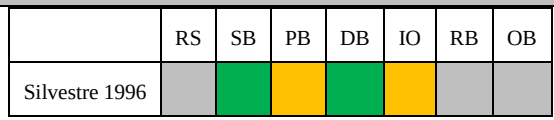
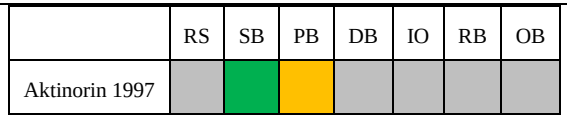
Outcome assessed: Weight gain (other measure)	Outcome measured	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Abiramalatha *et al.* 2019b[156]:	Number of infants with weight < 10th percentile at discharge	Re-feeding vs discard gastric residuals	59 (1 study) 29 vs 30	RR 1.29 [0.38, 4.34]	n/a	*Low certainty [Overview]:* • *Serious risk of bias* • *Serious imprecision*	
		Subgroup: Infants fed only human milk	41 (1 study) 20 vs 21	RR 1.58 [0.29, 8.46]	n/a	Not assessed	
		Subgroup: Infants fed formula milk	18 (1 study) 9 vs 9	RR 1.0 [0.18, 5.63]	n/a	Not assessed	
		Test for subgroup differences: Chi^2=0.14, df=1 (P=0.93) I^2=0%					

Outcome assessed: Weight gain (other measure)	Outcome measured	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Watson *et al.* 2013[135]	Change in weight (g/week) Change in weight (g/week) (continued)	Transpyloric versus gastric tube feeding	93 (4 studies) 45 vs 48	MD -5.50 [-26.88, 15.89] 7 trials* reported no statistically significant differences. 2 trials** found statistically significant higher rates of weight gain in transpyloric fed group.	0%	*Low certainty [Overview]* • *Very serious risk of bias (unblinded, 2 trials quasi-randomised)*	
	Weight (kg) measured at Expected Date of Delivery	Transpyloric versus gastric tube feeding	Unclear:assumed 36 (1 study) 21 vs 15	MD -0.3 [-0.6, -0.03]	n/a	*Low certainty [Overview]* • *Very serious risk of bias (quasi-randomised)*	

Outcome assessed: Weight gain (other measure)	Outcome measured	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Dawson *et al.* 2012[132]	Days to regain birthweight	Push versus gravity bolus tube feeding	No eligible studies	n/a	n/a	n/a	n/a
Premji *et al.* 2011[133]	Days to regain birthweight	Continuous vs intermittent bolus milk feeding (all infants, NG and OG)	401 (4 studies) 186 vs 215	MD -0.46 [-1.48, 0.55]	0%	*Low certainty [Overview]:* • *Serious risk of bias* • *Serious imprecision*	
		Continuous vs intermittent bolus milk feeding (all infants, NG only)	206 (3 studies) 103 vs 103	MD -0.31 [-1.65, 1.03]	0%		
		Continuous vs intermittent bolus milk feeding in infants	120 (3 studies) 46 vs 74	MD -0.13 [-2.11, 1.84]	0%		

<table>
<tr><th>Outcome assessed: Weight gain (other measure)</th><th>Outcome measured</th><th>Comparison</th><th>Number of subjects (studies)</th><th>Measure of effect (95% CI)</th><th>I^2</th><th>GRADE assessment</th><th>Risk of Bias</th></tr>
<tr><td rowspan="4">Premji et al. 2011[133]</td><td rowspan="4"></td><td><1000g (NG or OG)</td><td></td><td></td><td></td><td></td><td></td></tr>
<tr><td>Continuous vs intermittent bolus (NG) milk feeding in infants <1000g</td><td>70 (2 studies) 29 vs 41</td><td>MD -0.46 [-3.51, 2.60]</td><td>0%</td><td></td><td><table><tr><th></th><th>RS</th><th>AC</th><th>PB</th><th>DB</th><th>IO</th><th>RB</th><th>OB</th></tr><tr><td>Aktinorin 1997</td><td>Yes</td><td></td><td>No</td><td>Can’t tell</td><td>No</td><td></td><td></td></tr><tr><td>Silvestre 1996</td><td>Yes</td><td></td><td>No</td><td>Yes</td><td>Partial</td><td></td><td></td></tr></table></td></tr>
<tr><td>Continuous vs intermittent bolus milk feeding (NG) in infants >1000g and <1249g</td><td>71 (2 studies) 39 vs 32</td><td>MD -0.40 [-2.45, 1.66]</td><td>66%</td><td></td><td><table><tr><th></th><th>RS</th><th>AC</th><th>PB</th><th>DB</th><th>IO</th><th>RB</th><th>OB</th></tr><tr><td>Aktinorin 1997</td><td>Yes</td><td></td><td>No</td><td>Can’t tell</td><td>No</td><td></td><td></td></tr><tr><td>Silvestre 1996</td><td>Yes</td><td></td><td>No</td><td>Yes</td><td>Partial</td><td></td><td></td></tr></table></td></tr>
<tr><td>Continuous vs intermittent (NG) bolus milk feeding in infants >1250g and <1499g</td><td>32 (1 study) 16 vs 16</td><td>MD 0.0 [-3.53, 3.53]</td><td>n/a</td><td></td><td><table><tr><th></th><th>RS</th><th>AC</th><th>PB</th><th>DB</th><th>IO</th><th>RB</th><th>OB</th></tr><tr><td>Silvestre 1996</td><td>Yes</td><td></td><td>No</td><td>Yes</td><td>Partial</td><td></td><td></td></tr></table></td></tr>
</table>

Outcome assessed: Weight gain (other measure)	Outcome measured	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Premji *et al.* 2011[133]	Weight gain (g/week)	Continuous vs intermittent bolus milk feeding (all infants, NG and OG) & (NG only)	106 (2 studies) 54 vs 52	MD 6.27 [-1.28, 13.81]	0%	*Low certainty [Overview]:* • *Serious risk of bias* • *Serious imprecision*	
	Weight gain (g/day)	Continuous vs intermittent bolus milk feeding in infants <1000g	30 (1 study) 12 vs 18	MD 2.0 [0.54, 3.46]	n/a	*Overview assessment:* *Low certainty:* • *Serious risk of bias* • *Serious imprecision*	
	Weight gain (g/day)	Continuous vs intermittent bolus (NG) milk feeding in infants <1000g	30 (1 study) 12 vs 18	MD 2.0 [0.54, 3.46]	n/a		

Outcome assessed: Weight gain (other measure)	Outcome measured	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Premji *et al.* 2011[133]	Weight gain (g/day)	Continuous vs intermittent bolus milk feeding (NG) in infants >1000g and <1249g	31 (1 study) 17 vs 14	MD 2.0 [0.16, 3.84]	n/a		
	Weight gain (g/day)	Continuous vs intermittent (NG) bolus milk feeding in infants >1250g and <1499g	32 (1 study) 16 vs 16	MD 0.0 [-1.77, 1.77]	n/a		
	Days to 2040g	Continuous vs intermittent bolus milk feeding (NG or OG)	80 (1 study) 39 vs 41	MD -2.0 [-7.92, 3.92]	n/a	*Overview assessment:* *Very Low certainty:* • *Very serious risk of bias* • *Very serious imprecision*	

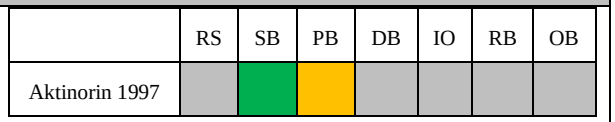
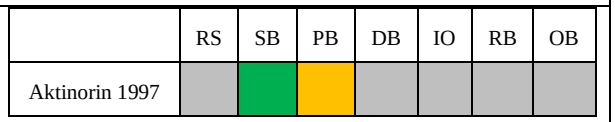
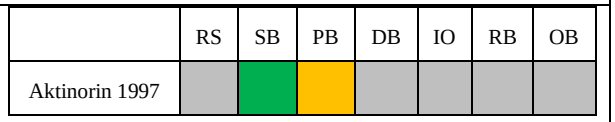
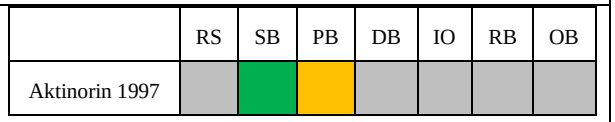
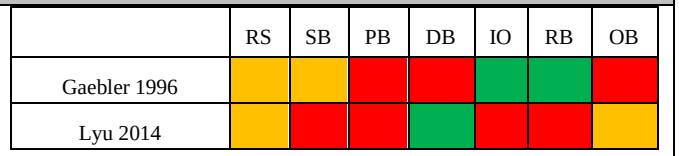
Outcome assessed: Weight gain (other measure)	Outcome measured	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Premji *et al.* 2011[133]	Days to 2040g	Continuous vs NG intermittent bolus milk feeding	80 (1 study) 39 vs 41	MD -2.0 [-7.92, 3.92]	n/a		
	Days to 2040g	Continuous vs intermittent bolus milk feeding in infants <1000g	40 (1 study) 17 vs 23	MD 0.0 [-5.85, 5.85]	n/a		
	Days to 2040g	Continuous vs NG intermittent bolus milk feeding in infants <1000g	40 (1 study) 17 vs 23	MD 0.0 [-5.85, 5.85]	n/a		
	Days to 2040g	Continuous vs NG intermittent bolus milk feeding in infants >1000g and <1249g	40 (1 study) 22 vs 18	MD 1.00 [-3.01, 5.01]	n/a		

Table 32:Appendix 6: Summary table of effects of the interventions on Weight gain (other measures): 'Adjunctive strategies''

Outcome assessed: Weight gain (other measure)	Outcome reported	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Muelbert *et al.* 2019[157]	Nil further weight outcomes	Exposure to smell and taste stimulation of milk with tube feeds versus no exposure	n/a	n/a	n/a	n/a	n/a
Greene *et al.* 2016[159]	Weight gain (g)	Oral stimulation vs standard care for promoting oral feeding	81 (2 studies) 41 vs 40	MD 0.73 [-1.05, 2.51] 2 further studies (reported weight change from 4 oral feeds/day-8 oral feeds/day, and % weight change) outcome not reported.	41%	Low quality: • High risk of selection bias, performance bias, attrition bias, reporting bias • Moderate heterogeneity (30-60%)	

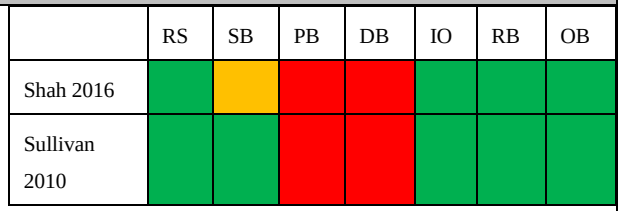
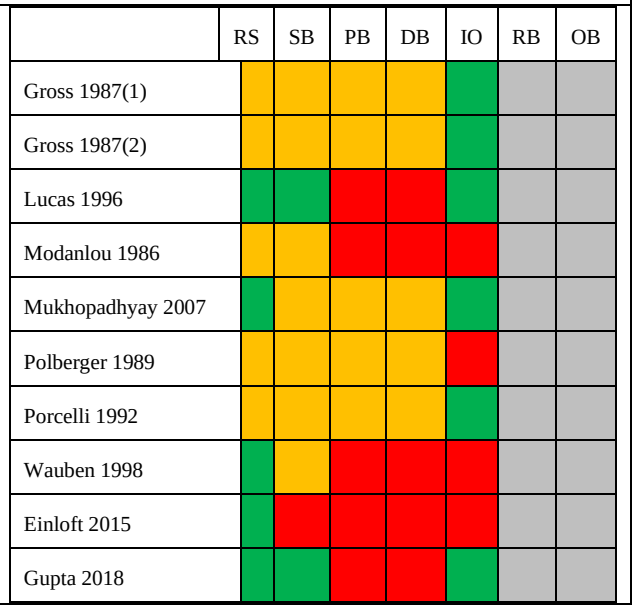
Outcome assessed: Weight gain (other measure)	Outcome reported	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Greene *et al.* 2016[159]	Weight gain (g)	Oral stimulation vs non-oral intervention for promoting oral feeding	n/a	Nil eligible. 3 trials reported "weight changes, "weight at discharge" and "weight at end of intervention" but data not reported in review.	n/a	n/a	n/a
Anabrees 2015 *et al.*[158]	Weight at discharge home (g).	Glycerin prophylaxis versus placebo/no intervention	81 (1 study) 39 vs 42	MD -62.0 [-317.49, 193.49]	n/a	*Low certainty [Overview]* • *Serious risk of bias* • *Serious imprecision*	

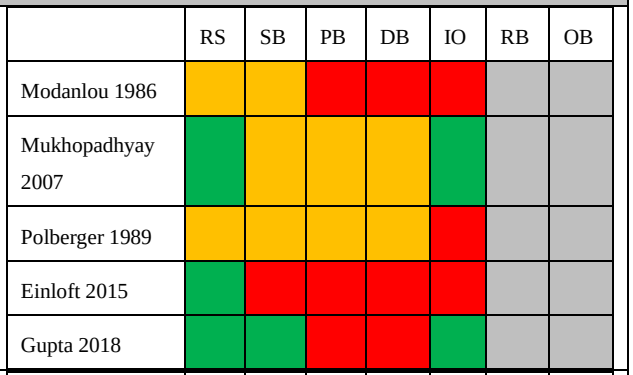
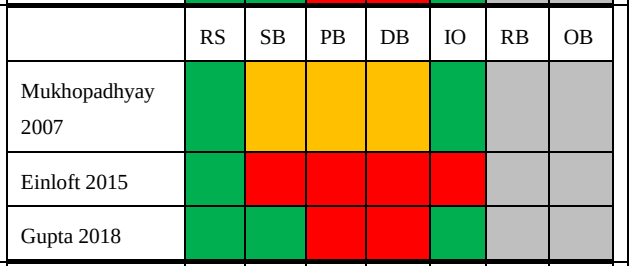
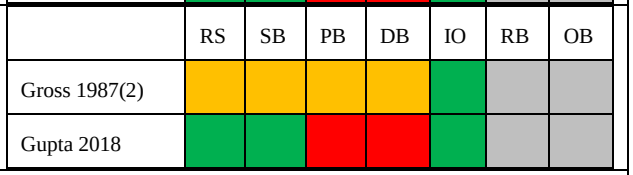
Outcome assessed: Weight gain (other measure)	Outcome reported	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Ng *et al.* 2008[134]	Time to regain birthweight (days)	Erythromycin vs placebo for prevention of feeding intolerance	1 study Oei 2001 (Total enrolled 22v21)	Oei 2011: Mean (SD) 14.9 (2.6) vs.15.3 (6.6) p = 0.83	n/a	*Low certainty [Overview]* • *Very serious imprecision*	Oei 2011: Blinding of Randomisation: Yes; Blinding of intervention: Yes; Complete follow up: Yes; Blinding of outcome measures: Cannot tell
		Erythromycin vs placebo for treatment of feeding intolerance (high dose)	2 studies Ng SC 2003 (Total enrolled 13v11) Nuntnaraumit 2006 (Total enrolled 23 v 23)	Ng SC 2003: Mean (SD) 12.8 (4.4) vs. 16.8 (6.2) p = 0.11 Nuntnarumit 2006: Median (IQR) 11 (10 to 14) vs. 12 (11 to 15) days p = 0.49]	n/a	*Very Low certainty [Overview]:* • *Very serious inconsistency* • *Very serious imprecision*	Ng SC 2003: Blinding of Randomisation: Yes; Blinding of intervention: Yes; Complete follow up: Yes; Blinding of outcome measures: Cannot tell Nuntnarumit 2006: Blinding of Randomisation: Yes; Blinding of intervention: Yes; Complete follow up: Yes; Blinding of outcome measures: Yes

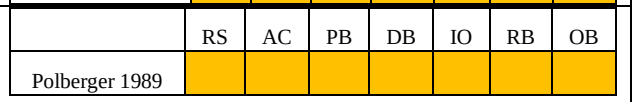
Outcome assessed: Weight gain (other measure)	Outcome reported	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Ng *et al.* 2008[134]	Weight gain (g) from birth to discharge	Erythromycin vs placebo for prevention of feeding intolerance	1 study (Total enrolled 36v37)	Median (IQR): 425 (162.5 to 1190) vs. 715 (450 to 1117) p = 0.15	n/a	*Low certainty [Overview]:* • *Very serious imprecision*	Patole 2000: Blinding of Randomisation: Yes; Blinding of intervention: Yes; Complete follow up: Yes; Blinding of outcome measures: Yes
	Weight at discharge (g)	Erythromycin vs placebo for treatment of feeding intolerance (high dose)	1 study (Total enrolled 23v23)	median (IQR) of 2170 (1987 to 2587) vs. 2560 (2130 to 3600) grams, p = 0.06	n/a	*Low certainty [overview]:* • *Serious risk of bias* • *Serious imprecision*	Nuntnarumit 2006: Blinding of Randomisation: Yes; Blinding of intervention: Yes; Complete follow up: Yes; Blinding of outcome measures: Yes
	Weight at study day 8	Erythromycin vs placebo for treatment of feeding intolerance (low dose).	1 study Total enrolled:15v12	Mean (SD) of 1625 (430) vs. 1611 (476) grams, p > 0.05	n/a	*Very low certainty [Overview]:* • *Serious risk of bias* • *Very serious imprecision*	El Henawy 2003: Blinding of Randomisation: Yes; Blinding of intervention: Yes; Complete follow up: Yes; Blinding of outcome measures: Yes

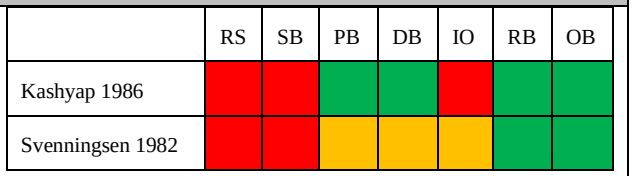
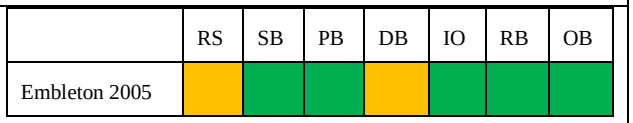
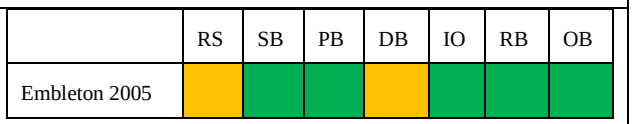
<table>
<tr><th>Outcome assessed: Weight gain (other measure)</th><th>Outcome reported</th><th>Comparison</th><th>Number of subjects (studies)</th><th>Measure of effect (95% CI)</th><th>I^2</th><th>GRADE assessment</th><th>Risk of Bias</th></tr>
<tr><td>Ng et al. 2008[134]</td><td>weight gain since enrollment</td><td>Erythromycin vs placebo for treatment of feeding intolerance (low dose).</td><td>1 study:
Total enrolled 30v30</td><td>“no difference in either gestational age strata”</td><td>n/a</td><td>Low certainty [Overview]:
• Very serious imprecision</td><td>
<table>
<tr><th></th><th>Blinding of Randomisation</th><th>Blinding of intervention</th><th>Complete follow up</th><th>Blinding of outcome measures</th></tr>
<tr><td>Aly 2007</td><td>Yes</td><td>Yes</td><td>Yes</td><td>Yes</td></tr>
</table>
</td></tr>
</table>

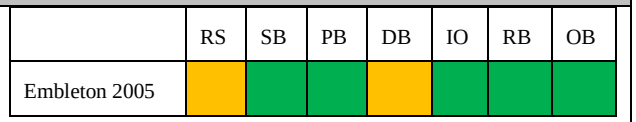
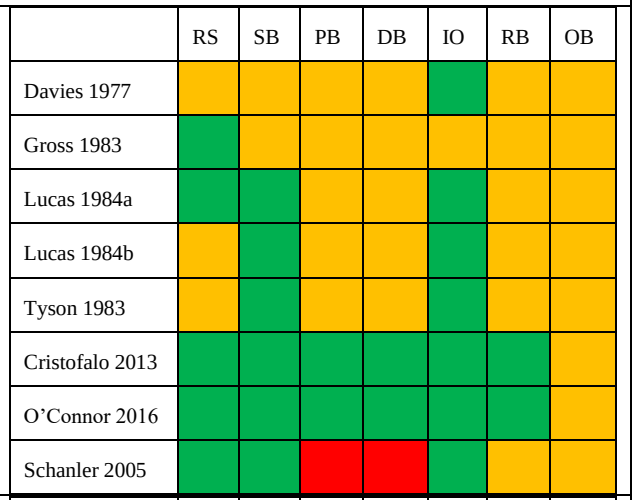
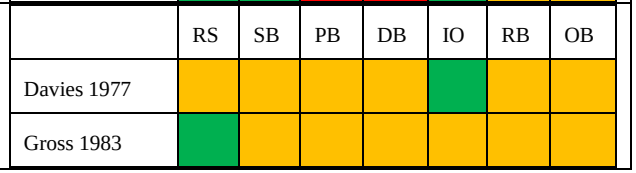
Table 33: Appendix 6: Summary table of effects of the interventions on Linear growth (crown-heel length mm or cm/week): 'What to feed'

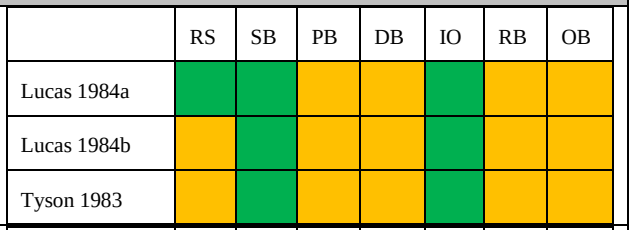
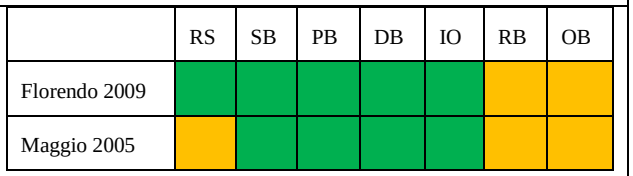
Outcome assessed: Linear growth (crown-heel length cm or mm/week)	**Comparison**	**Number of subjects (studies)**	**Measure of effect (95% CI)**	**I^2**	**GRADE assessment**	**Risk of Bias**
Thanigainathan *et al.* 2020[128] (cm/week)	Early versus late fortification *Subgroup analyses not possible.*	237 (2 studies) 120 vs 117	MD 0.10 [-0.03, 0.22]	0%	Low certainty: • Lack of blinding • Small sample size	
Brown *et al.* 2020[143]	fortified vs unfortified breast milk	741 (10 studies) 377 vs 364	0.11 [0.08, 0.15] cm/week	69%	Low certainty • Unexplained heterogeneity • Uncertain random sequence generation, conceal allocation and blind assessments in most trials	

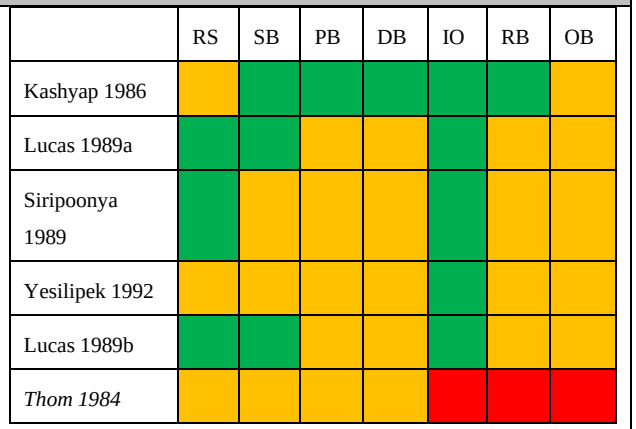
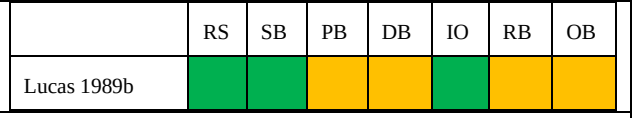
Outcome assessed: Linear growth (crown-heel length cm or mm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Brown *et al.* 2020[143]	Subgroup: trials recruiting only very preterm or VLBW infants	375 (5 studies) 191 vs 184	MD 0.16 [0.11, 0.20]	72%		
	Subgroup: trials conducted in low- or middle-income countries	343 (3 studies) 176 vs 167	MD 0.14 [0.09, 0.19]	75%		
	Subgroup: trials using preterm formula powder as fortifier	174 (2 studies) 92 vs 82	MD 0.07 [0.01, 0.14]	0%		
	Test for subgroup differences: Chi^2=4.41, df = 3 (P=0.22) I^2=32%					

Outcome assessed: Linear growth (crown-heel length cm or mm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Amissah *et al.* 2020[146]	Carbohydrate (prebiotic) supplement vs no supplement	Nil studies eligible	n/a	n/a	n/a	n/a
Amissah *et al.* 2020b[148] (cm/week)	Protein supplementation versus no supplementation *Subgroup analyses not possible.*	68 (4 studies) 35 vs 33	MD 0.12 [0.07, 0.17] cm/week	89%	Low certainty • Uncertain methodology • Moderate-to-high heterogeneity	
Amissah *et al.* 2020c[144]: (cm/week)	Fat supplemented human milk versus control	14 (1 study) 7 vs 7	MD 0.1 [–0.08, 0.3] cm/week	n/a	Very low certainty • Uncertain risk of bias (insufficient methodological detail). • Imprecision (very small sample size) • Wide confidence intervals spanning across benefits and harms.	

Outcome assessed: Linear growth (crown-heel length cm or mm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Fenton *et al.* 2020[149]	High versus low protein intake	48 (2 studies)25 vs 23	MD 0.16 [-0.02, 0.34]	42%	Low certainty • Serious risk of bias (selection, performance, detection attrition) Serious imprecision (Broad 95% CI, included both higher and lower length gain in estimate).	
	Very high versus high protein intake	77 (1 study) 25 vs 52	MD 0.0 [-0.14, 0.14]	n/a	Moderate certainty Imprecision	
Fenton *et al.* 2020[149]: Linear growth at term (cm/wk)	Very high versus high protein intake	74 (1 study) 24 vs 50	MD 0.10 [0.00, 0.20]	n/a	*High certainty [Overview]*	

Outcome assessed: Linear growth (crown-heel length cm or mm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Fenton *et al.* 2020[149] Linear growth at 12 weeks corrected age (cm/week)	Very high versus high protein intake	73 (1 study) 24 vs 49	0.0 [-0.49, 0.49]	n/a	*High certainty [Overview]*	
Quigley *et al.* 2019[145]	Formula milk vs Donor breast milk	820 (8 studies) 402 vs 418	MD 1.21 [0.77, 1.65] Plus 1 study did not detect a between-group difference in average length at 15 days after birth or at 36 weeks' post-menstrual age	68%	Moderate-certainty evidence: • High heterogeneity	
	Subgroup: Term formula versus unfortified DBM	128 (2 studies) 54 vs 74	MD 0.80 [0.10, 1.50]	0%		

Outcome assessed: Linear growth (crown-heel length cm or mm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Quigley *et al.* 2019[145] (continued)	Subgroup: Preterm formula versus unfortified DBM	147 (3 studies) 74 vs 73	MD 1.96 [1.10, 2.82]	65%		
	Subgroup: Preterm formula versus fortified DBM	545 (3 studies) 274 vs 271	MD 1.10 [0.33, 1.87]	83%		
	Test for subgroup differences: $Chi^{2=}$ 4.35, df=2 (P=0.11), I^2=54%					
Ng *et al.* 2019[150] (mm/week)	Hydrolysed versus non-hydrolysed formula,	97 (2 studies) 50 vs 47	MD -0.04 [-1.24, 1.15]	0%	*Overview assessment:* *Very Low certainty:* • *Serious risk of bias (including funding source)* • *Serious imprecision*	

Outcome assessed: Linear growth (crown-heel length cm or mm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Walsh *et al.* 2019[127] (mm/week)	Nutrient-enriched formula versus standard formula	386 (5 studies) 197 vs 189	MD 0.22 [-0.70, 1.13] + 1 trial reported "no difference" in average daily rate of length gain	67%	Low certainty: • Uncertain random sequence generation • Uncertain allocation concealment • Uncertain blinding. • Moderate to high heterogeneity.	
	Subgroup: sole diet	184 (4 studies) 94 vs 91	MD 1.72 [0.23, 3.20]	47%		
	Subgroup: supplement to human milk	201 (1 study) 103 vs 98	MD -0.70 [-1.86, 0.46]	67%		
	Test for subgroup differences: Chi^2=6.31, df = 1 (P=0.02) I^2=84.1%					

Outcome assessed: Linear growth (crown-heel length cm or mm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Dempsey *et al.* 2019[142]	Banked preterm versus banked term human milk	Nil studies eligible	n/a	n/a	n/a	n/a
Brown *et al.* 2019[126]	Formula versus maternal breastmilk	Nil studies eligible	n/a	n/a	n/a	n/a
Basuki *et al.* 2019[147]	Dilute versus full-strength formula	Nil studies eligible	n/a	n/a	n/a	n/a
Premkumar *et al.* 2019[151]	Human milk-derived fortifier versus bovine milk-derived fortifier	Nil studies eligible (only absolute change in head circumference during intervention)	n/a	n/a	n/a	n/a
Nasuf *et al.* 2018[152]	Oropharyngeal colostrum (OPC) compared to control (water, saline or no intervention) in preterm infants	Not assessed	n/a	n/a	n/a	n/a

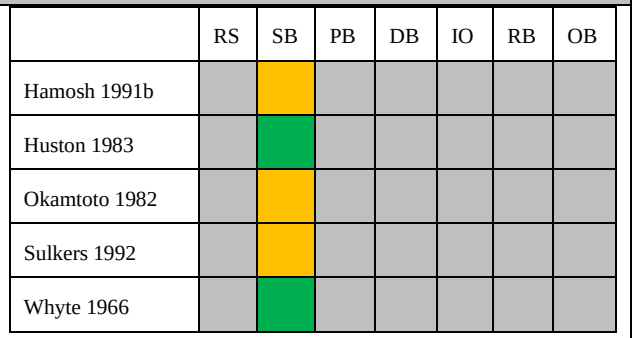
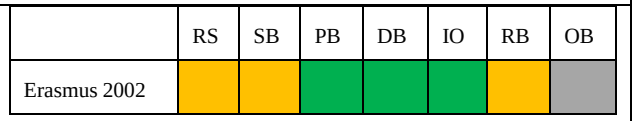
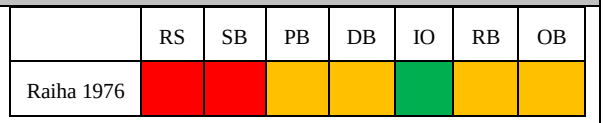
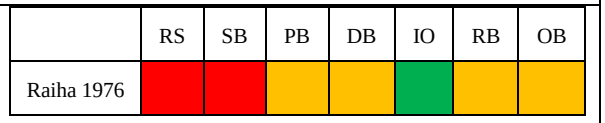
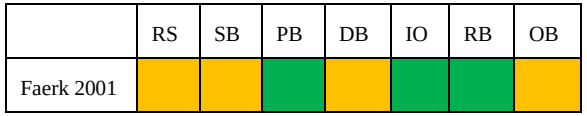
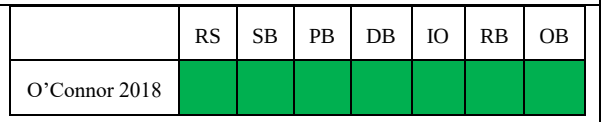
Outcome assessed: Linear growth (crown-heel length cm or mm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Abiramalatha *et al.* 2017[59]	High-volume vs standard-volume feeds	Nil eligible studies for this outcome	n/a	n/a	n/a	n/a
Nehra *et al.* 2002[129] Length gain, cm/wk	High MCT formula versus low MCT formula	109 (5 studies) 53 vs 56	MD 0.14 [-0.04, 0.31]	50%	*Very Low certainty [Overview]:* • *Serious risk of bias* • *Serious inconsistency* • *Serious imprecision*	
Tan-Dy *et al.* 2013[130] Length gain (cm/week) at 14 days after study entry/study exit if occurred earlier	Lactase treated feeds vs placebo	130 (1 study) 66 vs 64	MD 0.30 [-0.13, 0.73]	n/a	*Low certainty [Overview]* • *Serious risk of bias from unclear selection and reporting bias* • *Serious imprecision*	

Table 34:Appendix 6: Summary table of effects of the interventions on Linear growth (other measures): 'What to feed'

Outcome assessed: Length (other measures)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Quigley *et al.* 2019[145] Crown-rump length (mm/week)	Formula (term or preterm) versus donated breast milk (DBM)	106 (1 study) 84 vs 22	MD 0.59 [0.08, 1.10]	n/a	*Very low certainty [Overview]:* • *Very serious risk of bias (including funding source)* • *Serious imprecision*	
Quigley *et al.* 2019[145] Femoral length (mm/week)	Formula (term or preterm) versus donor breast milk (DBM)	106 (1 study) 84 vs 22	MD 0.34 [0.13, 0.55]	n/a	*Very low certainty [Overview]:* • *Very serious risk of bias (including funding source)* • *Serious imprecision*	
Amissah *et al.* 2020b[148]: Length at term-equivalent age (cm)	Protein supplementation versus no supplementation	76 infants (1 study) 36 vs 40	MD -0.5 [-1.65, 0.65]	n/a	*Low certainty [Overview]:* • *Serious risk of bias* • *Serious imprecision*	
Premkumar *et al.* 2019[151] : Change in length during the intervention (cm)	Human milk derived fortifier vs bovine milk-derived fortifier	118 (1 study) 61 vs 57	MD -0.10 [-2.25, 0.65]	n/a	Low certainty • wide confidence intervals • single study	

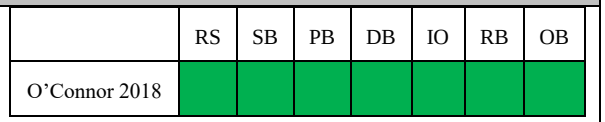
Outcome assessed: Length (other measures)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Premkumar *et al.* 2019[151] : Length-for-age z score		118 (1 study) 61 vs 57	MD 0.10 [-0.57, 0.77]	n/a	*Moderate certainty [Overview]:* • *Serious imprecision*	

Table 35: Appendix 6: Summary table of effects of the interventions on Linear growth (crown-heel length mm or cm/week): 'When to feed'

Outcome assessed: Linear growth (crown-heel length mm/week)	**Comparison**	**Number of subjects (studies)**	**Measure of effect (95% CI)**	**I^2**	**GRADE assessment**	**Risk of Bias**
Oddie *et al* 2017[136]	Slow versus faster rates of feed advancement	Nil eligible studies.	n/a	n/a	n/a	n/a
Yeo *et al* 2019[154]	Stopping feeds vs continuing feeds during transfusion	Not assessed	n/a	n/a	n/a	n/a
Morgan *et al* 2014[137]	Delayed versus early introduction of progressive feeding	Nil eligible studies	n/a	n/a	n/a	n/a
Abiramalatha *et al* 2019a[155]	Routine monitoring of gastric residuals vs no routine monitoring of gastric residuals	Nil eligible studies	n/a	n/a	n/a	n/a
	Routine monitoring of gastric residuals quality vs routine monitoring of gastric residuals quality and quantity					

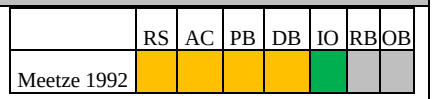
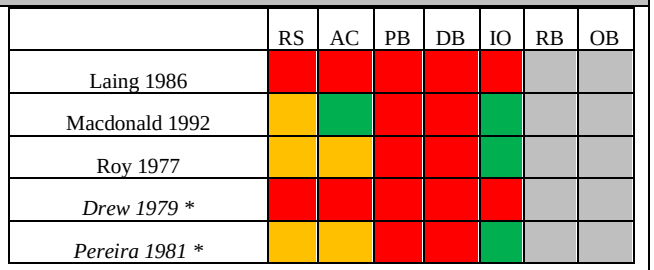
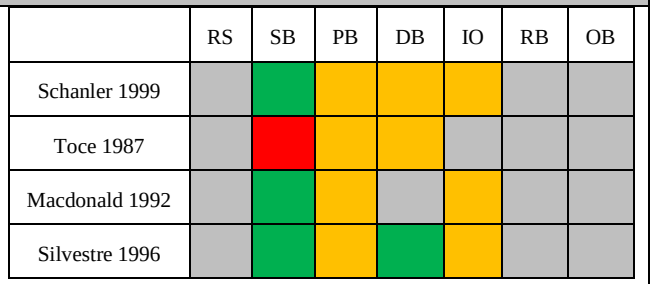
Outcome assessed: Linear growth (crown-heel length mm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Morgan 2013[131]	Early trophic feeding vs enteral fasting	1 study	Maetze 1992: "similar for both groups	n/a	*Very low certainty [Overview]:* • *Serious risk of bias* • *Very serious imprecision (narrative)*	

Table 36: Appendix 6: Summary table of effects of the interventions on Linear growth (crown-heel length mm or cm/week): 'How to feed'

Outcome assessed: Linear growth (crown-heel length mm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Abiramalatha *et al.* 2019b[156]	Re-feeding vs discard gastric residuals	No eligible studies	n/a	n/a	n/a	n/a
Watson *et al.* 2013[135] (mm/week)	Transpyloric versus gastric tube feeding	93 (3 studies) 35 vs 58	MD -0.67 [-2.36, 1.02] **2 further studies did not find a statistically significant difference.*	0%	*Low certainty [Overview]:* *Very serious risk of bias (includes quasi-randomsied trial and statistically significant baseline differences)*	
Dawson *et al.* 2012[132]	Push versus gravity bolus tube feeding	Not assessed	n/a	n/a	n/a	n/a
Premji *et al.* 2011[133] (cm/week)	Continuous vs intermittent bolus milk feeding	330 (4 studies) 167 vs 163	MD 0.08 [-0.01, 0.17]	0%	*Low certainty [Overview]:* • *Very serious risk of bias (including quasi-randomised trials)*	

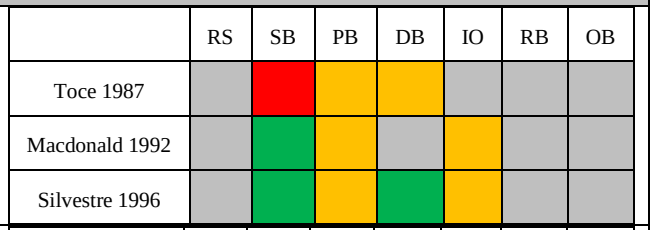
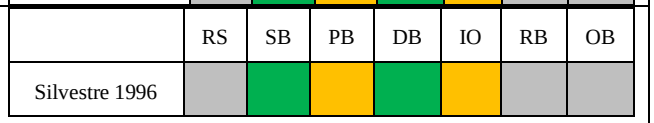
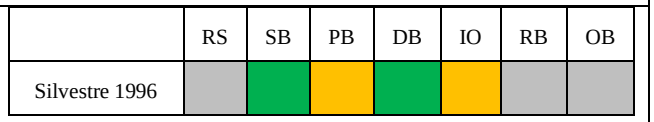
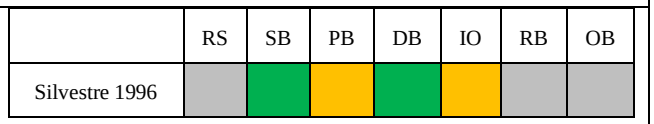
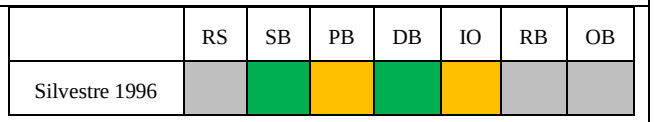
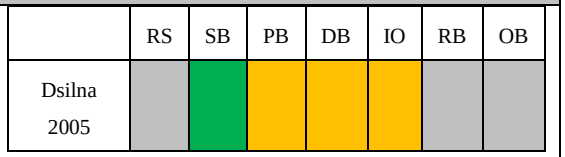
Outcome assessed: Linear growth (crown-heel length mm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Premji *et al.* 2011[133] (cm/week)	Continuous vs NG intermittent bolus milk feeding	159 (3 studies) 84 vs 75	MD 0.07 [-0.04, 0.18]	0%	No GRADE assessment performed	
	Continuous vs intermittent bolus milk feeding in infants <1000g	30 (1 study) 12 vs 18	MD 0.07 [-0.08, 0.22]	n/a	No GRADE assessment performed	
	Continuous vs NG intermittent bolus milk feeding in infants <1000g	30 (1 study) 12 vs 18	MD 0.07 [-0.08, 0.22]	n/a	No GRADE assessment performed	
	Continuous vs NG intermittent bolus milk feeding in infants >1000g and <1249g	31 (1 study) 17 vs 14	MD 0.0 [-0.15, 0.15]	n/a	No GRADE assessment performed	
	Continuous vs NG intermittent bolus milk feeding in infants >1250g and <1499g	32 (1 study) 16 vs 16	MD 0.14 [-0.08, 0.36]	n/a	No GRADE assessment performed	

Table 37: Appendix 6: Summary table of effects of the interventions on Linear growth (other measures): 'How to feed'

Outcome assessed: Length (other measures)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Premji *et al.* 2011[133] Growth rate (mm/day) of the lower leg from birth to 32 weeks PMA	Continuous vs intermittent bolus milk feeding	1 study	"significantly faster growth rate in infants in the continuous nasogastric feeding method group" "p=0.002"	n/a	*Very low certainty [Overview]:* • *Serious risk of bias* • *Very serious imprecision*	
	Continuous vs intermittent bolus milk feeding (NG only)	1 study	Data not reported but assumed from "A significant difference in growth rate of the lower leg, in favor of continuous nasogastric feeding method, was reported only for the birth to 36 weeks postmenstrual age time period." no significant difference was found	n/a		

<table>
<tr><th>Outcome assessed: Length (other measures)</th><th>Comparison</th><th>Number of subjects (studies)</th><th>Measure of effect (95% CI)</th><th>I^2</th><th>GRADE assessment</th><th>Risk of Bias</th></tr>
<tr><td>Premji et al. 2011[133]

(continued)</td><td>Continuous versus intermittent bolus mild feeding (NG or OG) in infants 1000g</td><td></td><td>“significant difference in growth rate of the lower leg … in favor of the continuous nasogastric feeding method”</td><td></td><td></td><td></td></tr>
<tr><td rowspan="2">Premji et al. 2011[133]

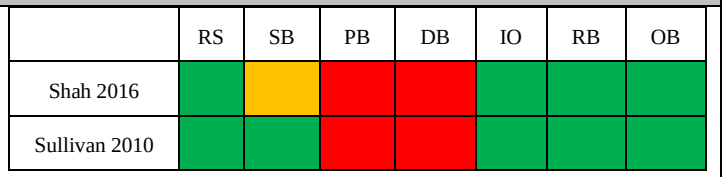
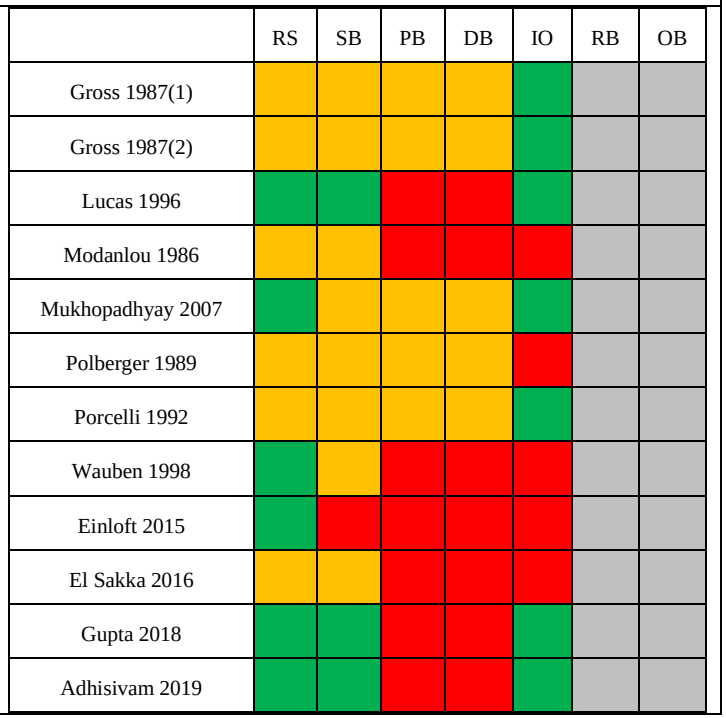
Growth rate (mm/day) of the lower leg from birth to 36 weeks PMA</td><td>Continuous vs intermittent bolus milk feeding</td><td>1 study</td><td>“significantly faster growth rate in infants in the continuous nasogastric feeding method group” “p=0.012”</td><td>n/a</td><td>Very low certainty [Overview]:
• Serious risk of bias
• Very serious imprecision</td><td rowspan="2"><table><tr><th></th><th>RS</th><th>SB</th><th>PB</th><th>DB</th><th>IO</th><th>RB</th><th>OB</th></tr><tr><td>Dsilna 2005</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr></table></td></tr>
<tr><td>Continuous vs NG intermittent bolus milk feeding (NG only)</td><td>1 study</td><td>“A significant difference in growth rate of the lower leg, in favor of continuous nasogastric feeding method, was reported only for the birth to 36 weeks postmenstrual age time period”</td><td>n/a</td><td></td></tr>
</table>

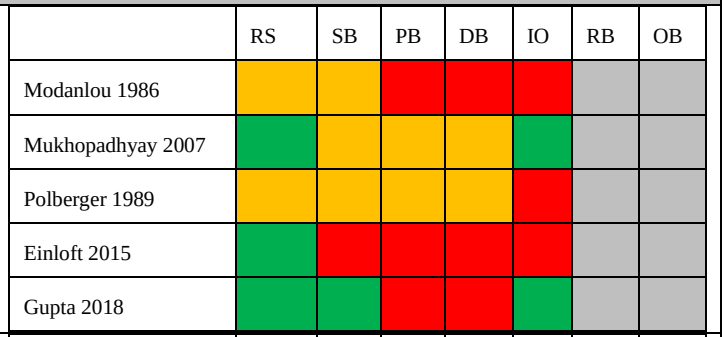
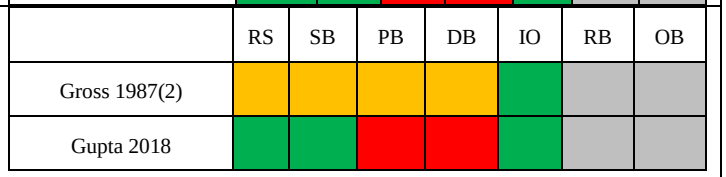
Outcome assessed: Length (other measures)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Premji *et al.* 2011[133] (continued)	Continuous versus intermittent bolus mild feeding (NG or OG) in infants 1000g		"significant difference in growth rate of the lower leg … in favor of the continuous nasogastric feeding method"	n/a		

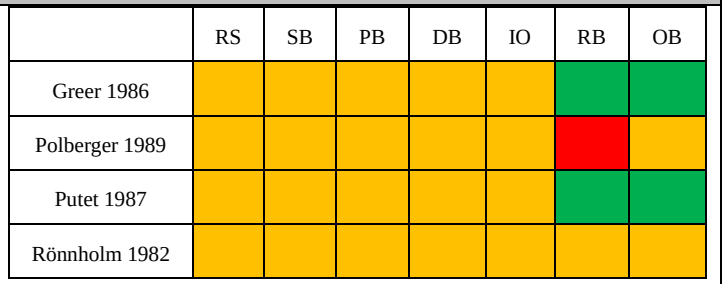
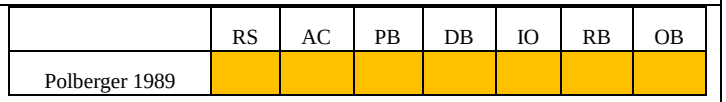
Table 38: Appendix 6: Summary table of effects of the interventions on Linear growth (any): 'Adjunctive strategies'

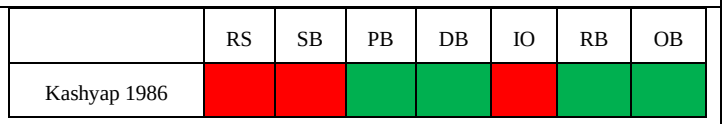
Outcome assessed: Linear growth	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Muelbert *et al.* 2019[157]	Exposure to smell and taste stimulation of milk with tube feeds versus no exposure	No eligible studies	n/a	n/a	n/a	n/a
Greene *et al.* 2016[159]	Oral stimulation vs standard care/non-oral intervention for promoting oral feeding	No eligible trials	n/a	n/a	n/a	n/a
Anabrees *et al.* 2015[158]	Glycerin prophylaxis versus placebo/no intervention for prevention/treatment of feeding intolerance	Not assessed	n/a	n/a	n/a	n/a
Ng *et al.* 2008[134]	Erythromycin vs placebo for prevention/treatment of feeding intolerance	Not assessed	n/a	n/a	n/a	n/a

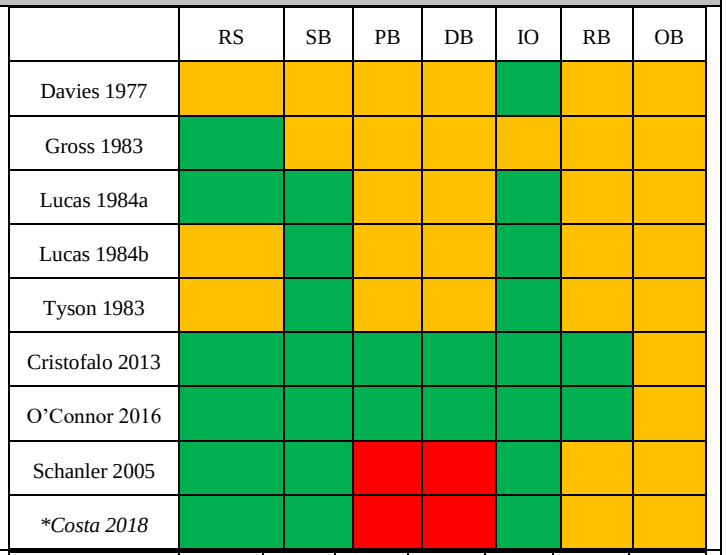
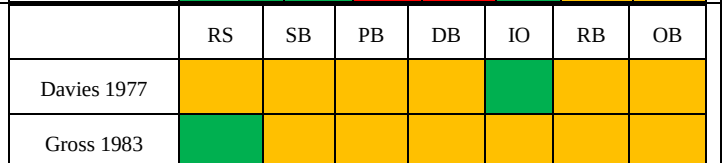
Table 39: Appendix 6: Summary table of effects of the interventions on Head growth (mm or cm/week): 'What to feed'

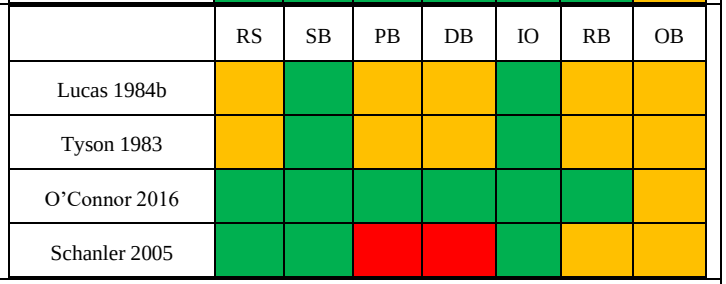
Outcome assessed: Head growth (cm or mm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Thanigainathan *et al.* 2020[128] (cm/week)	Early versus late fortification	237 (2 studies) 120 vs 117	MD -0.01 [-0.07, 0.06]	27%	Low certainty: • Lack of blinding • Small sample size	
Brown *et al.* 2020[143]	Multi-nutrient fortified vs unfortified breast milk	555 (8 studies) 283 vs 272	MD 0.06 [0.03, 0.08] Cm/week	42%	Moderate High risk of bias (uncertainty about methods used to generate random sequence, conceal allocation and blind assessments) in most trials.	

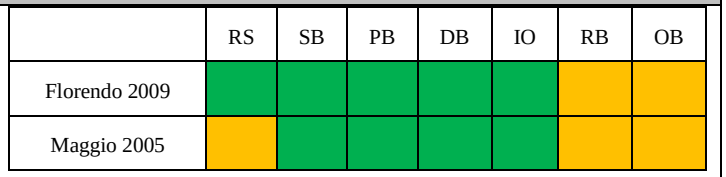
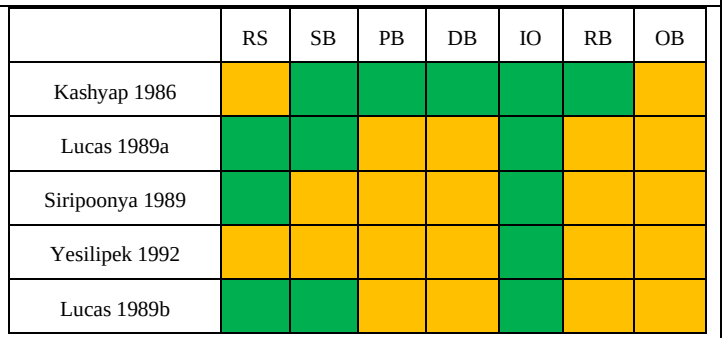
Outcome assessed: Head growth (cm or mm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Brown *et al.* 2020[143]	Subgroup: trial recruiting only very preterm or VLBW infants	375 (5 trials) 191 v 184	MD 0.07 [0.03, 0.11]	69%		
	Subgroup: Trials conducted in low- or middle-income countries	423 (4 trials) 216v207	MD 0.04 [0.01, 0.08]	53%		
	Subgroup: Trials using preterm formula powder as fortifier	174 (2 trials) 92v82	MD 0.05 [-0.00, 0.11]	8%		
	Test for subgroup differenced: Chi^2=1.25, df=3 (P=0.74), I^2=0%					

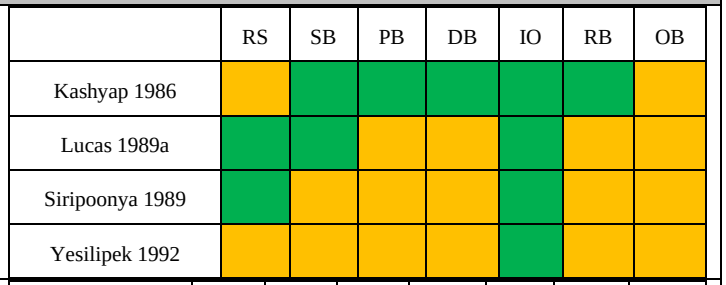
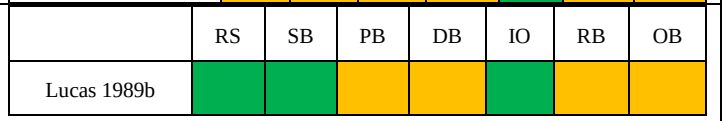
Outcome assessed: Head growth (cm or mm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Amissah *et al.* 2020a[146]	Carbohydrate (prebiotic) supplement vs no supplement	Nil eligible	n/a	n/a	n/a	n/a
Amissah *et al.* 2020b[148] (cm/week)	Protein supplementation versus no supplementation	68 (4 studies) 35 vs 33	MD 0.06 [0.01, 0.12] *Subgroup analyses not possible.*	84%	Low quality • Uncertain methodology • Moderate-to-high heterogeneity	
Amissah *et al.* 2020c[144] (cm/week)	Fat supplemented human milk versus control	14 (1 study) 7 vs 7	MD 0.2 [CI –0.07, 0.4] n.b. cm/week *Subgroup analyses not possible*	n/a	Very low quality • Uncertain risk of bias (insufficient methodological detail). • Imprecision (very small sample size)	

Outcome assessed: Head growth (cm or mm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Amissah *et al.* 2020c[144] (cm/week) (continued)					• Wide confidence intervals spanning across benefits and harms.	
Fenton *et al.* 2020[149] Head growth (cm/wk)	High versus low protein intake	18 (1 study) 9 vs 9	MD 0.37 [0.16, 0.58] 3 studies "no significant differences" (no data)	n/a	Very low • Risk of bias (selection, performance, detection, attrition). • Serious imprecision (single trial,very small)	

Outcome assessed: Head growth (cm or mm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Quigley *et al.* 2019[145] (mm/week)	Formula milk vs Donor breast milk	894 (8 studies) 438 vs 456	MD 0.85 [0.47, 1.23] Costa 2018 "did not detect a between-group difference in average head circumference at 15 days after birth or at 36 weeks' post-menstrual age."*	74%	Moderate-certainty evidence • high heterogeneity	
	Subgroup: Term formula versus unfortified formula	128 (2 studies) 54 v 74	MD 0.81 [0.15, 1.47]	0%		
	Subgroup: Preterm formula versus unfortified DBM	221 (3 studies) 110 v 111	MD 2.01 [1.21, 2.81]	84%		

Outcome assessed: Head growth (cm or mm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Quigley *et al.* 2019[145] (mm/week)	Subgroup: Preterm formula versus fortified DBM	545 (3 studies) 274v271	MD 0.30 [-0.27, 0.86]	0%		
	Test for subgroup differences: Chi^2=11.78, df=2 (P=0.00) I^2=83%					
	Sole diet	305 (5 studies) 145v160	MD 1.36 [0.85, 1.88]	77%		
	Supplement to human milk	589 (3 studies) 293 v 296	MD 0.24 [-0.32, 0.80]	0%		
	Test for subgroup differences: Chi^2=8.37 df=1 (P=0.00) I^2=88%					

Outcome assessed: Head growth (cm or mm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Ng *et al.* 2019[150] (mm/week)	Hydrolysed versus non-hydrolysed formula,	97 (2 studies) 50 vs 47	MD 0.27 [-0.39, 0.94]	82%	*Very Low certainty [Overview]:* • *Serious risk of bias (including funding source)* • *Serous inconsistency (high heterogeneity)* • *Serious imprecision*	
Walsh *et al.* 2019[127] (mm/week)	Nutrient-enriched formula versus standard formula	399 (5 studies) 199 vs 200	MD 1.04 [0.18, 1.89]	57%	Low risk •Uncertainty about methods used to generate random sequence, conceal allocation, and mask assessments in trials. •Moderate to high heterogeneity.	

Outcome assessed: Head growth (cm or mm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Walsh *et al.* 2019[127] (mm/week)	Subgroup: Sole diet	184 (4 studies) 94v90	MD 2.26 [1.00, 3.52]	85%		
	Subgroup: Supplemental to human milk	115 (1 study) 105v110	MD 0.00 [-1.6, 1.16]	n/a		
	Test for subgroup difference: Chi^2=6.67, df=1 (P=0.010), I^2=85%					
Dempsey *et al.* 2019[142]	Banked preterm versus banked term human milk	Nil studies eligible	n/a	n/a	n/a	n/a
Brown *et al.* 2019[126]	Formula versus maternal breastmilk	Nil studies eligible	n/a	n/a	n/a	n/a
Basuki *et al.* 2019[147]	Dilute versus full-strength formula	Nil studies eligible	n/a	n/a	n/a	n/a

Outcome assessed: Head growth (cm or mm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Premkumar *et al.* 2019[151]	Human milk-derived fortifier versus bovine milk-derived fortifier	Nil studies eligible (only absolute change in head circumference during intervention)	n/a	n/a	n/a	n/a
Nasuf *et al.* 2018[152]	Oropharyngeal colostrum vs control	Not assessed	n/a	n/a	n/a	n/a
Abiramalatha *et al.* 2017[59]	High-volume vs standard-volume feeds	Nil eligible studies for this outcome	n/a	n/a	n/a	n/a

<table>
<tr><th>Outcome assessed: Head growth (cm or mm/week)</th><th>Comparison</th><th>Number of subjects (studies)</th><th>Measure of effect (95% CI)</th><th>I^2</th><th>GRADE assessment</th><th>Risk of Bias</th></tr>
<tr>
<td>Nehra et al. 2002[129]

Head circumference gain, cm/wk</td>
<td>High MCT formula versus low MCT formula

Subgroup analysis not done for this outcome.</td>
<td>109 (5 studies) 53 vs 56</td>
<td>MD -0.03 [-0.15, 0.08]</td>
<td>0%</td>
<td>Very low certainty [Overview]:
• Serious risk of bias
• Serious inconsistency
• Serious imprecision</td>
<td>
<table>
<tr><th></th><th>RS</th><th>SB</th><th>PB</th><th>DB</th><th>IO</th><th>RB</th><th>OB</th></tr>
<tr><td>Hamosh 1991b</td><td>Can’t tell</td><td></td><td>Clinical - yes
Researchers - Can’t tell</td><td>Can’t tell</td><td>Complete follow up</td><td></td><td></td></tr>
<tr><td>Huston 1983</td><td>Can’t tell</td><td>Yes</td><td>Can’t tell</td><td>Can’t tell</td><td>Complete follow up</td><td></td><td></td></tr>
<tr><td>Okamoto 1982</td><td>Can’t tell</td><td>Can’t tell</td><td>Clinical - yes
Researchers - no</td><td>Can’t tell</td><td>Complete follow up</td><td></td><td></td></tr>
<tr><td>Sulkers 1992</td><td>Can’t tell</td><td>Can’t tell</td><td>Can’t tell</td><td>Can’t tell</td><td>Complete follow up</td><td></td><td></td></tr>
<tr><td>Whyte 1966</td><td>Can’t tell</td><td>Yes</td><td>Yes</td><td>Can’t tell</td><td>Complete follow up</td><td></td><td></td></tr>
</table>
</td>
</tr>
<tr>
<td>Tan-Dy et al. 2013[130]

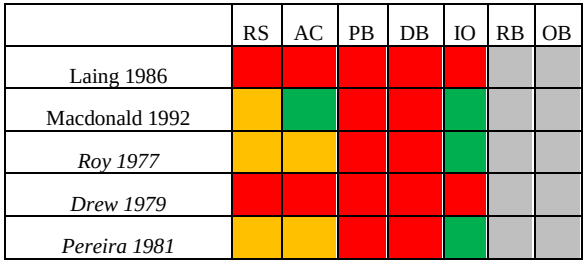
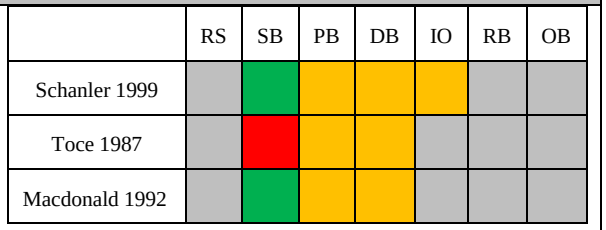
Head circumference gain (cm/week) at 14 days after study entry /study exit if occurred earlier</td>
<td>Lactase treated feeds vs placebo</td>
<td>130 (1 study) 66 vs 64</td>
<td>MD 0.10 [-0.18, 0.38]</td>
<td>n/a</td>
<td>Low certainty [Overview]
- Serious risk of bias
- Serious imprecision</td>
<td>
<table>
<tr><th></th><th>RS</th><th>SB</th><th>PB</th><th>DB</th><th>IO</th><th>RB</th><th>OB</th></tr>
<tr><td>Erasmus 2002</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
</table>
</td>
</tr>
</table>

Table 40: Appendix 6: Summary table of effects of the interventions on Head growth (mm or cm/week): 'When to feed'

Outcome assessed: Head circumference gain (cm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Yeo *et al.* 2019[154]	Stopping feeds vs continuing feeds during transfusion	Head growth not assessed	n/a	n/a	n/a	n/a
Abiramalatha *et al.* 2019a[155]	Routine monitoring of gastric residuals vs no routine monitoring of gastric residuals	No eligible studies	n/a	n/a	n/a	n/a
	Routine monitoring of gastric residuals quality vs routine monitoring of gastric residuals quality and quantity					

Outcome assessed: Head circumference gain (cm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Oddie *et al.* 2017[136]	Slow versus faster rates of feed advancement	Nil studies eligible "Head circumference z-score at discharge" only.	n/a	n/a	n/a	n/a
Morgan *et al.* 2014[137]	Delayed versus early introduction of progressive feeding	No eligible studies	n/a	n/a	n/a	n/a
Morgan *et al.* 2013[131]	Early trophic feeding vs enteral fasting	2 studies Total enrolled 100 (48v52) and 47 (22v25). Unclear numbers analysed.	McClure 2000: MD 0.7 cm/week (95% CI 0.1 to 1.3) Maetze 1992: "increases in head circumference…similar for both groups"	n/a	*Very low certainty [Overview]:* • *Serious risk of bias* • *Serious inconsistency* • *Very serious imprecision (narrative).*	

Table 41: Appendix 6: Summary table of effects of the interventions on Head growth (mm or cm/week): 'How to feed'

Outcome assessed: Head circumference gain (cm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Abiramalatha *et al.* 2019b[156]	Re-feeding vs discard gastric residuals	No eligible studies	n/a	n/a	n/a	n/a
Watson *et al.* 2013[135] (mm/week)	Transpyloric versus gastric tube feeding	75 (2 studies) 26 vs 49	MD 0.56 [-0.95, 2.08] *3 further trials did not find a statistically significant difference*	0%	*Low certainty [Overview]:* • *Very serious risk of bias (unblinded, unclear selection, includes trial with alternate assignment, baseline differences)*	
Dawson *et al.* 2012[132]	Push versus gravity bolus tube feeding	Not assessed	n/a	n/a	n/a	n/a
Premji *et al.* 2011[133]	Continuous vs intermittent bolus milk feeding	248 (3 studies) 125 vs 123	MD -0.03 [-0.09, 0.04]	0%	*Low certainty [Overview]:* • *Very serious risk of bias (unblinded, attrition, includes trial with alternate assignment)*	

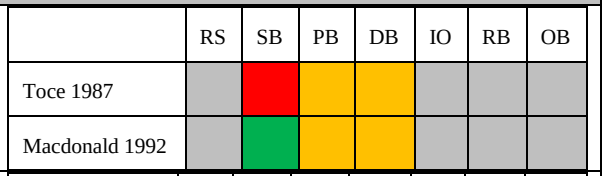
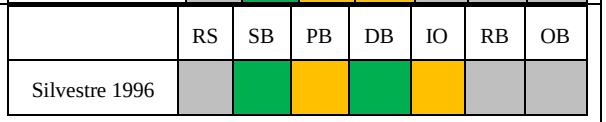
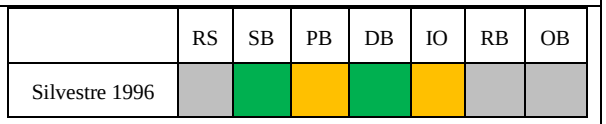
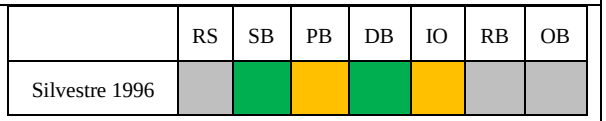
Outcome assessed: Head circumference gain (cm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Premji *et al.* 2011[133]	Continuous vs NG intermittent bolus milk feeding	77 (2 studies)	MD 0.01 [-0.12, 0.13]	0%	Not performed	
	Continuous vs intermittent bolus milk feeding in infants <1000g	30 (1 study) 12 vs 18	MD 0.07 [-0.03, 0.17]	n/a	Not performed	
	Continuous vs NG intermittent bolus milk feeding in infants <1000g	30 (1 study) 12 vs 18	MD 0.07 [-0.03, 0.17]	n/a	Not performed	
	Continuous vs NG intermittent bolus milk feeding in infants >1000g and <1249g	31 (1 study) 17 vs 14	MD 0.0 [-0.52, 0.52]	n/a	Not performed	
	Continuous vs NG intermittent bolus milk feeding in infants >1250g and <1499g	32 (1 study) 16 vs 16	MD 0.0 [-0.10, 0.10]	n/a	Not performed	

Table 42: Appendix 6: Summary table of effects of the interventions on Head growth (mm or cm/week): 'Adjunctive strategies'

Outcome assessed: Head circumference (cm/week)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Muelbert *et al.* 2019[157]	Exposure to smell and taste stimulation of milk with tube feeds versus no exposure	Nil eligible studies	n/a	n/a	n/a	n/a
Anabrees *et al.* 2015[158]	Glycerin prophylaxis versus placebo/no intervention	Not assessed	n/a	n/a	n/a	n/a
Ng *et al.* 2008[134]	Erythromycin vs placebo for prevention of feeding intolerance	Not assessed	n/a	n/a	n/a	n/a
Greene *et al.* 2016[159]	Oral stimulation vs standard care for promoting oral feeding	Nil eligible	n/a	n/a	n/a	n/a
	Oral stimulation vs non-oral intervention for promoting oral feeding					

Table 43: Appendix 6:Summary table of effects of the interventions on Head growth (other measures): 'What to feed'

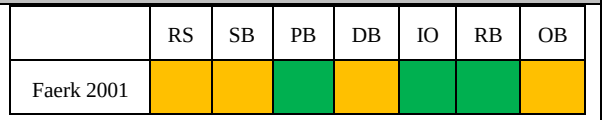
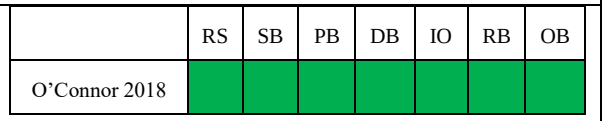
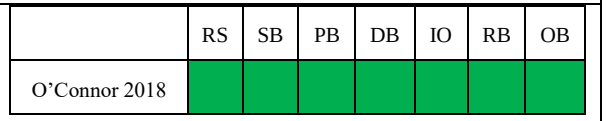
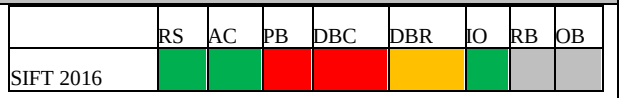
Outcome assessed: Head growth (other measures)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Amissah *et al.* 2020b[148]: Head circumference at term-equivalent age	Protein supplementation versus no supplementation	76 (1 study) 36 vs 40	MD 0.30 [-0.24, 0.84]	n/a	*Moderate certainty [Overview]:* *Serious risk of bias*	
Premkumar *et al.* 2019[151] : Change in head circumference (cm)	Human milk derived fortifier vs bovine milk-derived fortifier	118 (1 study) 61 vs 57	MD -0.60 [-1.52 to 0.32]		low certainty • imprecision • single study	
Premkumar *et al.* 2019[151] Head circumference –for-age z score	Human milk derived fortifier vs bovine milk-derived fortifier	118 (1 study) 61 vs 57	MD 0.00 [-0.49 to 0.49]		low certainty • imprecision • single study	

Table 44: Appendix 6: Summary table of effects of the interventions on Head growth (other measures): 'When to feed'

Outcome assessed: Head growth (other measures)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Yeo *et al.* 2019[154]	Stopping feeds vs continuing feeds during transfusion	Not assessed	n/a	n/a	n/a	n/a
Abiramalatha *et al.* 2019a[155]	Routine monitoring of gastric residuals vs no routine monitoring of gastric residuals	Nil eligible studies	n/a	n/a	n/a	n/a
	Routine monitoring of gastric residuals quality vs routine monitoring of gastric residuals quality and quantity					
Oddie *et al.* 2017[136]: Head circumference z-score at hospital discharge	Slow versus faster rates of feed advancement	2286 (1 trial) 1156 vs 1130	MD 0.0 [-0.13, 0.13]	n/a	*Moderate certainty [Overview]:* • *Serious risk of bias*	

Outcome assessed: Head growth (other measures)	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Morgan 2014[137]	Delayed versus early introduction of progressive feeding	Nil eligible studies	n/a	n/a	n/a	n/a
Morgan 2013[131]	Early trophic feeding vs enteral fasting	Nil further assessments	n/a	n/a	n/a	n/a

Table 45: Appendix 6: Summary table of effects of the interventions on Head growth (other measures): 'How to feed'

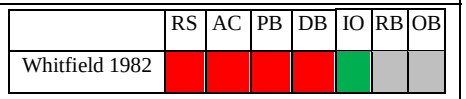
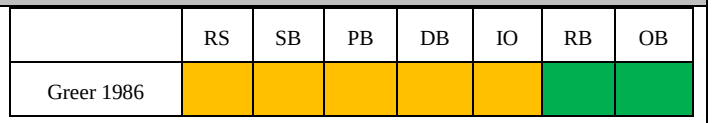
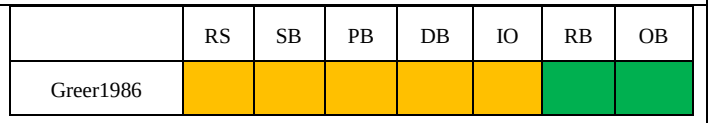
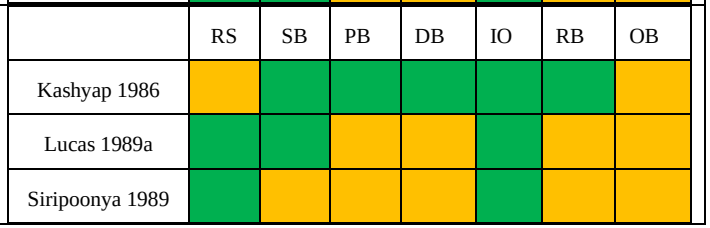
Outcome assessed: Head growth (other measures)	Additional head growth outcomes measured	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Abiramalatha *et al.* 2019b[156]	n/a	Re-feeding vs discard gastric residuals	None additional assessed	n/a	n/a	n/a	n/a
Watson *et al.* 2013[135]:	Head circumference measured at Expected Date of Delivery	Transpyloric versus gastric tube feeding	1 study (assumed 36 infants 21v15)	MD -1.0 [-1.7, -0.3]	n/a	*Low certainty [Overview]:* • *Very serious risk of bias (quasi-randomised trial)*	
Dawson 2012[132]	n/a	Push versus gravity feeds	Not assessed	n/a	n/a	n/a	n/a
Premji 2011[133]	n/a	Continuous vs intermittent bolus milk feeding	No additional outcomes	n/a	n/a	n/a	n/a

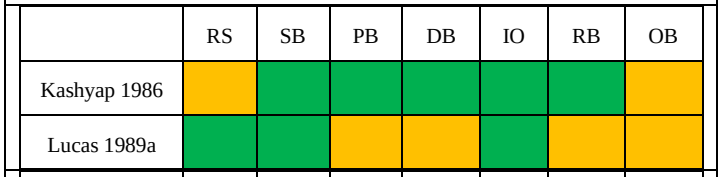
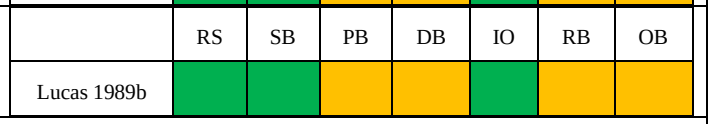
Table 46: Appendix 6: Summary table of effects of the interventions on Head growth (other measures): 'Adjunctive strategies'

Outcome assessed: Head circumference (other measures)	Additional head growth outcomes measured	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Muelbert *et al.* 2019[157]	None	Exposure to smell and taste stimulation of milk with tube feeds versus no exposure	n/a	n/a	n/a	n/a	n/a
Anabrees *et al.* 2015[158]	None	Glycerin prophylaxis versus placebo/no intervention	n/a	n/a	n/a	n/a	n/a
Ng *et al.* 2008[134]	None	Erythromycin vs placebo for prevention of feeding intolerance	n/a	n/a	n/a	n/a	n/a
Greene *et al.* 2016[159]	None	Oral stimulation vs standard care for promoting oral feeding	n/a	n/a	n/a	n/a	n/a
	None	Oral stimulations vs other non-oral intervention					

Table 47: Appendix 6:Summary table of effects of the interventions on other short term growth measures: 'What to feed'

Outcome assessed: Other short term growth measures	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Thanigainathan *et al.* 2020[128]	Early versus late fortification	Nil others reported	n/a	n/a	n/a	n/a
Brown 2020[143]	Multi-nutrient fortified vs unfortified breast milk	Nil others reported	n/a	n/a	n/a	n/a
Amissah *et al* 2020a[146]	Carbohydrate (prebiotic) supplementation vs no supplementation of human milk	Nil other reported	n/a	n/a	n/a	n/a
Amissah *et al* 2020b[148]: Growth: skinfold thickness triceps (mm/week)	Protein supplementation versus no supplementation of human milk	20 (1 study) 10 vs 10	MD 0.06 [–0.09 to 0.21]	n/a	*Low certainty [Overview]:* • *Serious risk of bias* *Serious imprecision*	
Amissah *et al* 2020b[148] Growth: skinfold thickness subscapular (mm/week)	Protein supplementation versus no supplementation of human milk	20 (1 study) 10 vs 10	MD 0.0 [-0.17 to 0.17]	n/a	Low *certainty [Overview]:* • *Serious risk of bias* *Serious imprecision*	

Outcome assessed: Other short term growth measures	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Amissah *et al* 2020c	Fat supplementation versus no supplementation of human milk	No eligible studies	n/a	n/a	n/a	n/a
Fenton *et al.* 2020[149]	High versus low protein intake	Nil others reported	n/a	n/a	n/a	n/a
Quigley *et al.* 2019[145]	Formula milk vs Donor breast milk	Nil reported				
Ng *et al.* 2019[150]	Hydrolysed versus non-hydrolysed formula	Nil reported				
Walsh *et al* 2019[127]: Rate of skinfold thickness gain - triceps (mm/week)	Nutrient-enriched formula versus standard formula.	364 (4 studies) 186 vs 178	MD 0.12 [0.07, 0.17]	0%	*Very low certainty [Overview]:* • *Very serious risk of bias* • *Serious imprecision*	
	Subgroup: sole diet	163 (3 studies) 83 vs 80	MD 0.16 [0.05, 0.27]	0%		

Outcome assessed: Other short term growth measures	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Walsh *et al* 2019[127]: Rate of skinfold thickness gain - triceps (mm/week)	Subgroup: supplement to human milk	201 (1 study) 103 vs 98	MD 0.12 [0.07, 0.17]	n/a		
	Test for subgroup differences: $Chi^2 = 0.57$, df = 1 (P=0.45), I^2=0%					
Walsh *et al* 2019[127]: Rate of skinfold thickness gain - subscapular (mm/week)	Nutrient-enriched formula versus standard formula	339 (3 studies) 173 vs 166	MD 0.10 [0.04, 0.16]	25%	*Very low certainty [Overview]:* • *Very serious risk of bias* • *Serious imprecision*	
	Subgroup: sole diet	138 (2 studies) 70 vs 68	MD 0.15 [0.17, 0.24]	0%		
	Subgroup: supplement to human milk	201 (1 study) 103 vs 98	MD 0.06 [-0.02, 0.14]	n/a		
	Test for subgroup differences: $Chi^2 = 2.39$, df = 1 (P=0.12), I^2=58.1%					
Dempsey *et al* 2019[142]	Banked preterm versus banked term human milk	Nil other reported	n/a	n/a	n/a	n/a

Outcome assessed: Other short term growth measures	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Brown *et al* 2019[126]	Formula versus maternal breastmilk	Nil eligible studies	n/a	n/a	n/a	n/a
Basuki *et al* 2019[147]	Dilute versus full-strength formula	Nil other reported	n/a	n/a	n/a	n/a
Premkumar *et al* 2019[151]	Human milk-derived fortifier versus bovine milk-derived fortifier	Nil other assessed	n/a	n/a	n/a	n/a
Nasuf *et al* 2018[152]	Oropharyngeal colostrum (OPC) compared to control (water, saline or no intervention) in preterm infants	Not assessed	n/a	n/a	n/a	n/a
Abiramalatha *et al* 2017[59]	High-volume vs standard-volume feeds	Not assessed	n/a	n/a	n/a	n/a
Tan-Dy *et al* 2013[130]		Nil further assessed	n/a	n/a	n/a	n/a

<table>
<tr><th>Outcome assessed: Other short term growth measures</th><th>Comparison</th><th>Number of subjects (studies)</th><th>Measure of effect (95% CI)</th><th>I^2</th><th>GRADE assessment</th><th>Risk of Bias</th></tr>
<tr>
<td>Nehra et al 2002[129]

Skin fold thickness gain, mm/wk</td>
<td>High MCT formula versus low MCT formula</td>
<td>14 (1 study) 4 vs 10</td>
<td>MD -0.15 [-0.41, 0.11]</td>
<td>n/a</td>
<td>Low certainty [Overview]:

• Serious risk of bias
• Serious imprecision</td>
<td>
<table>
<tr><th></th><th>RS</th><th>SB</th><th>PB</th><th>DB</th><th>IO</th><th>RB</th><th>OB</th></tr>
<tr><td>Okamtoto 1982</td><td></td><td></td><td></td><td></td><td></td><td></td><td></td></tr>
</table>
Sequence generation: “can’t tell”

Allocation concealment: “can’t tell” Blinding of intervention: Clinical caretakers -yes/Research team – no

Blinding of outcome: “can’t tell”

Complete follow up
</td>
</tr>
</table>

Table 48: Appendix 6:Summary table of effects of the interventions on other short term growth measures: 'When to feed'

Outcome assessed: Other short term growth measures	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Yeo *et al* 2019[154]	Stopping feeds vs continuing feeds during transfusion	Not assessed	n/a	n/a	n/a	n/a
Abiramalatha *et al* 2019a[155]	Routine monitoring of gastric residuals vs no routine monitoring of gastric residuals	Not assessed	n/a	n/a	n/a	n/a
	Routine monitoring of gastric residuals quality vs routine monitoring of gastric residuals quality and quantity					
Oddie *et al* 2017[136]	Slow versus faster rates of feed advancement	Nil further outcomes reported	n/a	n/a	n/a	n/a
Morgan *et al* 2014[137]	Delayed versus early introduction of progressive feeding	Nil further outcomes reported	n/a	n/a	n/a	n/a

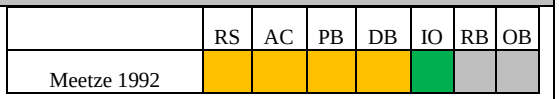
Outcome assessed: Other short term growth measures	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Morgan *et al* 2013[131]: Increase in mid-arm circumference	Early trophic feeding vs enteral fasting	1 study Total enrolled 47 (22vs 25)	Maetze 1992: "similar for both groups	n/a	*Very low certainty [Overview]:* • *Serious risk of bias* • *Very serious imprecision (narrative report)*	

Table 49: Appendix 6: Summary table of effects of the interventions on other short term growth measures: 'How to feed'

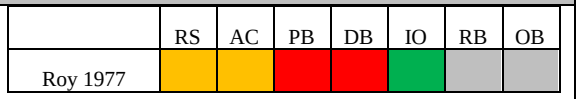
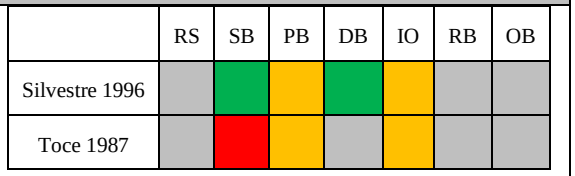
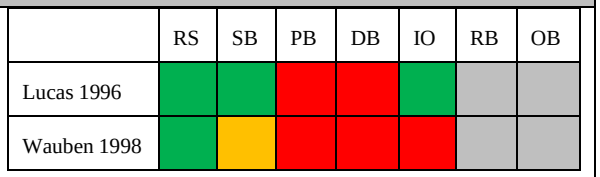
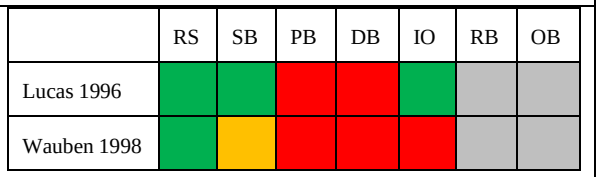
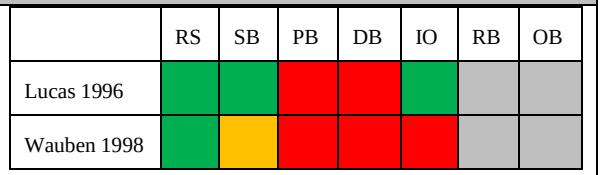
Outcome assessed: Other short term growth measures	Outcome measured	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Abiramalatha *et al* 2019b[156]	None additional assessed	Re-feeding vs discard gastric residuals	n/a	n/a	n/a	n/a	n/a
Watson *et al* 2013[135135]	Change in skinfold thickness (mm/week)	Transpyloric versus gastric tube feeding	18 (1 study) 9 vs 9	MD -0.20 [-1.18, 0.78]	n/a	*Low certainty [Overview]:* • *Very serious risk of bias*	
Dawson *et al* 2012[132]	None additional assessed	Push versus gravity bolus tube feeding	n/a	n/a	n/a	n/a	n/a
Premji *et al* 2011[133]	Change in triceps skinfold thickness (mm/week)	Continuous vs intermittent bolus milk feeding *(=Continuous vs NG intermittent bolus milk feeding)*	135 (2 studies) 72 vs 63	MD 0.0 [-0.06, 0.06]	0%	*Low certainty [Overview]:* • *Very serious risk of bias*	

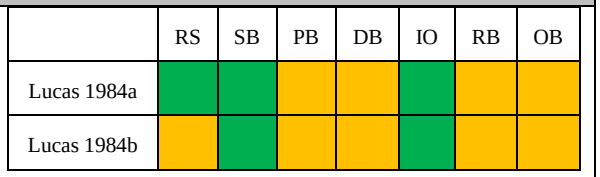
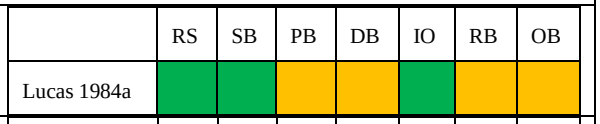
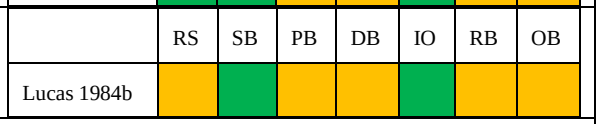
Table 50: Appendix 6: Summary table of effects of the interventions on other short term growth measures: 'Adjunctive strategies'

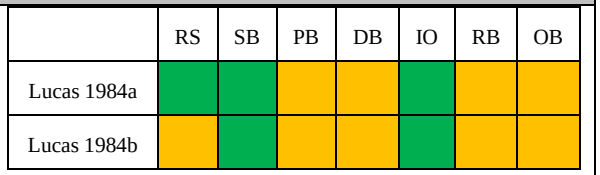
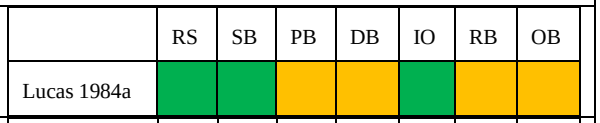
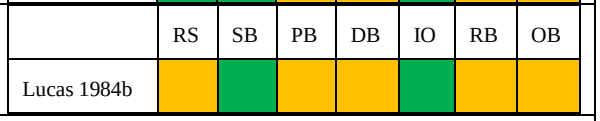
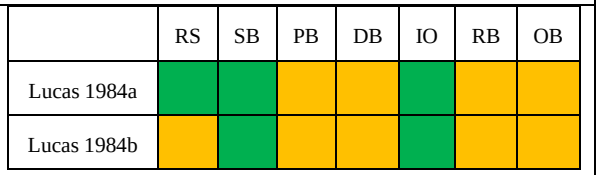
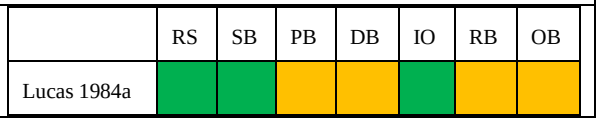
Outcome assessed: Other short term growth measures	**Comparison**	**Number of subjects (studies)**	**Measure of effect (95% CI)**	**I^2**	**GRADE assessment**	**Risk of Bias**
Muelbert *et al* 2019[157]	Exposure to smell and taste stimulation of milk with tube feeds versus no exposure	Nil other short term growth outcomes reported	n/a	n/a	n/a	n/a
Anabrees *et al* 2015[158]	Glycerin prophylaxis versus placebo/no intervention	Nil further assessed	n/a	n/a	n/a	n/a
Ng *et al* 2008[134]	Erythromycin vs placebo for prevention of feeding intolerance	Nil further assessed	n/a	n/a	n/a	n/a
Greene *et al* 2016[159]	Oral stimulation vs standard care for promoting oral feeding	Nil further assessed	n/a	n/a	n/a	n/a
	Oral stimulation vs other non-oral intervention to promote oral feeding					

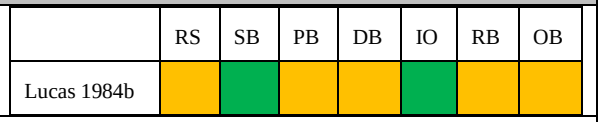
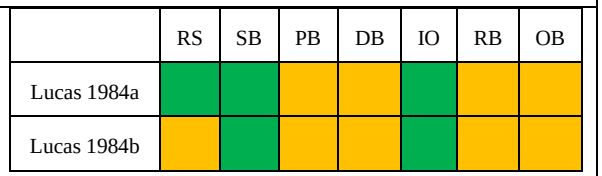
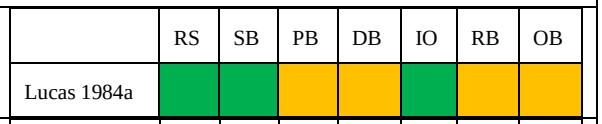
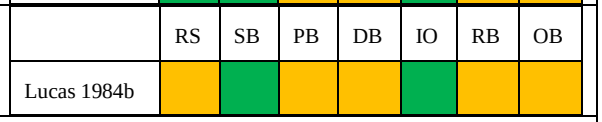
Table 51: Appendix 6: Summary table of effects of the interventions on long term growth measures: 'What to feed'

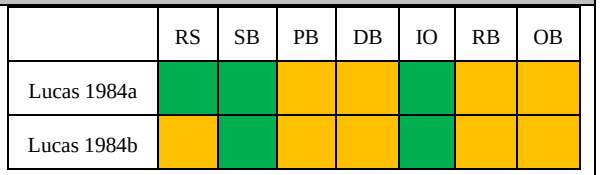
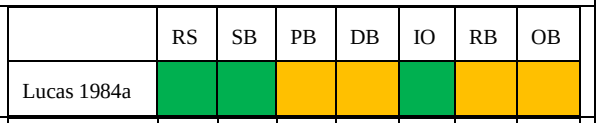
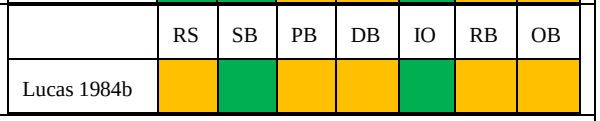
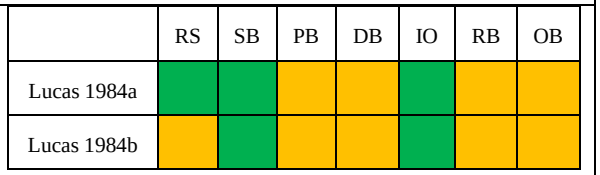
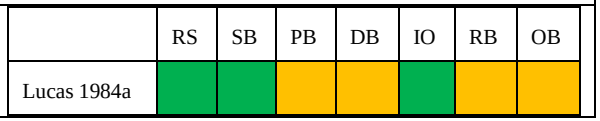
Outcome assessed: Long-term growth	**Outcome assessed**	**Comparison**	**Number of subjects (studies)**	**Measure of effect (95% CI)**	**I^2**	**GRADE assessment**	**Risk of Bias**
Thanigainathan *et al.* 2020[128]	No longterm growth data	Early versus late fortification	n/a	n/a	n/a	n/a	n/a
Brown *et al.* 2020[143]	Weight at 12 to 18 months (kg)	fortified vs unfortified breast milk	270 (2 trials) 137 vs 133	MD -0.03 [-0.31, 0.25]	0%	*Moderate certainty [Overview]: Serous risk of bas*	
	Length at 12 to 18 months (cm)	fortified vs unfortified breast milk	270 (2 trials) 137 vs 133	MD -0.19 [-0.98, 0.60]	0%	*Moderate certainty [Overview]: Serous risk of bas*	

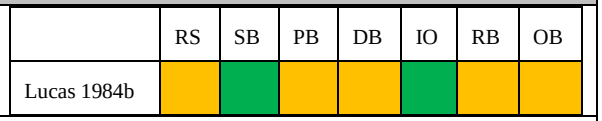
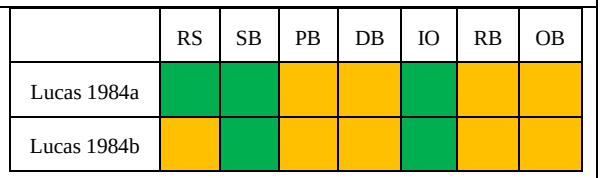
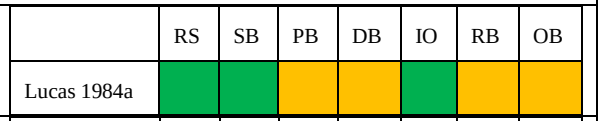
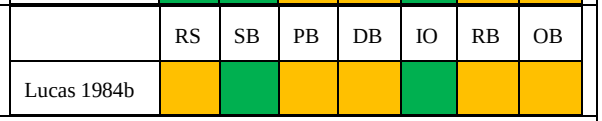
Outcome assessed: Long-term growth	**Outcome assessed**	**Comparison**	**Number of subjects (studies)**	**Measure of effect (95% CI)**	**I^2**	**GRADE assessment**	**Risk of Bias**
Brown *et al.* 2020[143]	Head circumference at 12 to 18 months (cm)	fortified vs unfortified breast milk	270 (2 trials) 137 vs 133	MD -0.10 [-0.37, 0.18]	0%	*Moderate certainty [Overview]:* *Serous risk of bas*	
Amissah *et al.* 2020a[146]	No longterm growth data	Carbohydrate (prebiotic) supplementation vs no supplementation of human milk	Nil studies eligible	n/a	n/a	n/a	n/a
Amissah *et al.* 2020b[148]	No longterm growth data	Protein supplementation versus no supplementation of human milk	Nil studies eligible	n/a	n/a	n/a	n/a
Amissah *et al.* 2020c[144]	No longterm growth data	Fat supplementation versus no supplementation of human milk	Nil eligible studies	n/a	n/a	n/a	n/a

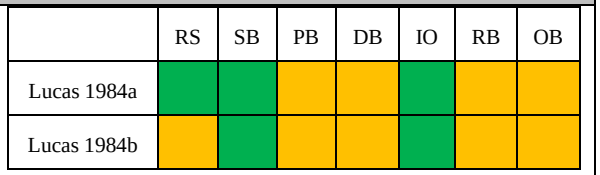
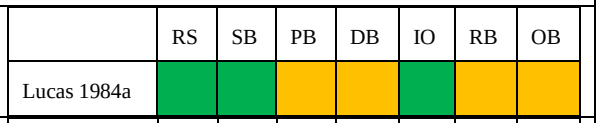
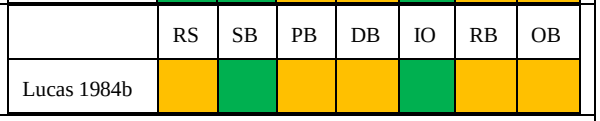
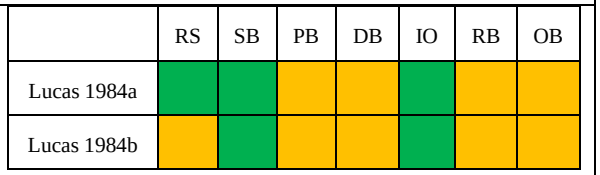
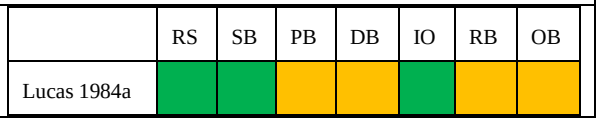
Outcome assessed: Long-term growth	**Outcome assessed**	**Comparison**	**Number of subjects (studies)**	**Measure of effect (95% CI)**	**I^2**	**GRADE assessment**	**Risk of Bias**
Fenton *et al.* 2020[149]	No longterm growth data	High versus low protein formula	Nil eligible studies	n/a	n/a	n/a	n/a
Quigley *et al* 2019[145]	Weight (kg) at 9 months post-term	Formula versus donated breast milk *= subgroup preterm formula vs unfortified donor breast milk*	369 (2 studies) 174 vs 195	MD -0.03 [-0.26, 0.21]	14%	*Moderate certainty [Overview]* • *Serious risk of bias*	
		Subgroup: sole diet	110 (1 study) 48 vs 62	MD 0.20 [-0.27, 0.67]	n/a		
		Subgroup: supplement to human milk	259 (1 study) 126 vs 133	MD -0.10 [-0.37, 0.17]	n/a		
		Test for subgroup differences	Chi^2=1.17, df=1 (P=0.28), I^2=14%				

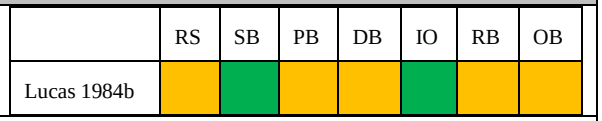
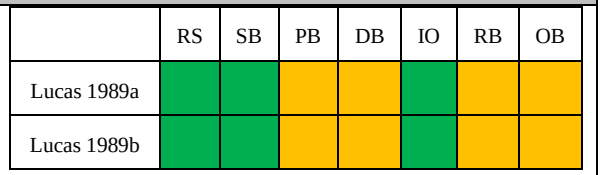
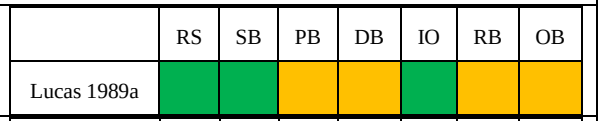
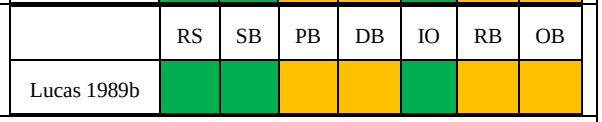
Outcome assessed: Long-term growth	**Outcome assessed**	**Comparison**	**Number of subjects (studies)**	**Measure of effect (95% CI)**	**I^2**	**GRADE assessment**	**Risk of Bias**
Quigley *et al* 2019[145]	Length (cm) at 9 months post-term	Formula versus donated breast milk *= subgroup preterm formula vs unfortified donor breast milk*	369 (2 studies) 174 vs 195	MD 0.03 [-0.64, 0.70]	0%	*Moderate certainty [Overview]* • *Serious risk of bias*	
		Subgroup: sole diet	110 (1 study) 48 vs 62	MD 0.40 [-0.93, 1.73]	n/a		
		Subgroup: supplement to human milk	259 (1 study) 126 vs 133	MD -0.10 [-0.88, 0.68]	n/a		
		Test for subgroup differences: Chi^2=0.40 df=1 (P=0.54), I^2=0%					
	Head circumference (cm) at 9 months post-term	Formula versus donated breast milk *= subgroup preterm formula vs unfortified donor breast milk*	369 (2 studies) 174 vs 195	MD 0.20 [-0.13, 0.53]	0%	*Moderate certainty [Overview]:* • *Serious risk of bias*	
		Subgroup: sole diet	110 (1 study) 48 vs 62	MD 0.20 [-0.45, 0.85]	n/a		

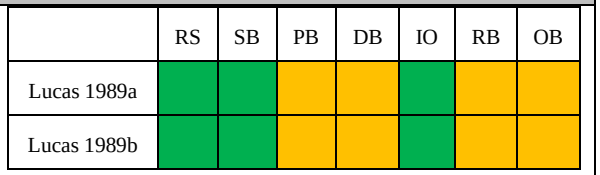
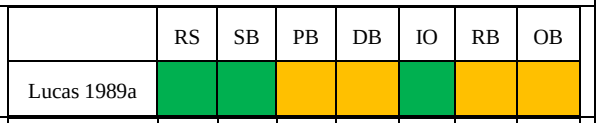
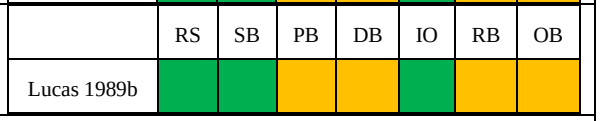
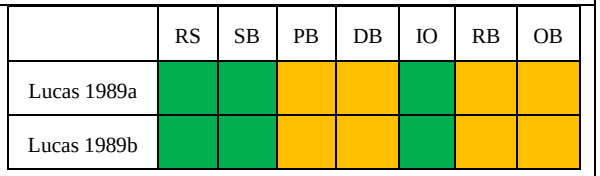
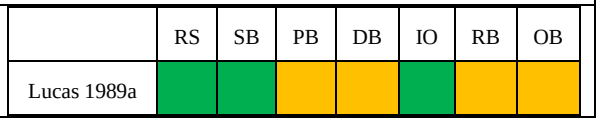
Outcome assessed: Long-term growth	Outcome assessed	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Quigley *et al* 2019[145]		Subgroup: supplement to human milk	259 (1 study) 126 vs 133	MD 0.20 [-0.18, 0.58]	n/a		
		Test for subgroup differences: Chi^2=0.00 df=1 (=1.00), I^2=0%					
	Weight (kg) at 18 months post-term	Formula versus donated breast milk *= subgroup preterm formula vs unfortified donor breast milk*	438 (2 studies) 217 vs 221	MD 0.10 [-0.15, 0.35]	0%	*Moderate certainty [Overview]:* • *Serious risk of bias*	
		Subgroup: sole diet	136 (1 study) 64 vs 72	MD 0.10 [-0.37, 0.57]	n/a		
		Subgroup: supplement to human milk	302 (1 study) 153 vs 149	MD 0.10 [-0.19, 0.39]	n/a		
		Test for subgroup differences: Chi^2=0.00 df=1 (=1.00), I^2=0%					

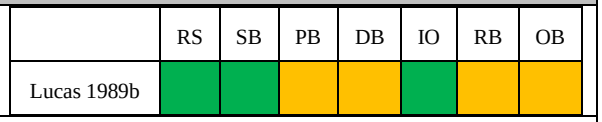
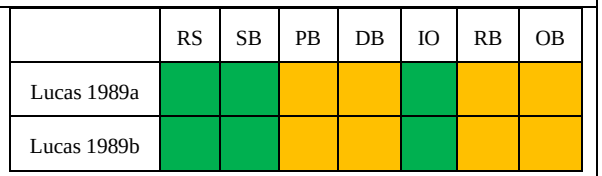
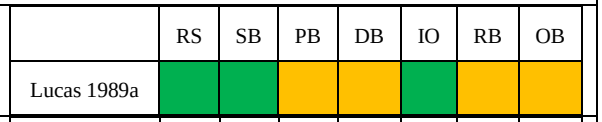
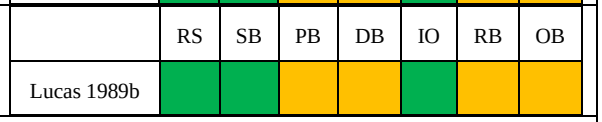
Outcome assessed: Long-term growth	**Outcome assessed**	**Comparison**	**Number of subjects (studies)**	**Measure of effect (95% CI)**	**I^2**	**GRADE assessment**	**Risk of Bias**
Quigley *et al* 2019[145]	Length (cm) at 18 months post-term	Formula versus donated breast milk *= subgroup preterm formula vs unfortified donor breast milk*	438 (2 studies) 217 vs 221	MD 0.53 [-0.15, 1.20]	0%	*Moderate certainty [Overview]:* • *Serious risk of bias*	
		Subgroup: sole diet	136 (1 study) 64 vs 72	MD 0.10 [-0.37, 0.57]	n/a		
		Subgroup: supplement to human milk	302 (1 study) 153 vs 149	MD 0.10 [-0.19, 0.39]	n/a		
		Test for subgroup differences: Chi^2=0.02 df=1 (=0.90), I^2=0%					
	Head circumference (cm) at 18 months post-term	Formula versus donated breast milk *= subgroup preterm formula vs unfortified donor breast milk*	438 (2 studies) 217 vs 221	MD 0.10 [-0.19, 0.39]	0%	*Moderate certainty [Overview]:* • *Serious risk of bias*	
	Head circumference	Subgroup: sole diet	136 (1 study) 64 vs 72	MD 0.10 [-0.44, 0.64]	n/a		

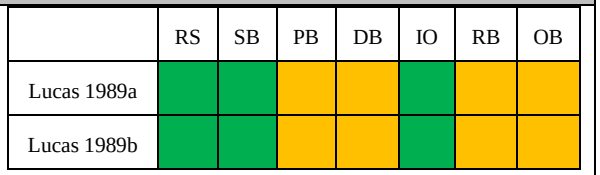
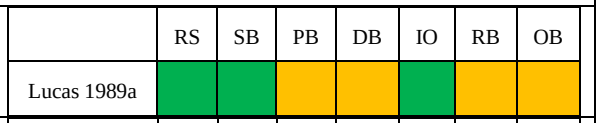
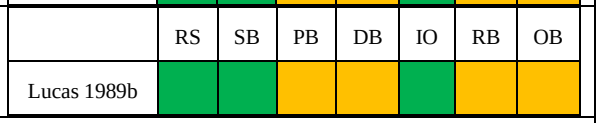
Outcome assessed: Long-term growth	Outcome assessed	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Quigley *et al* 2019[145]	(cm) at 18 months post-term	Subgroup: supplement to human milk	302 (1 study) 153 vs 149	MD 0.10 [-0.25, 0.45]	n/a		
		Test for subgroup differences:Chi^2=0.00 df=1 (=1.00), I^2=0%					
	Weight (kg) at 7.5-8 years of age	Formula versus donated breast milk *= subgroup preterm formula vs unfortified donor breast milk*	420 (2 studies) 213 vs 207	MD -0.56 [-1.42, 0.29]	47%	*Moderate certainty [Overview]:* • *Serious risk of bias*	
		Subgroup: sole diet	130 (1 study) 62 vs 68	MD 0.5 [-1.24, 2.24]	n/a		
		Subgroup: supplement to human milk	290 (1 study) 151 vs 139	MD -0.90 [-1.88, 0.08]	n/a		
		Test for subgroup differences: Chi^2=1.89 df=1 (=0.17), I^2=47%					

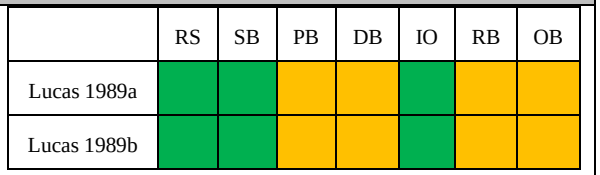
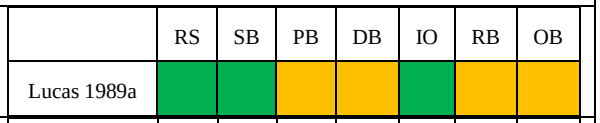
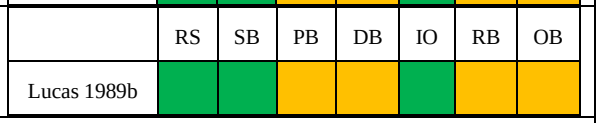
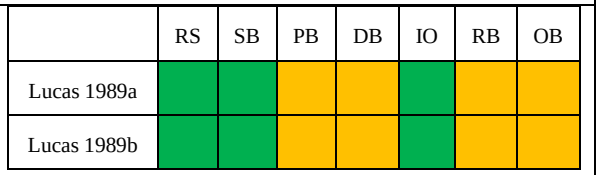
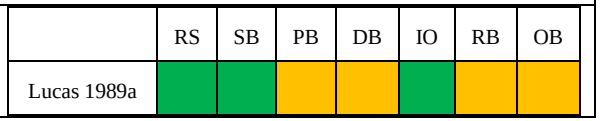
Outcome assessed: Long-term growth	**Outcome assessed**	**Comparison**	**Number of subjects (studies)**	**Measure of effect (95% CI)**	**I^2**	**GRADE assessment**	**Risk of Bias**
Quigley *et al* 2019[145]	Length (cm) at 7.5-8 years of age	Formula versus donated breast milk *= subgroup preterm formula vs unfortified donor breast milk*	420 (2 studies) 213 vs 207	MD 0.05 [-1.12, 1.23]	0%	*Moderate certainty [Overview]:* • *Serious risk of bias*	
		Subgroup: sole diet	130 (1 study) 62 vs 68	MD 1.0 [-1.26, 3.26]	n/a		
		Subgroup: supplement to human milk	290 (1 study) 151 vs 139	MD -0.30 [-1.68, 1.08]	n/a		
		Test for subgroup differences: Chi^2=0.93 df=1 (=0.34), I^2=0%					
	Head circumference (cm) at 7.5-8 years of age	Formula versus donated breast milk *= subgroup preterm formula vs unfortified donor breast milk*	420 (2 studies) 213 vs 207	MD -0.19 [-0.54, 0.16]	1%	*Moderate certainty [Overview]:* • *Serious risk of bias*	
		Subgroup: sole diet	130 (1 study) 62 vs 68	MD 0.10 [-0.56, 0.76]	n/a		

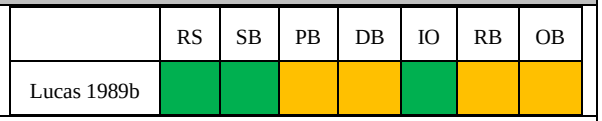
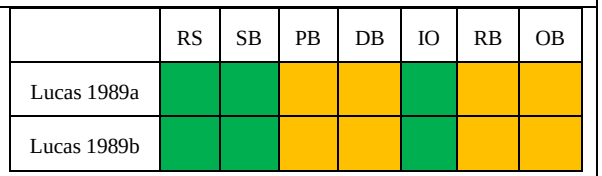
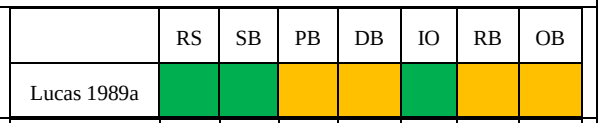
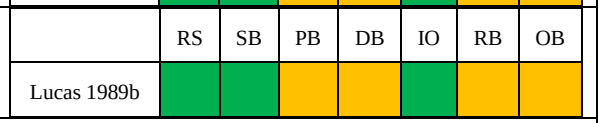
Outcome assessed: Long-term growth	Outcome assessed	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Quigley *et al* 2019[145]		Subgroup: supplement to human milk	290 (1 study) 151 vs 139	MD -0.30 [-0.71, 0.11]	n/a		
		Test for subgroup differences: Chi^2=1.01 df=1 (=0.32), I^2=1%					
Ng *et al.* 2019[150]	Not assessed	Hydrolysed versus non-hydrolysed formula,	n/a	n/a	n/a	n/a	n/a
Walsh *et al* 2019[127]	Weight (kg) at 18 months post term	Nutrient-enriched formula versus standard formula	334 (2 studies) 166 vs 168	MD 0.06 [-0.21, 0.33]	0%	*Moderate certainty [Overview]:* • *Serious risk of bias*	
		Subgroup: sole diet	119(1 study) 61 vs 58	MD 0.20 [-0.31, 0.71]	n/a		
		Subgroup: supplement to human milk	215 (1 study) 105 vs 110	MD 0.0 [-0.32, 0.32]	n/a		
		Test for subgroup differences	Chi^2=0.43, df=1 (P=0.51), I^2=0%				

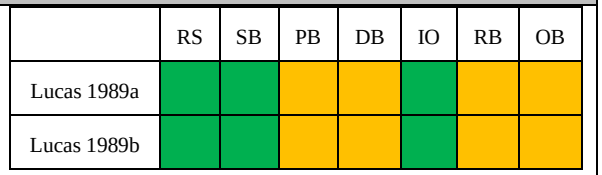
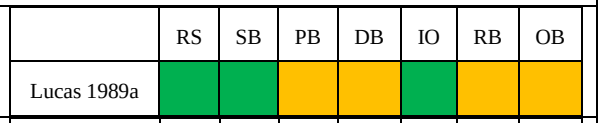
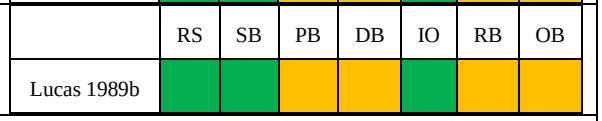
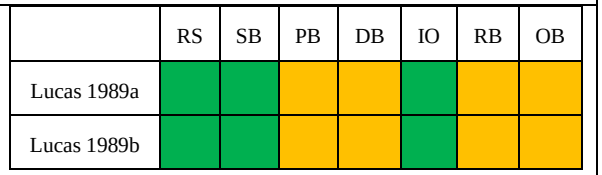
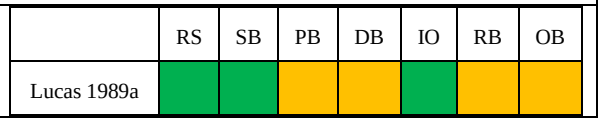
Outcome assessed: Long-term growth	Outcome assessed	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Walsh *et al* 2019[127]	Height (cm) at 18 months post term	Nutrient-enriched formula versus standard formula	334 (2 studies) 166 vs 168	MD 0.31 [-0.43, 1.06]	48%	*Moderate certainty [Overview]:* • *Serious risk of bias*	
	Height (cm) at 18 months post term	Subgroup: sole diet	119 (1 study) 61 vs 58	MD 1.20 [-0.26, 2.66]	n/a		
		Subgroup: supplement to human milk	215 (1 study) 105 vs 110	MD 0.00 [-0.87, 0.87]	n/a		
		Test for subgroup differences: Chi^2=1.91, df=1 (P=0.17), I^2=47.7%					
	Head circumference (cm) at 18 months post term	Nutrient-enriched formula versus standard formula	334 (2 studies) 166 vs 168	MD 0.09 [-0.26, 0.43]	0%	*Moderate certainty [Overview]:* • *Serious risk of bias*	
		Subgroup: sole diet	119 (1 study) 61 vs 58	MD 0.20 [-0.32, 0.72]	n/a		

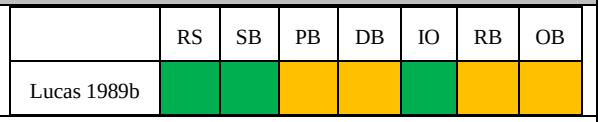
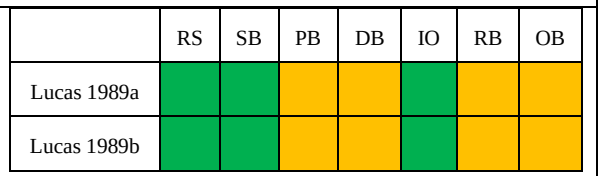
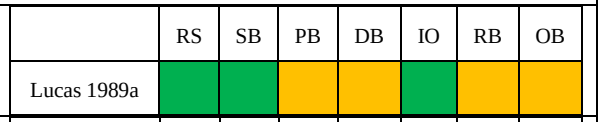
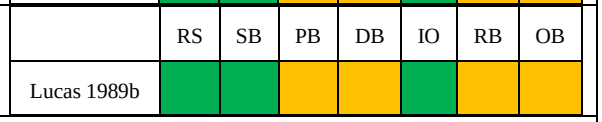
Outcome assessed: Long-term growth	Outcome assessed	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Walsh *et al* 2019[127]		Subgroup: supplement to human milk	215 (1 study) 105 vs 110	MD 0.00 [-0.45, 0.45]	n/a		
		Test for subgroup differences: Chi^2=0.32, df=1 (P=0.57), I^2=0%					
	Triceps skinfold thickness (mm) at 18 months post term	Nutrient-enriched formula versus standard formula	334 (2 studies) 166 vs 168	MD 0.01 [-0.42, 0.45]	0%	*Moderate certainty [Overview]:* • *Serious risk of bias*	
	Triceps skinfold thickness (mm) at 18 months post term	Subgroup: Sole diet	119 (1 study) 61 vs 58	MD 0.20 [-0.50, 0.90]	n/a		
		Subgroup: Supplement to human milk	215 (1 study) 105 vs 110	MD -0.10 [-0.65, 0.45]	n/a		
		Test for subgroup differences:Chi^2=0.43, df=1 (P=0.51), I^2=0%					

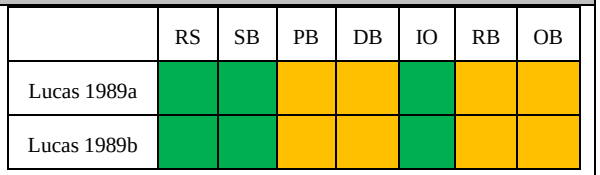
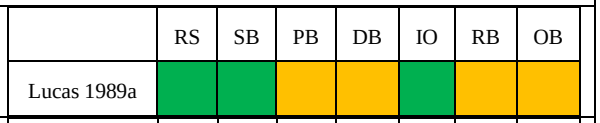
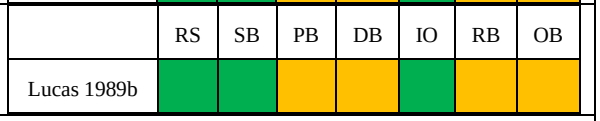
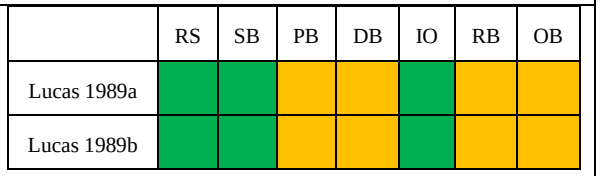
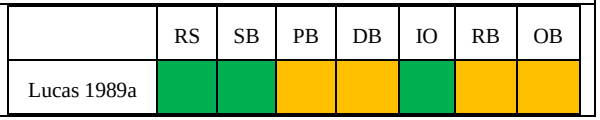
Outcome assessed: Long-term growth	**Outcome assessed**	**Comparison**	**Number of subjects (studies)**	**Measure of effect (95% CI)**	**I^2**	**GRADE assessment**	**Risk of Bias**
Walsh *et al* 2019[127]	Subscapular skinfold thickness (mm) at 18 months post term	Nutrient-enriched formula versus standard formula	334 (2 studies) 166 vs 168	MD -0.14 [-0.40, 0.13]	0%	*Moderate certainty [Overview]:* • *Serious risk of bias*	
		Subgroup: Sole diet	119 (1 study) 61 vs 58	MD 0.00 [-0.42, 0.47]	n/a		
		Subgroup: Supplement to human milk	215 (1 study) 105 vs 110	MD 0.20 [-0.52, 0.12]	n/a		
	Subscapular skinfold thickness (mm) at 18 months post term	Test for subgroup differences: Chi^2=0.48, df=1 (P=0.49), I^2=0%					

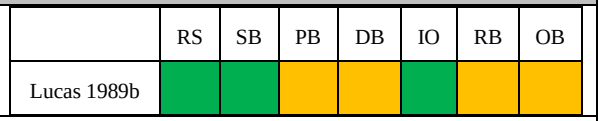
Outcome assessed: Long-term growth	**Outcome assessed**	**Comparison**	**Number of subjects (studies)**	**Measure of effect (95% CI)**	**I^2**	**GRADE assessment**	**Risk of Bias**
Walsh *et al* 2019[127]	Body mass index (kg/m²) at 18 months post term	Nutrient-enriched formula versus standard formula	334 (2 studies) 166 vs 168	MD -0.10 [-0.43, 0.23]	0%	*Moderate certainty [Overview]:* • *Serious risk of bias*	
		Subgroup: Sole diet	119 (1 study) 61 vs 58	MD -0.10 [-0.70, 0.50]	n/a		
		Subgroup: Supplement to human milk	215 (1 study) 105 vs 110	MD -0.10 [-0.50, 0.30]	n/a		
		Test for subgroup differences: Chi^2=0.00, df=1 (P=1.00), I^2=0%					
	Weight (kg) at 7.5 to 8 years post term	Nutrient-enriched formula versus standard formula	359 (2 studies) 178 vs 181	MD 0.30 [-0.55, 1.15]	0%	*Moderate certainty [Overview]:* • *Serious risk of bias*	
		Subgroup: Sole diet	135 (1 study) 67 vs 68	0.30 [-0.99, 1.59]	n/a		

Outcome assessed: Long-term growth	Outcome assessed	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Walsh *et al* 2019[127]		Subgroup: Supplement to human milk	224 (1 study) 111 vs 113	0.30 [-0.84, 1.44]	n/a		
		Test for subgroup differences:	Chi^2=0.00, df=1 (P=1.00), I^2=0%				
	Height (cm) at 7.5 to 8 years post term	Nutrient-enriched formula versus standard formula	359 (2 studies) 178 vs 181	MD 0.93 [-0.30, 2.16]	0%	*Moderate certainty [Overview]:* • *Serious risk of bias*	
	Height (cm) at 7.5 to 8 years post term	Subgroup: Sole diet	135 (1 study) 67 vs 68	MD 1.30 [-0.69, 3.29]	n/a		
		Subgroup: Supplement to human milk	224 (1 study) 111 vs 181	MD 0.70 [-0.86, 2.26]	n/a		
		Test for subgroup differences:	Chi^2=0.22, df=1 (P=0.64), I^2=0%				

Outcome assessed: Long-term growth	Outcome assessed	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Walsh *et al* 2019[127]	Head circumference (cm) at 7.5 to 8 years post term	Nutrient-enriched formula versus standard formula	359 (2 studies) 178 vs 181	MD -0.12 [-0.45, 0.21]	0%	*Moderate certainty [Overview]:* • *Serious risk of bias*	
		Subgroup: Sole diet	135 (1 study) 67 vs 68	MD 0.99 [-0.52, 0.52]	n/a		
		Subgroup: Supplement to human milk	224 (1 study) 111 vs 113	MD -0.20 [-0.62, 0.22]	n/a		
		Test for subgroup differences: Chi^2=0.34, df=1 (P=0.56), I^2=0%					
	Triceps skinfold thickness (mm) at 7.5 to 8 years post term	Nutrient-enriched formula versus standard formula	359 (2 studies) 178 vs 181	MD -0.16 [-0.91, 0.60]	6%	*Moderate certainty [Overview]:* • *Serious risk of bias*	
		Subgroup: Sole diet	135 (1 study) 67 vs 68	MD 0.30 [-0.85, 1.45]	n/a		

Outcome assessed: Long-term growth	Outcome assessed	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Walsh *et al* 2019[127]		Subgroup: Supplement to human milk	224 (1 study) 111 vs 113	MD -0.50 [-1.50, 0.50]	n/a		
		Test for subgroup differences: Chi^2=1.06, df=1 (P=0.30, I^2=6.1%					
	Subscapular skinfold thickness (mm) at 7.5 to 8 years post term	Nutrient-enriched formula versus standard formula	359 (2 studies) 178 vs 181	MD -0.05 [-0.67, 0.57]	0%	*Moderate certainty [Overview]:* • *Serious risk of bias*	
		Subgroup: Sole diet	135 (1 study) 67 vs 68	MD -0.10 [-0.96, 0.76]	n/a		
	Subscapular skinfold thickness (mm) at 7.5 to 8 years post term	Subgroup: Supplement to human milk	224 (1 study) 111 vs 113	MD 0.00 [-0.90, 0.90]	n/a		
		Test for subgroup differences: Chi^2=0.02, df=1 (P=0.88), I^2=0%					

Outcome assessed: Long-term growth	Outcome assessed	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Walsh *et al* 2019[127]	Waist-to-hip ratio at 7.5 to 8 years post term	Nutrient-enriched formula versus standard formula	359 (2 studies) 178 vs 181	MD -0.02 [-0.16, 0.12]	0%	*Moderate certainty [Overview]:* • *Serious risk of bias*	
		Subgroup: Sole diet	135 (1 study) 67 vs 68	MD -0.03 [-0.20, 0.14]	n/a		
		Subgroup: Supplement to human milk	224 (1 study) 111 vs 113	MD 0.01 [-0.26, 0.28]	n/a		
		Test for subgroup differences: Chi^2=0.06, df=1 (P=0.80), I^2=0%					
	Body mass index (kg/m²) at 7.5 to 8 years post term	Nutrient-enriched formula versus standard formula	359 (2 studies) 178 vs 181	MD 0.06 [-0.33, 0.44]	0%	*Moderate certainty [Overview]:* • *Serious risk of bias*	
		Subgroup: Sole diet	135 (1 study) 67 vs 68	MD 0.0 [-0.57, 0.57]	n/a		

Outcome assessed: Long-term growth	Outcome assessed	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Walsh *et al.* 2019[127]		Subgroup: Supplement to human milk	224 (1 study) 111 vs 113	MD 0.10 [-0.41, 0.61]	n/a		
		Test for subgroup differences: Chi^2=0.07, df=1 (P=0.80), I^2=0%					
Dempsey *et al.* 2019[142]	No longterm growth data	Banked preterm versus banked term human milk	Nil studies eligible	n/a	n/a	n/a	n/a
Brown *et al.* 2019[126]	No longterm growth data	Formula versus maternal breastmilk	Nil studies eligible	n/a	n/a	n/a	n/a
Basuki *et al.* 2019[147]	No longterm growth data	Dilute versus full-strength formula	Nil studies eligible	n/a	n/a	n/a	n/a
Premkumar *et al.* 2019[151]	No longterm growth data	Human milk-derived fortifier versus bovine milk-derived fortifier	Nil studies eligible	n/a	n/a	n/a	n/a

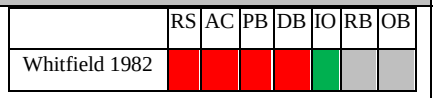
Outcome assessed: Long-term growth	Outcome assessed	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Nasuf *et al.* 2018[152]	No longterm growth data	Oropharyngeal colostrum (OPC) compared to control (water, saline or no intervention) in preterm infants	Not assessed	n/a	n/a	n/a	n/a
Abiramalatha *et al.* 2017[59]	No longterm growth data	High-volume vs standard-volume feeds	Nil eligible studies for this outcome	n/a	n/a	n/a	n/a
Tan-Dy *et al.* 2013[130]	No longterm growth data	Lactase treated feeds vs placebo	n/a	n/a	n/a	n/a	n/a
Nehra *et al.* 2002[129]	No longterm growth data	High MCT formula versus low MCT formula	n/a	n/a	n/a	n/a	n/a

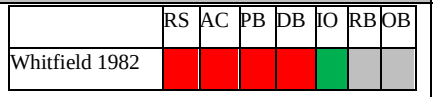
Table 52: Appendix 6: Summary table of effects of the interventions on long term growth measures: 'When to feed'

Outcome assessed: Long term growth	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Yeo *et al.* 2019[154]	Stopping feeds vs continuing feeds during transfusion	Not assessed	n/a	n/a	n/a	n/a
Abiramalatha *et al.* 2019a[155]	Routine monitoring of gastric residuals vs no routine monitoring of gastric residuals	Nil studies eligible	n/a	n/a	n/a	n/a
	Routine monitoring of gastric residuals quality vs routine monitoring of gastric residuals quality and quantity					
Oddie *et al.* 2017[136]	Slow versus faster rates of feed advancement	Nil studies eligible	n/a	n/a	n/a	n/a
Morgan *et al.* 2014[137]	Delayed versus early introduction of progressive feeding	Nil eligible studies	n/a	n/a	n/a	n/a

Outcome assessed: Long term growth	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Morgan *et al.* 2013[131]	Early trophic feeding vs enteral fasting	Nil studies eligible	n/a	n/a	n/a	n/a

Table 53: Appendix 6: Summary table of effects of the interventions on long term growth measures: 'How to feed'

Outcome assessed: Long term growth	**Outcome**	**Comparison**	**Number of subjects (studies)**	**Measure of effect (95% CI)**	**I^2**	**GRADE assessment**	**Risk of Bias**
Abiramalatha *et al.* 2019b[156]	Growth measures following discharge from hospital to latest follow up	Re-feeding vs discard gastric residuals	No eligible studies	n/a	n/a	n/a	n/a
Watson *et al.* 2013[135]	Weight measured at 3 months	Transpyloric versus gastric tube feeding	1 study	"no statistically significant differences"		*Very low certainty [Overview]* • *Very serious risk of bias (Quasi-randomised allocation, unblinded, unequal and considerable loss to follow up at EDD)* • *Very serious imprecision (narrative only)*	
	Head circumference measured at 3 months	Transpyloric versus gastric tube feeding	1 study	"no statistically significant differences"		*Very low certainty [Overview]* • *Very serious risk of bias (Quasi-randomised allocation, unblinded, unequal and considerable loss to follow up at EDD)* • *Very serious imprecision (narrative only)*	

Outcome assessed: Long term growth	Outcome	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Watson *et al.* 2013[135]	Weight measured at 6 months	Transpyloric versus gastric tube feeding	1 study	"no statistically significant differences"		*Very low certainty [Overview]* • *Very serious risk of bias (Quasi-randomised allocation, unblinded, unequal and considerable loss to follow up at EDD)* • *Very serious imprecision (narrative only)*	
	Head circumference measured at 6 months	Transpyloric versus gastric tube feeding	1 study	"no statistically significant differences"		*Very low certainty [Overview]* • *Very serious risk of bias (Quasi-randomised allocation, unblinded, unequal and considerable loss to follow up at EDD)* • *Very serious imprecision (narrative only)*	
Dawson *et al.* 2012[132]	None assessed	Push versus gravity bolus tube feeding	n/a	n/a	n/a	n/a	n/a

Outcome assessed: Long term growth	Outcome	Comparison	Number of subjects (studies)	Measure of effect (95% CI)	I^2	GRADE assessment	Risk of Bias
Premji *et al.* 2011[133]	None assessed	Continuous vs intermittent bolus milk feeding	Not assessed	n/a	n/a	n/a	n/a

Table 54: Appendix 6: Summary table of effects of the interventions on long term growth measures: 'Adjunctive strategies'

Outcome assessed: Long term growth	**Comparison**	**Number of subjects (studies)**	**Measure of effect (95% CI)**	**I^2**	**GRADE assessment**	**Risk of Bias**
Muelbert 2019[157]	Exposure to smell and taste stimulation of milk with tube feeds versus no exposure	Not assessed	n/a	n/a	n/a	n/a
Anabrees 2015[158]	Glycerin prophylaxis versus placebo/no intervention	Not assessed	n/a	n/a	n/a	n/a
Ng 2008[134]	Erythromycin vs placebo for prevention of feeding intolerance	Not assessed	n/a	n/a	n/a	n/a
	Erythromycin vs placebo for treatment of feeding intolerance	Not assessed	n/a	n/a	n/a	n/a
Greene 2016[159]	Oral stimulation vs standard care for promoting oral feeding	Not reported	n/a	n/a	n/a	n/a

Outcome assessed: Long term growth	**Comparison**	**Number of subjects (studies)**	**Measure of effect (95% CI)**	**I^2**	**GRADE assessment**	**Risk of Bias**
Greene 2016[159]	Oral stimulation vs non-oral intervention for promoting oral feeding	Not reported	n/a	n/a	n/a	n/a