Animal milks compared to follow-on formula, low-fat milk, plant-based milk or fortified milk and its associated outcomes in children 12-23 months of age

Ana C Fernández-Gaxiola,[1] Luz María De-Regil,[2] Sandra Gallegos-Lecona[1]

[1] National Institute of Public Health, Mexico City, Mexico
[2] LMD Inc, Toronto, Canada

Table of Contents

Abstract

Background: Milk is an important food for children during their second year of life. At 12 months of age, children that are no longer breastfeeding are usually given some type of animal milk, full-fat or lower-fat, as a standard practice for regular feeding and for the prevention of undernutrition. There are also other types of milks available for young children such as fortified milks that contain extra nutrients and plant-based milks are alternatives for children who have lactose intolerance or are allergic to human's or cow's milk proteins. Follow-on formulas marketed to children +1y are also available to partially satisfy nutritional requirements. However, evidence about their benefits to justify their proper use is lacking to recommend them for young children, and therefore, to ensure their optimal growth, health, and nutrition during their second year of life.

Objectives: To assess the effects of animal milk intake and its associated beneficial or harmful outcomes in young children 12-23 months of age, compared to follow-on formula marketed for children 1y+, lower-fat milk, plant-based milk alternatives or fortified milk.

Search methods: In September 2020, we searched CENTRAL, MEDLINE, Embase, eleven other databases, including CINAHL, IBECS and IMBIOMED. In addition, we examined reference lists, and contacted authors and known experts to identify additional studies that report a control group.

Selection criteria: We included randomised controlled trials (RCTs) with either individual or cluster randomisation, non-randomised controlled trials and comparative observational studies that report a control group. Participants were children 12-23 months of age, healthy with and without low birth weight, and without any health condition that impedes their normal growth The intervention was animal milk (full-fat or lower fat) compared with no other milk, follow-on milk, plant-based milk or fortified milk. We also compared full-fat animal milk versus lower-fat animal milk.

Data collection and analysis: Two review authors independently screened studies for inclusion, extracted data and assessed the risks of bias of included studies. We synthesised results narratively and conducted meta-analyses for outcomes relating to four intervention types. We assessed our confidence in the certainty of effect estimates with the GRADE framework as very low, low, moderate or high, and presented 'Evidence profile' tables.

Main Results: We included 19 studies (5,579 participants in the review, with two randomised at the cluster level, one RCT for which data could not be extracted, and five studies contributing to qualitative analysis. The main limitations of the studies were lack of blinding and selection bias. Studies were performed in high- and middle-income countries, eight studies were funded by private industry. Nine studies did not report any declaration of interests.

1. Animal milk (full-fat or lower fat) versus no other milk

A single study reported on this comparison but did not allow for data extraction and therefore was not included in the meta-analysis.

2. Animal milk (full-fat or lower-fat) versus follow-on formula

Available evidence showed that giving children 12-23 mo follow-on formula probably makes little or no difference to body weight compared to animal milk, when given during 4 to 12 mo and exposed to 300 ml or more per day (MD 0.13, 95% CI -0.11 to 0.36; 3 studies; 604 participants; moderate certainty evidence) nor to height compared to children given animal milk, when given during 4 to 12 mo and exposed to 300 ml or more per day (MD 0.20, 95% CI -0.31 to 0.72; 3 studies; 604 participants; moderate certainty evidence). Authors also found evidence that giving children 12-23 mo follow-on formula probably improves serum vitamin D concentrations compared to animal milk, by 16.27 more nmol of serum vitamin D per litre, when given during 5 to 12 mo and exposed to more than 150 ml/d (95% CI -21.23, -11.31; 455 participants; moderate certainty evidence)

Giving children 12-23 mo follow-on formula may also improve haemoglobin and ferritin serum concentrations, by 2.61 more grams of haemoglobin (95% CI -4.86 to -0.37; 663 participants; 5 studies; low-certainty evidence) and by 9.87 more µg of ferritn, per litre (95% CI -15.02 to -4.72; 1098 participants; 5 studies; low-certainty evidence) compared to children given animal milk.

A cross-sectional food consumption study was consistent documenting that daily consumption of 250 ml of cow milk in children 12-23 mo increased the risk of insufficient intakes for several nutrients which might be prevented by giving them follow-on formula instead. While total energy and macronutrient intakes were similar in the two groups, except protein intake of cow milk group which was much higher in the later compared to recommended intakes and significantly higher than children receiving follow-on formula. Two RCTs with non-randomized reference groups were also consistent showing not beneficial effect of follow-on formula on children´s growth and positive impact on haemoglobin concentrations.

3. Full-fat animal milk versus lower-fat milk

Results from a single study showed that giving children 12-23 mo low-fat milk may make little or no difference to serum cholesterol (MD -0.17 95% CI -0.92 to 0.58; 17 participants; 1 study; very low-certainty evidence), serum low-density lipoproteins (MD -0.25 95% CI -0.94 to 0.44; 17 participants;

1 study; very low-certainty evidence), serum high-density lipoproteins, serum triglycerides or LDL/HDL ratio compared to full-fat milk. A retrospective case-control study found that children on lower-fat milks were significantly less rasping; from 18 (SEM =1.5) to 9 (SEM=1.8) days a month (p<0.000). Also, days with fever (1.0 to 0.0 days a month, p< 0.000), coughing (18 to 10 days a month, p<0.000) and runny/blocked nose decreased significantly (18 to 11 days a month, p=0.008) in the intervention group.

4. Animal milk (full-fat or lower-fat) versus plant-based milk

Results from a single trial showed that giving children 12-23 mo plant-based milk may make little or no difference to serum cholesterol (MD -0.16 95% CI -0.76 to 0.44; 21 participants; 1 study; very low-certainty evidence), serum low-density lipoproteins (MD 0.03 95% CI -0.48 to 0.54; 21 participants; 1 study; very low-certainty evidence), serum high-density lipoproteins, serum triglycerides or LDL/HDL ratio compared to animal milk.

5. Animal milk (full-fat or lower-fat milk) versus fortified milk

We found evidence pooling studies in meta-analysis suggesting that giving children 12-23 mo fortified milk were less likely to have anaemia (RR 2.29 95% CI 1.12 to 4.69; 1324 participants; 3 studies; low-certainty evidence) and less likely to have iron deficiency (RR 1.21 95% CI 0.57, 2.56; 349 participants; 1 study; very low-certainty evidence) compared to those who received animal milk. Results showed that giving children 12-23 mo fortified milk may reduce plasma zinc concentrations, by 0.43 more µmol per litre than those given animal milk (95% CI 0.11 to 0.76; 115 participants; 2 studies; very low-certainty evidence). Giving children 12-23 mo fortified milk may also improve haemoglobin and ferritin concentrations, by 5.91 more grams per litre (95% CI -9.84 to -1.99; 1354 participants, 6 studies; low certainty evidence), and by 5.70 more µg of ferritin per litre (95% CI -7.49 to -3.92; 852 participants, 3 studies, low certainty evidence) and were 8 times less likely to have iron deficiency anaemia compared to those given animal milk.

Fortified milk may make little or no difference to children´s growth, body composition, serum iron, oral health (mean DMFS and caries free) and morbidity (respiratory episodes and diarrhoea episodes) compared to animal milk. A prospective cohort study documented an improvement at the end of the six months in the incidence of diarrhoea among the children receiving fortified milk (monthly average n= 70 children; 30.4 vs 25.5 episodes/100 children/month, P < 0.025) compared with children receiving animal milk.

Conclusions: Feeding children 12-23 mo with at least 250 ml of follow-on formula or fortified milk daily during at least 4-5 months can have a positive effect on anaemia and haemoglobin, and can improve vitamin D, serum iron and zinc. Providing children 12-23 mo with follow-on formula or fortified milk may have the same effect as animal milk on children´s growth, body composition and child development, and may make little or no difference on the nutrient status of other nutrients.

1. Background

Milk[a] is an important food during childhood for many children as it is one of the highest sources of energy, protein, calcium, phosphorus, and other micronutrients. Inappropriate feeding practices between the 12 and 23 months of age can increase mortality as well as result in immediate and lifelong growth and developmental shortfalls. Most children continue breastfeeding for at least two years. However, in practice, many children are fed with several types of milk or substitutes -animal or plant-based- fortified milks and follow-on formulas have been used to replace breastfeeding for children 12-23 months of age, including follow-on formulas marketed for children over 1y of age without a proper justification of their benefits over appropriate dietary practices. Current obesity trends and vegan practices have opened the discussion on whether it is appropriate to feed children in this age group with low-fat animal milks or plant-based milk substitutes.

Between 12 and 23 months of age, milk should cover around one third of the children's energy needs (Dewey 2003). Proteins, which serve as major structural component, are needed for new tissue growth and maintenance (i.e., protein synthesis and turnover). Proteins should contribute between 5 and 20% of the energy intake ("percentage of energy as protein, [PE%]") in young children 12-23 months of age (Hoppe 2005) (Michaelsen 2007). Fats and fatty acids during the first years of life are major determinants of growth (main energy source), infant development and long-term health, and provide essential fatty acids (Uauy 2009). Carbohydrates, another crucial energy source, should provide 9-14g/100 kcal (around 30%) of the energy of children 1-2 years of age (Stephen 2012) (Suthutvoravut 2015). These include mainly lactose corresponding to the high energy demands of the human brain, but also monosaccharides and disaccharides such as fructose, sucrose, galactose, as well as starch. Breastmilk is also a key source of vitamin A, calcium and other minerals supporting growth.

Failure to achieve appropriate feeding practices between the 12 and 23 months of age can increase mortality (Sankar 2015) as well as result in immediate and lifelong growth and developmental shortfalls. Globally, 144.0 million children under 5 years of age are stunted and 47.0 million children under 5 are wasted, of which 14.3 million were severely wasted. Stunting before the age of 2 years predicts poorer cognitive and educational outcomes in later childhood and adolescence and has significant educational and economic consequences at the individual, household, and community levels. For women, particularly, stunting in early life has been associated with a lower age at first birth and a higher number of pregnancies and children (WHO 2014a). After a child reaches two years of age, it is very difficult to reverse stunting that has occurred earlier (Victora 2008). Wasting, in turn, has been shown to increase the risk of death in childhood from infectious diseases such as diarrhoea, pneumonia and measles, and is an important precursor of stunting (WHO 2014b). Risk of anaemia is also greater in children who have had a late introduction to iron-rich foods or have been fed animal milk (low in iron content). Anaemia has been shown to be a public health problem for many decades in young children. Globally, 41.7% of children under-five have anaemia (World Bank 2017), with variations among different areas and local conditions (Stevens 2013). Approximately

[a] Milk is the normal mammary secretion of milking animals obtained from one or more milkings without either addition to it or extraction from it, intended for consumption as liquid milk or for further processing (FAO 1999). The composition of milk is unique to each species and is expected to meet the nutritional needs of each species.

half of cases of anaemia result from iron deficiency. However, anaemia may also be caused by other micronutrient deficiencies.

Recent evidence shows that young children who are undernourished in the first 2 years of life and who put on weight rapidly later in childhood and in adolescence are at higher risk of developing chronic diseases related to nutrition (Victora 2008).

In addition to the nutritional affects, young children's feeding practices may also influence gut and oral health. For example, feeding bottles increase the risk of oral diseases, such as mouth breathing, malocclusion, alteration of bite, and tooth decay (Avila 2015). Gastrointestinal reflux frequency also seems to be higher in bottle-fed infants compared to breasted children (Chen 2017).

The World Health Organization recommends that children receive appropriate complementary feeding with continued breastfeeding from the 12 to 23 months of life. Most children (>50%) continue breastfeeding for at least two years in 41 out of the 130 countries with data in the UNICEF Infant and Young Child Feeding database. However, for multiple reasons, several types of milk or substitutes -animal or plant-based- have been used to replace breastfeeding for children 12-23 months of age (Box 1).

Box 1. Characteristics of milks and milks substitutes use to feed children 12-23 mo

Human milk -our gold standard- contains 9 g protein/l to be compared with 34 g/l in cow's milk, and lactose content differs with 70 g/l in human and 48 g/l in cow's. Fat content is similar (about 38 g/l) but saturated fatty acids is considerably higher in cow milk compared with human milk. Milk protein is dominated by the whey fraction, which constitutes 60% of total protein, while the casein protein fraction constitutes 40%, IgA is by far the major immunoglobulin fraction in human milk and in general is higher than in other species. (Jensen 1995, Hernell 2011, Michaelsen 2007).

Animal milk -includes mainly that of cows or bovines, but also goat, sheep, camel, buffalo, donkey, and yak milk have been used, among others. However, cow's milk has traditionally been the first choice for regular feeding when young children are not breastfed and for the prevention and treatment of moderate and severe malnutrition in children as it is a rich and cheap source of protein, calcium and vitamin D. Free consumption of full-fat animal milk favours linear growth of children as potassium, magnesium, phosphorus and zinc, and the high lactose content also seems to support growth due to improved absorption of minerals. However, unmodified cow´s milk is a strong negative determinant of iron status. (Treck 2013, Agostoni 2011, Turck 2013, Mølgaard 2011, Martin 2011, Ziegler 1983, Ziegles 2007, Woldu 2014).

Low-fat milk -refers to the animal milk with reduced fat content (between 75-99-5% fat reduction). In cow milk with reduced fat, the PE% is very high (28 PE% in 2% milk and 39 PE% in skimmed milk), which is one of the reasons that reduced-fat milks are typically not recommended for infants and young children. While skimmed (non-fat) milk is not recommended as a major food source for children under two because it does not contain essential fatty acids, is deficient in fat-soluble vitamins and has a high potential renal solute load in relation to energy. Semi-skimmed milk may be acceptable after 12 months of age (WHO 2005, WHO 2011, Michaelsen 2014).

Follow-up or follow-on formulas -are milk or plant protein-based formulas intended to partially satisfy the nutritional requirements of young children aged 1-3 years. Their intake can increase the supply of some micronutrients in this specific age group. Since follow up formulas marketed for children 1y+ (in Codex terms) or young child formulae, in practice, frequently replace breastfeeding they are classified as breast milk substitutes. According to the CODEX standards, when prepared in accordance with the instructions for use, 100 ml of follow-up formula marketed for 6-36m shall provide not less than 60 kcal (or 250 kJ) and not more than 85 kcal (or 355 kJ). Should also contain between 1.6 to 2.7 g per 100/kcal of protein of nutritional quality equivalent to that of casein or a greater quantity of other protein in inverse proportion to its nutritional quality; between 3 and 6 g per 100 kcal of lipids with a level of linoleic acid not less than 300 mg per 100 kcal. Many formulas still present great differences in composition when compared to human milk (EFSA 2016, WHO 2018, FAO 1987, Mendonça 2017, Mazzocchi 2018).

Plant-based formulas -are an alternative for babies who have lactose intolerance or are allergic to human's or cow's milk proteins. The most common sources for these formulas are soy and rice hydrolysates. Many nutritional deficiencies with these formulas have been reported in the past, including a poor amino acid profile and are lower digestibility than animal protein. However, current soy formulas are supplemented with appropriate quantities of amino acids such as methionine, taurine, and carnitine and contain appropriate amounts of iron, zinc, calcium, phosphorus. There are concerns with the use of soy formulas and possible hormonal effects on the reproductive system presumed due to isoflavones present in soy protein. (Malunga 2014, Vandenplas 2014a, Verduci 2019).

Fortified milk -is that which contains extra nutrients that are not naturally found in milk in significant amounts and is not necessarily marketed to infants and children. Although there are compositional differences, fortified milk usually contains nutrients that infants and young children are at particular risk for deficiency. Fortified milk is a complementary feeding option.

Milk is an important food during childhood for many children as it is one of the highest sources of energy, protein, calcium, phosphorus, and other micronutrients (Campbell 2017) (Fox 2006). WHO maintains that breastfeeding remains the most appropriate liquid part of a progressively diversified diet for the vast majority of children between 6 and 24 months of age, once complementary feeding has begun. Some paediatric organizations adhere to this position and encourage continued breastfeeding, but also offer as alternative the provision of 500 mL per day of homogenized (3.25% M.F.) cow's milk (Critch 2014). However, in practice, many children are fed with follow-on formulas marketed for children over 1y of age without a proper justification of their benefits over appropriate dietary practices and current obesity trends and vegan practices have opened the discussion on whether it is appropriate to feed children in this age group with low-fat animal milks or plant-based milk substitutes. Evidence is needed for adequate recommendations to ensure adequate growth, health, and nutrition of young children during their second year of life, and fulfilment of the first 1,000 days period.

The objective of this review was to assess the effects of animal milk intake and its associated beneficial or harmful outcomes in young children 12-23 months of age, compared to follow-on formula marketed for children 1y+, lower-fat milks, plant-based milk alternatives or fortified milk.

2. Methods

2.1 Criteria for considering studies for this review

We aimed to include: RCTs, with randomisation at either the individual or cluster level; non-randomised controlled trials (where allocation of treatment has been made, for example, by alternate allocation, date of birth, or alphabetical order); comparative observational studies that report a control group; controlled cohort studies (prospective and retrospective); controlled before-and-after studies with at least two control and two intervention sites; interrupted time series with at least three measure points both before and after the intervention.

2.2 Type of interventions

We included interventions involving the consumption of some type of milk, including all types of animal milks or milk substitute as part of young children´s daily diet and performed the following comparisons:

1. Animal milk (full-fat or lower fat milk) versus no other milk;
2. Animal milk (full-fat or lower fat milk) versus follow-on formula marketed for children 1y+;
3. Full-fat animal milk versus lower fat milk;

4. Animal milk (full-fat or lower fat milk) versus plant-based milk alternatives;
5. Animal milk (full-fat or lower fat milk) versus fortified milk (full-fat or lower fat)

We included studies with co-interventions, such as counselling or nutritional advice, if they were the same in both the intervention and comparison groups, either by design or statistical analyses.

2.3 Outcomes

Outcomes were selected based on the relevance to the intervention and prioritized in consultation with the guideline development group (GDG) to determine the critical and important outcomes. According to outcomes reported in included trials and to GRADE framework, seven critical outcomes were included in the evidence profile tables.

Critical outcomes included:

1. Growth (weight (kg), height (cm), head circumference (cm), WAZ, HAZ)
2. Body composition (% fat body, or as defined by trialists)
3. Long-term food preferences/dietary patterns (as defined by trialists)
4. Longer-term outcomes (NCDs) (as defined by trialists)
5. Nutrient status (including fatty acids, triglycerides and lipoproteins, and micronutrients concentrations as defined by trialists)
6. Child development (as defined by trialists)
7. Anaemia (haemoglobin concentration below a cut-off defined by the trialists, adjusted by altitude, as appropriate, or as defined by trialists)
8. Iron deficiency anaemia (IDA, as defined by trialists)
9. Iron deficiency (ID, as serum ferritin <12 μg/L or as defined by trialists)
10. Mean haemoglobin (g/L)
11. Ferritin (μg/L)

Important outcomes included:

1. Nutrient intakes (sufficient, excessive as defined by trialists)
2. Feeding practices (e.g., use of feeding bottles / cup, feeding methods, eating habits, or as defined by trialists)
3. Oral health (dental caries)
4. Morbidity (e.g., diarrhoea, gastroesophageal reflux, as defined by trialists)
5. Dietary diversity (proportion of children 12-23 mo. of age receiving foods from 4 or more food groups: 1) cereals, roots and tubers, 2) nuts and legumes, 3) dairy products, 4) animal meats, 5) eggs, 6) fruits and vegetables rich in Vitamin A and 7) other fruits and vegetables (UNICEF 2007); or as defined as trialists
6. Gut health (e.g., intolerance, regular bowel movements, gastrointestinal discomfort, as defined by trialists)
7. Allergy (gastrointestinal milk allergy IgE mediated, as defined by trialists)
8. Phyto-oestrogen related outcomes (e.g., plasma concentrations of isoflavones such as genistein and daidzein, as defined by trialists.

2.4 Search strategy and selection of studies

We searched the following electronic databases: Cochrane Central Register of Controlled Trials (CENTRAL), MEDLINE, Embase, CINAHL EBSCOhost, Web of Science (ISI) SCI, SSCI, CPCI-exp & CPCI-SSH, ClinicalTrials.gov, WHO International Clinical Trials Registry Platform (ICTRP), IBECS, ScIELO, African Index Medicus, WHOLOS, IndMED, Native Health Research Database. We searched available studies with no limits and up to September 2020 (see search strategy in Appendix 1). We searched available studies with no limits and up to September 2020.

Secondary reference searching was conducted on all studies included in the review as well as on previously published review articles on related topics. In addition, experts in the field were contacted to identify unpublished articles and accepted conference abstracts.

Screening and selection of studies

All records were screened using Covidence systematic review software (Covidence 2017). Two reviewers independently screened titles and abstracts of all records yielded by the searches against the selection criteria. Full-text reports of all relevant or potentially relevant studies that seemed to meet the inclusion criteria were assessed for eligibility. Disagreements were resolved through discussion and consensus. Authors recorded their decisions of the selection process in a PRISMA diagram (Moher 2009).

2.5 Data extraction and management

Data from eligible studies was extracted using a form designed to collect detailed data for this review. Authors resolved any discrepancies through discussion. If the information regarding any of the studies was unclear, we attempted to contact the authors of the original reports, to ask them to provide further details. We completed the data collection form electronically and recorded information on:

1. Study identification: authors, year of publication, references to study
2. Study design: description of setting and participants, description of interventions, exposure details, assessment of risk of bias;
3. Outcomes: details of how and when measured, and results, conclusions, and limitations.
4. Others: funding source.

2.6 Quality assessment

Two reviewers independently assessed the risk of bias in each included study using a simple contingency form that follows the domain-based evaluation (ROB 2.0) (Sterne 2019) (sequence generation; allocation concealment; blinding; incomplete outcome data; selective reporting bias; and other sources of bias, etc.). If there was insufficient information to assess the risk of bias, we rated the domain at 'unclear risk of bias', until further information is published or made available to us. If there was sufficient information, we categorised the domain as being either at 'low risk of bias' or 'high risk of bias' accordingly. We resolved any disagreements by discussion.

The risk of bias of observational studies was be assessed using the ROBINS-I tool (Sterne 2016). This tool considers bias due to confounding, in selection of participants into the study, in classification of interventions, due to deviations from intended interventions, due to missing data, in the measurement of outcomes, and in the selection of the reported results.

2.7 Data analysis and reporting

We included cluster-randomised studies with individually randomised studies in the analyses. Cluster-randomised studies were labelled with a (C). For studies with more than two intervention groups (multi-arm studies), we included the directly relevant arms only. When we identify studies with various relevant arms, we combined the groups into a single pair-wise comparison (Higgins 2011b) and included the data in the corresponding subgroup category. If the control group was shared by two or more study arms, we divided the control group (events and total population) over the number of relevant subgroup categories to avoid double counting the participants. We carried out analyses, as far as possible, on an intention-to-treat basis, i.e., by attempting to include all participants randomised to each group in the analyses.

We analysed the results from controlled non-randomised and comparative observational study designs separately from randomised study designs.

Reviewers carried out statistical analyses using RevMan 5.4.1 (Review Manager 2014). We used random-effects meta-analyses due to possible heterogeneity in the interventions, populations and methods used in different trials. We used Mantel-Haenszel weighting for dichotomous outcomes and inverse variance for continuous outcomes, to adjust the effect measure according to the extent of its variation both between and within studies. We combined adjusted estimates using the generic inverse variance (GIV) option in RevMan 5.4. We presented dichotomous outcome data as average risk ratios (RRs) with 95% confidence intervals (CIs). We presented continuous outcome data as mean differences (MD) with 95% CIs, measured at the end of the intervention.

Assessment of heterogeneity and of reporting biases

We examined the forest plots from meta-analysis to look for heterogeneity among studies and used the I^2 and T^2 statistics to quantify the level of heterogeneity among the trials in each analysis. If we identified substantial heterogeneity (I^2 greater than approximately 50%) we noted this in the text and explored, it by pre-specified subgroup analysis (see Subgroup analysis and investigation of heterogeneity). We would advise caution in the interpretation of those results where there were high levels of unexplained heterogeneity.

When reviewers suspect reporting bias (see Assessment of risk of bias in included studies), we attempted to contact the study authors to ask them to provide missing outcome data. We investigated reporting biases (such as possible publication bias) using funnel plots, assessing asymmetry visually.

Assessing the certainty of the evidence

We used the GRADE approach to interpret findings (Balsheim 2011). For assessments of the overall certainty of evidence for each outcome that included pooled data from included studies from RCTs only, we downgraded the evidence from 'high certainty' by one level for serious (or by two for very serious) study limitations (risk of bias), indirectness of evidence, serious inconsistency, imprecision of effect estimates (the number of people analysed, for example) or potential publication bias. Data from observational studies started at low certainty. This assessment was limited only to the studies included in this review and as we did not consider there was a serious risk of publication bias, we

did not downgrade in this domain. We used the GRADE profiler (GRADEpro GDT 2015) to create 'Evidence Profile' tables.

2.8 Subgroup analysis and investigation of heterogeneity

When data was available (at the study or national level) and it was appropriate, we carried out the following subgroup analyses on all primary outcomes, to look for possible differences between studies:

1. Stunting (prevalence at the study level or national level) (de Onis 2018)
    - Non-specified
    - Very low prevalence <2.5%
    - Low prevalence (2.5-9.9%)
    - Medium prevalence (10-19.9%)
    - High prevalence (20-30%)
    - Very high prevalence (>30%)

2. Wasting (prevalence at the study or national level) (de Onis 2018)
    - Non-specified
    - Very low prevalence <2.5%
    - Low prevalence (2.5-4.9%)
    - Medium prevalence (5-9.9%)
    - High prevalence (>10%)

3. Type of feeding before 12 mo:
    - Non-specified
    - Breastmilk only
    - Breastfeeding + complementary feeding
    - Breastmilk substitute + complementary feeding
    - Breastfeeding +breast milk substitute + complementary feeding

4. Anaemia (prevalence at study or national level) (WHO 2001)
    - Non-specified
    - No public health problem (0-4.9%)
    - Mild prevalence (5 -19.9%)
    - Moderate prevalence (20-39.9%)
    - Severe prevalence (40% or higher)

5. Funding source
    - Private industry
    - Other

6. Authors' declaration of interests

2.9 Sensitivity analysis

We conducted sensitivity analyses ad hoc to examine the potential effect of clustering on the CI of the summary estimates, by removing cluster-RCTs from the analyses and comparing the effects. For cluster-randomised trials, we carried out sensitivity analysis using a range of ICC on overall effect

estimate and have reported these effects. We also conducted sensitivity analyses to assess the impact of including studies with high levels of missing data in the overall assessment of treatment effect.

3. Results

3.1 Search results

Our search strategy identified 4,065 references (4,033 after removing duplicates) for possible inclusion. We screened 88 articles in full text for potential inclusion. We included 19 studies (36 references) and all were reported in English. We excluded 50 studies (52 records) with reasons and identified one ongoing study and three studies awaiting classification. Of the 19 included studies, 13 RCTs contributed to meta-analysis. We have summarised the study selection process in Figure 1.

Figure 1. Study selection

3.2 Characteristics of included studies

We included 19 studies involving 5,579 participants all of which met pre-established inclusion criteria. All studies were published between 1986 and 2018. Two studies: Rivera 2010 (C) and Stecksén-Blicks 2009 (C) were randomised at cluster level. We have only included the estimated effective sample size in the analysis, after adjusting the data to account for the clustering effect which both authors reported to have adjusted for in their results.

Table 1. Characteristics of included RCTs and their interventions

Study	Country	Setting	Total sample randomized	Study population age*	Intervention
Akkermans 2016	Germany, UK & Netherlands	Paediatric clinics and children's hospitals	318	12-36 mo	Follow-up for 20 weeks. Follow-on formula contained 1.2 mg Fe/100 ml (form not specified) and 1.7 mg vitamin D/100 ml per day. Control product was a non-fortified cow milk that contained 0.02 mg Fe/100 ml and no vitamin D. The energy concentrations of both products were comparable (46 kcal/100 ml for cow milk and 50 kcal/100 ml for follow-on formula). Both products were supplied in powdered form with instructions for preparing the milk by diluting the powder with water.
Bhatnagar 1996	India	Diarrhoea units of the All-India Institute of Medical Sciences (AIIMS) and Kasturba Hospital	96	3-24 mo, with persistent diarrhoea for 14 days to 12 weeks	Children were assigned to 1 of 2 groups: 1 received milk-based cereal dietary regimen, and 2 received milk-free cereal dietary regimen. Both diets were isocaloric (86.9 calories/100 g for 9 months; 95.6 kcal/100g for >9 months) consisting of puffed rice cereal, sugar, and oil differing in only their source of protein, which was either milk or egg white, respectively. Both diets were offered at the rate of 150 kcal/kg per day.
Sazawal 2007		Peri-urban community in New Delhi	633	12-36 mo, both anaemic and non-anaemic	Follow-up for 1 year. Children were randomly assigned to 1 of 2 groups: fortified milk vs control milk. Single serving 32 g sachets were used to fortify milk. Fortified milk provided additional 7.8 mg zinc, 9.6 mg iron, 4.2 g selenium, 0.27 mg copper, 156 g vitamin A, 40.2 mg vitamin C, 7.5 mg vitamin E per day (three feeds). Assistants delivered 21 sachets each week to each home and advised that the child should consume up to three sachets a day.
Daly 1996	UK	Inner city area	100	6-18 mo	Follow-up until 24 mo. Children were assigned to 1 of 2 groups: Group 1: received iron fortified milk, supplied free of charge. Group 2: would continue with cow milk and recruited from a single care centre. Parents received an equivalent monthly payment to purchase it.
Morley 1999		Children's centres	493	9 mo	Follow-up until 18 mo. The formula milks were supplied in powdered form; tins of iron fortified formula were labelled "formula 28" and tins of unfortified formula were labelled "formula 61". This code was not revealed by the manufacturers until the study was completed and all data had been entered and checked. Powdered milk was supplied ad libitum to the subjects' homes, and parents were given written and verbal information on how to make up the milk.
Maldonado 2007	Spain	Faculty of Medicine, University of Granada	33	12-30 mo	Follow-up for 4 mo. Children were randomly assigned to 1 of 2 groups: 1 received 500 ml iron follow-on formula and 2 received 500 ml cow milk. All included children had received breastfeeding for 6 to 8 months and follow-on formula thereafter until the first year of life. Intervention milk contained 1.2 mg/100 ml Fe (form not specified).
Lovell 2018	Australia & New Zealand	Urban children centres	160	12 mo	Follow-up until 24 mo. Intervention milk "GUMLi" had a reduced energy and protein content when compared to standard, commercial follow-on formula on the market, 60 kcal/100ml vs. 71 kcal/100ml and 1.7 g/100ml protein vs. 2.2 g/100ml. GUMLi was also fortified with iron (1.7 mg Fe/100ml form not specified), vitamin C and vitamin D, probiotics, and prebiotics. Control cow milk was energy matched with a protein content of 3.1 g/100ml.
Symleck-Gay 2009	New Zealand	Urban centres	225	12-20 mo	Follow-up for 5 mo. Fortified cow milk had iron as ferrous sulphate, calcium, magnesium, zinc, vitamin C, vitamin E, niacin, vitamin A, vitamin D, vitamin B-6, thiamine, and folate (Heinz Nurture Toddler Enriched Milk

					Drink; Heinz Wattie's Ltd, Hastings, New Zealand. Non fortified (Standard Instantized Whole Milk Powder with required A and D added, Fonterra, Auckland, New Zealand. The milks were packaged (Sutton Group Ltd, Auckland, New Zealand) into identical white 900-g cans (Canpac International, Hamilton, New Zealand) along with identical 17-mL scoops.
Rivera 2000 (C)	México	Marginal and periurban communities, beneficiaries of social programs	567	12-30 mo	Follow-up for 6 mo. -400 ml of fortified milk (48 g powder) provided 5.28 mg iron, 5.28 mg zinc, 48 mg vitamin C and 32.1 mcg folic acid. -400 ml of unfortified milk (48 g powder) provided 0.16 mg iron, 1.6 mg zinc, 6.8 mg vitamin C and 24 mcg folic acid. Other nutrients content was similar.
Villalpando 2009			115	10-30 mo	Fortified cow´s milk contained 5.8 mg/400 mL of iron as ferrous gluconate, 5.28 mg/400mL of zinc as zinc oxide, and 48 mg/400mL of ascorbic acid. Unfortified cow´s milk contained 0.2 mg iron/400 mL, 1.9 mg zinc/400 mL, and 6.8 mg ascorbic acid/400 ml. Units of 220 g of the product were packed in metallic foil sachets. The packages of fortified milk and unfortified milks were undistinguishable, except for a colour-coded band in the upper corner of the sachet.
Stekel 1986	Chile	Community clinics of the National Health Service	510	3 mo	Follow-up until 15 mo. Fortified milk had 15 mg iron as ferrous sulphate per 100 gr or milk powdered. Fortified milk powdered was distributed in cans. Mothers received 3 kg milk product per month until child was 6 mo old, and 2 kg per month until the child was 24 mo old. Mean breastfeeding duration for the participant children was 4 months.
Stekel 1988			554	3 mo	Follow-up until 15 mo. Fortified milk was full fat (26%) powdered milk with 15 mg of Fe as ferrous sulphate, 100 mg of ascorbic acid, 1500 IU of vitamin A, and 400 IU of vitamin D per 100 g of powder. Children were clinically followed every 15 days by the same group of physicians.
Stecksén-Blick 2009 (C)	Sweden	Small communities with less than 10,000 inhabitants	248	12-60 mo	Follow-up for 21 mo. Children were served 150 ml medium-fat milk (1.5%) at lunch. Milk was prepared by the day care staff by adding one colour-coded capsule (10 ml) to each litre of milk. The capsules were kept frozen and contained fluoride and probiotic bacteria in skim milk to give a final concentration of 2.5 mg fluoride and 10 7 CFU/ml rhamnosus LB21 per litre in the intervention group. The capsules in the control group contained only skimmed milk and were identical in appearance except in colour code.
Svahn 1999		Day care centres	38	11 mo	Follow-up from 12 to 18 mo. Children were randomly assigned to 1 of 4 groups: Group 1: received low-fat milk (LF) (1.0 g fat/dl, 3.3 g protein/dl); and products including low-fat yogurt, sour milk, cream, margarine, cheese (17% fat), and butter (for frying) could be used. Group 2: received standard-fat milk (SF) (3.5 g fat/dl, 3.3 g protein/dl); and products including full-fat dairy products Group 3: received partially vegetable fat and protein-reduced milk (PVF) (3.5 g fat/dl, 50% vegetable; 2.2 g protein/dl); and low-fat dairy products. This milk was fortified with 7.0 mg Fe l-1 as ferrous gluconate. Group 4: received full-vegetable-fat milk (FVF) (3.5 g fat/dl, 100% vegetable; 3.0 g protein/dl); and dairy substitutes with vegetable fats. This milk was fortified with 14.9 mg F l-1 as ferrous lactate.

*Participants were children age between 12 and 24 months at the time of intervention. For trials with children outside this age range, we included studies where we were able to disaggregate the data for children aged 12 to 24 months, or when mean age was <24 months. We included trials with apparently healthy children, born at term or ≥ 36 weeks of gestation, with healthy weight.

Characteristics of milks in the included RCTs

Animal milk

Animal milk was specified as "cow´s milk" in all trials. One study specified the use of full-fat or low-fat cow´s milk (Van der Gaag 2015). Daly 1996 specified using pasteurised milk. Ghisolfi 2012 specified 70% of the children used semi-skimmed milk. Stecksén-Blicks 2009 referred to "standard

milk" and "medium fat milk" (i.e., 1.5% fat). Lovell 2018 specified using pasteurised and homogenized cow milk. Villapando 2006 used "whole" cow milk for control.

Twelve studies used powdered milks, whether the intervention or control groups were given cow milk (Akkerman 2016, Chatchatee 2014, Gill 1997, Lovell 2018, Maldonado 2007, Morley 1999, Rivera 2010 (C), Sazawal 2007, Stekel 1986, Stekel 1988, Van de Gaag 2009, and Villalpando 2006) in their study groups. Two studies did not specify the form of milk used (Daly 1996 and Ghisolfi 2012). In Svahn 1999, low-fat animal-milk was 1.0 g fat/dl, standard fat animal-milk was 3.5 g fat/dl, as well as the partially vegetable fat and full-vegetable fat milks. Bhatnagar 1996 used milk-based cereal and milk-free cereal dietary regimens, that consisted of puffed rice cereal+sugar+oil differing only in the source of protein (milk vs. egg white). Nutrient profile of animal milk of the included studies is given in Table 2.

Follow-on formula

All included studies used powdered follow-on formula and provided them free of charge on a periodic basis. Chatchatee 2014 used a follow-on formula with added short-chain galacto-oligosaccharides and polyunsaturated fatty acids. In Lovel 2018, the follow-on formula had a reduced protein content ("GUMLi") with pre- and probiotics. Each of the included studies had different levels of micronutrient concentrations per 100 ml of reconstituted milk, and we have given details of the micronutrient profile in Table 3.

Akkerman 2016 provided the minimum dose used among all the included studies providing a dose of 150 ml of follow-on formula per day. Chatchatee 2014 provided the maximum dose used among all the included studies providing a dose of 400-750 ml of follow-on formula per day. In Daly 1996 and Maldonado 2007 dose used was 500 ml per day. Daly 1996 adapted the recommended dose depending on children´s age. Ghisolfi 2012 provided a minimum of 250 ml of follow-on formula per day. Gill 1997 and Morley 1999 did not specify the dose of follow-on formula (nor animal milk) used. Lovell 2018 provided 300 ml per day.

The duration of the interventions providing follow-on formulas was 4 months in Maldonado 2007; 5 months in Akkerman 2016; 9 months in Gill 1997 and Morley 1999; 12 months in Chatchatee 2014, Daly 1996 and Lovell 2018. Ghisolfi 2012 cross sectional study lasted 3 days.

Plant-based milks

A single trial provided plant-based milk. Svahn 1999 provided partially-vegetable fat and full-vegetable fat milk diets. The partially- vegetable fat milk diet consisted on a full-fat milk (fat 3.5 g/dl; 50% vegetable fat from low euric acid rapeseed oil and 50% cow's milk fat; LA 0.41 g/100 ml, ALA 0.20 g/100 ml) and the full-vegetable fat diet consisted on a full-fat milk (fat 3.5 g/dl; 100% vegetable fat from palm, coconut and soybean oil; LA 0.62 g/100 ml, ALA acid 0.05 g/100 ml) and were specially prepared for the study. The duration of the intervention in Svahn 1999 was 6 months.

Fortified milk

Five studies provided fortified milk in powder form (Rivera 2010 (C), Stekel 1986, Stekel 1988, Szymlek-Gay 2009 and Villalpando 2006). Two studies used milk in liquid form (Sazawal 2007 and Stecksén-Blicks 2009 (C)) and were fortified at the time they were used. Svahn 1999 used milks ready

to be feed. Stekel 1986 and Stekel 1988 used semi skimmed milk (12% fat) and full-fat milk, respectively. Svahn 1999 used both full-fat and low-fat milk. Villapando 2006 used "whole" fortified milk and Rivera 2010 (C) used full-fat cow milk.

Five studies used fortified milk with multiple micronutrients (Rivera 2010 (C), Sazawal 2007, Stekel 1988, Szymlek-Gay 2009 and Villalpando 2006). Stekel 1986 used milk fortified with iron, vitamin C, and was acidified without additives. Iron fortifiers were ferrous sulphate in Stekel 1986, Stekel 1988 and Szymlek-Gay 2009, and ferrous gluconate in Villalpando 2006. Svahn 1999 used milks with different total fat, saturated fat, monounsaturated and long-chain polyunsaturated fatty acids (LC-PUFAs), linoleic acid and trans fatty acids contents. In Sazawal 2007, milk was fortified using 32g single serve sachets which were delivered every week. Milk assistants delivered 21 sachets at home and advised the mother to feed the child 3 sachets a day. Stecksén-Blicks 2009 (C) used colour-coded capsules, that were kept frozen, and medium fat milk (1.5% fat) to prepare the fortified milk every day. Capsules contained fluoride and probiotic lactobacilli. We have given details of the micronutrient profile of the fortified milks of the included studies in Table 4.

Rivera 2010 provided 400 ml of fortified milk per day, suggesting 200 ml in the morning and 200 ml at night. Stecksén-Blicks 2009 (C) used a daily dose of 150 ml and Villalpando 2006 used 400 ml of fortified milk. Szymlek-Gay 2009 asked participants to replace their regular milk intake with either commercially available iron-fortified powdered cow milk fortified or non-fortified powdered cow milk. Stekel 1986, Stekel 1988 and Svahn 1999 did not specified any dosage of milk used or provided. Sazawal 2007 provided 21 sachets per week to provide the child 2-3 sachets per day.

Funding sources

The private industry supported eight studies. Three studies (Akkerman 2006, Chatchatee 2014 and Lovell 2018) were financed by Danone Nutrition Research, two studies (Gill 1997 and Morley 1999) by Wyeth Laboratories, and one study (Daly 1996) by Farley Health Products (acquired by Heinz in 1994). Ghisfoli 2012 declared receiving no financial support from any baby food company or from the French Association of Baby Food Industries or any other public or private support.

Five studies were financed by the academy, government, and non-governmental organizations (Bhatnagar 1996, Rivera 2010 (C), Stekel 1986, Stekel 1988 and Villalpando 2006). Bhatnagar 1996 was financed by the Diarrheal Diseases Control Programme, World Health Organization. Stekel 1986 and Stekel 1988 were financed by the Research Corporation USA and by the Chilean Ministry of Health, The Consejo Nacional para la Alimentacion y Nutricion, the United Nations University, and the Departamento de Investigacion y Bibliotecas de la Universidad de Chile, respectively. Rivera 2010 (C) received a grant from Secretary of Social Development and Liconsa (Leche Liconsa) and Villalpando 2006 was financed by the Ministry of Social Development and the National Institute of Public Health.

In Akkerman 2006, the study products were produced, provided, and coded (for blinding purposes) by Nutricia Cuijk (commissioned by Danone Nutricia Research) and in Szymlek-Gay 2009 the Health Research Council of New Zealand and Heinz Watt provided the formulas. In Stecksén-Blicks 2009 (C) the probiotic strain was provided by Essum AB, Umeå, Sweden, the fluoride solution was prepared at the university biochemical laboratory and the capsules were produced at the local dairy (Norrmejerier, Umeå, Sweden). The study was supported financially by the County Council of

Västerbotten (TUA) and the Borrow Foundation, UK. Norrmejerier Ekonomisk Förening, Umeå, Sweden supported the study by preparation and distribution of the milk. Svahn 1999 was supported by The Albert Pahlsson Foundation. Funding sources were not mentioned in one trial (Van de Gaag 2015).

Lovell 2018 specified that the founders had no role in the design of the study, nor in the collection, analyses, or interpretation of data; nor in the writing of the manuscript; or in the decision to publish the results. Gill 1997 stated that all statistical analyses were performed by Wyeth Laboratories.

Declarations of interests

In Akkerman 1996 two out of the five authors declared no conflicts of interest. In Chatchatee 2014 two out of nine authors declared no conflicts of interest. In Ghisolfi 2012 three out of the six authors declared no conflicts of interest. In Lovell 2018 four out of the six authors declared no conflicts of interest. All authors in Gill 1997 received a study grant and formula from SMA Nutrition.

Three studies declared no competing interests for all authors (Rivera 2010 (C), Sazawal 2007, Szymlek-Gay 2009). Nine studies did not report any declaration of interests (Bhatnagar 1996, Daly 1996, Maldonado 2007, Morley 1999, Stecksén-Blicks 2009 (C), Stekel 1986, Stekel 1988, Van der Gaag 2015 and Villalpando 2006).

3.3 GRADE systematic review findings

See Appendix 2 for full risk of bias assessment of included RCTs. See Appendix 3 complete data and analysis. See Appendix 4 for complete GRADE evidence profiles. See Appendix 5 for complete characteristics of included studies and risk of bias tables.

3.3.1 Comparison 1: Animal milk (full-fat or lower-fat milk) versus no other milk

One RCT (Bhatnagar 1996) looked at animal milk versus no other milk on children with persistent diarrhoea but did not provided data to extract and we could not undertake meta-analysis for comparison 1.

3.3.2 Comparison 2: Animal milk (full-fat or lower-fat milk) versus follow-on formula

There were five studies (796 participants) included in this comparison (Akkerman 2016, Daly 1996, Lovell 2018, Maldonado 2007, and Morley 1999). These studies comprise all the data included in the synthesis of this review. We included one non-randomised study in this comparison (Ghisolfi 2012) and two RCTs (Chatchatee 2014 and Gill 1997) with non-randomized animal milk group for qualitative assessment.

Two studies met the prespecified criteria mentioned above for being at lower risk of bias (Akkerman 2016 and Morley 1999) and in sensitivity analyses these trials were retained in the analysis whilst trials at higher risk of bias (Daly 1996 and Maldonado 2007) were temporarily removed to examine whether they had any impact on the overall pattern of results.

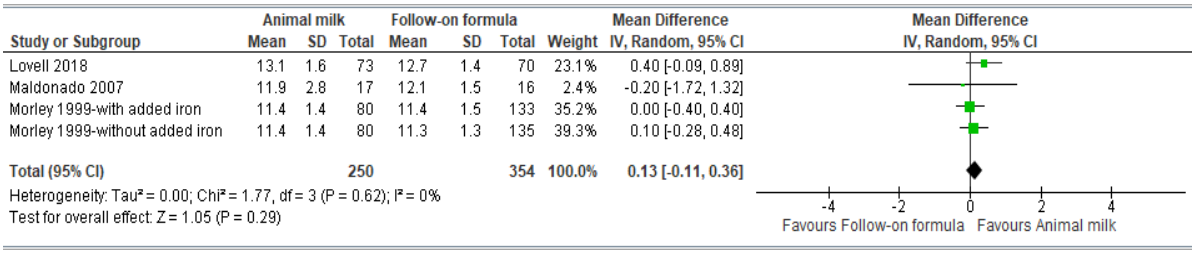
Critical outcomes

Growth

Weight (kg)

Three studies reported on this outcome (Lovell 2018, Maldonado 2007, Morley 1999). We pooled these studies in a meta-analysis and results showed that giving children 12-23 mo follow-on formula probably makes little or no difference to body weight compared to animal milk, when given during 4 mo and exposed to 300 ml or more per day (MD 0.13, 95% CI -0.11 to 0.36; 3 studies; 604 participants; moderate certainty evidence) (Figure 2). The heterogeneity was low (T2 = 0.00; Chi2 = 1.77 (P = 0.62); I2 = 0%).

Figure 2. Animal milk vs Follow-on formula on weight of children 12-23 mo

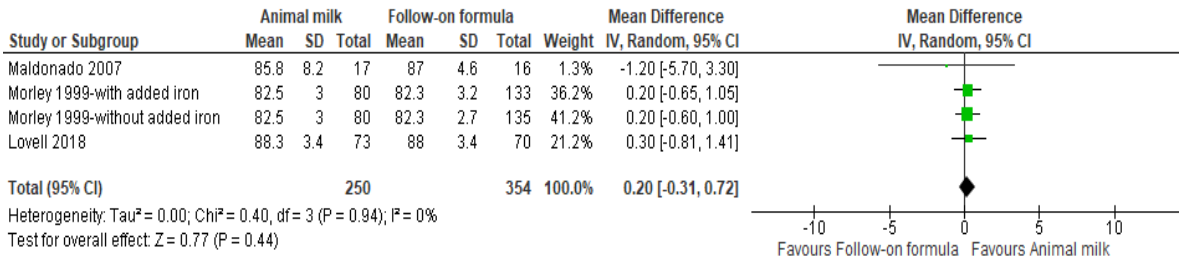


We conducted a subgroup analysis and found evidence that giving children 12-23 mo follow-on formula probably makes little or no difference to body weight compared to animal milk by the prevalence of stunting or wasting, the type of feeding before 12 mo or the prevalence of anaemia. We found evidence to suggest that follow-on formula makes little or no difference to body weight compared to animal milk by the funding source.

Height (cm)

Three studies reported on this outcome (Lovell 2018, Maldonado 2007, Morley 1999). We pooled these studies in a meta-analysis and found evidence that giving children 12-23 mo follow-on formula probably makes little or no difference to height compared to animal milk, when given during 4 to 12 mo and exposed to 300 ml or more per day (MD 0.20, 95% CI -0.31 to 0.72; 3 studies; 604 participants; moderate certainty evidence) (Figure 3). The heterogeneity was low (T2 = 0.00; Chi2 = 0.40 (P = 0.94); I2 = 0%).

Figure 3. Animal milk vs Follow-on formula on height of children 12-23 mo



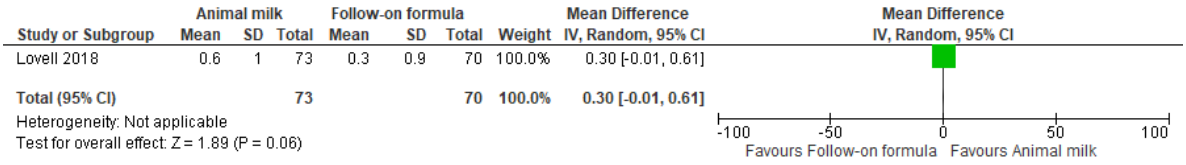
We conducted a subgroup analysis and found evidence that giving children 12-23 mo follow-on formula probably makes little or no difference to height compared to animal milk by the prevalence of stunting or wasting, the type of feeding before 12 mo or the prevalence of anaemia. We found

evidence to suggest that follow-on formula probably makes little or no difference to height compared to animal milk by the funding source.

Weight-for-height (WHZ)

A single study reported on this outcome (Lovell 2018). It found evidence that giving children 12-23 mo follow-on formula may make little or no difference to weight-to-height compared to animal milk, when given during 12 mo and exposed to 300 ml/d (MD 0.30, 95% CI -0.01 to 0.61; 143 participants; low certainty evidence) (Figure 4).

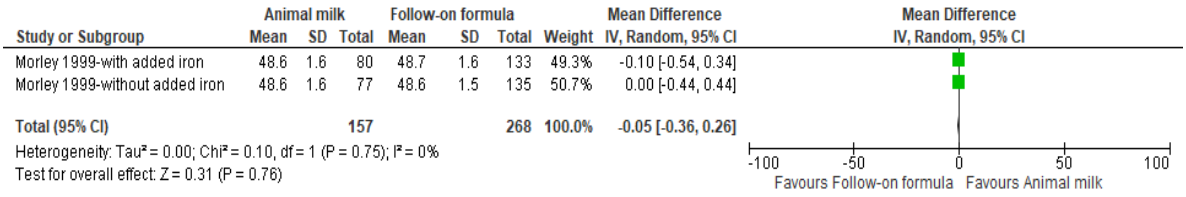
Figure 4. Animal milk vs Follow-on formula on WHZ of children 12-23 mo



Head circumference (cm)

One study reported data on this outcome (Morley 1999). It found evidence that giving children 12-23 mo follow-on formula may make little or no difference to head circumference compared to animal milk, when given during 9 mo (dose of exposure not specified) (MD -0.05, 95% CI −0.36 to 0.26; 425 participants; low certainty evidence) (Figure 5).

Figure 5. Animal milk vs Follow-on formula on head circumference of children 12-23 mo

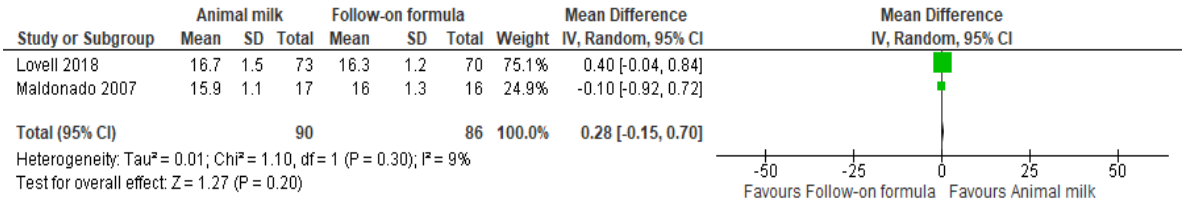


Body composition

Body Mass Index

Two studies reported on Body Mass Index (Lovell 2018 and Maldonado 2007). We pooled these studies in a meta-analysis and found evidence that giving children 12-23 mo follow-on formula may make little or no difference to body mass index compared to animal milk, when given during 4 to 12 mo and exposed to 300 ml or more per day (MD 0.28, 95% CI −0.15 to 0.70; 176 participants; low certainty evidence) (Figure 6).

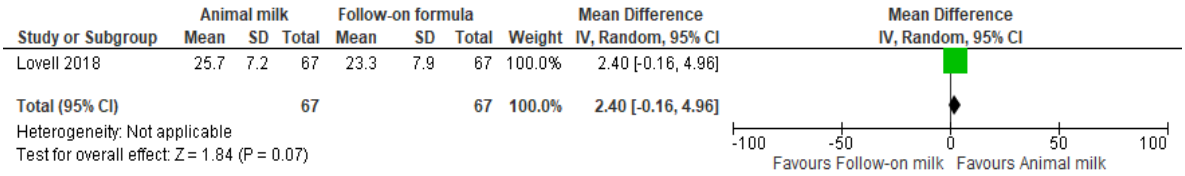
Figure 6. Animal milk vs Follow-on formula on Body Mass Index of children 12-23 mo



Body fat %

One study reported on percentage of body fat (Lovell 2018). We found evidence that giving children 12-23 mo follow-on formula may make little or no difference to body fat % compared to animal milk, when given during 12 mo and exposed to 300 ml/d (MD 2.40, 95% CI −0.16 to 4.96; 134 participants; low certainty evidence) (Figure 7).

Figure 7. Animal milk vs Follow-on formula on body fat % of children 12-23 mo



Long-term food preferences/dietary patterns

No data was found on this outcome.

Longer-term outcomes (NCDs)

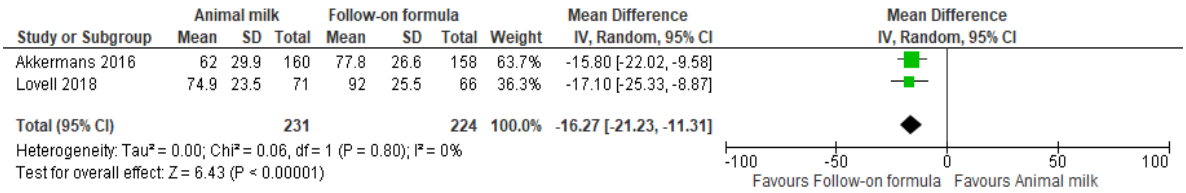
No data was found on this outcome.

Nutrient status

Serum vitamin D

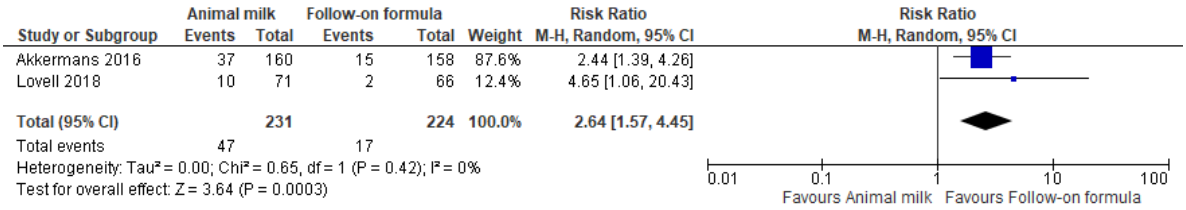
Two studies reported on serum vitamin D (as serum 25-hydroxyvitamin D [25(OH)D]) and vitamin D deficiency (Akkermans 2016 and Lovell 2018). We pooled these studies in a meta-analysis and found evidence that giving children 12-23 mo follow-on formula probably improves serum vitamin D concentrations compared to animal milk, by 16.27 more nmol of serum vitamin D per litre, when given during 5 to 12 mo and exposed to more than 150 ml/d (95% CI -21.23, -11.31; 455 participants; moderate certainty evidence) (Figure 8). The heterogeneity was low (T2 = 0.00; Chi2 = 0.06 (P = 0.80); I2 = 0%).

Figure 8. Animal milk vs Follow-on formula on serum vitamin D of children 12-23 mo



We also found evidence that giving children 12-23 mo follow-on formula probably reduces vitamin D deficiency compared to animal milk, when given during 5 to 12 mo and exposed to more than 150 ml/d (RR 2.64, 95% CI 1.57 to 4.45; 455 participants; moderate certainty evidence) (Figure 9). The heterogeneity was low (Tau² = 0.00; Chi² = 0.65 (P = 0.42); I² =0%).

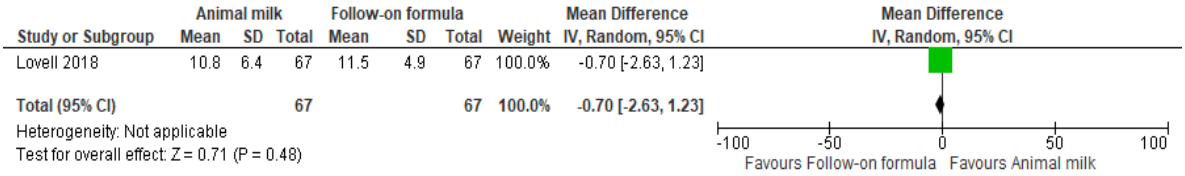
Figure 9. Animal milk vs Follow-on formula on vitamin D deficiency in children 12-23 mo



Serum iron

A single study reported on serum iron (Lovell 2018). It found evidence that giving children 12-23 mo follow-on formula may make little or no difference to serum iron concentrations compared to animal milk, when given during 12 mo and exposed to 300 ml/d (MD -0.70, 95% CI -2.63, 1.23; 134 participants; low certainty evidence) (Figure 10).

Figure 10. Animal milk vs Follow-on formula on serum iron in children 12-23 mo



Child development

Bayley mental index and psychomotor development index

One study reported on Bayley mental index (MDI) and psychomotor development index (PDI) (Morley 1999). It found evidence that giving children 12-23 mo follow-on formula may make little or no difference to Bayley mental index (MD -1.55, 95% CI -0.64, 3.73; 395 participants; low certainty evidence) and psychomotor development index (MD -1.15, 95% CI -3.07, 0.77; 395 participants; low certainty evidence) compared to animal milk, when given during 9 mo (dose of exposure not specified) (Figures 11 and 12).

Figure 11. Animal milk vs Follow-on formula on Bayley Mental Index in children 12-23 mo

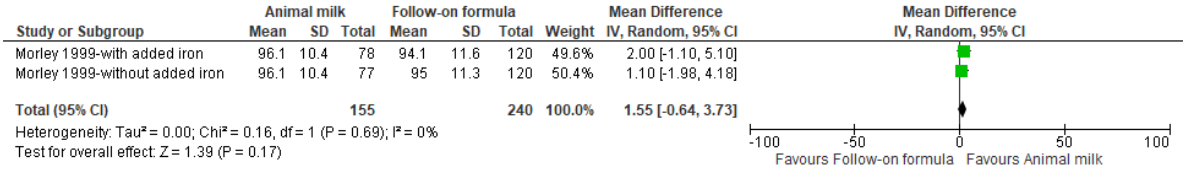
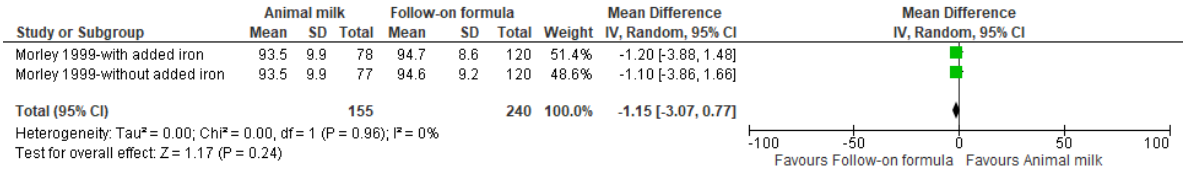


Figure 12. Animal milk vs Follow-on formula on psychomotor development index in children 12-23 mo



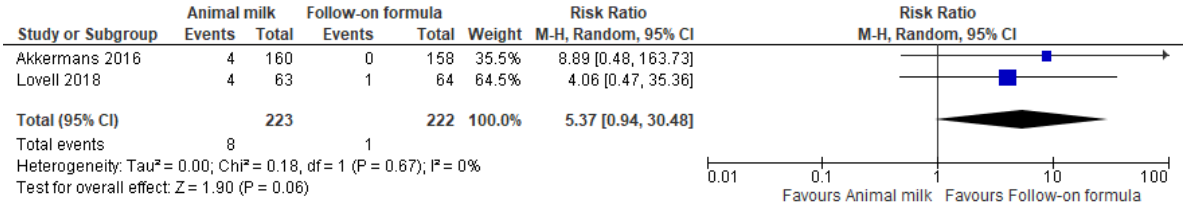
Anaemia

No data was found in this outcome.

Iron deficiency anaemia

Two studies reported on iron deficiency anaemia (IDA) (Akkermans 2016 and Lovell 2018). We pooled these studies in a meta-analysis and found inconclusive evidence that giving children 12-23 mo follow-on formula probably reduces iron deficiency anaemia compared to animal milk, when given during 5 to 12 mo and exposed to more than 150 ml/d (RR 5.37, 95% CI 0.94 to 30.48; 445 participants; moderate certainty evidence) (Figure 13). The heterogeneity was low (Tau² = 0.00; Chi² = 0.18 (P = 0.67); I² =0%).

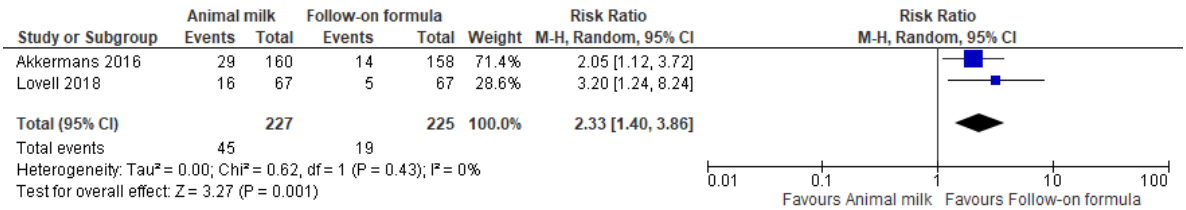
Figure 13. Animal milk vs Follow-on formula on iron deficiency anaemia in children 12-23 mo



Iron deficiency

Two studies reported on iron deficiency (ID) as serum ferritin <12 g/l (Akkermans 2016 and Lovell 2018). We pooled these studies in a meta-analysis and found evidence that giving children 12-23 mo follow-on formula probably reduces iron deficiency compared to animal milk, when given during 5 to 12 mo and exposed to more than 150 ml/d (RR 2.33 95% CI 1.40 to 3.86; 452 participants; moderate certainty evidence) (Figure 14). Heterogeneity between studies was low: Tau² = 0.00; Chi² = 0.62, (P = 0.43); I² = 0%.

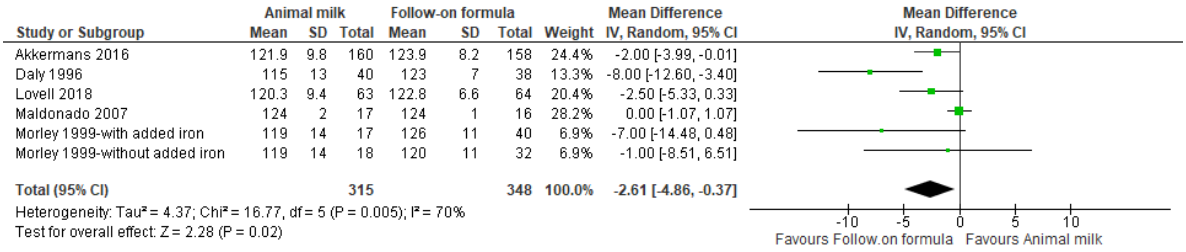
Figure 14. Animal milk vs Follow-on formula on iron deficiency in children 12-23 mo



Haemoglobin

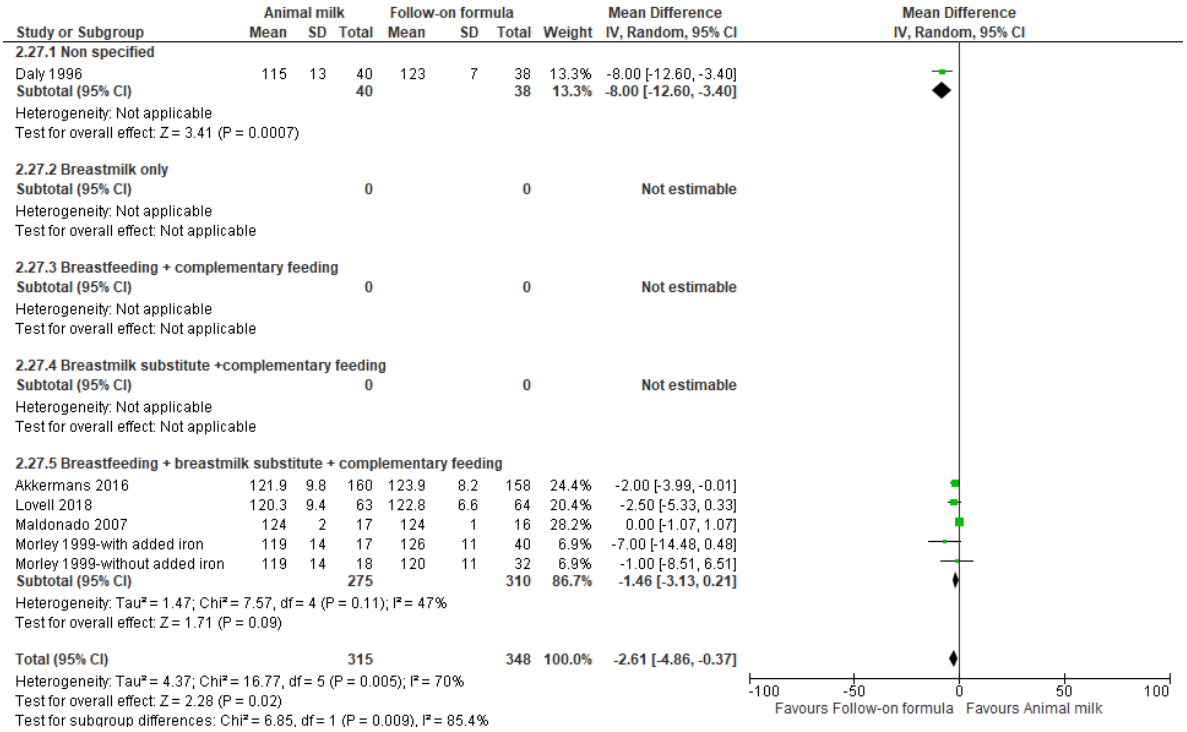
Five studies reported on this outcome (Akkermans 2016, Daly 1996, Lovell 2018, Maldonado 2007 and Morley 1999). We pooled these studies in a meta-analysis and found evidence that giving children 12-23 mo follow-on formula probably improves haemoglobin concentrations by 2.61 more grams of haemoglobin per litre, compared to animal milk, when given during 4 to 12 mo and exposed to more than 150 ml/d (95% CI -4.86 to -0.37; 663 participants; moderate certainty evidence) (Figure 15).The heterogeneity was moderate (Tau^2 = 4.37; Chi^2 = 16.77, (P = 0.005); I^2 = 70%. The effect remained similar even after excluding the trials at higher risk of bias (MD -2.32; 95% CI -3.88 to -0.77).

Figure 15. Animal milk vs Follow-on formula on haemoglobin in children 12-23 mo



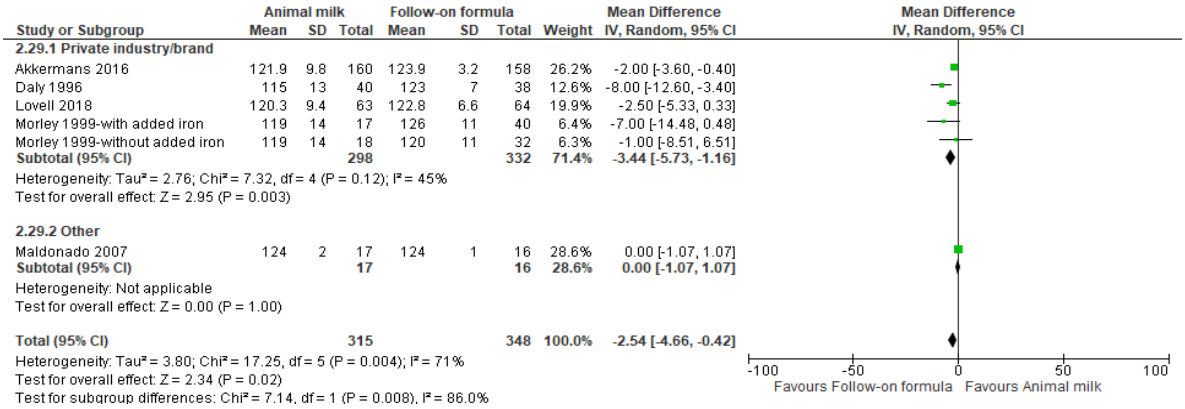
In subgroup analysis, all studies reported non-specified prevalence of stunting and wasting. Further subgroup analysis and showed evidence that giving children follow-on formula probably improves haemoglobin concentrations by 8.00 more grams of haemoglobin per litre, compared to animal milk, when type of feeding before 12 mo was not specified, given during 4 to 12 mo and exposed to 378 ml/d (95% CI -12.60 to -3.40; 78 participants; moderate certainty evidence) (Figure 16).

Figure 16. Animal milk vs Follow-on formula on haemoglobin concentration in children 12-23 mo



Further subgroup analysis showed evidence that giving children 12-23 mo follow-on formula probably improves haemoglobin concentrations compared to animal milk, when the study was funded by the private industry (MD -3.44, 95% CI -5.73 to -1.16; 630 participants; moderate certainty evidence) (Figure 17).

Figure 17. Animal milk vs Follow-on formula on haemoglobin concentration in children 12-23 mo by funding source

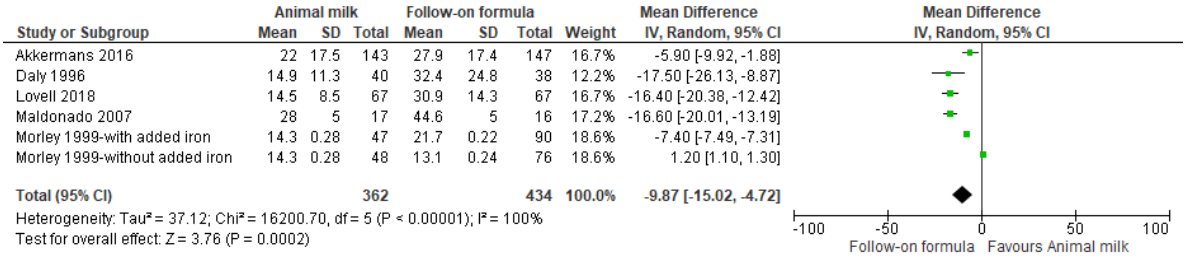


Ferritin

Five studies reported on this outcome (Akkermans 2016, Daly 1996, Lovell 2018, Maldonado 2007 and Morley 1999). We pooled these studies in a meta-analysis and found evidence that giving children 12-23 mo follow-on formula probably increase ferritin concentrations by 9.87 µg per litre compared to animal milk, when given during 4 to 12 mo and exposed to more than 150 ml/day (95%

CI -15.02 to -4.72; 796 participants; moderate certainty evidence) (Figure 18). There was considerable heterogeneity: Tau^2 = 37.12; Chi^2 = 16200.70, (P < 0.00001); I^2 =100%. The effect remained similar even after excluding the trials at higher risk of bias (MD -6.91; 95% CI -13.03 to -0.79).

Figure 18. Animal milk vs Follow-on formula on ferritin concentration in children 12-23 mo



In subgroup analysis, all studies reported non-specified prevalence of stunting and wasting. Further subgroup analysis showed evidence that giving children 12-23 mo follow-on formula probably increase ferritin concentration by 17.50 more µgrams per litre compared to animal milk, when type of feeding before 12 mo was not specified (95% CI -26.13 to -8.87; 78 participants; moderate certainty evidence) (Figure 19). We also found evidence suggesting that giving children 12-23 mo follow-on formula probably increase ferritin concentrations by 17.50 µgrams more of ferritin per litre, compared to animal milk, when anaemia prevalence was not specified (MD 95% CI -19.11 to -13.92; 78 participants; moderate certainty evidence) (Figure 20). In further subgroup analysis, we found evidence to suggest that giving children 12-23 mo follow-on formula probably increase ferritin concentrations compared to animal milk, whether the study was funded by private industry or not (Figure 21).

Figure 19. Animal milk vs Follow-on formula on ferritin concentration in children 12-23 mo by type of feeding before 12 mo

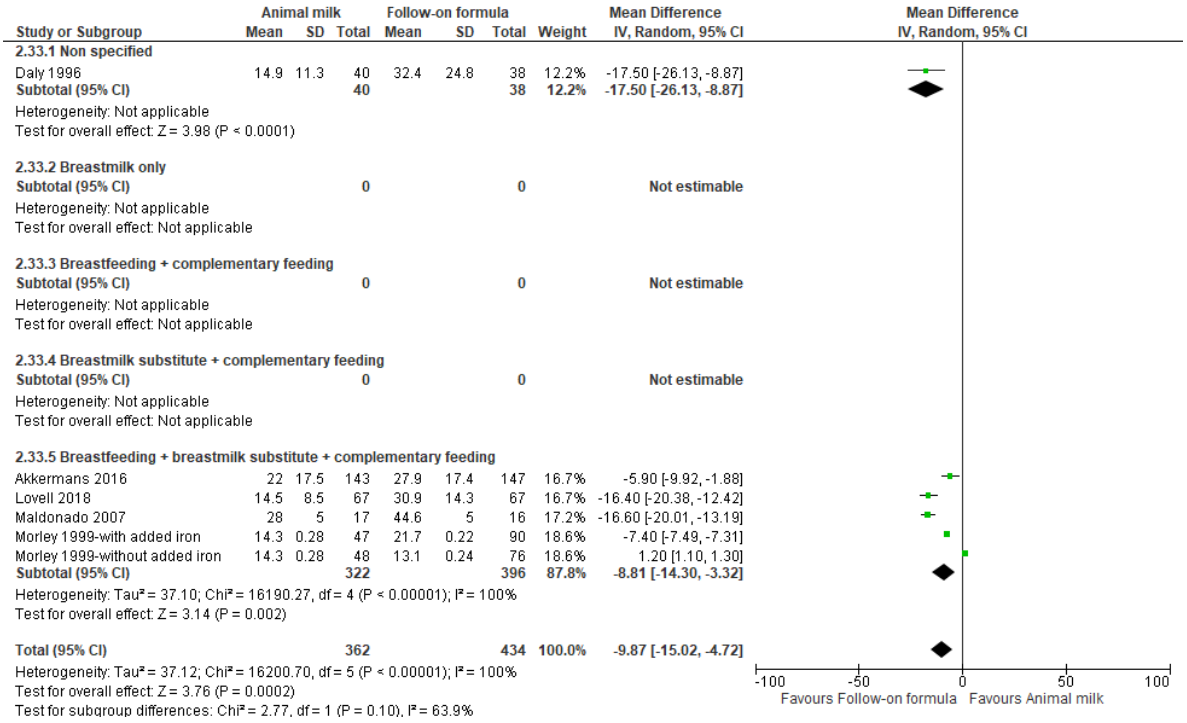


Figure 20. Animal milk vs Follow-on formula on ferritin concentration in children 12-23 mo by prevalence of anaemia

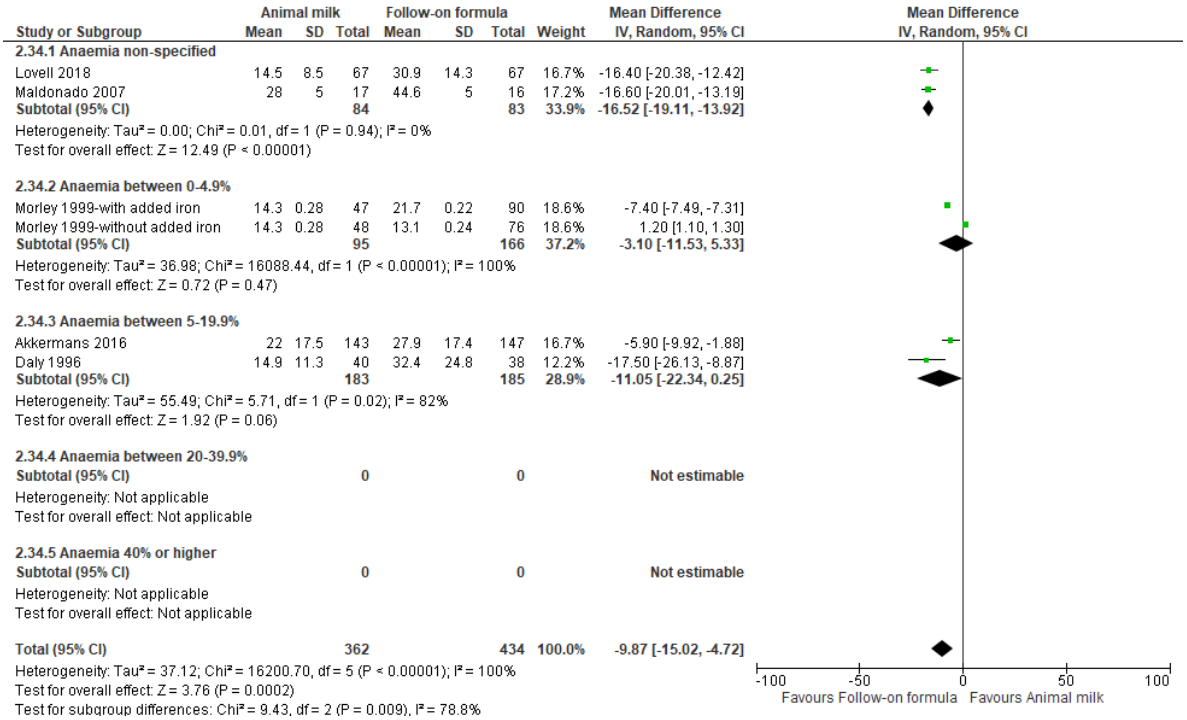
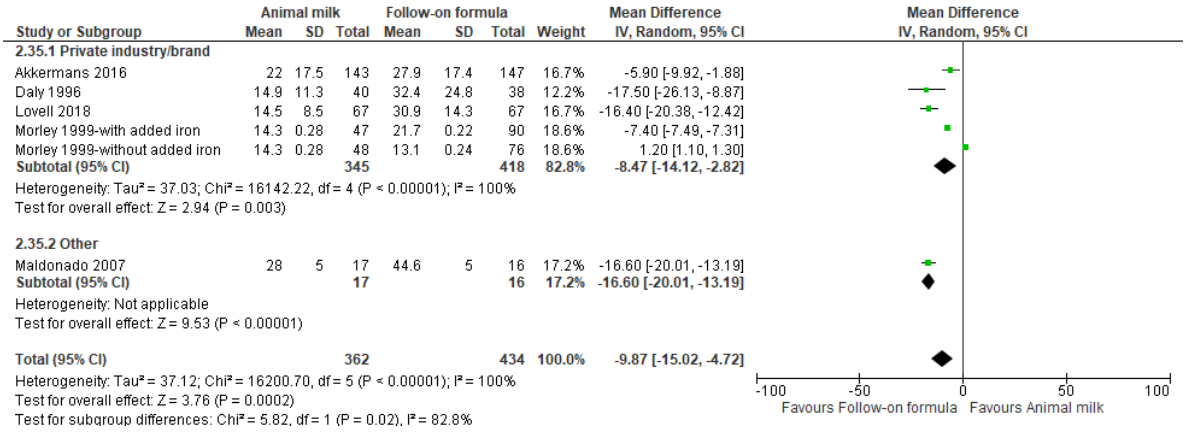


Figure 21. Animal milk vs Follow-on formula on ferritin concentration in children 12-23 mo by funding source



Important outcomes

Nutrient intakes (sufficient, excessive)

No data was found on this outcome.

Feeding practices

No data was found on this outcome.

Oral health

No data was found on this outcome.

Morbidity

No data was found on this outcome.

Dietary diversity

No data was found on this outcome.

Gut health

Stool frequency and consistency

A single study reported on this outcome (Akkermans 2016). Authors found inconclusive evidence that giving children 12-23 mo follow-on formula may reduce stool frequency per day compared to animal milk, when given during 5 mo and exposed to more than 150 mL/d (MD 2.00, 95% CI 0.18 to 20.83; 306 participants; low certainty evidence) (Figure 22). The same study found that giving children 12-23 mo follow-on formula may make little or no difference in stool consistency measured as on an ordered 5-point scale with pictures (1: watery; 2: soft, pudding-like; 3: soft-formed; 4: dry-formed; 5: dry hard pellets). compared to animal milks, when given for 5 mo and exposed to more than 150 mL/d (RR 1.22, 95& CI 0.96, 1.52; 306 participants; low certainty evidence) (Figure 23).

Figure 22. Animal milk vs Follow-on formula on stool frequency in children 12-23 mo

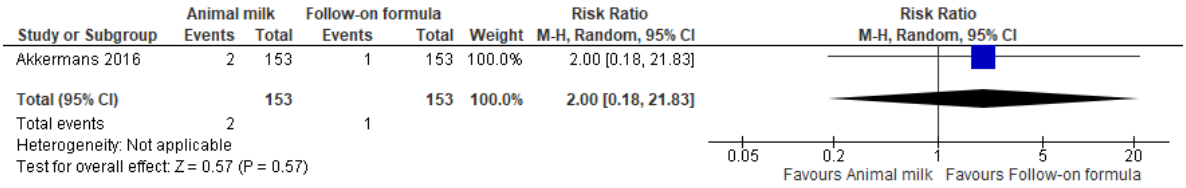
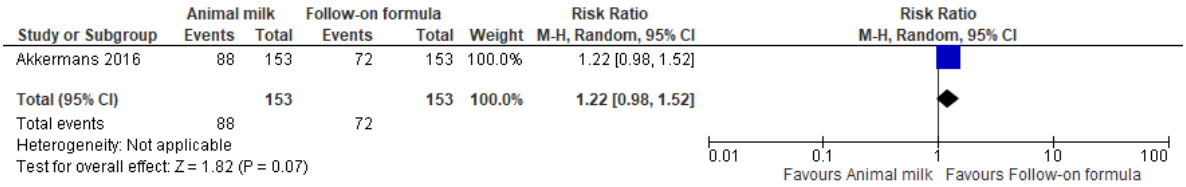


Figure 23. Animal milk vs Follow-on formula on stool consistency in children 12-23 mo



Allergy

No data was found on this outcome.

Phyto-oestrogen related outcomes

No data was found on this outcome.

3.3.3 Comparison 3: Full-fat animal milk versus lower-fat milk

One study (17 participants) was included in this comparison (Svahn 1999). The study was assessed as being at low risk of bias for confounding, attrition bias and performance bias; and was assessed as being at high risk of bias for detection bias. We included one non-randomised study (Van der Gaag 2015) in this comparison for qualitative assessment.

Critical outcomes

Growth

No data was found on this outcome.

Body composition

No data was found on this outcome.

Long-term food preferences/dietary patterns

No data was found on these outcomes.

Nutrient status

Serum lipid profile

One study reported on this outcome (Svahn 1999). The study found evidence that giving children 12-23 mo lower-fat milk may make little or no difference to serum cholesterol compared to full-fat milk, when given during 6 mo and exposed to milk (dose not specified) and 10% of energy from dairy products (MD -0.17 95% CI -0.92 to 0.58; 17 participants; low certainty evidence) (Figure 24). Test for overall effect: Z = 0.44 (P = 0.66). Authors also found evidence that giving children 12-23 mo lower-fat milk may make little or no difference to serum low-density lipoproteins (MD -0.25 95% CI -0.94 to 0.44; 17 participants; low certainty evidence), serum high-density lipoproteins (MD -0.10 95% CI -0.30 to 0.10; 17 participants; low certainty evidence), serum triglycerides (MD 0.34 95% CI -0.12 to 0.80; 17 participants; low certainty evidence), or LDL/HDL (MD -0.05 95% CI -0.83 to 0.73; 17 participants; low certainty evidence) compared to full-fat milk, when given during 6 mo (Figures 25 to 28).

Figure 24. Full-fat milk vs Lower-fat milk on serum cholesterol in children 12-23 mo

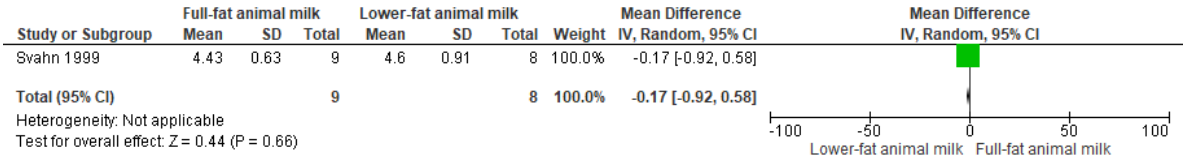


Figure 25. Full-fat milk vs Lower-fat milk on serum low-density lipoproteins in children 12-23 mo

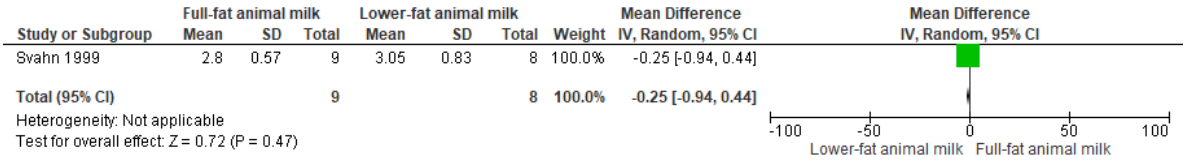


Figure 26. Full-fat milk vs Lower-fat milk on serum high-density lipoproteins in children 12-23 mo

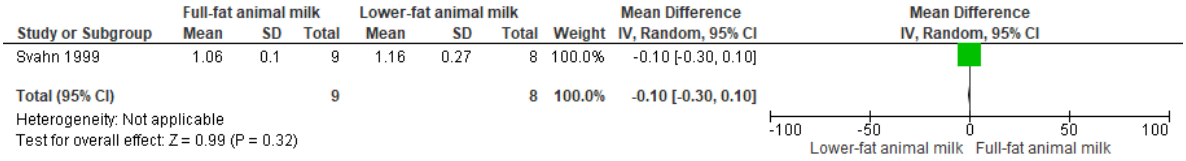


Figure 27. Full-fat milk vs Lower-fat milk on serum triglycerides in children 12-23 mo

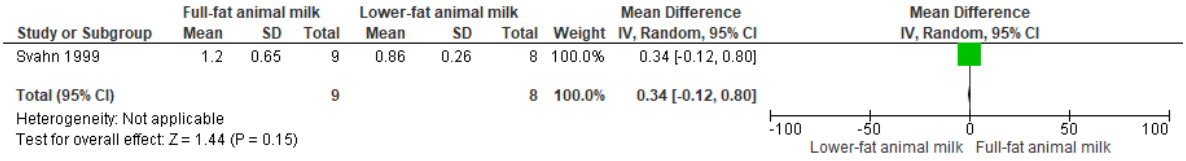
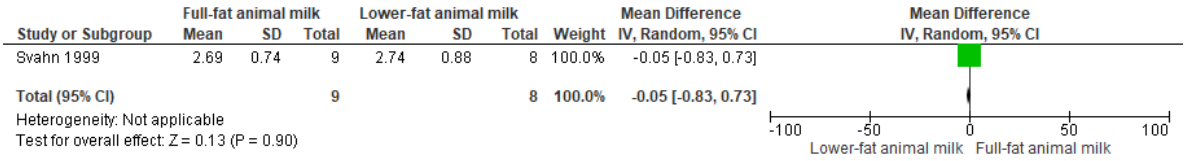


Figure 28. Full-fat milk vs Lower-fat milk on LDL/HDL ratio in children 12-23 mo



Child development

No data was found on this outcome.

Anaemia

No data was found on this outcome.

Iron deficiency anaemia

No data was found on this outcome.

Iron deficiency

No data was found on this outcome.

Haemoglobin

No data was found on this outcome.

Ferritin

No data was found on this outcome.

Important outcomes

Nutrient intakes (sufficient, excessive)

No data was found on this outcome.

Feeding practices

No data was found on this outcome.

Oral health

No data was found on this outcome.

Morbidity

No data was found on this outcome.

Dietary diversity

No data was found on this outcome.

Gut health

No data was found on this outcome.

Allergy

No data was found on this outcome.

Phyto-oestrogen related outcomes

No data was found on this outcome.

3.3.4 Comparison 4: Animal milk (full-fat or lower-fat milk) versus plant-based milk

A single study (21 participants) was included in this comparison (Svahn 1999). The study was assessed as being at low risk of bias for confounding, attrition bias and performance bias; and was assessed as being at high risk of bias for detection bias.

Critical outcomes

Growth

No data was found on this outcome.

Body composition

No data was found on this outcome.

Long-term food preferences/dietary patterns

No data was found on these outcomes.

Nutrient status

Serum lipid profile

One study reported on this outcome (Svahn 1999). The study found evidence that giving children 12-23 mo plant-based milk may make little or no difference to serum cholesterol compared to animal milk, when given during 6 mo and exposed milk (dose not specified) and 10% of energy from dairy products (MD -0.16 95% CI -0.76 to 0.44; 21 participants; low certainty evidence) (Figure 29). Test for overall effect: Z = 0.53 (P = 0.60). Authors also found evidence that giving children 12-23 mo plant-based milk may make little or no difference to serum low-density lipoproteins (MD 0.03 95% CI -0.48 to 0.54; 21 participants; low certainty evidence), serum high-density lipoproteins (MD -0.18 95% CI -0.85 to 0.49; 21 participants; low certainty evidence), serum triglycerides (MD -0.08 95% CI -0.63 to 0.47; 21 participants; low certainty evidence), or LDL/HDL (MD 0.33 95% CI -0.36 to 1.02; 21 participants; low certainty evidence) compared to animal milk, when given during 6 mo (Figures 30 to 33).

Figure 29. Animal milk vs. Plant-based milk on serum cholesterol in children 12-23 mo

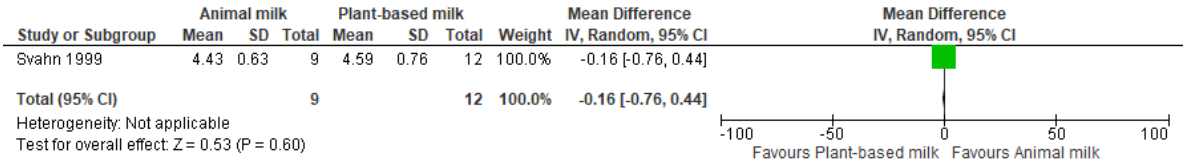


Figure 30. Animal milk vs. Plant-based milk on serum low-density lipoproteins in children 12-23 mo

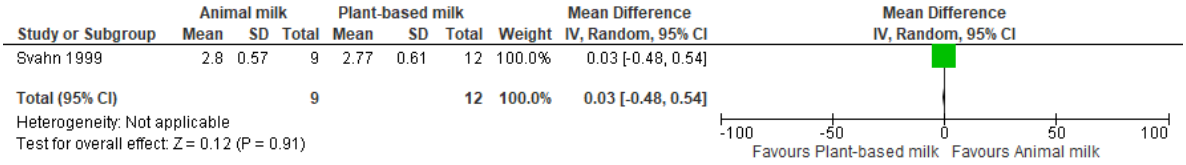


Figure 31. Animal milk vs. Plant-based milk on serum high-density lipoproteins in children 12-23 mo

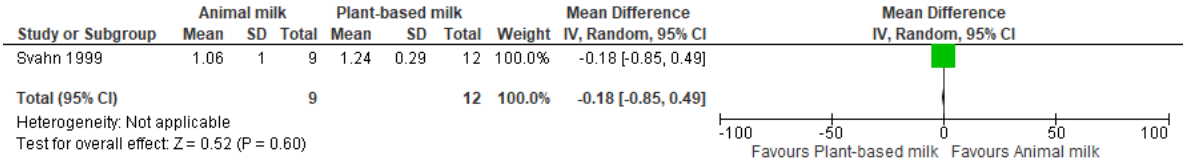


Figure 32. Animal milk vs. Plant-based milk on serum triglycerides in children 12-23 mo

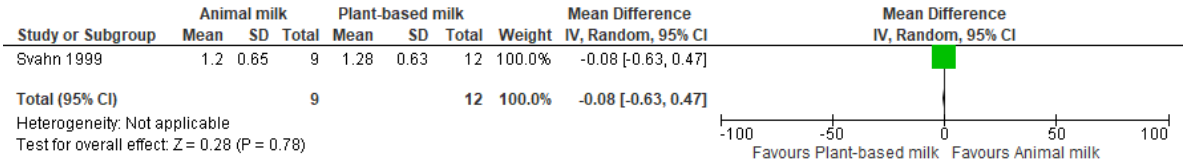
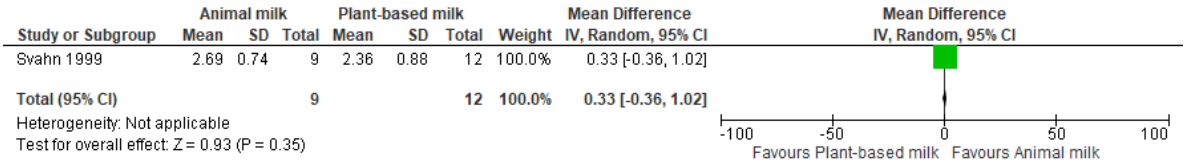


Figure 33. Animal milk vs. Plant-based milk on LDL/HDL ratio in children 12-23 mo



Child development

No data was found on this outcome.

Anaemia

No data was found on this outcome.

Iron deficiency anaemia

No data was found on this outcome.

Iron deficiency

No data was found on this outcome.

Haemoglobin

No data was found on this outcome.

Ferritin

No data was found on this outcome.

Important outcomes

Nutrient intakes (sufficient, excessive)

No data was found on this outcome.

Feeding practices

No data was found on this outcome.

Oral health

No data was found on this outcome.

Morbidity

No data was found on this outcome.

Dietary diversity

No data was found on this outcome.

Gut health

No data was found on this outcome.

Allergy

No data was found on this outcome.

Phyto-oestrogen related outcomes

No data was found on this outcome.

3.3.5 Comparison 5: Animal milk (full-fat or lower-fat milk) versus fortified milk

There were eight studies (2,905 participants) included in this comparison (Rivera 2010 (C), Sazawal 2007, Stecksén-Blicks 2009 (C), Stekel 1986, Stekel 1988, Svahn 1999, Szymlek-Gay 2009 and Villalpando 2006). These studies comprise all the data included in the synthesis of this review. We included one non-randomised study in this comparison (Brusner 1993) for qualitative assessment.

Two of the studies met the prespecified criteria mentioned above for being at lower risk of bias (Sazawal 2007 and Stecksén-Blicks 2009 (C)) and in sensitivity analyses these trials were retained in the analysis whilst trials at higher risk of bias (Rivera 2010 (C), Svahn 1999 and Villalpando 2006) were temporarily removed to examine whether they had any impact on the overall pattern of results.

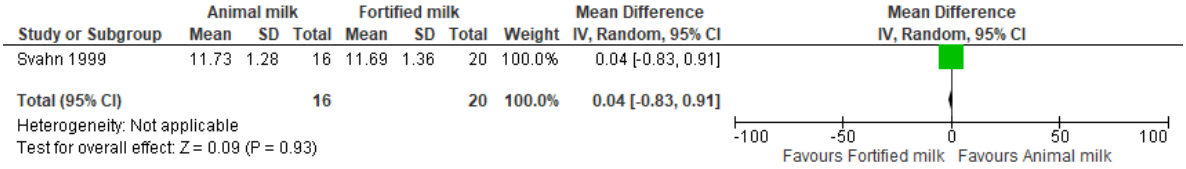
Primary outcomes

Growth

Weight (kg)

A single study reported on this outcome (Svahn 1999). Results showed that giving children 12-23 mo fortified milk may make little or no difference to body weight compared to animal milk, when given during 6 mo and exposed to milk (dose nor specified) and 10% of energy from dairy products (MD 0.04 95% CI -0.83, 0.91; 36 participants; low certainty evidence) (Figure 34).

Figure 34. Animal milk vs. Fortified milk on weight in children 12-23 mo



Undernutrition

One study reported on stunting and wasting (Sazawal 2007). Results showed that giving children 12-23 mo fortified milk makes little or no difference to stunting compared to animal milk, when given during 12 mo and exposed to 32 g sachets, 3 per day everyday (RR 0.98 95% CI 0.74, 1.28; 378 participants; low certainty evidence) (Figure 35). Results also showed that giving children 12-23 mo fortified milk makes little or no difference to stunting nor to stunting and wasting together, compared to animal milk, when given during 12 mo and exposed to 32 g sachets, 3 per day everyday (RR 1.06 95% CI 0.78, 1.44; 378 participants; low certainty evidence) and (RR1.14 95% CI 0.86, 1.50; 378 participants; low certainty evidence), respectively (Figures 36 to 37).

Figure 35. Animal milk vs. Fortified milk on stunting in children 12-23 mo

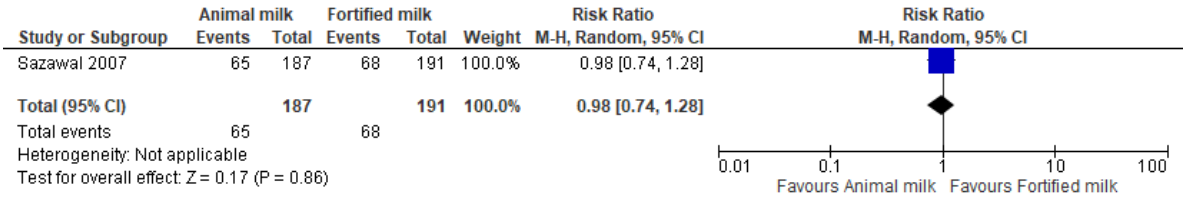


Figure 36. Animal milk vs. Fortified milk on wasting in children 12-23 mo

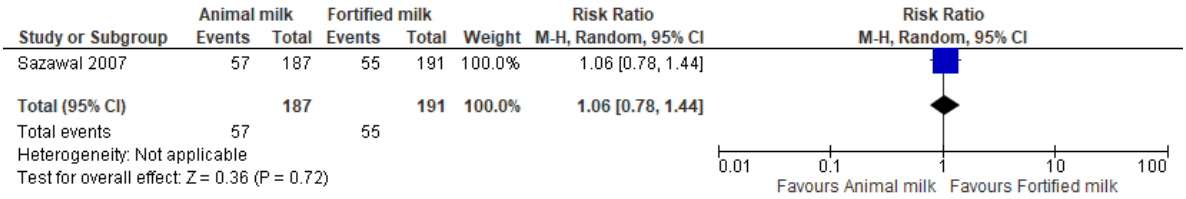
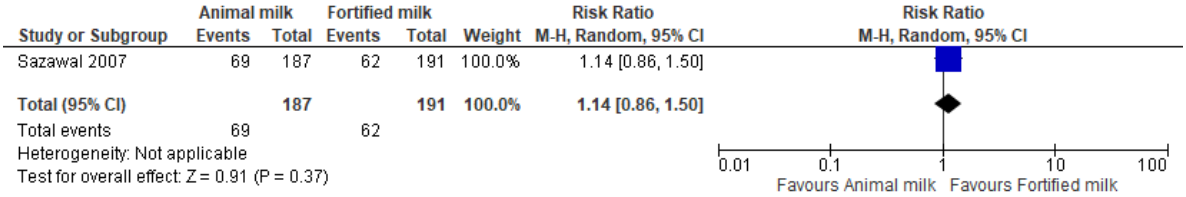


Figure 37. Animal milk vs. Fortified milk on stunting and wasting in children 12-23 mo



Body composition

No data was found on this outcome.

Long-term food preferences/dietary patterns

No data was found on this outcome.

Longer-term outcomes (NCDs)

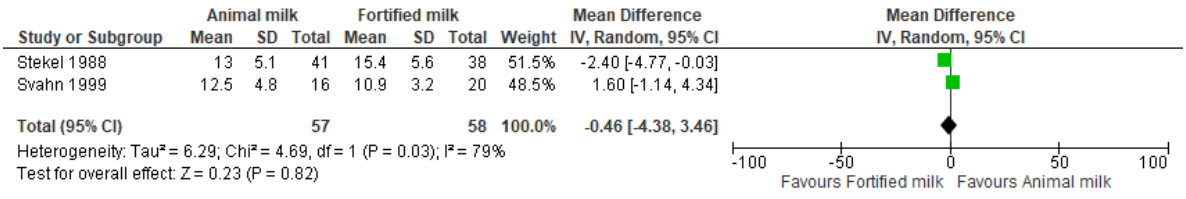
No data was found on this outcome.

Nutrient status

Serum iron

Two studies reported on serum iron concentrations (Stekel 1988 and Svahn 1999). We pooled these studies in a meta-analysis and found evidence suggesting that giving children 12-23 mo fortified milk makes little or no difference to serum iron concentrations compared to animal milk, when given during 12 mo (dose of exposure not specified) (MD -0.46 95% CI -4.38 to 3.46; 115 participants; low certainty evidence). The heterogeneity was moderate ($Tau^2 = 6.29$; $Chi^2 = 4.69$, (P = 0.03); $I^2 = 79\%$) (Figure 38).

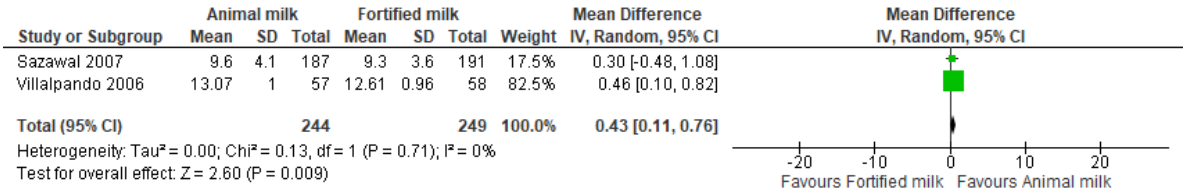
Figure 38. Animal milk vs. Fortified milk on serum iron in children 12-23 mo



Plasma zinc

Two other studies reported on plasma zinc concentrations (Sazawal 2007 and Villalpando 2006). We pooled these studies in a meta-analysis and found evidence suggesting that giving children 12-23 mo fortified milk may reduce plasma zinc concentrations by 0.43 more µmol per litre compared to animal milk, when given during 6 m and exposed to 400 mL/d (95% CI 0.11 to 0.76; 493 participants; low certainty evidence) (Figure 39). Heterogeneity between studies was low: Tau² = 0.00; Chi² = 0.13, (P = 0.71); I² = 0%. The effect did not remain similar after excluding the trial at higher risk of bias (MD 0.30; 95% CI -0.48 to 1.08).

Figure 39. Animal milk vs. Fortified milk on plasma zinc in children 12-23 mo



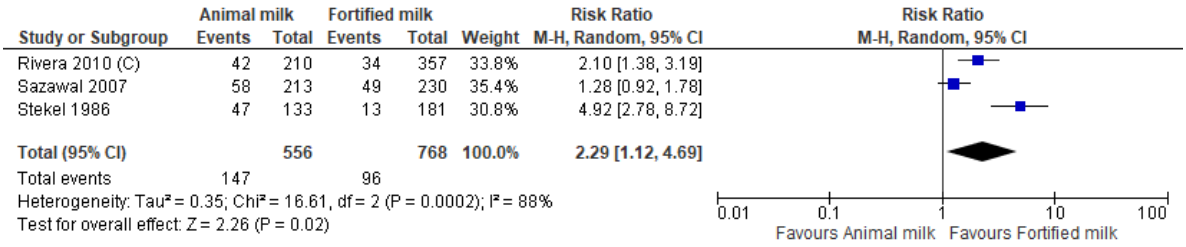
Child development

No data as found on this outcome.

Anaemia

Three studies reported on this outcome (Rivera 2010 (C), Sazawal 2007 and Stekel 1986). We pooled these studies in a meta-analysis and found evidence suggesting that giving children 12-23 mo fortified milk may reduce anaemia, by 2.29 times, compared to animal milk, when given during 12 mo and exposed to 400 mL/d (95% CI 1.12 to 4.69; 1324 participants; low certainty evidence) (Figure 40). We found considerable heterogeneity between studies: Tau² = 0.35; Chi² = 16.61 (P = 0.0002); I²= 88%. The effect did not remain similar after excluding the trials at higher risk of bias (MD 2.46; 95% CI 0.65 to 9.35).

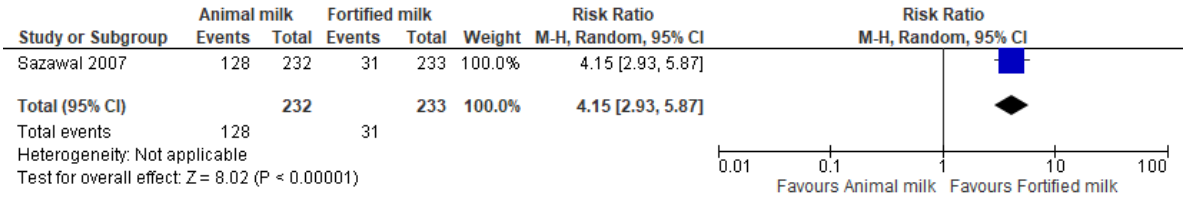
Figure 40. Animal milk vs. Fortified milk on anaemia in children 12-23 mo



Iron deficiency anaemia

One study reported on iron deficiency anaemia (Sazawal 2007). Authors found evidence that giving children 12-23 mo fortified milk may reduce iron deficiency anaemia compared to animal milk, when given during 12 mo and exposed to 32 g sachets, 3 per day everyday (RR 4.15 95% CI 2.93, 5.87; 465 participants; low certainty evidence) (Figure 41).

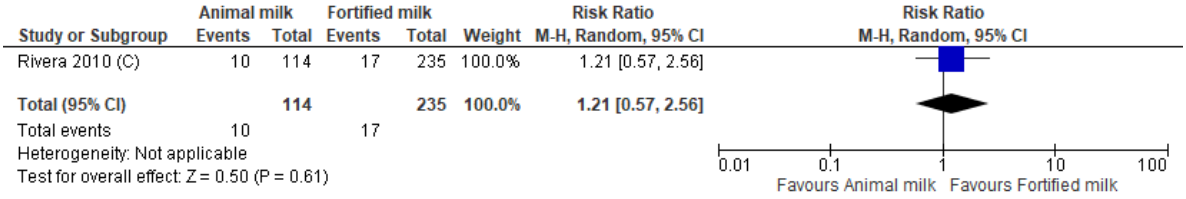
Figure 41. Animal milk vs. Fortified milk on iron deficiency anaemia in children 12-23 mo



Iron deficiency

One study reported on iron deficiency (Rivera 2010 (C)). Authors found inconclusive evidence suggesting that giving children 12-23 mo fortified milk may reduce iron deficiency compared to animal milk, when given during 12 mo and exposed to 400 mL/d (RR 1.21 95% CI 0.57, 2.56; 349 participants; low-certainty evidence) (Figure 42).

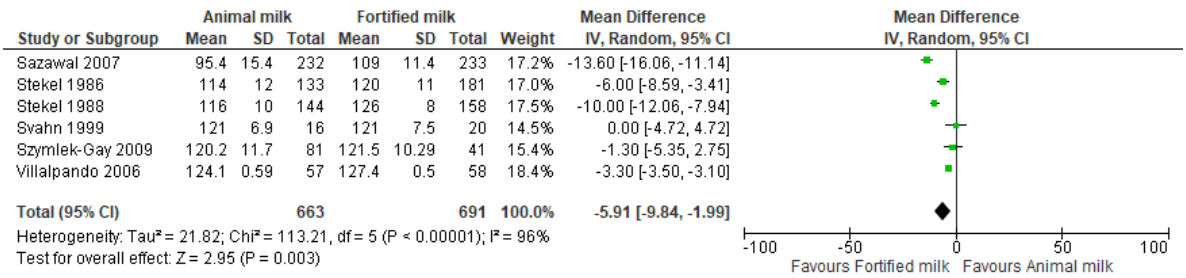
Figure 42. Animal milk vs. Fortified milk on iron deficiency in children 12-23 mo



Haemoglobin

Six studies reported on this outcome (Sazawal 2007, Stekel 1986, Stekel 1988, Svahn 1999, Szymlek-Gay 2009 and Villalpando 2006). We pooled these studies in a meta-analysis and found evidence that giving children fortified milk probably improves haemoglobin concentrations with 5.91 more grams per litre, compared to animal milk, when given during 6 mo and exposed to 400 mL/d (95% CI -9.84 to -1.99; 1354 participants, moderate certainty evidence) (Figure 43). There was considerable heterogeneity between studies: $Tau^2 = 21.82$; $Chi^2 = 113.21$, (P = 0.0002); $I^2 = 96\%$. The effect remained similar even after excluding the trials at higher risk of bias (MD -7.94; 95% CI -12.36 to 3.52).

Figure 43. Animal milk vs. Fortified milk on haemoglobin concentration in children 12-23 mo



We conducted a subgroup analysis and found evidence that giving children 12-23 mo fortified milk probably improves haemoglobin concentration compared to animal milk, when stunting prevalence

was very low (<2.5%) (MD -8.08; 95% CI -12.00 to -4.16; 616 participants; moderate certainty evidence), medium (10-19.9%) (MD -3.30; 95% CI -3.50 to -3.10; 115 participants; moderate certainty evidence) or high (>20%) (MD -13.60; 95% CI -16.06 to -11.14; 465 participants; moderate certainty evidence) (Figure 43). Further subgroup analysis showed that giving children 12-23 mo fortified milk probably improves haemoglobin concentration when wasting prevalence was very low (<2.5) (MD -6.36; 95% CI -10.80 to -1.92; 731 participants; moderate certainty evidence) or high (>10%) (MD -13.60; 95% CI -16.06 to -11.14; 465 participants; moderate certainty evidence) (Figure 44).

Figure 43. Animal milk vs. Fortified milk on haemoglobin concentration in children 12-23 mo by stunting

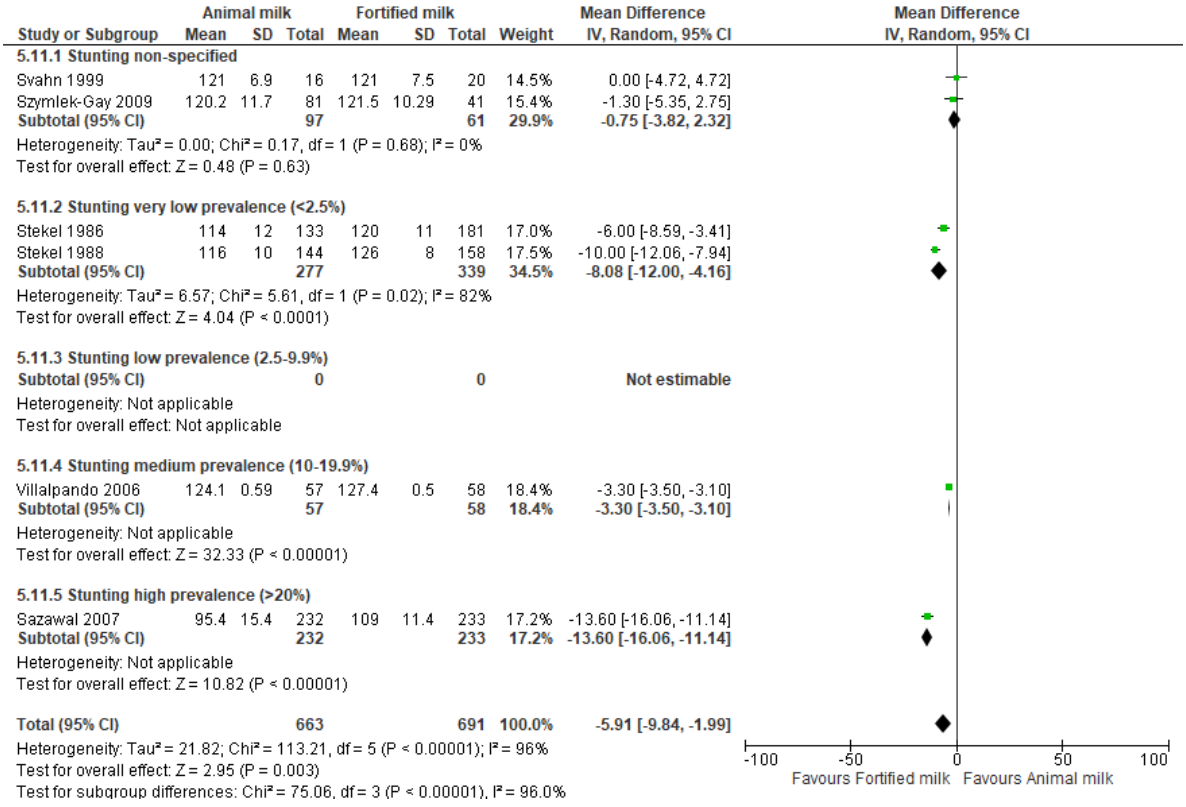
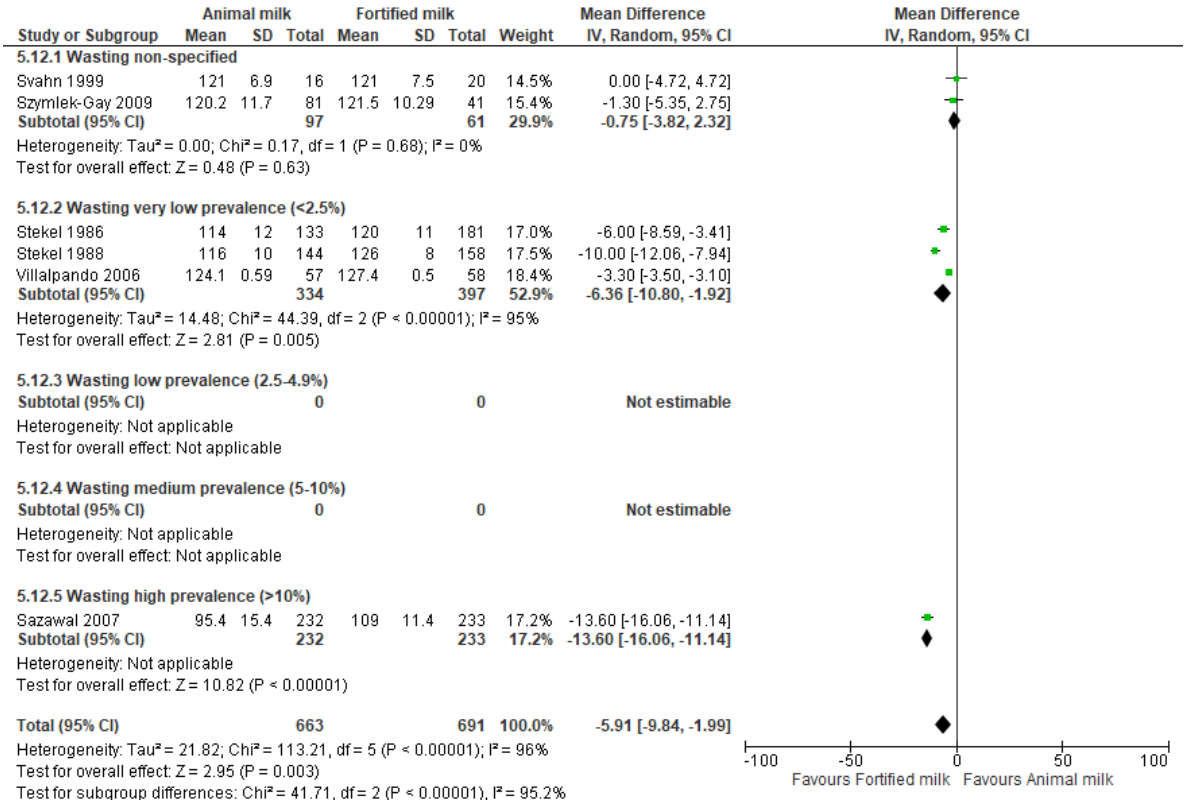
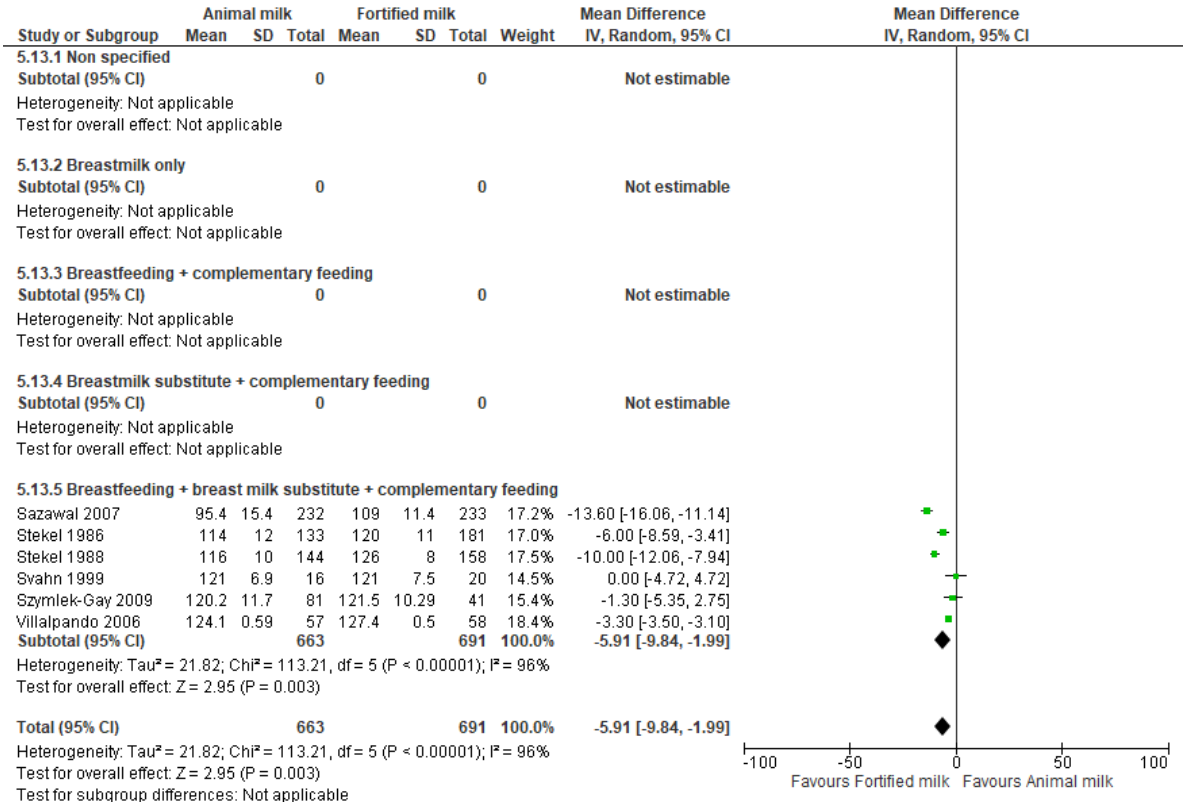


Figure 44. Animal milk vs. Fortified milk on haemoglobin concentration in children 12-23 mo by wasting



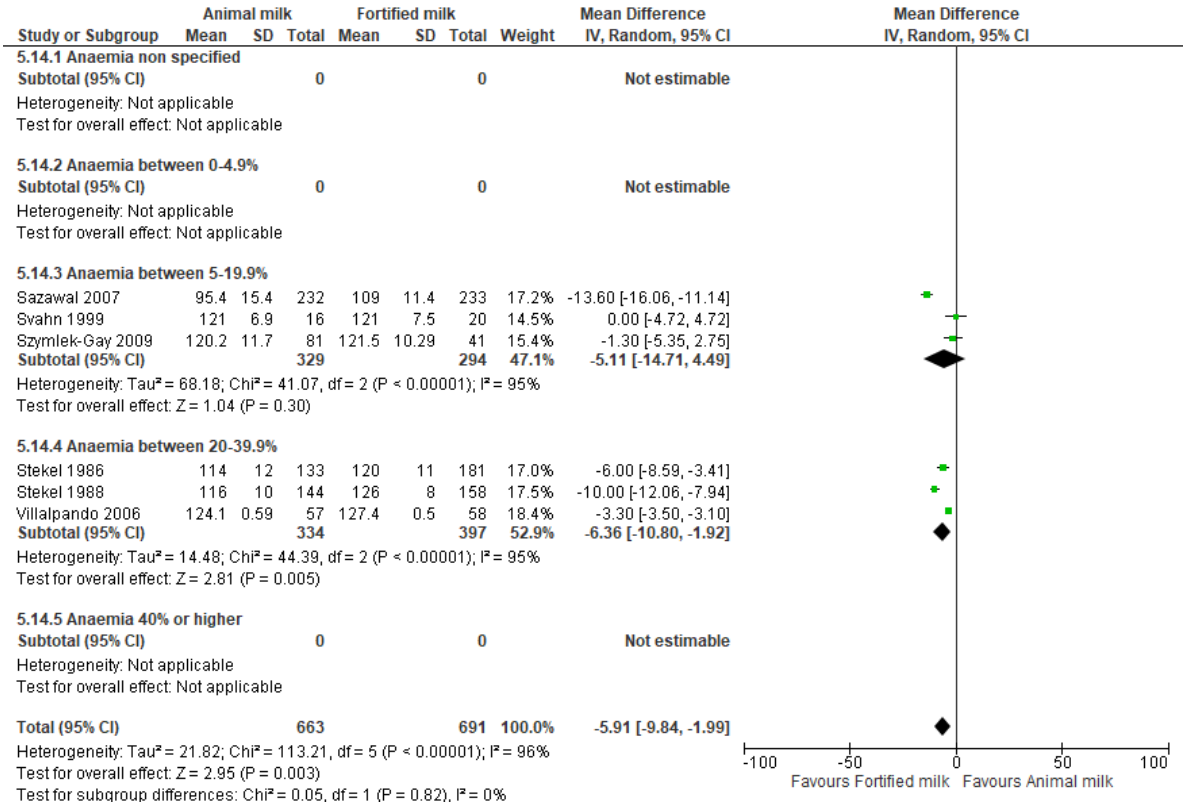
In further subgroup analysis we found evidence suggesting that giving children 12-2-3 mo fortified milk probably improves haemoglobin concentration by 5.91 more grams per litre compared to animal milk, when given during 6 mo and exposed to 400 mL/d when type of feeding before 12 mo included breastfeeding + breastmilk substitute + complementary feeding (95% CI -9.84 to -1.99; 1354 participants; moderate certainty evidence) (Figure 45).

Figure 45. Animal milk vs. Fortified milk on haemoglobin concentration in children 12-23 mo by type of feeding before 12 mo



Subgroup analysis also showed that giving children 12-23 mo fortified milk may increase haemoglobin concentration by 6.36 more grams per litre when the prevalence of anaemia was 20-39.9% compared to animal milk (95% CI -10.80 to -1.92), 731 participants; moderate certainty evidence) (Figure 46).

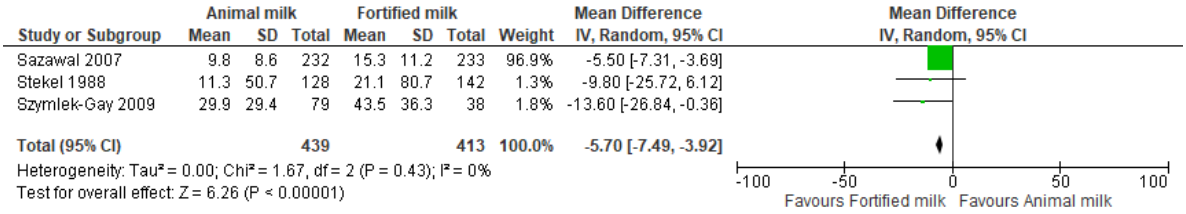
Figure 46. Animal milk vs. Fortified milk on haemoglobin concentration in children 12-23 mo by type of feeding before 12 mo



Ferritin

Three studies reported on this outcome (Sazawal 2007, Stekel 1988 and Szymlek-Gay 2009). We pooled these studies in a meta-analysis and evidence showed that giving children 12-23 mo fortified milk probably improves ferritin concentration by 5.70 more µg per litre compared to animal milk, when given during 6 mo (dose of exposure not specified) (95% CI -7.49 to -3.92; 852 participants; moderate certainty evidence) (Figure 47). There was low heterogeneity between studies: Tau² = 0.00; Chi² = 1.67 (P = 0.43); I² = 0%.

Figure 47. Animal milk vs. Fortified milk on ferritin concentration in children 12-23 mo



Important outcomes

Nutrient intakes (sufficient, excessive)

No data was found on this outcome.

Feeding practices

No data was found on this outcome.

Oral health

Mean decayed, missing and filled surfaces index and caries

One study reported on oral health (Stecksén-Blicks 2009 (C). Authors assessed mean dmfs (mean decayed, missing, and filled surfaces index in molars and canines) index and caries-free (dmfs =0). Results showed that giving children 12-23 mo fortified milk may reduce mean dmfs compared to animal milk, when given for 21 mo and exposed to 150 mL/d on weekdays (MD 1.30 95% CI 0.37 to 2.23; 186 participants; low certainty evidence). However, it found evidence that giving children 12-23 mo fortified milk may reduce caries-free in molars and canines compared to animal milk, when given during 21 mo and exposed to 150 mL/d on weekdays (RR 0.73 95% CI 0.59 to 0.91; 186 participants; low certainty evidence) (Figures 48 and 49).

Figure 48. Animal milk vs. Fortified milk on mean decayed, missing, and filled surfaces index in molars and canines (mean dmfs) in children 12-23 mo

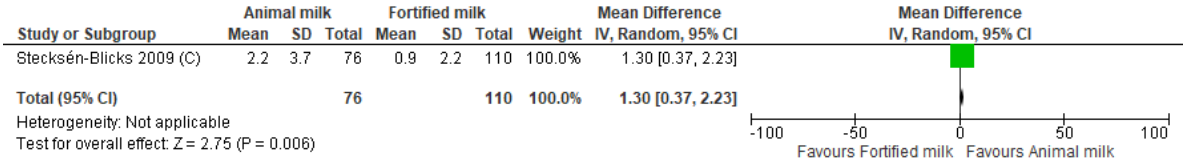
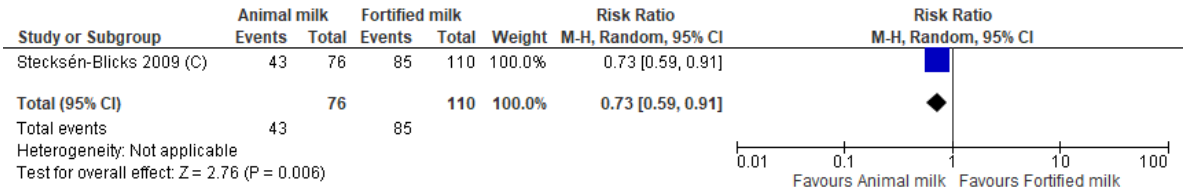


Figure 49. Animal milk vs. Fortified milk on caries free (dmfs in molars and canines = 0) in children 12-23 mo



Morbidity

Respiratory and diarrhoea episodes

A single study reported on respiratory and diarrhoea episodes per child per year (Stekel 1988). It found inconclusive evidence suggesting that giving children 12-23 mo fortified milk maymake little or no difference on respiratory and diarrhoea episodes compared to animal milk, when given during 12 mo (dose not specified) (MD 0.03 95% CI -0.14 to 0.20; 554 participants; low certainty evidence) and (MD 0.80 95% CI 0.27 to 1.33; 554 participants; low certainty evidence) (Figures 50 and 52).

Figure 50. Animal milk vs. Fortified milk on respiratory episodes in children 12-23 mo

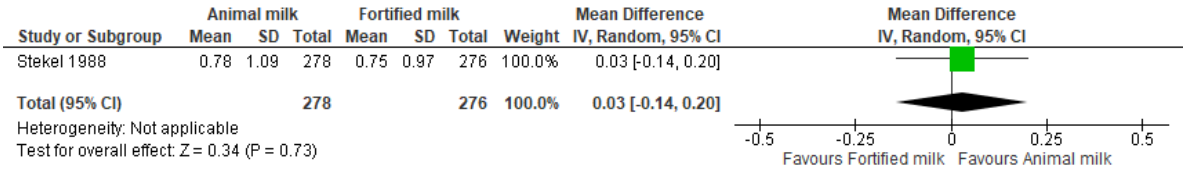
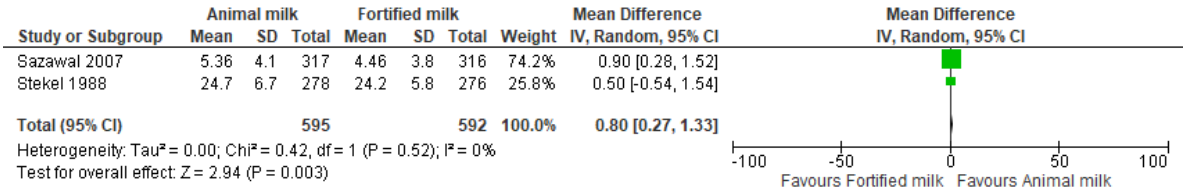


Figure 51. Animal milk vs. Fortified milk on respiratory episodes in children 12-23 mo



Dietary diversity

No data was found on this outcome.

Gut health

No data was found on this outcome.

Allergy

No data was found on this outcome.

Phyto-oestrogen related outcomes

No data was found on this outcome.

3.3.6 Qualitative analysis of non-randomized studies

Three studies complemented comparison 2 (Animal milk versus follow-on formula). Ghisolfi 2012, a cross-sectional food consumption study, reported on the estimated proportion of children at risk for nutrient excess or insufficiency receiving animal milk or follow-on formula. The study documented that daily consumption of 250 ml of cow milk in children 12-23 mo increased the risk of insufficient intakes for several nutrients which might be prevented by giving them follow-on formula instead. While total energy and macronutrient intakes were similar in the two groups, except protein intake of cow milk group which was much higher in the later compared to recommended intakes and significantly higher than children receiving follow-on formula. A high percentage of children receiving cow´s milk had intake of linoleic acid (51%) and a-linolenic acid (84%) below the lower limit of the adequate Intake, and intake of iron (59%) vitamin C (49 %) and alimentary vitamin D (100%) less than the Estimated Average Requirement. Significant differences were observed in the proportions of children with a risk of dietary inadequacy between the two groups for all the mentioned nutrients (P <0.001). In children receiving follow-on formula, this imbalance was only observed for vitamin D. Intake of foods other than milk and dairy products could not account for these discrepancies. Groups were balanced in terms of gender, but significantly lower in terms of age in the children receiving follow-on formula at 534 (SD 15) days (17.5 (SD 0.5) months; P=0.003). Diet was similar between groups in terms of the total mass of food, energy, carbohydrates, lipids, sodium, calcium, phosphorous and magnesium.

The other two studies were RCTs where controlled groups were not randomized (they were "reference groups"). Chatchatee 2014 compared a follow-on formula with 1.2 g/100 mL of scGOS/ lcFOS (9:1) (Immunofortis) and 19.2 mg/100 mL of n-3 LCPUFAs (EPA + DHA, 4:6) with animal milk, and documented decreased a risk of developing at least 1 infection (299/388 [77%] compared to children who received animal milk (313/379 [83%], respectively (RR 0.93, 95% CI 0.87–1.00; P=0.03). There were no significant differences in weight and height between study arms was observed. However, it is not clear if there were significant differences in weight and height compared with the reference group that received animal milk. Gill 1997, the other RCT, compared a follow-on formula

with added iron, to a regular formula and animal milk. Authors documented that haemoglobin levels were <110g/l in 33% of infants fed animal milk compared with 13% and 11% in those receiving non-iron-fortified and iron-fortified formula respectively. Iron deficiency in children who received animal milk was 43% compared to 22% in those children who received non-iron fortified follow-on formula, and 6% in those children who received iron fortified follow-on formula. No statistically significant differences were observed between the 2 groups in both studies, regarding age, sex, length, and weight at birth and any of the other baseline characteristics analysed.

Van der Gaag 2015 complemented comparison 3 as authors compared dietary advice on full-fat milk and dairy products versus lower-fat milk and dairy products. Authors hypothesizing that respiratory tract complaints will decrease in the full-fat group because they contain vitamin A, E and C, all vitamins with anti-oxidative capacities in a retrospective case-control study. After three months, children on lower-fat milk and dairy products were significantly less rasping; from 18 (SEM =1.5) to 9 (SEM=1.8) days a month ($p<0.000$). Also, days with fever (1.0 to 0.0 days a month, $p< 0.000$), coughing (18 to 10 days a month, $p<0.000$) and runny/blocked nose decreased significantly (18 to 11 days a month, $p=0.008$) in the intervention group. Wheezing was not affected by the dietary advice (1 to 0 days a month, $p>0.05$). The body mass index was not altered after the advice; it changed from 16.4 to 16.6 ($p=0.570$). The study was rated as unclear risk of bias.

The value of the prospective cohort (Brusner 1993) study was reporting on the incidence of diarrhoea and bowel movements in children receiving fortified milk compared to animal milk (comparison 5). There was an improvement at the end of the six months in the incidence of diarrhoea among the children receiving fortified milk (monthly average n= 70 children; 30.4 vs 25.5 episodes/100 children/month, $P < 0.025$) compared with children receiving animal milk (monthly average n=85). Children receiving fortified milk had more bowel movements on day 1 ($P < 0.03$) and liquid or semi-liquid stools were passed for more than 15 days more frequently ($P < 0.05$) compared with children receiving animal milk. Groups were balanced according to baseline characteristics.

4. Discussion

4.1 Summary of main results

Available data indicated that feeding children 12-23 mo with follow-on formula has a similar effect on growth assessed by body weight and height compared to animal milk when given during 4 to 12 mo and exposed to 300 ml or more per day. This positive response does not differ when children were breast fed and then given breast milk substitutes and complementary feeding during their first year of life, when anaemia status was not specified or when the study was funded by a brand. Data also indicates that feeding children 12-23 mo follow-on formula increased their nutrient status measured by vitamin D, iron serum concentrations and were less likely to have iron deficiency anaemia and iron deficiency, compared to children fed with animal milk. The positive effect on vitamin D can be seen when follow-on formula was given during 5 to 12 mo and children were exposed to more than 150 ml/d, for a positive effect on iron the exposure may have to increase to 400 ml per day. Children 12-23 mo that were fed follow-on formula had higher haemoglobin and ferritin concentrations compared to children that were fed animal milk, when given 150 ml per day for at least 4 mo. Results from the cross-sectional (Ghisolfi 2012) study were consistent showing that feeding children with follow-on formula may reduce the risk for nutrient insufficiency, except for protein intake (for which animal milk showed being more effective). Results from the two RCTs

studies with non-randomized controlled groups (Chatchatee 2014 and Gill 1997) were also consistent and showed no differences on weight or height.

Information on lower-fat milks and plant-based milks was scarce and limited to nutrient status assessed by lipid serum profile. There was no response of feeding children 12-23 mo with lower-fat milk on serum lipid profile, and response does not differ when feeding them with full-fat milk or animal milk. Results from the prospective case-control (Van der Gaag 2015) study added that children who had full-fat milk (and full-fat dairy products) for 2 months significantly decreased respiratory tract complaints compared to children who had low-fat milk (and low-fat dairy products).

Data also indicated that children 12-23 mo receiving fortified milk, whether full-fat or lower-fat, had similar growth and nutrient status measured by iron serum concentrations. Children 12-23 mo receiving fortified milk also had higher haemoglobin and ferritin concentrations when given during 6 mo and exposed to 400 mL/d, higher dose, and exposure to that of follow-on formula on same outcomes. Children 12-23 mo that received fortified milk had higher serum zinc concentrations when given during 6 mo and 400 ml per day; and were more likely to be caries-free compared to those that received animal milk, when children were given 150 ml on weekdays, respectively. Data also indicated that children 12-23 mo receiving fortified milk were less likely to have anaemia and iron deficiency anaemia compared to those receiving animal milk when given during 12 mo and exposed to 400 mL/d, also higher dose, and exposure to that of follow-on formula on same outcomes. Data was scarce for most outcomes. As the quality of the evidence was, on average, mixed, the confidence in the effect estimate is limited and the true effect may be substantially different from the estimate of the effect. The prospective cohort (Brusner 1993) was significant in showing the first evidence of fortified milk in reducing gastrointestinal symptoms in terms of diarrhoea and bowel movements and improved stool consistency. The limitations of unbalanced groups, most likely due to the absence of random allocation into treatment groups, limits our interpretation of the study.

There was no data available for long term food preferences, longer term outcomes, feeding practices, dietary diversity, allergies, or phyto-oestrogen related outcomes for any of the comparisons and for most subgroup analysis.

4.2 Completeness and applicability of the evidence

This review included a total of 19 studies of which 16 were RCTs, involving 5,579 children 12-23 mo. Half of the studies were done in high-income countries and the other half in middle income (middle and upper middle) countries. The first ones are countries with low prevalence of anaemia, stunting and wasting among children 12-23 mo; while the second ones are likely to have a varied prevalence of all forms of malnutrition (UNICEF 2017). The studies that reported on anaemia and related outcomes, including ID, IDA, haemoglobin, and ferritin serum concentrations were conducted in both high- and middle-income countries, where prevalence are mixed. All studies with growth outcomes were conducted in high income countries where the prevalence of under nutrition is low. Nutrient composition of animal milks, follow-on formula and fortified milks fed to children 12-23 mo varied between studies (see Table 2, Table 3 and Table 4). The included RCTs were conducted in a varied of settings: urban cities and peri-urban areas, inner cities, day care centres and child health

centres. Most RCTs from middle-income countries were developed in the context on social programs. All these are important determinants of malnutrition and overall included outcomes.

Table 2. Nutrient content of animal milks (100 ml) in included studies

Study	Energy kJ	Kcal	Protein (g)	Carb (g)	Fat (g)	Vitamin A (µg)	Folic acid (µg)	Vitamin B12 (µg)	Vitamin C (mg)	Vitamin D (µg)	Iron (mg)	Zinc (mg)
Akkermans, 2016	-	-	3.5	5.2	1.7	13	1.6	142	0.55	0	0.02	0.4
Daly 1996[1]	-	67	3.2	4.8	3.9	52	6.0	0.4	1	0.03	0.05	0.4
Ghisolfi 2012	-	-	-	-	-	-	-	-	-	-	-	-
Lovell 2018	245	-	3.1	4.5	3.1[2]	-	-	-	-	0.06	0.02	-
Maldonado 2006[3]	270	64	3.1	4.7	3.6[4]	30	5.8	0.38	0.9	0.18	0.02	0.4
Morley 1999	-	-	-	-	-	-	-	-	-	-	0.005	-
Rivera 2010[5]	-	59.2	3.1	4.65	3.1	54	6	0.11	1.7	0.45	0.04	0.4
Sazawal 2007[6]	-	45.14	2.01	4.89	1.89	17.4	-	0.27	0.78	0.36	0	0.18
Stecksén-Blicks, 2009	-	-	-	-	-	-	-	-	-	-	-	-
Stekel, 1986[7]	-	-	-	-	-	-	-	-	-	-	-	-
Stekel, 1988	-	-	-	-	-	-	-	-	-	-	-	-
Svahn, 2000[8]	-	66	-	-	3.5	-	-	-	-	-	-	-
Szymlek-Gay, 2009[9]	-	-	-	-	-	N/A	-	-	-	N/A	0.01	-
Villalpando, 201910	248	-	3.12	4.67	3.12	5.4	1.3	-	1.7	-	0.05	0.47

N/A – quantity not specified

[1] Sodium (mg) 55, Potassium (mg) 140, Calcium (mg) 115, Magnesium (mg) 11, Phosphorus (mg) 92, Chloride (mg) 100, Carotene (µg) 21, Thiamin B1 (mg) 0.04, Riboflavin B2 (mg) 0.17, Nicotinamide (mg) 0.1, Vitamin B6 (mg) 0.06, Pantothenic acid (mg) 0.35, Biotin (µg) 1.9.

[2] Saturated (g) 1.9

[3] Cholesterol (mg) 14, Calcium (mg) 115, Phosphorus (mg) 100, Copper (µg) 7, Vitamin E (mg) 0.12, Thiamin B1 (µg) 38, Vitamin B2 (µg) 180, Vitamin B6 (µg) 32, Acid pantothenic (mg) 0.35, Biotin (µg) 2.5, Nicotinamide (µg) 95, Traces of Iodine.

[4] Fat (g): Saturated 2.3, Monounsaturated 1.2, Polyunsaturated 0.1.

[5] Riboflavin (mg) 0.13

[6] Taurine (mg): 4.8, Vitamin E (mg) 0.06, Thiamin (mg) 0.06, Riboflavin (mg) 0.18, Niacin (mg) 0.45, Vitamin B6 (mg) 0.06, Pantothenic acid (mg) 0.27, Folate (µg) 11.4, Biotin (µg) 2.49, Choline 11.4, Phosphorus (mg) 60, Magnesium (mg) 8.4, Iodine (µg) 3.6, Selenium (µg) 0.24, Copper (mg) 0.003, Sodium (mg) 36, Potassium (mg) 126, Chloride (mg) 90.

[7] Low fat milk without fortification

[8] Saturated fat (g/100ml) 2.17, Monounsaturated fat (g/100ml) 0.73, Polyunsaturated fat (g/100ml) 0.07, Cholesterol (mg/100ml) 9.1. Fatty acids (% wt/wt):
Myristic acid (C14:0) 10.8, Palmitic acid (C16:0) 30.6, Stearic acid (C18:0) 12.0, Palmitoleic acid (C16:1n-7) 1.6, Oleaic acid (C18:1n-9) 22.7, Linoleic acid (C18.2n-6) 1.5, α-Linolenic acid (C18:3n-3) 0.5.

[9] This animal milk had mandatory quantities of required vitamins A and D, quantities not specified.

[10] Vitamin A as retinol palmitate, vitamin C as sodium ascorbate and zinc as zinc oxide.

Table 3. Nutrient content of follow-on formulas (100 ml) in included studies

Study	Type of formula	Energy Kcal	Protein (g)	Carb (g)	Fat (g)	Vitamin A (µg)	Vitamin C (mg)	Vitamin D (µg)	Calcium (mg)	Iron (mg)	Zinc (mg)	Copper (µg)	Iodine (µg)
Akkermans 2016	Follow-on formula	-	-	-	-	-	-	X	-	-	-	-	-
Daly 1996[1]	Follow-on formula	67	2	8	3	80	10	1.1	72	1.2	0.4	41	-
Ghisolfie 2012[2]	Growing up milk (GUM)	-	-	-	-	-	X	X	X	X	X	-	-
Lovell 2018[3]	Reduced-energy growing-up milk (GUMLi)	60	1.7	7.8	1.9	-	-	1.3	-	1.7	-	-	-
Maldonado 2007[4]	Iron-supplemented toddler formula	66	2.4	7.4	3	65	4	1.3	108	1.2	0.6	42	15
Morley 1999	lower iron formula	-	-	-	-	-	-	-	-	0.09	-	-	-
	Follow on formula	-	-	-	-	-	-	-	-	0.12	-	-	-

N/A- quantity not specified

[1] Sodium 30 mg, Potassium 100 mg, Magnesium 7.1 mg, standard, Phosphorus 59 mg, Chloride: 65 mg. Thiamin B1 0.04 m, Riboflavin B1 0.15 mg, Nicotinamide 0.65 mg, Vitamin E 0.48 mg, Vitamin B6 0.04 mg, Vitamin B12 0.2 mg, Free folic acid 7 µg, Folic acid 0, Pantothenic acid 0.36 mg, Biotin 3 µg.

[2] Vitamin E, Magnesium

[3] Cholecalciferol (1.3 µg)

[4] Cholesterol (1.2 mg), Phosphor (77 mg), Vitamin E (0.8), Vitamin B1 (52 µg), Vitamin B2 (170 µg), Vitamin B6 (40 µg). Vitamin B12 (0.3 µg), Folic acid (5.3 µg), Pantothenic acid (0.35 mg), Biotin (2 µg), Nicotinamide (180 µg), Choline (10 mg).

Table 4. Nutrient content of fortified milks (100 ml) in included studies

Study	Type of milk	Energy Kcal	Protein (g)	Carb (g)	Fat (g)	Vitamin A (µg)	Vitamin C (mg)	Vitamin D (µg)	Folic acid (µg)	Iron (mg)	Calcium (mg)	Zinc (mg)	Fluoride (mg)
Brusner 1992	Fortified powder	-	-	-	-	-	-	-	-	12[1]		-	-
Rivera 2010 (C)[2]	Fortified powder	59.2	3.1	4.65	3.1	54	12	0.45	8.02	1.32	-	1.32	-
Sazawal 2007[3]	Fortified milk	45.14	2.01	4.89	1.89	33	4.8	0.36	-	0.96	72	0.96	-
Stecksén-Blicks 2009[4]	Milk Supplemented with Probiotic	-	-	-	-	-	-	-	-	-	-	-	250

	Lactobacilli and Fluoride												
Stekel 1986	Low fat milk (12% fat)	-	-	-	-	-	-	-	-	15[1]	-	-	-
Stekel 1988[5]	Full fat powdered milk	-	-	-	-	1500 (IU)	100	400 (IU)	-	15[1]	-	-	-
Svahn 2000 [6]	Low fat milk	42	-	-	1.0	-	-	-	-	-	-	-	-
	Partially vegetable-fat milk[7]	68	-	-	3.5	-	-	-	-	-	-	-	-
	Full vegetable fat milk[8]	80	-	-	3.5	-	-	-	-	-	-	-	-
Szymlek-Gay 2009[9]	Iron-fortified powdered cow milk	-	-	-	-	N/A	-	N/A	-	1.5[1]	N/A	N/A	-
Villalpando 2006[10]	Powdered milk	59.2	3.12	4.67	3.12	5.4	12	-	8.025	1.32	-	1.32	-

N/A – not quantity specified
[1] As ferrous sulphate
[2] Riboflavin (mg) 0.13, Sodium (mg) 44.5
[3] Taurine (mg): 4.8, Vitamin E (mg) 0.81, Thiamin (mg) 0.06, Riboflavin (mg) 0.18, Niacin (mg) 0.45, Vitamin B6 (mg) 0.06, Pantothenic acid (mg) 0.27, Folate (µg) 11.4, Vitamin B12 (µg) 0.27, Biotin (µg) 2.49, Choline 11.4, Phosphorus (mg) 60, Magnesium (mg) 8.4, Iodine (µg) 3.6, Selenium (µg) 0.66, Copper (mg) 0.03, Sodium (mg) 36, Potassium (mg) 126, Chloride (mg) 90.
[4] 10^7 CFU/ml L. rhamnosus LB21
[5] Streptococcus lactis (total acidity was 2.4 g lactic acid/l00 g powder and vacuum packaged in tin cans with a shelf life of 2 y)
[6] Saturated fat (g/100 ml) 0.62). Monounsaturated fat (g/100 ml) 0.21, Polyunsaturated fat (g/100ml) 0.02, Cholesterol (mg/100ml) 2.6. Fatty acids (% wt/wt): Myristic acid (C14:0) 10.8, Palmitic acid (C16:0) 30.6, Stearic acid (C18:0) 12.0, Palmitoleic acid (C16:1n-7) 1.6, Oleaic acid (C18:1n-9) 22.7, Linoleic acid (C18.2n-6) 1.5, α-Linolenic acid (C18:3n-3) 0.5.
[7] Saturated fat (g/100ml) 1.19, Monounsaturated fat (g/100ml) 1.35, Polyunsaturated fat (g/100ml) 0.6, Cholesterol (mg/100ml) 4.6, Fatty acids (% wt/wt): Myristic acid (C14:0) 6.6, Palmitic acid (C16:0) 19.1, Stearic acid (C18:0) 7.9, Palmitoleic acid (C16:1n-7) 1.1, Oleaic acid (C18:1n-9) 37.1, Linoleic acid (C18.2n-6) 11.8, α-Linolenic acid (C18:3n-3) 5.8.
[8] Saturated fat (g/100ml) 1.67, Monounsaturated fat (g/100ml) 1.16, Polyunsaturated fat (g/100ml) 0.67, Cholesterol (mg/100ml) 0, Fatty acids (% wt/wt): Myristic acid (C14:0) 3.8, Palmitic acid (C16:0) 31.4, Stearic acid (C18:0) 3.9, Palmitoleic acid (C16:1n-7) 1.3, Oleaic acid (C18:1n-9) 28.3, Linoleic acid (C18.2n-6) 17.7, α-Linolenic acid (C18:3n-3) 1.4.
[9] Magnesium, Niacin (Vitamin B3), Thiamin (Vitamin B1), Vitamin B6, Folate.
[10] Vitamin A as retinol palmitate, vitamin C as sodium ascorbate, iron as ferrous gluconate, and zinc as zinc oxide.

Studies that reported on nutrient status only assessed iron, zinc, vitamin D and lipids serum profile, and those that reported on children growth only assessed weight and height. Data was only available for subgroup analysis in growth and haematological outcomes. None of the included studies had co-interventions in the intervention and comparison groups. Another potential aspect of the evidence is the duration of the milk feeding period and its effects on measured outcomes. Four studies began the intervention when children were younger and continue follow-up

until children had 15 and 18 months of age and two studies continue until children had 36 mo or more. The children´s age range included in this review (12-23 mo) is part of the first 1,000 days, the window of opportunity for optimal children growth and development, however, for many of the outcomes measured, a greater impact can be expected in younger children (i.e., 12-18 mo). The included studies had a total duration of follow-up from 4 months to 12 months. Most trials considered a dose of milk of at least 250 ml per day, whether animal milk (full-fat or lower-fat), follow-on formula, plant-based milk or fortified milk; but only few studies considered milk consumed as part of healthy complementary feeding practices in children under 2 years of age, which may include consumption of milk in other presentations like cereal with milk or dairy products.

We are uncertain about the effect of follow-on formula and fortified milk on growth outcomes regarding the prevalence of stunting or wasting compared to animal milk. Results also suggest that feeding children 12-23 mo with follow-on formula or fortified milk every day is enough to produce a positive effect on nutrient status assessed by serum concentrations of vitamin D, iron, and zinc. The efficacy of these interventions is higher than the efficacy of animal milk in different settings.

Regarding anaemia and haematological outcomes, according to our review feeding children 12-23 mo with follow-on formula containing at least 1.2 mg of iron per 100 ml and other micronutrients every day for 5 months, is enough to produce a positive effect on iron deficiency in the context of high-income countries. Results do not suggest the same positive effect when feeding children 12-23 mo with fortified milk (1 study) with a similar iron and micronutrient content in the context of a social program targeting vulnerable population. Results suggest that feeding children 12-23 mo with follow-on formula has a similar effect than animal milk on iron deficiency anaemia. However, feeding children with fortified milk may be effective in reducing anaemia and iron deficiency anaemia. We are uncertain if follow-on formula has a similar effect on anaemia as no studies reported on this outcome.

Children 12-23 mo fed with fortified milk were more likely to be caries-free compared to those fed with animal milk. Fortified milk had similar effects than animal milk on dfms index (decayed, missing and filled surfaces) on molars and canines in children. There was only one study comparing the effects of feeding children with animal milk or no other milk on number of stools, stools weight, fluid, and energy intake. However, it did not provide data to abstract for the analysis.

Information regarding the benefits of plant-based milks and lower-fat milks on the prespecified outcomes was scarce and there was no information on the effects of animal milk versus no other milk on children 12-23 mo. One study (21 participants) reported on nutrient status measured by lipids serum concentrations of children fed with plant-based milk and lower-fat milk (Svahn 1999). The study found no evidence that the effects of these milks on these indicators were different from that produced by animal milk. None of the other included studies reported data on these outcomes. There was only one study comparing the effects of feeding children with follow-on formula compared to animal milk (full-fat or lower-fat) on child development index scores (Bayley mental index (MDI) and psychomotor development index (PDI). It found no effect in feeding children 12-23 mo with follow-on formula.

4.3 Quality of the evidence

We assessed 5 RCTs, out of the 16 RCTs that contributed to meta-analysis, as having overall low risk of bias and seven studies as having a high or unclear risk of bias in some of the domains. Excluding these studies during sensitivity analysis, while reduced heterogeneity, did not alter the results nor the conclusions of the evidence we found in the meta-analysis.

Study limitations and risk of bias of included studies

Some studies did not describe the randomisation method used. Many studies did not described blinding of participants, care providers and outcome assessors or it was not attempted at all, although some studies reported that technical staff and field workers were unaware of group allocation as milks were provided in similar packages or tins in most studies. We assessed two studies as being at risk for confounding because baseline characteristics of the participants were not specified. One study had a small sample size.

Imprecision

Imprecision due to very small sample sizes or few events in the included studies was unlikely with exception of one study (Svahn 1999). However, we considered imprecision in continuous outcomes (i.e. haemoglobin and ferritin measurements) an important factor in the overall assessment of the evidence. There was considerable variation in the effects of the interventions among participants for continuous outcomes, as results showed wide CIs around the effect estimate (Ryan 2016).

Inconsistency

We considered that clinical inconsistency was unlikely for our outcomes. Variability in participants characteristics, interventions, and outcomes across the included studies was likely to be low (Ryan 2016). However, methodological inconsistency was a potentially important factor in the overall assessment of evidence for our outcomes. We found differences between studies in terms of methodological factors, specifically blinding and allocation concealment, that may have led to differences in the observed intervention effects (Higgins 2011b). We found substantial heterogeneity in some outcomes, that could be partly explained by subgroup analyses. However, we did not perform subgroup analysis for secondary outcomes (i.e., haemoglobin and ferritin). Although this does not necessarily mean that the true intervention effect varies, results should be interpreted with some caution.

Indirectness

We found no indirectness regarding interventions, or outcomes assessed across studies. We found indirectness regarding study populations, as studies included a variety of different settings which paralleled those under real conditions, but vulnerable populations were almost not included.

The evidence summarised in the review comes from studies addressing many of the main review questions. However, several of our primary and secondary outcomes were almost not addressed.

4.4 Potential bias in the review process

Two review authors independently carried out the review process, with the same data extraction sheet and tools to assess risk of bias in the included studies. Both review authors independently assessed the eligibility of studies for inclusion, participated in data extraction, and conducted the 'Risk of bias' assessments. One review author entered the data into a form design for the review and the other checked the data for accuracy.

Many studies had minimum information regarding the allocation concealment and blinding. In the absence of precise details, we considered mutual discussion among review authors as final in this review since these include subjective components. Many studies reported some outcomes in a non-extractable way such as figures or statistics not compatible for use in meta-analysis. Authors did not set any language limit in the search for studies in this review and we obtained all full-text articles. We extensively searched grey literature and trials registries, along with contacting agencies involved in carrying out RCTs and subject experts, thus minimising publication bias in this review.

4.5 Agreements and disagreements with other reviews

To our knowledge, this is the first review to compare animal milks with follow-on formula, plant-based milks, lower-fat milks, and fortified milks on outcomes in children 12-23 mo. The body of evidence in this review demonstrates beneficial effects of follow-on formula and fortified milk on children 12-23, both similar and higher to those of animal milk depending on the outcome.

A review on the use of follow-on formulas (Vandenplas 2014b) concluded based on the evidence they reviewed, that follow-on formulas are effective in improving micronutrient status compared to animal milk and decreasing protein intake unlike fortified milks. Although we only had data to assess the effect of follow-on formulas on two micronutrients (i.e., iron and vitamin D), our findings are consistent with this review. Follow-on formula was also effective in reducing iron deficiency anaemia and improving haemoglobin and ferritin concentrations in our review. A cross-sectional observation study with food survey and blood sampling found similar results in children 24 mo of age and found association with follow-on formula during 5.7 mo (volume consume between <100 and >500 ml per day) and a better overall iron status, including increased haemoglobin and ferritin concentrations and a significantly decreased frequency in iron deficiency (Sacri 2021).

A review by Matsuyama 2016 that assessed the effect of fortified milk on growth and nutrient status in children 6-47 months of age, included eight studies that reported positive changes in haemoglobin concentration (g/l) in children 6 to 47 mo fed with fortified milk. Six out of those eight studies were included in our review. However, we excluded Stevens 1995 because authors compared two follow-on formulas, and Xuan 2013 did not compare against animal milk. According to our definitions of interventions, two other studies compared animal milk versus follow-on formula (Morley 1999 and Daly 1996) not fortified milk. Consistent with our results, authors did not find any effect of fortified milk on weight of children

and concluded that fortified milk is an effective source of complementary nutrition to supplement children in need when consumed in appropriate amounts in addition to a normal diet.

Consistent with our findings, a previous review on fortified milk and cereal foods on infants and children aged 6 months to 5 years found that iron+micronutrients fortification increased haemoglobin concentrations compared to the control group by 0.62 g/dl (95%-CI: 0.34 to 0.89) and ferritin by 11.3 µg per litre (95%-CI: 3.3 to 19.2) compared to control groups (Eichler 2012). Also, fortified milk (or cereals) reduced the risk of suffering from anaemia by 50% (risk ratio 0.50, 95%-CI: 0.33 to 0.75). This evidence was result of pooling 13 and 11 trials of a mix of fortified milk, follow-on formulas, and fortified products, respectively. Several of the studies include in Eichler 2012 were also included in this review too. The authors of Eichler 2012 concluded that micronutrients fortified milk and cereal products can be an effective alternative to reduce anaemia of children up to three years of age in developing countries, however, the authors concluded the evidence for functional health outcomes was still inconclusive.

5. Conclusions

Feeding children 12-23 mo with other milks has been recommended when breast milk is no longer provided. When given on a regular basis, feeding children 12-23 mo with at least 250 ml of follow-on formula with added iron is enough to have some benefits on their micronutrient status (evidence limited to vitamin D), improve their haemoglobin and ferritin concentrations and decrease their risk of having iron deficiency more effectively than animal milk. Evidence also shows, that feeding children 12-23 mo regularly with at least 250 ml of fortified milk also improves micronutrient status (evidence limited to zinc), improve their haemoglobin concentrations, and reduces the risk for anaemia and iron deficiency anaemia. The effect of follow-on formula and fortified milk seems to be context specific as fortified milk has mostly been tested in the context of social programs while follow-on formula has been tested in clinics, day care centres and other community settings.

Fortified milk has a similar response to that of animal milk with respect to morbidity (respiratory and diarrhoea episodes), and oral health (mean decayed, missing, and filled surfaces index in molars and canines and caries free).

The efficacy of follow-on formula and fortified milk on growth indicators, including body weight, height, head circumference and body mass index, is similar to that of animal milk. There is very limited evidence on the efficacy and effectiveness of plant-based milks and lower-fat milks to recommend feeding children 12-23 with them and confer similar benefits.

Current evidence shows no indication that any of these interventions has adverse effects on children health and nutrient status. This review may provide enough evidence supporting the effectiveness of follow-on formula and fortified milk on some micronutrients, anaemia, and haematological status. However, there was heterogeneity between studies using follow-on formulas and fortified milks and we could not draw reliable conclusions from various subgroup analyses due to a limited number of studies in each subgroup. More evidence is needed for growth,

long term food preferences, longer term outcomes, child development, as well as other micronutrients outcomes. Particularly, evidence on plant-based milk and lower-fat milk is scarce.

Implications for research

Overall, we found information about animal milk, follow-on formula, plant-based milk, lower-fat milk, and fortified milk confusing as these interventions are mixed in many studies (for example, some use different terms without defining them as milk-based beverages, cow milk formula, standard formula). Lack of methodological rigor in some RCTs included in this review has resulted in medium to low quality evidence in the review. Improving the quality of primary studies is needed.

This review has highlighted the need for further research in this area. Studies in settings where children are vulnerable and different forms of malnutrition are prevalent are needed, particularly on:

1. the effects of not proving any milk on children's growth, body composition, longer-term outcomes, nutrient intake and status, oral health, and early childhood development.

2. the effects of animal milk on children's growth, body composition, long-term food preferences and longer-term outcomes, nutrient intake and status, and oral health, considering local regulations on the mandatory or voluntary addition of certain vitamins, especially vitamin D and A.

3. the effects of follow-on formula on children's long-term food preferences, longer-term outcomes, nutrient status, oral health, child development and feeding practices, considering their nutrient profile.

3. the effects of lower-fat milk on children's growth, body composition, long-term food preferences, longer-term outcomes, nutrient intake, and status, feeding practices, oral health, and early child development.

4. the effects of plant-based milk on children's growth, body composition, long-term food preferences, nutrient intake and status, phyto-oestrogen related outcomes, feeding practices, oral health, allergies, and early child development, considering their nutrient profile.

5. the effects of fortified milk on children's growth, body composition, long-term food preferences, longer-term outcomes, feeding practices, oral health, and early child development, considering their nutrient composition+.

6. References

6.1 Included studies

Akkermans 2016

Akkermans M, Eussen S, Van Der Horst-Graat J, Van Elburg R, Van Goudoever H, Brus, F. A micronutrient-fortified young child formula improves the iron and vitamin D status of healthy young European children: A randomised double-blind controlled trial. J Pediatric Gastroenterol Nutr 2017;62:663.
Akkermans MD, van der Horst-Graat J, Eussen S, van Goudoever J,Brus F. Iron and vitamin D deficiency in healthy young children in Western Europe despite current nutritional recommendations. J Pediatr Gastroenterol Nutr 2016;62:635-42.

Bhatnagar 1996

Bhatnagar S, Bhan MK, Singh KD, Saxena SK, Shariff M. Efficacy of milk-based diets in persistent diarrhea: a randomized, controlled trial. Pediatrics 1996;98(6 part 1):1122-26.

Brunser 1993

Brunser, O.; Espinoza, J.; Araya, M.; Pacheco, I.; Cruchet, S. Chronic iron intake and diarrhoeal disease in infants. A field study in a less-developed country. Eur J Clin Nutr 1993;47(5):317-26.

Chatchatee 2014

Chatchatee P, Lee WS, Carrilho E, Kosuwon P, Simakachorn N, Yavuz Y, Schouten B, Graaff PL, Szajewska H. Effects of growing-up milk supplemented with prebiotics and LCPUFAs on infections in young children. J Pediatric Gastroenterol Nutr 2014;58(4):428-37.

Daly 1996

Daly A, MacDonald A, Aukett A, Williams J, Wolf A, Davidson J, Booth IW. Prevention of anaemia in inner city toddlers by an iron supplemented cows' milk formula. Archives Dis Childh 1996;75(1):9-16.
Daly A. Prevention of anaemia in inner-city toddlers by the use of a follow-on formula. Professional Care Mother Child 1997;7(5):141-46.
Williams J, Wolff A, Daly A, MacDonald A, Aukett A, Booth IW. Iron supplemented formula milk related to reduction in psychomotor decline in infants from inner city areas: randomised study. BMJ (clinical research ed.) 1999;318(7185):693-697.

Ghisolfi 2012

Fantino M, Gourmet E. Nutrient intakes in 2005 bynon breast-fed French children of less than 36 months. Arch Pediatr 2008;15:446-455.
Ghisolfi J, Fantino M, Turck D, de Courcy GP, Vidailhet M, Ghisolfi J, et al. Nutrient intakes of children aged 1-2 years as a function of milk consumption, cows' milk or growing-up milk. Public Health Nutrition 2012;16(3):524-34.

Gill 1997

Gill DG, Vincent S, Segal DS. Follow-on formula in the prevention of iron deficiency: a multicentre study. Acta Paediatr 1997;86(7):683-89.

Lovell 2018

Davies PSW, Wall CR, Hill RJ, Lovell AL, Matsuyama M, Milne T, et al. Growth and body composition in children randomised to cow's milk or a growing up milk between 1 and 2 years of age. J Pediatric Gastroenterol Nutr 2018;66(Suplement 2):1137.

Lovell AL, Davies PSW, Hill RJ, Milne T, Matsuyama M, Jiang Y, et al. Compared with cow milk, a growing-up milk increases vitamin D and iron status in healthy children at 2 years of age: the Growing-Up Milk-Lite (GUMLi) Randomized Controlled Trial. J Nutr 2018;148(10):1570-79.

Lovell AL, Davies PSW, Hill RJ, Milne T, Matsuyama M, Jiang Y, et al. A comparison of the effect of a Growing Up Milk - Lite (GUMLi) v. cows' milk on longitudinal dietary patterns and nutrient intakes in children aged 12-23 months: the GUMLi randomised controlled trial. Br J Nutr 2019;121(6):678-87.

Lovell AL, Milne T, Jiang Y, Chen RX, Grant CC, Wall CR. Evaluation of the Effect of a Growing up Milk Lite vs. Cow's Milk on Diet Quality and Dietary Intakes in Early Childhood: the Growing up Milk Lite (GUMLi) Randomised Controlled Trial. Nutrients 2019;11(1):203.

Wall CR, Hill RJ, Lovell AL, Matsuyama M, Milne T, Grant CC, et al. A multicenter, double-blind, randomized, placebo-controlled trial to evaluate the effect of consuming Growing Up Milk "Lite" on body composition in children aged 12-23 mo. Am J Clin Nutr 2019;109(3):576-85.

Maldonado 2007

Maldonado LJ, Baró L, Ramírez-Tortosa MC, Gil F, Linde J, López-Huertas E, et al. Intake of an iron-supplemented milk formula as a preventive measure to avoid low iron status in 1-3 year-olds. Anales de Pediatría 2007;66(6):591-96.

Morley 1999

Morley R, Abbott R, Fairweather-Tait S, MacFadyen U, Stephenson T, Lucas A. Iron fortified follow on formula from 9 to 18 months improves iron status but not development or growth: a randomised trial. Arch Dis Child 1999;81(3):247-252.

Singhal A, Morley R, Abbott R, Fariweather-Tait S, Stephenson T, Lucas A. Clinical safety of iron-fortified formulas. Pediatrics 2000;105(3):E38.

Singhal A, Morley R, Abbott R, Fariweather-Tait S, Stephenson T, Lucas A. Clinical safety of iron-fortified formulas. Pediatrics 2000;105(3):E38.

Rivera 2010 (C)

Mendez Gomez I, Shamah Levy T, Mendez Ramirez I. Three statistical approaches to the estimation of intervention effect on a randomized cluster trial to reduce anemia and iron deficiency in children aged 12-30 months. FASEB Journal 2014;28(supplement 1):810.27.

Rivera JA, Shamah LT, Villalpando S, Monterrubio E. Effectiveness of a large-scale iron-fortified milk distribution program on anemia and iron deficiency in low-income young children in Mexico. Am J Clin Nutr 2010;91(2):431-39.

Sazawal 2007

Sazawal S, Dhingra U, Dhingra P, Hiremath G, Kumar J, Sarkar A, et al. Effects of fortified milk on morbidity in young children in north India: community based, randomised, double masked placebo controlled trial. BMJ (Clinical research ed.) 2007;334(7585):140.

Sazawal S, Dhingra U, Dhingra P, Hiremath G, Sarkar A, Dutta A, et al. Micronutrient fortified milk improves iron status, anemia and growth among children 1-4 years: a double masked, randomized, controlled trial. PloS one 2010;5(8):e12167.
Sazawal S, Dhingra U, Hiremath G, Sarkar A, Dhingra P, Dutta A, Menon VP, Black R.. Effects of Bifidobacterium lactis HN019 and prebiotic oligosaccharide added to milk on iron status, anemia, and growth among children 1 to 4 years old. J Pediatric Gastroenter Nutr 2010;51(3):341-346.

Stecksén-Blicks 2009 (C)

Stecksén-Blicks C, Sjöström I, Twetman S. Effect of long-term consumption of milk supplemented with probiotic lactobacilli and fluoride on dental caries and general health in preschool children: a cluster-randomized study. Caries research 2009;43(5):374-81.

Stekel 1986

Stekel A, Olivares M, Pizarro F, Chadud P, Cayazzo M, López I, et al. Prevention of iron deficiency in infants by fortified milk. Field study of a low-fat milk. Arch Latinoam Nutr 1986;36(4):654-61.

Stekel 1988

Stekel A, Olivares M, Cayazzo M, Chadud P, Llaguno S, Pizarro F. Prevention of iron deficiency by milk fortification. II A field trial with a full-fat acidified milk. Am J Clin Nutr 1988;47(2):265-69.

Svahn 1999

Svahn JC, Axelsson IE, Räihä NC. Macronutrient and energy intakes in young children fed milk products containing different quantities and qualities of fat and protein. Journal of pediatric gastroenterology and nutrition. J Pedtr Gastroenterlo Nutr 1999;29(3):273-81.
Svahn JC, Feldl F, Räihä NC, Koletzko B, Axelsson IE. Fatty acid content of plasma lipid fractions, blood lipids, and apolipoproteins in children fed milk products containing different quantity and quality of fat.. J Pediatric Gastroenterol Nutr 2000;31(2):152-161.
Svahn JC, Feldl F, Raiha NCR, Koletzko B, Axelsson IEM. Different quantities and quality of fat in milk products given to youngchildren: effects on long chain polyunsaturated fatty acids and transfatty acids in plasma. Acta Paediatr 2002;91:20-29.
Virtanen MA, Svahn CJ, Viinikka LU, Räihä NC, Siimes MA, Axelsson IE. Iron-fortified and unfortified cow's milk: effects on iron intakes and iron status in young children. Acta Paediatrica 2001;90(7):724-731.

Szymlek-Gay 2009

Houghton LA, Gray AR, Szymlek-Gay EA, Heath AL, Ferguson EL. Vitamin D-fortified milk achieves the targeted serum 25-hydroxyvitamin D concentration without affecting that of parathyroid hormone in New Zealand toddlers. J Nutrition 2011;141(10):1840-1846.
Szymlek-Gay EA, Ferguson EL, Heath AL, Gray AR, Gibson R. Food-based strategies improve iron status in toddlers: a randomized controlled trial. Am J Clin Nutr 2009;90:1541-51.
Szymlek-Gay EA, Gray AR, Heath ALM, Ferguson EL, Edwards T, Skeaff SA. Iodine-fortified toddler milk improves dietary iodine intakes and iodine status in toddlers: a randomised controlled trial. Eur J Nutr 2020;59(3):909-919.

Van der Gaag 2015

Van Der Gaag, E.. Whole milk and butter; clinically significant reduction of respiratory tract complaints in children. Annals of Nutrition and Metabolism 2015;67():305 2015;67:305.

Villalpando 2006

Villalpando S, Shamah LT, Rivera JA, Lara Y, Monterrubio E. Fortifying milk with ferrous gluconate and zinc oxide in a public nutrition program reduced the prevalence of anemia in toddlers. J Nutr 2006;136(10):2633-37.

6.2 Other references

Agostoni 2000. Agostoni C, Carratu B, Boniglia C, Riva E, Sanzini E. Free amino acid content in standard infant formulas: comparison with human milk. J Am Coll Nutr 2000;19:434-8.

Agostoni 2011. Agostoni C, Turck D.. Is cow's milk harmful to a child´s health? J Pediatr Gastroent Nutr 2011;53(6):594-600.

Allen 2001. Allen LH, Gillespie S. What works? A review of the efficacy and effectiveness of nutrition interventions. United Nations Administrative Committee on Coordination Sub-Comittee on Nutrition edition. Vol. Nutrition policy paper 19. Geneva in collaboration with the Asian Development Bank, Manila: ACC/SCN (2001). SR. ACC/SCN, 2001.

Andreas 2015. Andreas NJ, Kampmann B, Le-Doare KM. Human breast milk: a review on its composition and bioactivity. Early Human Dev 2015;91(11):629-35.

Avila 2015. Avila W, Pordeus I, Paiva S, Martins C. Breast and Bottle Feeding as Risk Factors for Dental Caries: A Systematic Review and Meta-Analysis. PLoS One 2015;10(11):e0142922.

Balarajan 2011. Balarajan Y, Ramakrishnan U, Ozaltin E, Shankar AH, Subramanian SV. Anaemia in low-income and middle-income countries. Lancet 2011;378:2123-35.

Balsheim 2011. Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: Rating the quality of evidence. J Clin Epidemiol 2011;64(4):401-6.

Borenstein 2008. Borenstein M, Hedges LV , Higgins JPT , Rothstein HR. Introduction to Meta-Analysis (Statistics in Practice). Chichester (UK): John Wiley & Sons, 2008.

Campbell 2017. Campbell KJ, Abbott G, Zheng M, McNaughton SA. Early life protein intake: Food sources, correlates, and tracking across the first 5 years of life. J Acad Nutr Diet 2017;117(8):1188-97.

Chen 2017. Chen PL, Soto-Ramirez N, Zhang H, Karmaus W. Association Between Infant Feeding Modes and Gastroesophageal Reflux: A Repeated Measurement Analysis of the Infant Feeding Practices Study II. J Hum Lact 2017;33(2):267-277.

Covidence 2017. Covidence systematic review software. Melbourne, Australia: Veritas HealthInnovation. Available at www.covidence.org., 2017.

Critch 2014. Critch JN. Nutrition for healthy term infants, six to 24 months: An overview. Paediatr Child Health 2014;Dec; 19(10):547–549.

Deeks 2011. Deeks JJ, Higgins JPT , Altman DG , editor(s). Chapter 9: Analysing data and undertaking meta-analyses. In: In: Higgins JPT, Green S, editor(s). Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0 (updated March 2011). The Cochrane Collaboration. Available from handbook.cochrane.org 2011.

de Onis 2018. de Onis M, Borghi E, Arimond M, Webb P, Croft T, Saha K, et al. Prevalence thresholds for wasting, overweight and stunting in children under 5 years. Public Health Nutrition 2018;doi:10.1017/S1368980018002434.

Dewey 2003. Dewey K, Brown K. Update on technical issues concerning complementary feeding of young children in developing countries and implications for developing countries. Food Nutr Bull 2003;24(1):5-28.

EFSA 2016. European Food Safety Authority. Report from the commission to the European Parliament and the council on young-child formulae. EFSA, 2016.

Eichler 2012. Eichler K, Wieser S, Ruthermann I, Brugger U. Effects of micronutrient fortified milk and cereal food for infants and children: a systematic review. BMC Public Health 2012;12:506.

FAO 1987. FAO. CODEX standard for follow-up formula. CODEX STAN 156 1987; Amended 1989, 2011, 2017:2-4.

FAO 1999. FAO. CODEX general standard for the use of dairy terms. Vol. CODEX STAN 206-1999. Rome: FAO, 1999.

Fox 2006. Fox MK, Reidy K, Novak T, Ziegler P. Sources of energy and nutrients in the diets of infants and toddlers. J Am Diet Assoc 2006;106(1):28 e1-e5.

GRADEpro GDT 2015. McMaster University. GRADEpro GDT:GRADEpro GuidelineDevelopment Tool. McMaster University [Computer program]. (developed by Evidence Prime, Inc)., 2015.

Hernell 2011. Hernell O. Human milk vs cow´s milk and the evolution of infant formulas. Nestlé Nutr Inst Workshop Ser Pediatr Program 2011;67:17-28.

Higgins 2011a. Higgins JPT , Deeks JJ , Altman DG , editor(s). Chapter 16: Special topics inStatistics. In: Higgins JPT, Green E, editor(s). Cochrane Handbook forSystematic Reviews of Interventions Version 5.1.0 (updated March 2011). In: The Cochrane Collaboration. Available from handbook.cochrane.org. Cochrane, 2011.

Higgins 2011b. Higgins JPT, Green S, editor(s). The Cochrane Collaboration, 2011. Vol. Available from handbook.cochrane.org. Cochrane, 2011.

Higgins 2019. Higgins JPT, Savovic J, Page MJ, Elbers RG, Sterne JAC. Chapter 8: Assessing risk of bias in a randomized trial. In: Higgins JPT, Thomas J, Chandler J, Cumpston M, Li T, Page MJ, Welch VA, editor(s). Cochrane Handbook for Systematic Reviews of Interventions version 6.0. Cochrane, (updated July 2019).

Hoppe 2005. Hoppe C, Molgaard C, Vaag A, Barkholt V, Michaelsen KF. High intakes of milk, but not meat, increase s-insulin and insulin resistance in 8-year-old boys. Eur J Clin Nutr 2005;59:393-8.

Høst A 2002. Høst A, Halken S, Jacobsen HP, et. al. Clinical course of cow's milk protein allergy/intolerance and atopic diseases in childhood. Pediatr Allergy Immunol 2002;13(15):23–28.

Jensen 1995. Jensen RG. Handbook of Milk Composition. San Diego: Academic Press, 1995.

Koletzko 2012. Koletzko B, Brands B, Poston L, Godfrey K, Demmelmair H. Early nutrition programming of long-term health. Proc Nutr Soc 2012;71:371-378.

Le Huërou-Luron 2018. Le Huërou-Luron I, Bouzerzour K, Ferret-Bernard S, et al. A mixture of milk and vegetable lipids in infant formula changes gut digestion, mucosal immunity and microbiota composition in neonatal piglets. Eur J Nutr 57;2:463-476.

Malunga 2014. Malunga LN, Bar-El Dadon S, Zinal E, Berkovich Z, Abbo S, Reifen R. The potential use of chickpeas in development of infant follow-on formula. Nutr J 2014;13:8.

Martin 2011. Martin RM, Holly JM, Gunnell D. Milk and linear growth: programming of the IGF-I axis and implication for health in adulthood. Nestle Nutr Workshop Ser Pediatr Program 2011;67:79-97.

Mazzocchi 2018. Mazzocchi A, D'Oria V, De Cosmi V, Bettocchi S, Milani GP, Silano M, et. al. The role of lipids in human milk and infant formulae. Nutrients 2018;10:567.

Mendonça 2017. Mendonça MA, Araújo WMC, Borgo LA, Alencar ER. Lipid profile of different infant formulas for infants. PLoS One 2017;12(6):e0177812.

Michaelsen 2007. Michaelsen KF, Hoppe C, Lauritzen L, Molgaard C. Whole cow’s milk: why, what and when? Nestle Nutr Workshop Ser Pediatr Program 2007;60:1-16 discussion 216-9.

Michaelsen 2014. Michaelsen K, Greer F. Proteins needs early in life and long term health. Am J Clin Nutr 2014;99:718S-22S.

Moher 2009. Moher D , Liberati A , Tetzlaff J , Altman DG , PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLOS Medicine 2009;6(7):e1000097.

Mølgaard 2011. Mølgaard C, Larnkjær A, Arnberg K, Michaelsen KF. Milk and growth in children: effects of whey and casein. Nestle Nutr Workshop Ser Pediatr Program 2011;67:67-78.

Nejrup 2017. Nejrup RG, Licht TR, Hellgren LI. Fatty acid composition and phospholipid types used in infant formulas modifies the establishment of human gut bacteria in germ-free mice. Fatty acid composition and phospholipid types used in infant formulas modifies the establishment of human gut bacteria in germ-free mice 2017;7(1):3975.

Review Manager 2014. Review Manager 5 (RevMan 5). Version 5.3 [Computer program]. Copenhagen: Nordic Cochrane Centre: The Cochrane Collaboration, 2014.

Ryan 2016. Ryan R, Hill S. Cochrane Consumers and Communication Group. How to GRADE the quality of the evidence Version 3.0. Available at: cccrg.cochrane.org/author-resources (accessed January 20 2021) 2016;December:1-25.

Sacri 2021. Sacri AS, Bocquet A, Montalembert M, Hercberg S, Gouya L, Blondel B, et al. Young children formula consumption and iron deficiency at 24 months in the general population: A national-level study. Clin Nutr 2021;40(1):166-173.

Sankar 2015. Sankar MJ, Sinha B, Chowdhury R, et. al. Optimal breastfeeding practices and infant and child mortality: a systematic review and meta-analysis. Acta Paediatr 2015;104(467):3-13.

Stephen 2012. Stephen A, Alles M, de Graaf C, Fleith M, Hadjilucas E, Isaacs E, et al. The role and requirements of digestible dietary carbohydrates in infants and toddlers. Eur J Clin Nutr 2012;66:765-779.

Sterne 2016. Sterne J, Hernan MA, Reeves BC, Savović J, Berkman N, Viswanathan M, et al. ROBINS-I: a tool for assessing risk of bias in non-randomised studies of interventions. BMJ 2016;355:i4919.

Sterne 2019. 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;366:14898.

Stevens 2013. Stevens GA, Finucane MM, De-Regil LM, Paciorek CJ, Flaxman SR, Branca F et al. Global, regional, and national trends in haemoglobin concentration and prevalence of total and severe anaemia in children and pregnant and non-pregnant women for 1995–2011: a systematic analysis of population-representative data. Lancet Glob Health 2013;1(1):E16-E25.

Stoltzfus 2004. Stoltzfus RJ, Mullany L, Black RE. Iron deficiency anaemia. In: Ezzati M, Lopez Ad, Rodgers A, Murray CJL, editors. Comparative quantification of health risks: global and regional burden of disease attributable to selected major risk factors. Geneva: World Health Organization, 2004.

Tolentino 2007. Tolentino K, Friedman JF. An update on anemia in less developed countries. Am J Trop Med Hyg 2007;77:44-51.

Turck 2013. Turck D. Cow's milk and goat's milk. World Rev Nutr Diet 2013;108:56-62.

Uauy 2009. Uauy R, Dangour A. Fat and fatty acid requirements and recommendations for infants of 0-2 years and children of 2-18 years. Ann Nutr Metab 2009;55:76-96.

UNICEF 2007. United Nations Children Fund. Indicators for assessing infant and young child feeding practices Part 1 Definitions. Conclusions of a consensus meeting held 6–8 November 2007 in Washington, DC, US. Washington: Unicef, 2007.

UNICEF 2017. United Nations Children Fund. The State of the World Children. UNICEF 2017.

Vandenplas 2014a. Vandenplas Y, Castrellon PG, Rivas R, et al. Safety of soya-based infant formulas in children. Br J Nutr 2014;11(8):1340-1360.

Vandenplas 2014b. Vandenplas Y, De Ronne N, Van De Sompel A, Huysentruyt K, Robert M, Rijo J, et al. A Belgian consensus-statement on growing-up milks for children12–36 months old. Eur J Pediatr 2014;173(10):1365-71.

Verduci 2019. Verduci E, D'Elios S, Cerrato L, et al. Cow's Milk Substitutes for Children: Nutritional Aspects of Milk from Different Mammalian Species, Special Formula and Plant-Based Beverages. Nutrients. 2019;11(8):1739. Published 2019 Jul 27. doi:10.3390/nu11081739. Cow's Milk Substitutes for Children: Nutritional Aspects of Milk from Different Mammalian Species, Special Formula and Plant-Based Beverages. Nutrients 2019;11(8):1739.

Victora 2008. Victora CG, Adair L, Fall C, Hallal PC, Martorell R, Ritcher L, Sachdev HS, Maternal and Child Undernutrition Study Group. Maternal and child under nutrition: consequences for adult health and human capital. Lancet 2008;371(9609):340-57.

WHO 2001. World Health Organization. Iron deficiency anaemia: assessment, prevention, and control. A guide for programme managers. Vol. (WHO/NHD/01.3). Geneve: World Health Organization, 2001.

WHO 2003. WHO. Global strategy for infant and young child feeding. Geneva: World Health Organization, 2003.

WHO 2005. World Health Organization. Guiding principles for feeding non-breastfed children 6-24 months of age. Geneva: WHO Library Cataloguing.in-Publication Data, 2005.

WHO 2011. World Health Organization, Food and Agriculture Organization of the United Nations. Milk and milk products. second edition edition. Codex Alimentarius, 2011.

WHO 2014a. WHO. Global nutrition targets 2025: stunting policy brief (WHO/NMH/NHD/14.3). Geneva: World Health Organization, 2014.

WHO 2014b. WHO. Global Nutrition Targets 2025: Wasting policy brief. Geneva: World Health Organization, 2014.

WHO 2018. WHO. Information note: clarification on the classification of follow up formulas for children 6-36 months as breastmilk substitutes. Geneva: World Health Organization, 2018.

Woldu 2014. Woldu M, Mezgebe HB, Lekisa J. Consumption of unmodified cow´s milk and the risk of iron deficiency anemia in infants and toddlers and its management. IJPSR 2014;5(1):51-59.

World Bank 2017. World Bank. Prevalence of anemia among children (% of children under 5). World Bank, 2019.

Xuan 2013. Xuan NN, Wang D, Grathwohl D, Thi Lan PN, Thi Kim HV, Goyer A, et al. Effect of a Growing-up Milk Containing Synbiotics on Immune Function and Growth in Children: A Cluster Randomized, Multicenter, Double-blind, Placebo Controlled Study. Clin Med Insights Pediatr 2013;7:49-56.

Ziegler 1983. Ziegler EE, Fomon SJ. Lactose enhances mineral absorption in infancy. J Pediatr Gastroenterol Nutr 1983;2:288-294.

Ziegler 2007. Ziegler EE. Adverse effects of cow's milk in infants. Nestle Nutr Workshop Ser Pediatr Program 2007;60:185-96.

Appendix 1. Search strategy

Search strategy in Medline (OVID)

1. Infant/
2. older infant*.tw.
3. 12 to 23 months old.tw.
4. (toddler* or boy* or girl* or child*).tw.
5. 1 or 2 or 3 or 4

6. Milk/
7. (milk* adj3 full fat).tw.
8. (whole fat adj3 milk*).tw.
9. 6 or 7 or 8
10. 5 and 9
11. Diet/
12. exp Nutrients/
13. growth/ or exp body size/
14. Child Development/
15. Food Preferences/
16. Anemia/
17. (nutrit* or grow* or develop* or an?emi* or diet* or eat*).tw.
18. health*.tw.
19. 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18
20. 10 and 19
21. randomized controlled trial.pt.
22. controlled clinical trial.pt.
23. randomized.ab.
24. placebo.ab.
25. drug therapy.fs.
26. randomly.ab.
27. trial.ab.
28. 21 or 22 or 23 or 24 or 25 or 26 or 27
29. exp animals/ not humans.sh.
30. 28 not 29
31. 20 and 30
32. Cross-Sectional Studies/
33. cross section* stud*.tw.
34. exp cohort studies/
35. cohort$.tw.
36. controlled clinical trial.pt.
37. epidemiologic methods/
38. limit 37 to yr=1966-1989
39. exp case-control studies/
40. (case$ and control$).tw.
41. 32 or 33 or 34 or 35 or 36 or 38 or 39 or 40

42. exp animals/ not humans.sh.
43. 41 not 42
44. 20 and 43
45. 31 or 44

Appendix 2. Risk of bias assessment of included RCTs

The risk of bias of observational studies was assessed using the ROBINS-I tool (Sterne 2016). This tool considers bias due to confounding, in selection of participants into the study, in classification of interventions, due to deviations from intended interventions, due to missing data, in the measurement of outcomes, and in the selection of the reported results.

Assessing risk of bias in randomised trials and non-randomised controlled trials

(1) Sequence generation (checking for possible selection bias)

Studies were assessed as:

low risk of bias if there was a random component in the sequence generation process (e.g., random number table; computer random number generator); high risk of bias if a non-random approach has been used (e.g. odd or even date of birth; hospital or clinic record number). Non-randomised studies should be scored 'high'; unclear risk of bias if not specified in the paper.

(2) Allocation concealment (checking for possible selection bias)

Studies were assessed as:

low risk of bias if participants and investigators enrolling participants could not foresee assignment because an appropriate method was used to conceal allocation (e.g., telephone or central randomisation; consecutively numbered sealed opaque envelopes). This rating was given to studies where the unit of allocation was by institution and allocation was performed on all units at the start of the study; high risk of bias if participants of investigators enrolling participants could possibly foresee assignments and potentially introduce selection bias (e.g. open random allocation; unsealed or non-opaque envelopes);unclear.

(3) Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection bias)

Studies were assessed as:

low risk of bias if outcomes were measured prior to the intervention, and no important differences were present across intervention groups; high risk of bias if important differences in outcomes between groups were present prior to intervention and were not adjusted for in the analysis; unclear risk of bias if there was no baseline measure of outcome (note: if 'high' or 'unclear' but there was sufficient information to do an adjusted analysis, the assessment should be 'low').

(4) Similarity of baseline characteristics (checking for confounding, a potential consequence of selection bias)

Studies were assessed as:

low risk of bias if baseline characteristics were reported and similar across intervention groups; high risk of bias if baseline characteristics we not reported or if there were differences across groups; unclear risk of bias if it was not clear (e.g., characteristics mentioned in text but no data presented).

(5) Incomplete outcome data (checking for possible attrition bias through withdrawals, dropouts, and protocol deviations)

Outcomes in each included study were assessed as:

low risk of bias due to incomplete outcome data, which could be either that there were no missing outcome data or the missing outcome data were unlikely to bias the results based on the following considerations: study authors provided transparent documentation of participant flow throughout the study, the proportion of missing data was similar in the intervention and control groups, the reasons for missing data were provided and balanced across the intervention and control groups, the reasons for missing data were not likely to bias the results (e.g. moving house).high risk of bias if missing outcome data was likely to bias the results. Studies also received this rating if an 'as-treated' (per protocol) analysis was performed with substantial differences between the intervention received and that assigned at randomisation, or if potentially inappropriate methods for imputation have been used; unclear risk of bias.

(6) Blinding (checking for possible performance and detection bias)

The risk of performance bias associated with blinding was assessed as:

low, high, or unclear risk of bias for participants; low, high or unclear risk of bias for personnel.

The risk of detection bias associated with blinding was assessed as:

low, high, or unclear risk of bias for outcome assessors.

Whilst assessed separately, we combined the results in a single evaluation of risk of bias associated with blinding as follows:

low risk of bias if there was blinding of participants and key study personnel and it was unlikely to have been broken, or the outcomes are objective. This rating was also given to studies where either participants and key study personnel were not blinded but outcome assessment was blinded, and the non-blinding of others was unlikely to introduce bias; high risk of bias if there was no blinding or incomplete blinding or if there was blinding that was likely to have been broken and the outcome or outcome assessment was likely to be influenced by a lack of blinding; unclear risk of bias.

(7) Contamination (checking for possible performance bias)

Studies were assessed as:

low risk of bias if allocation was by community, institution or practice and it was unlikely that the control group received the intervention; high risk of bias if it was likely that the control group received the intervention; unclear risk of bias if it was possible that contamination occurred but the risk of this happening was not clear.

(8) Selective reporting bias

Studies were assessed as:

low risk of bias if it was clear, either by availability of the study protocol or otherwise, that all prespecified outcomes that were of interest in the review have been reported; high risk of bias if it was clear that not all of the study's prespecified outcomes have been reported, or reported outcomes were not prespecified (unless justification for reporting is provided), or outcomes of interest are reported incompletely and cannot be used, or where one or more of the primary outcomes was reported using measurements or analysis methods that were not prespecified, or finally if the study report failed to include an important outcome that would be expected to have been reported; unclear risk of bias.

(9) Other sources of bias

Other possible sources of bias were described for each included study and a rating of low, high or unclear risk of bias was given for this item.

In addition to the above criteria, we also assessed cluster-RCTs with the following criteria:

(1) Recruitment bias

The studies were assessed as:

Low risk of bias if individuals were recruited to the trial before the clusters were randomised. High risk of bias if individuals were recruited to the trial after the clusters were randomised. Unclear risk of bias.

(2) Baseline imbalance

The studies were assessed as:

Low risk of bias if baseline characteristics were reported and were similar across clusters or if authors used stratified or pair matched randomisation of clusters. High risk of bias if baseline characteristics were not reported or if there were differences across clusters. Unclear risk of bias.

(3) Loss of clusters

The studies were assessed as:

Low risk of bias if no complete clusters were lost or omitted from the analysis. High risk of bias if complete clusters were lost or omitted from the analysis. Unclear risk of bias.

(4) Incorrect analysis

The studies were assessed as:

Low risk of bias if study authors appropriately accounted for clusters in the analysis or provided enough information for review authors to account for clusters in the meta-analysis. High risk of bias if study authors did not appropriately account for clusters in the analysis or did not provide enough information for review authors to account for clusters in the meta-analysis. Unclear risk of bias.

(5) Compatibility with individual RCTs

The studies were assessed as:

Low risk of bias if effects of the intervention were likely not altered by the unit of randomisation. High risk of bias if effects of the intervention were likely altered by the unit of randomisation. Unclear risk of bias

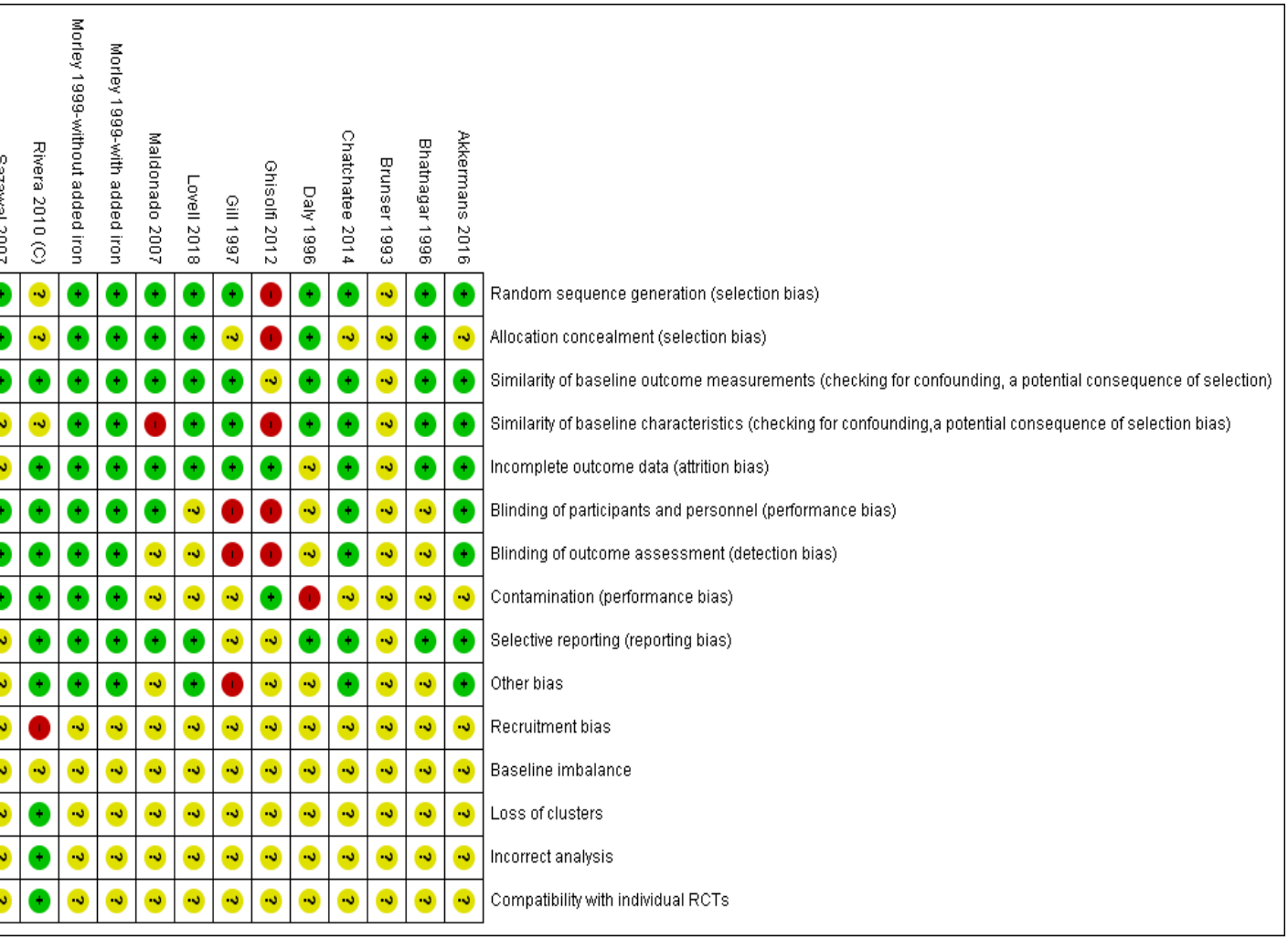
Risk of bias of included studies

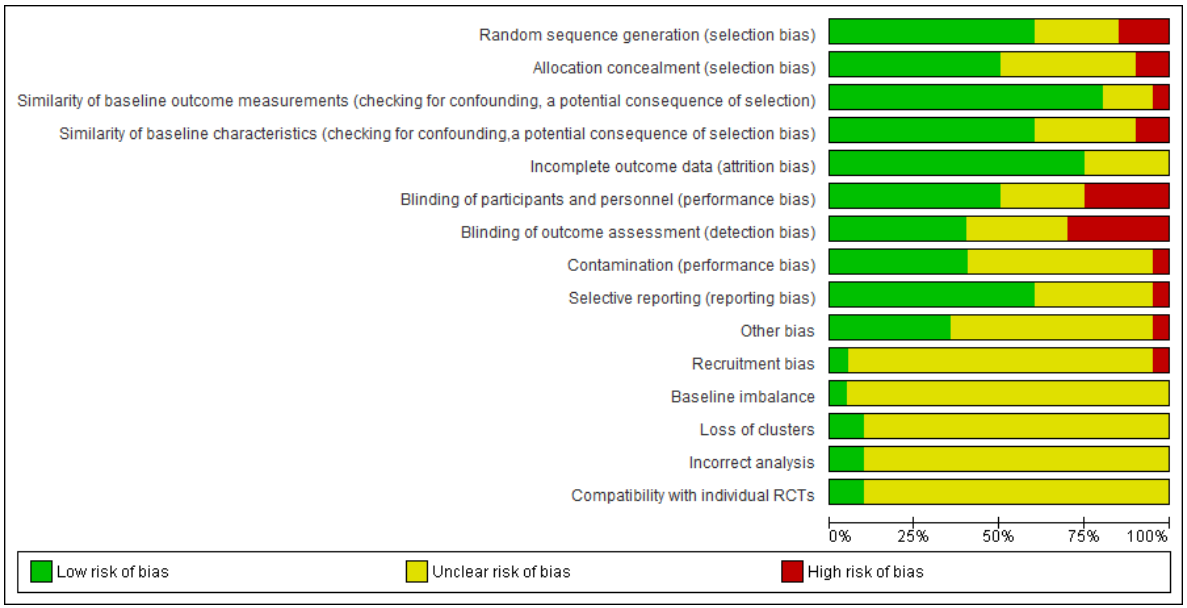
Both reviewers independently assessed the risk of bias in each included study using a simple contingency form that follows the domain-based evaluation (ROB 2.0) (Sterne 2019) (sequence generation; allocation concealment; blinding; incomplete outcome data; selective reporting bias; and other sources of bias, etc.). If there was insufficient information to assess the risk of bias, we rated the domain at 'unclear risk of bias', until further information is published or made available to us. If there was sufficient information, we categorised the domain as being either at 'low risk of bias' or 'high risk of bias' accordingly. We resolved any disagreements by discussion.

The risk of bias of observational studies was be assessed using the ROBINS-I tool (Sterne 2016). This tool considers bias due to confounding, in selection of participants into the study, in classification of interventions, due to deviations from intended interventions, due to missing data, in the measurement of outcomes, and in the selection of the reported results.

Figure A1 shows risk of bias' summary: review authors' judgements about each 'Risk of bias' item for each included study.

Figure A2 shows risk of bias' graph: review authors' judgements about each 'Risk of bias' item presented as percentages across all included studies


Akkermans 2016
Bhatnagar 1996
Brunser 1993
Chatchatee 2014
Daly 1996
Ghisolfi 2012
Gill 1997
Lovell 2018
Maldonado 2007
Morley 1999-with added iron
Morley 1999-without added iron
Rivera 2010 (C)
Random sequence generation (selection bias)
Allocation concealment (selection bias)
Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection)
Similarity of baseline characteristics (checking for confounding,a potential consequence of selection bias)
Incomplete outcome data (attrition bias)
Blinding of participants and personnel (performance bias)
Blinding of outcome assessment (detection bias)
Contamination (performance bias)
Selective reporting (reporting bias)
Other bias
Recruitment bias
Baseline imbalance
Loss of clusters
Incorrect analysis
Compatibility with individual RCTs


Random sequence generation (selection bias)
Allocation concealment (selection bias)
Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection)
Similarity of baseline characteristics (checking for confounding,a potential consequence of selection bias)
Incomplete outcome data (attrition bias)
Blinding of participants and personnel (performance bias)
Blinding of outcome assessment (detection bias)
Contamination (performance bias)
Selective reporting (reporting bias)
Other bias
Recruitment bias
Baseline imbalance
Loss of clusters
Incorrect analysis
Compatibility with individual RCTs
0%
25%
50%
75%
100%
Low risk of bias
Unclear risk of bias
High risk of bias

Appendix 3. Data and analyses

Data and analyses

1 Animal milk (full-fat or lower-fat milk) versus no other milk

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate

2 Animal milk (full-fat or lower-fat milk) versus follow-on formula

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
2.1 Weight (kg) (All)	4	604	Mean Difference (IV, Random, 95% CI)	0.13 [-0.11, 0.36]
2.2 Weight (kg) (subgroup by stunting)	4	604	Mean Difference (IV, Random, 95% CI)	0.13 [-0.11, 0.36]
2.2.1 Stunting non-specified	4	604	Mean Difference (IV, Random, 95% CI)	0.13 [-0.11, 0.36]
2.2.2 Stunting very low prevalence (<2.5%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.2.3 Stunting low prevalence (2-5-9.9%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.2.4 Stunting medium prevalence (10-19.9%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.2.5 Stunting high prevalence (>20%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.3 Weight (kg) (subgroup by wasting)	4	604	Mean Difference (IV, Random, 95% CI)	0.13 [-0.11, 0.36]
2.3.1 Wasting non-specified	4	604	Mean Difference (IV, Random, 95% CI)	0.13 [-0.11, 0.36]
2.3.2 Wasting very low prevalence (<2.5%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.3.3 Wasting low prevalence (2.5-4.9%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.3.4 Wasting medium prevalence (5-10%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.3.5 Wasting high prevalence (>10%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable

Appendix 4. GRADE Evidence Profiles

Author(s): Fernández-Gaxiola AC, De-Regil LM, Gallegos Lecona SC. Animal milks compared to follow-on formula, low-fat milk, plant-based milk or fortified milk and its associated outcomes in children 12-23 months of age.

Question: Should animal milk (full-fat or lower-fat milk) vs. follow-on formula be used in children 12-23 months of age.

Setting: Community settings

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Animal milk (full-fat or lower-fat milk)	follow-on formula	Relative (95% CI)	Absolute (95% CI)		
Weight (kg) (All)												
3 [1]	randomised trials	not serious	not serious	not serious	serious [a]	none	250	354	-	MD **0.13 kg higher** (-0.11 lower to 0.36 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Height (cm) (All)												
3 [1]	randomised trials	not serious	not serious	not serious	serious [a]	none	250	354	-	MD **0.20 cm higher** (-0.31 lower to 0.72 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Weight for height z score												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Animal milk (full-fat or lower-fat milk)	follow-on formula	Relative (95% CI)	Absolute (95% CI)		
1 [2]	randomised trials	serious	not serious	not serious	serious [a]	none	73	70	-	MD **0.3 higher** (0.01 lower to 0.61 higher)	⊕⊕◯◯ LOW	CRITICAL

Head circumference (cm)

2 [3]	randomized trial	not serious	not serious	not serious	serious [a]	none	157	268	-	MD **0.05 lower** (0.36 lower to 0.26 higher)	⊕⊕⊕.◯ MODERATE	CRITICAL

Body composition-Body Mass Index

2 [4]	randomized trial	not serious	not serious	not serious	serious [a]	none	90	86	-	MD **0.28 higher** (0.15 lower to 0.7 higher)	⊕⊕◯◯ LOW -	CRITICAL

Body composition-% Body fat

1 [2]	randomized trial	not serious	not serious	not serious	serious [a]	none	67	67	-	MD **2.4 higher** (0.16 lower to 4.96 higher)	⊕⊕◯◯ LOW -	CRITICAL

Nutrient status-Vit D as serum 25-hydroxyvitamin D [25(OH)D], nmol/L

2 [5]	randomised trials	not serious	not serious	not serious	serious [a]	none	231	224	-	MD **16.27 nmol/L lower** (-21.23 lower to 11.31 lower)	⊕⊕◯◯ LOW	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Animal milk (full-fat or lower-fat milk)	follow-on formula	Relative (95% CI)	Absolute (95% CI)		
Nutritional status-Vit D deficiency												
2 [5]	randomised trials	not serious	not serious	not serious	serious [a]	none	47/231 (20.3%)	17/224 (7.6%)	**RR 2.64** (1.57 to 4.45)	**124 more per 1000** (from 43 more to 262 more)	⊕⊕◯◯ LOW	CRITICAL
Nutritional status-Iron as serum iron (µmol/l)												
1 [2]	randomised trials	serious	not serious	not serious	serious [a]	none	67	67	-	MD **0.7 lower** (2.63 lower to 1.23 higher)	⊕⊕◯◯ LOW	CRITICAL
Child development-Bayley psychomotor development index (PDI)												
1 [3]	randomised trials	serious	not serious	not serious	serious [a]	none	155	240	-	MD **1.15 lower** (3.07 lower to 0.77 higher)	⊕⊕◯◯ LOW	CRITICAL
Child development-Bayley mental development index (MDI)												
1 [3]	randomised trials	not serious	not serious	not serious	serious [a]	none	155	240	-	MD **1.55 points higher** (0.64 lower to 3.73 higher)	⊕⊕⊕◯ MODERATE	CRITICAL
Iron deficiency anaemia (IDA)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Animal milk (full-fat or lower-fat milk)	follow-on formula	Relative (95% CI)	Absolute (95% CI)		
2 [5]	randomised trials	not serious	not serious	not serious	very serious [a]	none	9/223 (4.0%)	1/222 (0.5%)	**RR 6.16** (1.11 to 34.20)	**23 more per 1000** (from 0 fewer to 150 more)	⊕⊕◯◯ LOW	CRITICAL
Iron deficiency (ID, serum ferritin <12 µg/l)												
2 [5]	randomised trials	not serious	not serious	not serious	very serious [a]	none	45/227 (19.8%)	19/225 (8.4%)	**RR 2.33** (1.40 to 3.86)	**112 more per 1000** (from 34 more to 242 more)	⊕⊕◯◯ LOW	CRITICAL
Haemoglobin (g/L)												
5 [6]	randomised trials	serious	serious	not serious	not serious [a]	none	315	348	-	MD **2.61 g/L lower** (4.86 lower to 0.37 lower)	⊕⊕◯◯ LOW	CRITICAL
Ferritin (µg/L) (All)												
5 [6]	randomised trials	serious	serious	not serious	not serious [a]	none	362	434	-	MD **9.87 lower** (15.02 lower to 4.72 lower)	⊕⊕◯◯ LOW	CRITICAL
Gut health-Stool frequency (per day)												
1 [7]	randomised trials	not serious	not serious	not serious	very serious [a]	none	2/153 (1.3%)	1/153 (0.7%)	**RR 2.00** (0.18 to 21.83)	**7 more per 1000** (from 5 fewer to 136 more)	⊕⊕◯◯ LOW	IMPORTANT

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Animal milk (full-fat or lower-fat milk)	follow-on formula	Relative (95% CI)	Absolute (95% CI)		

CI: Confidence interval; **MD:** Mean difference; **RR:** Risk ratio

Explanations

a. Total number of participants is less than 400 (a "rule of thumb") for continuous outcomes and less than 300 for continuous outcomes. Information is likely to be insufficient to precise effect estimate.

1. Duration of the intervention varied between the 3 included studies, from 4, 9 and 12 months. Two trials were 2-arms and on was 3-arms. Evidence was downgraded one level. One trial was assessed as low risk of bias (Morley). Maldonado did not report on the similarity of baseline characteristics.

2. One trial reported on this outcome (Lovell). We downgraded on level for detection bias as final data analysis post data lock was not blinded by treatment group, and no interim analysis was planned for the trial. Trial was downgraded another level for imprecision.

3. One trial reported on this outcome and was downgraded one level due to small sample size (Morley). It is a 3-arm RTC and showed no clear signs of risk of bias.

4.Two trials reported on this outcome (Lovell and Maldonado). Lovell was assessed with unclear risk for performance bias and detection bias. Final data analysis post data lock was not blinded by treatment group, and no interim analysis was planned for the trial. The second trial was assessed with high risk of bias as similarity of baseline characteristics was not reported. Therefore, we downgraded one level the evidence. Both trials had serious imprecision due to small sample size and we downgraded the evidence another level.

5. Two trials reported on this outcome and in one the intervention lasted 5 months (Akkermans) while in the other lasted 12 months (Lovell). In the first one, there was unclear risk for contamination (performance bias) and allocation concealment (selection bias). Final data analysis post data lock was not blinded by treatment group, and no interim analysis was planned for the trial. There were no reported adverse reactions to the study milk and therefore the blinding procedure was maintained until the end of study. One trial (Akkermans) provided low-fat animal milk (1.7g/100 ml) and follow-up period was 5 months; while the other trial (Lovell) provided full-fat animal milk (3.1g/100ml) and follow up period was 12 months.

6. From the five trials reporting on this outcome, two were assessed at high risk of bias. Risk for performance bias in Daly as mothers from both groups on income support were still entitled to claim free cows' milk with milk tokens. However, as not all parents were in receipt of income support, and therefore not entitled to the cows' milk, the cows' milk group received funding to purchase 500 ml cows' milk per day. Risk for selection bias (confounding) as similarity of baseline characteristics was not reported. Age range of children varied between studies: 6-18 mo and followed until 24 mo in Daly; 9 mo and followed until 18 mo in Morley; 12-36 mo in Akkermans and followed for 5 mo; 12 mo and followed until 24 mo in Lovell; and 12-30 mo for 4 mo in Maldonado. There was substantial heterogeneity and therefore, evidence was downgraded another level for this.

7. One trial reported on this outcome (Akkermans). Multi-country, 2-arm study with children 12-24 mo of age visiting hospitals and clinics and the follow-up period was 5 months. Follow-on formula contained 1-2 mg of iron (form not specified) and 1.7 mg vitamin D, compared animal milk. The study was funded by the private industry but analyzed independently and before unblinding the study. Unclear risk for contamination (performance bias) and allocation concealment (selection bias). Results are inconclusive, but the size of the effect included potentially important benefits, test of overall effect (P=0.57). Evidence was downgraded to levels.

Author(s): Fernández-Gaxiola AC, De-Regil LM, Gallegos Lecona SC. Animal milks compared to follow-on formula, low-fat milk, plant-based milk or fortified milk and its associated outcomes in children 12-23 months of age.

Question: Should full-fat animal milk vs. lower-fat milk be given to children 12-23 months of age.

Setting: Community settings

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Full-fat animal milk	Lower-fat milk	Relative (95% CI)	Absolute (95% CI)		
Nutrient status-serum cholesterol (mmol/l)												
1[1]	randomized trials	serious	not serious	not serious	serious	none	9	8	-	MD **0.17 lower** (0.92 lower to 0.58 higher)	⊕⊕◯◯ LOW	CRITICAL
Nutrient status-serum low density lipoprotein (LDL) (mmol/l)												
1[1]	randomized trials	serious	not serious	not serious	serious	none	9	8	-	MD **0.25 lower** (0.94 lower to 0.44 higher)	⊕⊕◯◯ LOW	CRITICAL
Nutrient status-serum high density lipoprotein (HDL) (mmol/l)												
1[1]	randomized trials	serious	not serious	not serious	serious	none	9	8	-	MD **0.1 lower** (0.3 lower to 0.1 higher)	⊕⊕◯◯ LOW	CRITICAL
Nutrient status-serum triglycerides (mmol/l)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Full-fat animal milk	Lower-fat milk	Relative (95% CI)	Absolute (95% CI)		
1[1]	randomized trials	serious	not serious	not serious	serious	none	9	8	-	MD **0.34 higher** (0.12 lower to 0.8 higher)	⨁⨁◯◯ LOW	CRITICAL

Nutrient status-serum LDL/HDL

1[1]	randomized trials	serious	not serious	not serious	serious	none	9	8	-	MD **0.05 lower** (0.83 lower to 0.73 higher)	⨁⨁◯◯ LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference;

Explanations

1. One trial reported on outcomes for this comparison (Svahn). The intervention lasted 6 months. Evidence was downgraded two levels. It assessed as high risk for detection bias and reporting bias, and it had very small sample size so there was not enough information to detect a precise estimate of the effect. There was unclear risk for selection bias and for similarity of baseline characteristics.

Author(s): Fernández-Gaxiola AC, De-Regil LM, Gallegos Lecona SC. Animal milks compared to follow-on formula, low-fat milk, plant-based milk or fortified milk and its associated outcomes in children 12-23 months of age.

Question: Should animal milk (full-fat or lower fat) vs. plant-based milk be given to children 12-23 months of age.

Setting: Community settings

Evaluación de certeza							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Animal milk	Plant-based milk alternatives	Relative (95% CI)	Absolute (95% CI)		
Nutrient status-serum cholesterol (mmol/l)												
1	randomized trials	serious	not serious	not serious	serious	none	9	12	-	MD **0.16 lower** (0.76 lower to 0.44 higher)	⊕⊕◯◯ LOW	CRITICAL
Nutrient status-serum low density lipoprotein (LDL) (mmol/l)												
1	randomized trials	serious	not serious	not serious	serious	none	9	12	-	MD **0.03 lower** (0.48 lower to 0.54 higher)	⊕⊕◯◯ LOW	CRITICAL
Nutrient status-serum high density lipoprotein (HDL) (mmol/l)												

Evaluación de certeza							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Animal milk	Plant-based milk alternatives	Relative (95% CI)	Absolute (95% CI)		
1	randomized trials	serious	not serious	not serious	serious	none	9	12	-	MD **0.18 lower** (0.85 lower to 0.49 higher)	⨁⨁◯◯ LOW	CRITICAL

Nutrient status-serum triglycerides (mmol/l)

1	randomized trials	serious	not serious	not serious	serious	none	9	12	-	MD **0.08 lower** (0.63 lower to 0.47 higher)	⨁⨁◯◯ LOW	CRITICAL

Nutrient status-serum LDL/HDL

1	randomized trials	serious	not serious	not serious	serious	none	9	12	-	MD **0.33 higher** (0.36 lower to 1.02 higher)	⨁⨁◯◯ LOW	CRITICAL

CI: Confidence interval; **MD:** Mean difference; **Explanations**

1. One trial reported on outcomes for this comparison (Svahn). The intervention lasted 6 months. Evidence was downgraded two levels. It assessed as high risk for detection bias and reporting bias, and it had very small sample size so there was not enough information to detect a precise estimate of the effect. There was unclear risk for selection bias and for similarity of baseline characteristics.

Author(s): Fernández-Gaxiola AC, De-Regil LM, Gallegos Lecona SC. Animal milks compared to follow-on formula, low-fat milk, plant-based milk or fortified milk and its associated outcomes in children 12-23 months of age.

Question: Should animal milk (full-fat or lower-fat) vs. fortified milk (full-fat or lower-fat) be given to children 12-23 months of age.

Setting: Community settings

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Animal milk (full-fat or lower-fat)	fortified milk (full-fat or lower-fat)	Relative (95% CI)	Absolute (95% CI)		
Weight (kg)												
1[1]	randomised trials	serious	not serious	not serious	serious	none	16	20	-	MD **0.04 kg higher** (0.83 lower to 0.91 higher)	⊕⊕◯◯ LOW	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Animal milk (full-fat or lower-fat)	fortified milk (full-fat or lower-fat)	Relative (95% CI)	Absolute (95% CI)		
Undernutrition – Stunting and wasting												
1[2]	randomised trials	not serious	not serious	not serious	serious	none	65/187 (34.8%)	68/191 (35.6%)	**RR 0.98** (0.74 to 1.28)	**7 fewer per 1000** (from 93 fewer to 100 more)	⊕⊕⊕◯ LOW	CRITICAL
Undernutrition-Stunting												
1[2]	randomised trials	not serious	not serious	not serious	serious	none	65/187 (34.8%)	68/191 (35.6%)	**RR 0.98** (0.74 to 1.28)	**7 fewer per 1000** (from 93 fewer to 100 more)	⊕⊕⊕◯ LOW	CRITICAL
Undernutrition-Wasting												
1[2]	randomised trials	not serious	not serious	not serious	serious	none	57/187 (30.5%)	55/191 (28.8%)	**RR 1.06** (0.78 to 1.44)	**17 more per 1000** (from 63 fewer to 127 more)	⊕⊕⊕◯ LOW	CRITICAL
Nutrient status-Iron as serum iron (µmol/L)												
2[3]	randomised trials	serious	not serious	not serious	serious	none	57	58	-	MD **0.46 µmol/L lower** (4.38 lower to 3.46 higher)	⊕⊕◯◯ LOW	CRITICAL
Nutrient status-Zinc as plasma zinc (µmol/L)												

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Animal milk (full-fat or lower-fat)	fortified milk (full-fat or lower-fat)	Relative (95% CI)	Absolute (95% CI)		
2[4]	randomised trials	very serious	not serious	not serious	not serious	none	244	249	-	MD **0.43 higher** (0.11 higher to 0.76 higher)	⊕⊕◯◯ LOW	CRITICAL

Anaemia

3[5]	randomised trials	very serious	not serious	not serious	not serious	none	147/556 (26.4%)	96/768 (12.5%)	**RR 2.29** (1.12 to 4.69)	**161 more per 1000** (from 15 more to 461 more)	⊕⊕◯◯ LOW	CRITICAL

Iron deficiency anaemia (IDA)

1[2]	randomised trials	not serious	not serious	not serious	serious	none	128/232 (55.2%)	31/233 (13.3%)	**RR 4.15** (2.93 to 5.87)	**419 more per 1000** (from 257 more to 648 more)	⊕⊕◯◯ LOW	CRITICAL

Iron deficiency (ID, serum ferritin <12 µg/l)

1[6]	randomised trials	not serious	not serious	not serious	serious	none	10/114 (8.8%)	17/235 (7.2%)	**RR 1.21** (0.57 to 2.56)	**15 more per 1000** (from 31 fewer to 113 more)	⊕⊕◯◯ LOW	CRITICAL

Certainty assessment							№ of patients		Effect		Certainty	Importance
№ of studies	Study design	Risk of bias	Inconsistency	Indirectness	Imprecision	Other considerations	Animal milk (full-fat or lower-fat)	fortified milk (full-fat or lower-fat)	Relative (95% CI)	Absolute (95% CI)		
Haemoglobin (g/dL)												
6[7]	randomised trials	very serious	not serious	not serious	not serious	none	663	691	-	MD **5.91 g/L lower** (9.84 lower to 1.99 lower)	⊕⊕⊕◯ MODERATE	CRITICAL
Oral health -caries free (decayed, missing, and filled surfaces -dmfs- index in molars and canines= 0)												
1[8]	randomised trials	serious	not serious	not serious	serious	none	43/76 (56.6%)	85/110 (77.3%)	**RR 1,30** (0.37 to 2.23)	**209 fewer dmfs per 1000** (from 317 fewer to 70 fewer)	⊕⊕◯◯ LOW	IMPORTANT
Morbidity-Respiratory episodes per child per year												
1[9]	randomised trials	serious [e]	not serious	not serious	serious [d]	none	278	276	-	MD **0.03 respiratory episodes lower** (0.14 lower to 0.20 higher)	⊕⊕◯◯ LOW	IMPORTANT

Explanations

1. One trial reported on this outcome (Svahn) and the intervention lasted 6 months. Evidence was downgraded two levels. It assessed as high risk for detection bias and reporting bias, and it had very small sample size so there was not enough information to detect a precise estimate of the effect.

2. One trial reported on this outcome and the intervention lasted 12 mo (Sazawal). All children who had severe anaemia at baseline (Hb ≤ 70 g/l) were given a therapeutic dose of iron for three months in addition to their milk supplement. In the fortified milk group 9 children left the area, 2 children died, and 16 withdrawn consents. In the milk group 18 children left the area, 2 children died, and 16 withdrawn consents. Participants´ flow chart is not clear and therefore, incomplete outcome data is unclear too. Small sample sizes both in intervention and control groups. Evidence was downgraded two levels accordingly.

3. Two trials reported on this outcome (Stekel and Svahn). Randomization method and similarity of baseline outcome measurements not mentioned in one study. Small sample size both in control and intervention groups so there was not enough information to detect a precise estimate of the effect. Evidence was downgraded two levels.

4. Two trials reported on this outcome (Sazawal and Villalpando). Sazawal was assessed as unclear risk of bias for selection bias and attrition bias. In the fortified milk group 9 children left the area, 2 children died, and 16 withdrawn consents. In the milk group 18 children left the area, 2 children died, and 16 withdrawn consents. Participants' flow chart is not clear. Villalpando was assessed as high risk of bias for selection bias as the randomization procedure did not result in an even distribution of baseline anaemia in the 2 intervention groups (30.0% non-fortified milk, 41.4% fortified milk). Groups did not differ at 6 mo and the changes between baseline and 6 mo did not differ between the groups. However, baseline outcomes showed that fortification milk group was more susceptible to a larger improvement. Evidence was downgraded two levels.

5.Sazawal was assessed as unclear risk of bias for selection bias and attrition bias. In the fortified milk group 9 children left the area, 2 children died, and 16 withdrawn consents. In the milk group 18 children left the area, 2 children died, and 16 withdrawn consents. Participants' flow chart is not clear. Villalpando was assessed as high risk of bias for selection bias as the randomization procedure did not result in an even distribution of baseline anaemia in the 2 intervention groups (30.0% non-fortified milk, 41.4% fortified milk). Groups did not differ at 6 mo and the changes between baseline and 6 mo did not differ between the groups. However, baseline outcomes showed that fortification milk group was more susceptible to a larger improvement. Evidence was downgraded two levels. Stekel 1986 was assessed as high risk of bias as blinding was not reported.

6. One trial reported on this outcome (Villalpando). The trial was assessed as high risk of bias for selection bias as the randomization procedure did not result in an even distribution of baseline anaemia in the 2 intervention groups (30.0% non-fortified milk, 41.4% fortified milk). Groups did not differ at 6 mo and the changes between baseline and 6 mo did not differ between the groups.

7. Five trials reported on this outcome (Stekel1986, Stekel 1988, Svahn1999, Szymlek-Gay and Villalpando 2006). Evidence was downgraded two levels as randomization and blinding were not mentioned in some studies and risk of bias was assessed as high.8. One cluster RCT reported on this outcome (Stecksén-Blicks). Children were served 150 ml medium-fat milk (1.5%) at lunch. Blinding of outcome assessment was not mentioned and evidence was downgraded one level for this and another level for the sample size in both control and intervention groups.

9. Stekel 1988 reported on this outcome. The trial did not report on blinding of participants and personnel nor of outcome assesment and was downgraded one level for this. The trial had a small sample size in both intervention and control groups and was downgraded another level.

Appendix 5. Characteristics of included studies and risk of bias tables

Akkermans 2016

Methods	Double blind RCT, multi country trial with 2 arms
Participants	Children 12-36 months with a stable health status from 3 participating countries: 1. Germany (from 9 private paediatric clinics spread throughout the country). 2. Netherlands (Juliana Children's Hospital/Haga Teaching Hospital in The Hague, VU University Medical Center in Amsterdam, and Sophia Children's Hospital/Erasmus Medical Center in Rotterdam). 3. United Kingdom (Royal National Orthopedic Hospital in London and St Mary´s Hospital in Newport, Isle of Wight).
Interventions	Participants (n= 318) were randomly assigned to 1 of 2 groups: 1. group 1 (n= 158) received follow-on formula 2. group 2 (n= 160) received control cow milk A total of 264 children from Germany (83.0%), 42 from the Netherlands (13.2%) and 12 from the United Kingdom (3.8%) were included in the study sample.
Outcomes	Outcomes included serum measurement of ferritin, prevalence of iron deficiency (ID) and iron deficiency anaemia (IDA); nutrient intakes; weight, height, WAZ, HAZ, iron deficiency, and stools.
Notes	1. Follow-on formula contained 1.2 mg Fe/100 ml (form not specified) and 1.7 mg vitamin D/100 ml per day. 2. Control product was a non-fortified cow milk that contained 0.02 mg Fe/100 ml and no vitamin D. 3. The energy concentrations of both products were comparable (46 kcal/100 nl for cow milk and 50 kcal/100 ml for follow-on formula) 4. Both products were supplied in powdered form with instructions for preparing the milk by diluting the powder with water Source of funding: Trial was supported by Danone Nutrition Research. The study products were produced, provided, and coded (for blinding purposes) by Nutricia Cuijk (commissioned by Danone Nutricia Research. Data analysis interpretation were performed independently from Danone Nutricia Research.

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	A computer model was used for block randomization in which stratification was applied for country and sex.
Allocation concealment (selection bias)	Unclear risk	Not clear.
Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection)	Low risk	Baseline outcomes were comparable among study groups.
Similarity of baseline characteristics (checking for confounding, a potential consequence of selection bias)	Low risk	There were no differences in the baseline characteristics of the 2 groups except for parents´ education and working status, iron intake and vitamin D intakes.

Incomplete outcome data (attrition bias)	Low risk	Number of children that did not completed the study (i.e. 29.4% in control group and 27.8% in intervention group) and withdrawals y subject (i.e. 22 children in control group and 23 in intervention group) were similar between groups. There is a participants' flow throughout the study.
Blinding of participants and personnel (performance bias)	Low risk	Parents (and their children), investigators, and treating physicians were blinded to product allocation by coding the cans containing the study products.
Blinding of outcome assessment (detection bias)	Low risk	Statistical analyses were described in a statistical analysis plan, and it was finalized before the study was unblinded.
Contamination (performance bias)	Unclear risk	Not clear.
Selective reporting (reporting bias)	Low risk	The study was registered at the Clinical Trial Registry of the Netherlands, with registration no. 3609.
Other bias	Low risk	There are no other clear signs of other bias.
Recruitment bias	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Baseline imbalance	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Loss of clusters	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Incorrect analysis	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Compatibility with individual RCTs	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.

Bhatnagar 1996

Methods	RCT with 2 arms
Participants	All children 3-24 mo of age with diarrhoea for at least 14 days but less than 12 weeks and had passed three or more liquid stools in the preceding 24 hours, attending the diarrhoea treatment units of the All India Institute of Medical Sciences (AIIMS) and Kasturba Hospital (New Delhi, India) between February 1992 and January 1995 were enrolled in the study.
Interventions	Eligible children (n= 116), including 23 girls that were randomized separately in a similar manner) were randomly assigned to 1 of 2 groups: 1. group 1 (n= 60) received milk-based cereal dietary regimen 2. group 2 (n= 56) received milk-free cereal dietary regimen Both diets were isocaloric (86.9 calories/100 g for 9 months; 95.6 cal/100g for >9 months) consisting of puffed rice cereal, sugar, and oil differing in only their source of protein, which was either milk or egg white, respectively. Both diets were offered at the rate of 150 kcal/kg per day. Records available from this study do not allow for data extraction and thus do not contribute data for this review.
Outcomes	Number of stools, stools weight, fluid and energy intake.

Notes	1. Children in the milk-based cereal diet consumed 30% of calories from milk and 1.9g/kg lactose per day approximately. Source of funding: Supported by the Diarrheal Diseases Control Programme, World Health Organization.

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Stratified randomization scheme using permutation blocks of fixed length stratified by stool purge rates during stabilization period.
Allocation concealment (selection bias)	Low risk	Quote: "The randomization list prepared before the start of the study was kept at a central office at AIIMS, and the treatment assignment was read off the randomization list and given to the investigators at AIIMS and Kasturba Hospital on the phone by an independent person who was not concerned with the study."
Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection)	Low risk	All children were carefully screened for inclusion. The pre-randomization period included administration of fluid therapy for associated dehydration and two standard semisolid foods. Vomiting, number of stools and median stool weight were similar between groups.
Similarity of baseline characteristics (checking for confounding, a potential consequence of selection bias)	Low risk	Baseline comparisons of all outcomes between participating children were not significant.
Incomplete outcome data (attrition bias)	Low risk	Three children (two in group 1 and one in group 2) left against medical advice at 9, 61, and 92 hours, respectively, and their data have been included until the time of withdrawal from the study for outcomes other than treatment failure and weight change.
Blinding of participants and personnel (performance bias)	Unclear risk	Not clear.
Blinding of outcome assessment (detection bias)	Unclear risk	Not clear.
Contamination (performance bias)	Unclear risk	Not clear.
Selective reporting (reporting bias)	Low risk	Unlikely. Supported by the Diarrheal Diseases Control Programme, World Health Organization
Other bias	Unclear risk	Not clear.
Recruitment bias	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Baseline imbalance	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Loss of clusters	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.

Incorrect analysis	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Compatibility with individual RCTs	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.

Brunser 1993

Methods	Prospective cohort, 2 arms
Participants	Children in a community of low socioeconomic stratum in Santiago, Chile.
Interventions	Children were incorporated into each of two consecutive cohorts; each cohort was divided into two groups: 1. group 1 (monthly average= 70 children) received iron fortified milk 2. group 2 (monthly average= 83 children) received control milk Each cohort was followed up for 6 months.
Outcomes	Outcomes included incidence of diarrhoea as episodes/100 children/month, bowel movements on a day, liquid or semi-liquid stools and their aetiology, asymptomatic shedding of enteropathogens
Notes	1. Fortified milk contained 12mg/l of iron and control milk 1 mg/l of iron

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Not clear.
Allocation concealment (selection bias)	Unclear risk	Not clear.
Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection)	Unclear risk	Not clear.
Similarity of baseline characteristics (checking for confounding, a potential consequence of selection bias)	Unclear risk	Not clear.
Incomplete outcome data (attrition bias)	Unclear risk	Not clear.
Blinding of participants and personnel (performance bias)	Unclear risk	Not clear.
Blinding of outcome assessment (detection bias)	Unclear risk	Not clear.
Contamination (performance bias)	Unclear risk	Not clear.
Selective reporting (reporting bias)	Unclear risk	Not clear.
Other bias	Unclear risk	Not clear.
Recruitment bias	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Baseline imbalance	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.

Loss of clusters	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Incorrect analysis	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Compatibility with individual RCTs	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.

Chatchatee 2014

Methods	Double blind RCT, multi country intervention with 3 arms
Participants	Healthy children ages 11 to 29 months old attending a day care centre at least 2 times per week were recruited into the study. The study was carried out in private practices, children's hospitals, university hospitals, or site management organizations (organizations that provide clinical trial related services) located in 5 countries in Europe and Asia: Malaysia, the Netherlands, Poland, Portugal, and Thailand.
Interventions	Participants (n= 767) were randomly assigned to 1 of 3 groups: 1. group 1 (n= 388) received follow-on formula with added short-chain galacto-oligosaccharides and long-chain fructo oligosaccharides (scGOS/lcFOS) and long-chain polyunsaturated fatty acids (n-3 LCPUFAs) (EPA:DHA, 4:6). group 2 (n= 379) received control follow-on formula without added (scGOS/lcFOS) and long-chain polyunsaturated fatty acids (n-3 LCPUFAs) (EPA:DHA, 4:6). group 3 (n= 37) reference group of subjects receiving cow's milk already before the start of the study and which followed the same procedures as the intervention groups; however, this group was not randomized and consumed regular cow's milk for 52 weeks. A total of 135 children from Malaysia, 199 from the Netherlands,126 from Poland, 70 from Portugal and 167 from Thailand were included in the study sample.
Outcomes	Outcomes included number of episodes of respiratory tract and gastrointestinal infections; symptoms; type of medication used.
Notes	Follow-on formula was added 1.2 g/100 mL of scGOS/ lcFOS (9:1) (Immunofortis) and 19.2 mg/100 mL of n-3 LCPUFAs (EPAþDHA, 4:6). Regular follow-on formula (without scGOS/lcFOS/n-3 LCPUFAs was marketed for young children ages 1 to 3 years, enriched with key nutrients such as vitamins A, C, and D, iron, and calcium. 3. Both study products were packed in identical tins; they were of the same colour, weight, smell, and taste.

	Source of funding: Follow-on formulas were tested by Danone Research. As reference group using cow milk was not randomized, the study was not included in the meta-analysis.

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	A randomization code, developed using a computer random number generator was used. Control group was not randomized
Allocation concealment (selection bias)	Unclear risk	Not mentioned.
Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection)	Low risk	Weight and height differences were observed between the intervention and control groups, but were considered normal and therefore, not expected to influence study outcomes.
Similarity of baseline characteristics (checking for confounding, a potential consequence of selection bias)	Low risk	No statistically significant differences were observed between the 2 groups in regard to age, sex, length, and weight at birth and any of the other baseline characteristics analyzed, nor in their parents' education or professional status.
Incomplete outcome data (attrition bias)	Low risk	Number of early withdrawals was similar between groups (40 in intervention group and 30 in control group). Reasons for withdrawal were reported in both groups, and children that completed the study was similar between groups (348 in the intervention group and 349 in the control group).
Blinding of participants and personnel (performance bias)	Low risk	Researchers, parents and children were unaware of the real nature of the product
Blinding of outcome assessment (detection bias)	Low risk	Study was unblinded after the study was completed and after the statistical analyses were finalized.
Contamination (performance bias)	Unclear risk	Not clear.
Selective reporting (reporting bias)	Low risk	The study was registered at the Clinical Trial Registry of the Netherlands, with registration no. NTR1451.
Other bias	Low risk	There are no other clear signs of other bias.
Recruitment bias	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Baseline imbalance	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Loss of clusters	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Incorrect analysis	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Compatibility with individual RCTs	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.

Daly 1996

Methods	Double blind, randomized cohort trial, with 2 arms.
Participants	Children aged 6-8 months (567 identified) living in an inner-city area of Birmingham. A field researcher visited the families, and the parents of only those children whose mothers had already changed their children's diet to unmodified cows' milk (n= 116) were asked to consider including their children in the study.
Interventions	One hundred children (16 participants drop out from the study) were assigned to 1 of 2 groups: 1. group 1 (n= 38) received iron fortified milk, supplied free of charge.(n= 40). 2. group 2 (n= 40) would continue with cow milk and recruited froma single care centre. Parents received an equivalent monthly payment to purchase it. At 18 months, those infants on fortified milk were transferred back to cows' milk.
Outcomes	Outcomes included serum measurement of haemoglobin, ferritin, mean corpuscular volume, prevalence of iron deficiency anaemia (IDA), weight, length, nutrient intake.and developmental assessments at enrolment and at 18 and 24 months of age.
Notes	Source of funding: Farley Health Products.

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Random numbers in blocks of four.
Allocation concealment (selection bias)	Low risk	Randomization was kept blinded to all except for one field researcher that was unblinded. Five trained and experienced observers who performed the developmental scales were also blinded to group randomization.
Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection)	Low risk	No statistically significant differences in mean haemoglobin concentration, mean corpuscular volume and serum ferritin were observed at baseline between the two groups; 16% of the cows' milk group and 13% of the iron supplemented formula milk group were already anaemic.
Similarity of baseline characteristics (checking for confounding,a potential consequence of selection bias)	Low risk	No significant differences were present between the two groups with respect to race, number of single parents, smokers and nonsmokers, those receiving income support, car or telephone ownership, maternal age, family size, accommodation, and maternal education.
Incomplete outcome data (attrition bias)	Unclear risk	Quote: "Some data points were missing due to intercurrent illness in a participant, transiently being unable to locate children, or insufficient volume of blood for assay. Out of 269 contacts, a developmental score was unavailable on 11 occasions (3%)."
Blinding of participants and personnel (performance bias)	Unclear risk	Study was reported as double blinded. Researchers supplied the iron fortified milk free of charge, and gave those mothers whose infants remained on the cows' milk a monthly payment equivalent to the cost of 500 ml cows' milk daily. There was no report on whether the participants could tell the difference.

Blinding of outcome assessment (detection bias)	Unclear risk	Not mentioned.
Contamination (performance bias)	High risk	Quote: "Mothers from both groups on income support were still entitled to claim free cows' milk with milk tokens.However, as not all parents were in receipt of income support, and therefore not entitled to the cows' milk, the cows' milk group received funding to purchase 500 ml cows' milk per day."
Selective reporting (reporting bias)	Low risk	The study was approved by the South Birmingham Health Authority ethical committee.
Other bias	Unclear risk	Not clear.
Recruitment bias	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Baseline imbalance	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Loss of clusters	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Incorrect analysis	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Compatibility with individual RCTs	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.

Ghisolfi 2012

Methods	Non-randomized controlled study: cross-sectional study with 2 arms
Participants	Children aged from 15 d to 36 months (n= 713) were included in a food consumption survey but for the purpose of the study only children 12-24 months were included (n= 132). Parents participating in the survey were recruited from all regions of France (excluding overseas territories). Study researchers used a proportional sampling technique that took into account the population of each region, the age of the children, the professional status of the mother and the socioeconomic level of the family.
Interventions	Children (n=132) were divided into 2 groups defined according to their type of milk intake: 1. group 1 (n= 63) received cow milk, at least 250 ml (70% as semi-skimmed milk) and who did not receive growing-up milk or follow-on formula or dairy products based on growing-up milk or follow-on formula. 2. group 2 (n= 55) received follow-on formula, at least 250 ml.
Outcomes	The study measured energy and nutrient daily intake.
Notes	1. This minimal value of 250 ml per day for milk consumption was retained a priori since it corresponds to one daily bottle consumption. 2. Besides the total energy intake, the nutrients considered were: protein, lipids, total carbohydrates (excluding fibre), EFA (linoleic acid and alfa-linolenic acid), Na, Ca, P, Mg, Zn, Fe, vitamins B1, B2, B3, B5, B6, B9,

	B12, C, D (exclusively of food origin, referred to hereafter as ‘alimentary vitamin D’), E (expressed as a-tocopherol equivalents), total vitamin A (expressed as retinol equivalents), retinol and carotenoids (expressed as b-carotene equivalents). Source of funding: Syndicat français des aliments de l’enfance (French Association of Baby Food Industries),

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Not used.
Allocation concealment (selection bias)	High risk	Not used.
Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection)	Unclear risk	Not reported.
Similarity of baseline characteristics (checking for confounding,a potential consequence of selection bias)	High risk	Not reported.
Incomplete outcome data (attrition bias)	Low risk	Study researchers provided transparent documentation of participant flow throughout the study.
Blinding of participants and personnel (performance bias)	High risk	Study was designed without blinding.
Blinding of outcome assessment (detection bias)	High risk	Study was designed without blinding.
Contamination (performance bias)	Low risk	It is unlikely there was contamination in the study groups as children just continued to have their usual milk.
Selective reporting (reporting bias)	Unclear risk	According to French legislation, this type of nutritional survey does not need to be approved by an institutional review board.
Other bias	Unclear risk	Not clear.
Recruitment bias	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Baseline imbalance	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Loss of clusters	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Incorrect analysis	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Compatibility with individual RCTs	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.

Gill 1997

Methods	Randomized controlled trial, multi centre study with 3 arms

Participants	Healthy term infants, the product of a normal delivery, were recruited into the study at age 6 months. All infants were receiving formula or whole cow's milk at entry. Infants were recruited at 21 centres in the United Kingdom (n= 192) and Ireland (n= 214).
Interventions	Infants (n= 406) were a randomly assigned to 1 of 3 groups: 1. group 1 (n= 264) received iron fortified follow-on formula 2. group 2 (n= 85) received unfortified follow-on formula group 3 (n =57) reference group with chuldren that were receiving cow milk prior to the study, and thus were assigned to continue with it. Both formulas were provided free of charge. Mothers in the cow milk group (group 3) were remunerated for expenditure on milk and clinic attendances.
Outcomes	Infants were seen and assessed in local clinics (and its equipment) for anthropometric measurements including: weight, length, and head circumference; and laboratory measurements including: haemoglobin, serum ferritin, serum iron, total iron binding capacity, mean corpuscular volume, mean corpuscular haemoglobin, red blood cell count, platelet count and differential white cell count. All measurement were measured at 9- 10 months. at 12 months and at 15 months of age.
Notes	1. Iron fortified follow-on formula reconstituted contained 12.3 mg of iron per litre, in the form of anhydrous ferrous sulphate. 2. Unfortified follow-on formula was a matching identical formula. apart from its iron content of 1.4 mg per litre. 3. Use of non-dietary iron supplements was not permitted. Source of funding: SMA Nutrition (UK and Ireland). Each investigator received a study grant and formula from SMA Nutrition. Wyeth Laboratories provided the formulas. As reference group using cow milk was not randomized, study was not included in the meta-analysis.

Risk of bias table

Bias	**Authors' judgement**	**Support for judgement**
Random sequence generation (selection bias)	Low risk	Randomization was allocated on a ratio 3:1
Allocation concealment (selection bias)	Unclear risk	Not mentioned.
Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection)	Low risk	The study reported no significant difference in weight, length or head circumference at entry to the study. Mean haemoglobin values were significantly different at baseline between groups 1 and 3.
Similarity of baseline characteristics (checking for confounding, a potential consequence of selection bias)	Low risk	The two randomized groups were well balanced in terms of age and sex.
Incomplete outcome data (attrition bias)	Low risk	104 infants were withdrawn from analysis, 72 in group 1, 25 in group 2 and 7 in group 3. The major reasons for withdrawal were failure to return for assessments (n= 55) or protocol violation (n= 23), the latter including infants who were iron-deficient and/or anaemic (9 infants from group 1, 2 from group 2 and 9 from group 3).

Blinding of participants and personnel (performance bias)	High risk	No blinding was reported in the study.
Blinding of outcome assessment (detection bias)	High risk	No blinding was reported in the study.
Contamination (performance bias)	Unclear risk	Not clear.
Selective reporting (reporting bias)	Unclear risk	The study protocol was approved by the local ethical committees.
Other bias	High risk	All statistical analyses were performed by Wyeth Laboratories.
Recruitment bias	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies
Baseline imbalance	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies
Loss of clusters	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies
Incorrect analysis	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies
Compatibility with individual RCTs	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies

Lovell 2018

Methods	Double blind RCT, Multicenter intervention with 2 arms.
Participants	Healthy children 1 year of age (n=160) that lived in Auckland, New Zealand (n=108) and Brisbane, Australia (n=52).
Interventions	Participants were randomly assigned to 1 of 2 groups: 1. group 1 (n= 80) received 300 ml (or 6 scoops of powder in total) follow-on formula with reduced protein ("GUMLi" -growing-up milk lite-). 2. group 2 (n= 80) received 300 ml (or 6 scoops of powder in total) unfortified cow milk (homogenized and pasteurized). Dietary intakes were collected at baseline, months 3, 6, 9 and 12 post-randomization, using a validated food frequency questionnaire.
Outcomes	Primary outcome was to evaluate the effect of the intervention on body composition. Secondary outcomes included assessment of dietary patterns, measures of dietary intake (i.e. protein, B12, iron, vitamin D, vitamin C, zinc, sodium, PUFA, vitamin A, vitamin B-6, folate, magnesium, and selenium); anthropometry (weight, length, waist circumference); micronutrient status and cognitive development.

Notes	1. Intervention milk "GUMLi" had a reduced energy and protein content when compared to standard, commercial follow-on formula on the market, 60 kcal/100ml vs. 71 kcal/100ml and 1.7 g/100ml protein vs. 2.2 g/100ml. GUMLi was also fortified with iron (1.7 mg Fe/100ml form not specified), vitamin C and vitamin D, probiotics and prebiotics. 2. Control cow milk was energy matched with a protein content of 3.1 g/100ml. Source of funding: Study was funded by an investigator-initiated grant from Danone Pty.Ltd. (Danone Nutricia Research). Study milks were independently allocated by Danone Nutricia Research (Oceania).

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Children were randomized 1:1 to one of the two study groups using computer-generated randomisation sequences and stratified by study centre (Auckland and Brisbane).
Allocation concealment (selection bias)	Low risk	The randomisation list was prepared by a statistician working independently of the study team. These numbers were supplied to the milk manufacturer, and an independent person not involved in the research placed labels on milk tins. Researchers and participants were blinded to treatment allocation until completion of the trial.
Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection)	Low risk	Adjustments for baseline outcome and study location in the repeated measures analysis with interaction between treatment and time period showed no baseline differences between intervention and control groups.
Similarity of baseline characteristics (checking for confounding,a potential consequence of selection bias)	Low risk	Study reported no differences between treatment groups at baseline except for the father's current employment status, with 96% in the GUMLi group with full-time paid employment compared with 84% in the control group.
Incomplete outcome data (attrition bias)	Low risk	3 participants (2 cow milk and 1 GUMLi) were excluded from the available nutrient data at baseline (n = 157), and after 12 mo of the intervention 5 participants (1 cow milk and 4 GUMLi) were excluded from the available nutrient data (n= 136).
Blinding of participants and personnel (performance bias)	Unclear risk	The study was reported as double blinded. Milks were provided at no cost to participants, and were produced in powder form, packaged in plain, identical 900 g tins, with no additional nutrient information panels or nutrition-related statements.

Blinding of outcome assessment (detection bias)	Unclear risk	Final data analysis post data lock was not blinded by treatment group, and no interim analysis was planned for the trial. There were no reported adverse reactions to the study milk and therefore the blinding procedure was maintained until the end of study.
Contamination (performance bias)	Unclear risk	Not clear.
Selective reporting (reporting bias)	Low risk	The study received ethical approval from the Health and Disability Ethics Committee of the Ministry of Health, New Zealand (14/NTB/152). The trial was registered in the Australian New Zealand Clinical Trials Registry number: ACTRN12614000918628.
Other bias	Low risk	There are no other clear signs of other bias.
Recruitment bias	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Baseline imbalance	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Loss of clusters	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Incorrect analysis	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Compatibility with individual RCTs	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.

Maldonado 2007

Methods	Double blind RCT, with 2 arms.
Participants	Healthy young children between 12-30 months old, without any relevant pathology from birth until their inclusion in the trial in Granada, Spain. All should had a varied diet according to their age, included cow milk.
Interventions	Children (n= 33) were randomly assigned to 1 of 2 groups: 1. group 1 (n= 16) received 500 ml iron fortified follow-on formula (9 boys and 7 girls) 2. group 2 (n= 17) received 500 ml cow milk (10 boys and 7 girls) The duration of the intervention was 4 months. MIlks were provided in identical packages.
Outcomes	Outcomes included weight, length, body mass index, haemoglobin, serum ferritin, serum iron, serum transferrin, mean corpuscular volume, hematocrit and mean corpuscular haemoglobin.
Notes	1. All included children had received breastfeeding for 6 to 8 months and follow-on formula thereafter until the first year of life. 2. Intervention milk contained 1.2 mg/100 ml Fe (form not specified). Source of funding: Milks were provided by Puleva Food S.L.

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer random number generator.
Allocation concealment (selection bias)	Low risk	The study describes that participants and investigators enrolling participants could not foresee assignment, only personnel providing the milk packages knew assignment.
Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection)	Low risk	There were no significant differences in outcomes variables between study groups at baseline.
Similarity of baseline characteristics (checking for confounding, a potential consequence of selection bias)	High risk	Not reported.
Incomplete outcome data (attrition bias)	Low risk	All included participants completed the study and results are shown.
Blinding of participants and personnel (performance bias)	Low risk	Milks were provided in identical packages to keep study double blind (to researchers and participants). Only the personnel providing the milk packages to the parents knew to which study group child belonged to.
Blinding of outcome assessment (detection bias)	Unclear risk	Not reported.
Contamination (performance bias)	Unclear risk	It is possible that contamination had occurred, but the risk of this happening is not clear.
Selective reporting (reporting bias)	Low risk	Research protocol was approved by the Ethical committee of University of Granada, Spain.
Other bias	Unclear risk	Not clear.
Recruitment bias	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Baseline imbalance	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Loss of clusters	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.

Incorrect analysis	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Compatibility with individual RCTs	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.

Morley 1999

Methods	Randomized controlled trial, with 3 arms.
Participants	Children 9 mo old from 3 centres in the United Kingdom (Leicester, Norwich, and Nottingham). Children were healthy infants born after 36 or more completed weeks of gestation, weighing > 2500 g, and either singletons or sole survivors from a multiple pregnancy.
Interventions	Children (n=493) were assigned to 1 of 3 groups: 1. group 1 (n= 166) would continue with cow milk as before (estimated to contain 0.05 mg iron/litre) 2. group 2 (n= 165) were given unfortified follow-on formula (with 0.9 mg iron/litre) 3. group 3 (n= 162) were given fortified follow-on formula (with 1.2 mg iron/litre as ferrous sulphate) The formula milks were supplied in powdered form; tins of iron fortified formula were labelled "formula 28" and tins of unfortified formula were labelled "formula 61". This code was not revealed by the manufacturers until the study was completed and all data had been entered and checked. Powdered milk was supplied ad libitum to the subjects' homes, and parents were given written and verbal information on how to make up the milk. For the purpose of this review, in this reference groups 1 vs 3 were analyzed.
Outcomes	Outcomes measured included mental and psychomotor development; growth; and haematological indexes.
Notes	Source of funding: Wyeth Laboratories.

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Randomization schedules were prepared by an independent statistician using permuted blocks of random length. Each centre had a separate schedule.
Allocation concealment (selection bias)	Low risk	Subjects were randomized by the research nurse from consecutively numbered opaque sealed envelopes.
Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection)	Low risk	The groups were well balanced, especially in terms of pre randomization anthropometry and developmental status.
Similarity of baseline characteristics (checking for confounding, a potential consequence of selection bias)	Low risk	Demographic characteristics and pre randomization measures were similar between groups.
Incomplete outcome data (attrition bias)	Low risk	Missing data is similar between groups and participant flow throughout the study is provided.

Blinding of participants and personnel (performance bias)	Low risk	The formula milks were supplied in powdered form; tins of iron fortified formula were labelled “formula 28” and tins of unfortified formula were labelled “formula 61”. This code was not revealed by the manufacturers until the study was completed and all data had been entered and checked. Staff at 3 centres for follow-up were blind to dietary allocation.
Blinding of outcome assessment (detection bias)	Low risk	Code was not revealed until the study was completed, and all data had been entered and checked.
Contamination (performance bias)	Low risk	There are no other clear signs of other bias.
Selective reporting (reporting bias)	Low risk	The study was approved by the Research Ethics Committee in each of the three collaborating UK centres (Leicester, Norwich, and Nottingham)
Other bias	Low risk	There are no other clear signs of other bias.
Recruitment bias	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Baseline imbalance	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Loss of clusters	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Incorrect analysis	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Compatibility with individual RCTs	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.

Rivera 2010 (C)

Methods	Double-blind cluster RCT with 2 arms
Participants	Children 12-30 mo old, beneficiaries of a program that distributes whole milk at subsidized prices in distribution centres throughout Mexico to children aged 1−11 and other family members of households living in poverty.

Interventions	The study team randomized 12 milk distribution clusters (from a total of 542) to receive either non fortified milk or fortified milk, wherein 567 participants were enrolled and as beneficiaries were entitled to received 400 ml of milk per day: 1. group 1 (n = 210) received unfortified milk (5 clusters) 2. group 2 (n= 357) received fortified milk (7 clusters). This group was larger per request of program officials to provide fortified milk to as many children as possible. 1. 400 ml of fortified milk (48 g powder) provided 5.28 mg iron, 5.28 mg zinc, 48 mg vitamin C and 32.1 mcg folic acid. 2. 400 ml of unfortified milk (48 g powder) provided 0.16 mg iron, 1.6 mg zinc, 6.8 mg vitamin C and 24 mcg folic acid. 3. Other nutrients content was similar. 4. The protocol was reviewed and approved by the Human Subjects and Ethics Committee of the National Public Health Institute, Mexico. 5. Authors reported adjusting for cluster effects.
Outcomes	Outcomes included anaemia (based on haemoglobin) and iron deficiency (ID) (based on serum ferritin and serum soluble transferrin) improvement from baseline classification of mild anaemia, moderate anaemia or mild-to-moderate anaemia. Monthly average daily milk intake during intervention (mL).
Notes	1. 400 ml of fortified milk (48 g powder) provided 5.28 mg iron, 5.28 mg zinc, 48 mg vitamin C and 32.1 mcg folic acid. 2. 400 ml of unfortified milk (48 g powder) provided 0.16 mg iron, 1.6 mg zinc, 6.8 mg vitamin C and 24 mcg folic acid. 3. Other nutrients content was similar. 4. The protocol was reviewed and approved by the Human Subjects and Ethics Committee of the National Public Health Institute, Mexico. 5. Authors reported adjusting for cluster effects. Source of funding: Supported by a grant from the Mexican Secretary of Social Development and Liconsa.

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Reported as randomized but random sequence generation was not reported.
Allocation concealment (selection bias)	Unclear risk	Allocation concealment was not reported.
Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection)	Low risk	Baseline concentrations of haemoglobin and serum transferrin, prevalence of anaemia and iron deficiency (as indicated by serum transferrin), the anthropometric scores, and the socioeconomic index were similar between groups.

Similarity of baseline characteristics (checking for confounding, a potential consequence of selection bias)	Unclear risk	Quote: "The only exceptions were a statistically significant difference (P<0.05) in age in the fortified milk group (22.6 ± 6.4 mo in children included compared with 21.1 ± 6.7 mo in those not included) and a statistically significant difference (P<0.05) in the sex distribution (% of boys: 54.8% in children included and 40.5% in those not included)." Quote: "In contrast with the results of comparisons for those included in the anaemia model compared with those not included, several differences were found between children included and not included in the final models for serum transferrin and serum ferritin, particularly in the fortified milk treatment group. Statistically significant differences (P<0.05) were found for age, weight-for-height z score, haemoglobin, and serum transferrin concentrations in the serum transferrin sample (fortified milk group) and for age, sex, weight-for-age, height-for-age, haemoglobin concentration, and prevalence of anaemia in the serum ferritin sample (fortified milk group only).
Incomplete outcome data (attrition bias)	Low risk	Of the 767 participants, 635 (17% attrition) completed the first follow-up at 6 months and 584 (24% attrition) the second follow-up. From the fortified milk group, 138 were lost to follow-up and from the cow milk group, 73 were lost to follow-up. Reasons for attrition included refusal to participate, drop-outs without blood sample or haemoglobin measurement or other measurements.
Blinding of participants and personnel (performance bias)	Low risk	The study was reported as double blinded. Randomization was blinded to researchers, personnel working in the milk distribution centres, and technical staff involved in the study until the end of the trial. The colour code of milk packages was unknown to researchers, field workers, personnel in the distribution centres, and program beneficiaries and was not disclosed before data analysis. As the unit of randomization was the milk distribution clusters, personnel and program beneficiaries in each cluster were exposed to only one of the colour-coded bands and were unaware of the existence of a different colour-code band.
Blinding of outcome assessment (detection bias)	Low risk	Study was unblinded until the end of trial and colour code of milk packages was disclosed until data was analyzed.
Contamination (performance bias)	Low risk	It is unlikely that contamination occurred because all milk distributed in each centre was either fortified or unfortified, that is, with only one colour-code band.
Selective reporting (reporting bias)	Low risk	All outcomes stated in the methods section were reported. This study was approved by the Human Subjects and Ethics Committee of the National Public Health Institute.

Other bias	Low risk	There are no other clear signs of bias. Source of funding: Supported by a grant from the Mexican Secretary of Social Development and Liconsa.
Recruitment bias	High risk	Children were recruited to the trial after the clusters were randomized.
Baseline imbalance	Unclear risk	Not clear.
Loss of clusters	Low risk	No clusters were lost to follow-up.
Incorrect analysis	Low risk	Baseline characteristics of children and their families were compared between intervention groups. Regression and logistics models, adjusting for cluster effects, were used for the comparisons between interventions.
Compatibility with individual RCTs	Low risk	The cluster design did not seem to affect the findings of the study.

Sazawal 2007

Methods	Double-blind RCT, community-based trial with 4 arms.
Participants	Children 1-3 years old from a peri-urban population in New Delhi, India. All permanently resident families with children aged 1-3 years were invited to participate in the study.
Interventions	The 4-arms were divided in two trials. Two-arms were related to a clinical trial to evaluate the efficacy of a different milk preparation fortified with probiotic (compared with that reparation without fortification) in a non-factorial design with joint randomization. In the other two-armed trial, children (n = 633) were randomly assigned to 1 of 2 groups: 1. group 1 (n=316) received fortified milk 2. group 2 (n=317) received control milk Groups 1 and 2 were implemented concurrently with another two-armed clinical trial with groups 3 and 4: 3. group 3 received milk fortified with micronutrients and probiotics 4. group 4 received milk without probiotics For the purpose of this review only children from groups 1, 2 and 3, reported in this study were used. Irrespective of group allocation all children who had severe anaemia at baseline were given a therapeutic dose of iron for three months in addition to their milk supplement. Supplementation continued for 1 year.
Outcomes	Outcomes measured included days with severe illnesses, incidence and prevalence of diarrhoea, and acute lower respiratory illness.
Notes	1. Single serving 32 g sachets were used to fortify milk. 2. Fortified milk provided additional 7.8 mg zinc, 9.6 mg iron, 4.2 g selenium, 0.27 mg copper, 156 g vitamin A, 40.2 mg vitamin C, 7.5 mg vitamin E per day (three feeds). 3. Assistants delivered 21 sachets each week to each home and advised that the child should consume up to three sachets a day.Source of funding: Fonterra Brands, Auckland, New Zealand, funded the study and provided the milk powder used in the trial.

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	A random allocation sequence of group codes was generated with permuted blocks of length 16. Two separate randomization lists —one for children with baseline Hb > 70 g/l and another for children with baseline Hb ≤ 70 g/l were created, resulting in two serially numbered lists with allocated treatment codes before any children was enrolled.
Allocation concealment (selection bias)	Low risk	Allocation was not known to investigators or anyone in the field until the study was finished and the data analyzed.
Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection)	Low risk	Children in both groups were comparable at baseline for haematology and plasma zinc status.
Similarity of baseline characteristics (checking for confounding, a potential consequence of selection bias)	Unclear risk	Children in both the groups were comparable at baseline for sociodemographic variables. Not clear if breast fed children were included or not.
Incomplete outcome data (attrition bias)	Unclear risk	In the fortified milk group 9 children left the area, 2 children died, and 16 withdrawn consents. In the milk group 18 children left the area, 2 children died, and 16 withdrawn consents. Participants´ flow chart is not clear.
Blinding of participants and personnel (performance bias)	Low risk	Study was reported as double blinded, both to investigators and to participants.
Blinding of outcome assessment (detection bias)	Low risk	Study was unblinded until data was analyzed.
Contamination (performance bias)	Low risk	It is unlikely that contamination in the study occurred. The letter code of the supplementation box was stripped off and labelled with the child's identification information.
Selective reporting (reporting bias)	Unclear risk	The study was approved by the Human research and ethical review committee at the Johns Hopkins Bloomberg School of Public Health, and the Annamalai University, India. Participant's flow is not clear.
Other bias	Unclear risk	Not clear.
Recruitment bias	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Baseline imbalance	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Loss of clusters	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Incorrect analysis	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Compatibility with individual RCTs	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.

Stecksén-Blicks 2009 (C)

Methods	Double-blind cluster. placebo controlled RCT with 2 arms
Participants	Children 1−5 years of age attending 14-day care centres with 27 units in northern Sweden were invited and recruited through meetings with parents.
Interventions	Children (n= 248) were assigned 1 of 2 groups: 1. group 1 (n= 133, from 16-day care units)) received fortified milk (150 ml per day) with Lactobacillus rhamnosus LB21 (10 7 CFU/ml) and 2.5 mg fluoride per litre 2. group 2 (n= 115, from 10-day care units) received standard milk (150 ml per day). Children were served 150 ml medium-fat milk (1.5%) at lunch. Milk was prepared by the day care staff by adding one colour-coded capsule (10 ml) to each litre of milk. The capsules were kept frozen and contained fluoride and probiotic bacteria in skim milk to give a final concentration of 2.5 mg fluoride and 10 7 CFU/ml rhamnosus LB21 per litre in the intervention group. The capsules in the control group contained only skimmed milk and were identical in appearance except in colour code. Records available from this study do not allow for data extraction and thus do not contribute data for this review.
Outcomes	Primary outcome was caries increment and secondary outcome were measures of general health (days with respiratory and gastrointestinal symptoms, number of visits to doctor, days on antibiotic treatment, days with otitis media and days with sick leave).
Notes	1. The institutions were located in Nordmaling and Hörnefors, which are small communities with less than 10,000 inhabitants close to the city of Umeå in northern Sweden. 2. The intervention was served only during the weekdays (at school) and lasted 21 months. 3. Authors reported adjusting for cluster effects. Source of funding: The study was supported financially by the County Council of Västerbotten (TUA) and the Borrow Foundation, UK. Norrmejerier Ekonomisk Förening, Umeå, Sweden supported the study by preparation and distribution of the milk.

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Day care centres were randomly assigned to two parallel groups by a staff member at the local dairy by means of coin tossing.
Allocation concealment (selection bias)	Low risk	Day care units were referred to as blue or yellow units in order to conceal their allocation. The code was kept by an independent monitor and was not unveiled until all data were computerized.
Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection)	Low risk	There were no significant differences at baseline outcomes between groups.
Similarity of baseline characteristics (checking for confounding, a potential consequence of selection bias)	Low risk	There were no significant differences at baseline characteristics between groups.

Incomplete outcome data (attrition bias)	Low risk	After 3 months, 58 children left the project because they moved to primary schools and 4 dropped out for other reasons.
Blinding of participants and personnel (performance bias)	Low risk	Neither the researchers nor the clinicians, personnel or families at the day care centres knew whether the children received control or intervention milk during the course of the study.
Blinding of outcome assessment (detection bias)	Unclear risk	Not mentioned.
Contamination (performance bias)	Low risk	Since randomization was done at day care centre level, contamination was unlikely to have occurred.
Selective reporting (reporting bias)	Low risk	The study was approved by the research ethics committee at Umeå University (§562/03, dnr 03-475).
Other bias	Unclear risk	Not clear.
Recruitment bias	Low risk	Children were recruited to the trial before the clusters were randomized.
Baseline imbalance	Low risk	The intra-cluster correlation coefficient was estimated to reflect the homogeneity of the sample using long one-way ANOVA.
Loss of clusters	Low risk	No clusters were lost to follow-up.
Incorrect analysis	Low risk	Because of the clustered design, the outcomes in caries and general health were analyzed with age as covariate in a multilevel logistic regression.
Compatibility with individual RCTs	Low risk	The cluster design did not seem to affect the findings of the study.

Stekel 1986

Methods	Randomized controlled trial with 2 arms.
Participants	Children 3 months old, spontaneously weaned, from three peripheral community clinics of the National Health Service in Chile.
Interventions	Children (n= 510) were randomly assigned to 1 of 2 groups: 1. group 1 (n=276) received low-fat fortified powdered milk 2. group 2 (n=232) received low-fat non-fortified powdered milk Both milks had 12% fat. Both milk products were prepared by the mother and were utilized in a 10% dilution (weight/volume) plus 5% sucrose and usually 3% corn flour.
Outcomes	Outcomes included haemoglobin, hematocrit, serum iron, total iron binding capacity, total transferrin saturation and protoporphyrin. Weight, length and cephalic perimeter were also measured.

Notes	1. Fortified milk had 15 mg iron as ferrous sulphate per 100 gr or milk powdered. 2. Forified milk powdered was distributed in cans. 3. Mothers received 3 kg milk product per month until child was 6 mo old, and 2 kg per month until the child was 24 mo old. 4. Mean breastfeeding duration for the participant children was 4 months. Source of funding: Research Corporation USA.

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Method not reported.
Allocation concealment (selection bias)	Unclear risk	Method not reported.
Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection)	Low risk	Baseline outcomes were similar between groups except for ferritin values that were significantly lower in the intervention group.
Similarity of baseline characteristics (checking for confounding, a potential consequence of selection bias)	Unclear risk	Not mentioned.
Incomplete outcome data (attrition bias)	Low risk	Main causes of attrition were migration and violations of protocol which are reasons not likely to bias the results.
Blinding of participants and personnel (performance bias)	High risk	No blinding was reported.
Blinding of outcome assessment (detection bias)	High risk	No blinding was reported.
Contamination (performance bias)	Unclear risk	Not clear.
Selective reporting (reporting bias)	Unclear risk	Not clear.
Other bias	Unclear risk	Not clear.
Recruitment bias	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Baseline imbalance	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Loss of clusters	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Incorrect analysis	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Compatibility with individual RCTs	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.

Stekel 1988

Methods	Randomized controlled trial with 2 arms
Participants	Children 3 months old, spontaneously weaned, from two peripheral community clinics of the National Health Service in Chile.
Interventions	Children (n= 554) were randomly assigned to 1 of 2 groups: 1. group 1 (n= 276) received full-fat fortified powdered milk 2. group 2 (n= 278) received full-fat non-fortified powdered milk Both milk products were prepared by the mother and were utilized in a 10% dilution (weight/volume) plus 5% sucrose and usually 3% corn flour.
Outcomes	Outcomes included serum measurements of haemoglobin, haematocrit, serum iron, serum total binding capacity, serum ferritin, transferrin saturation; and anthropometric measurements (weight and length).
Notes	1. Fortified milk was full-fat (26%) powdered milk with 15 mg of Fe as ferrous sulfate, 100 mg of ascorbic acid, 1500 IU of vitamin A, and 400 IU of vitamin D per 100 g of powder. 2. Children were clinically followed every 15 days by the same group of physicians. Source of funding: Not mentioned.

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Unclear risk	Method not reported.
Allocation concealment (selection bias)	Unclear risk	Method not reported.
Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection)	Low risk	Baseline outcomes was similar between groups. Birth weight and weight were similar in both girls and boys as was iron status.
Similarity of baseline characteristics (checking for confounding, a potential consequence of selection bias)	Low risk	Baseline characteristics were similar between groups. Sex distribution within participants was similar with a slight predominance of males and no significant differences in mother's age, parity, or number of pregnancies were found. The socioeconomic level was identical between groups.
Incomplete outcome data (attrition bias)	Low risk	Main causes of attrition were migration and violations of protocol which are reasons not likely to bias the results.
Blinding of participants and personnel (performance bias)	High risk	No blinding was reported.
Blinding of outcome assessment (detection bias)	High risk	No blinding was reported.
Contamination (performance bias)	Unclear risk	Not clear.
Selective reporting (reporting bias)	Unclear risk	The study was approved by the Institute of Nutrition and Food Technology of the University Chile's Ethics in Human Research Committee. Anthropometric and nutritional survey outcomes for acceptability and compliance not shown.
Other bias	Unclear risk	Not clear.
Recruitment bias	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.

Baseline imbalance	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Loss of clusters	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Incorrect analysis	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Compatibility with individual RCTs	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.

Svahn 1999

Methods	Double blinded RCT with 4 arms
Participants	Healthy children 11 mo of age from child health centres in Malmö, Sweden.
Interventions	Children (n= 54, 33 girls, 21 boys) were randomly assigned to 1 of 4 groups: 1. group 1 (n= 8) received low-fat milk (LF) (1.0 g fat/dl, 3.3 g protein/dl); and products including low-fat yogurt, sour milk, cream, margarine, cheese (17% fat), and butter (for frying) could be used. 2. group 2 (n= 9) received standard-fat milk (SF) (3.5 g fat/dl, 3.3 g protein/dl); and products including full-fat yogurt, sour milk, cream, cheese (28-32% fat), and butter (for frying) could be used. 3. group 3 (n= 9) received partially vegetable fat and protein-reduced milk (PVF) (3.5 g fat/dl, 50% vegetable; 2.2 g protein/dl); and products including low-fat yogurt, sour milk, margarine with 100% vegetable fat, cheese, and cream with high content of vegetable fat could be used. This milk was fortified with 7.0 mg Fe l^{-1} as ferrous gluconate. 4. group 4 (n= 12) received full-vegetable-fat milk (FVF) (3.5 g fat/dl, 100% vegetable; 3.0 g protein/dl); and products including low-fat yogurt, sour milk, margarine with 100% vegetable fat, cheese, and cream with high content of vegetable fat could be used. These milks was fortified with 14.9 mg F l^{-1} as ferrous lactate. Milks were ready to feed and were provided free by the investigator. The intervention lasted 6 months.
Outcomes	Nutrient intake, blood lipids, growth, iron intake and iron status were measured at 12, 15, and 18 months.
Notes	1. All children were breast fed at birth; at 3 months 78% were still breast fed (partially or exclusively), at 6 months 50%, at 9 months 19%, and at 11 months 8%. None of the children was breast fed at 12 months of age. 2. No other milk or formula was fed to the children Source of funding: The LF, SF, and PVF milks were prepared specially for the study by Valio OY, Helsinki, Finland, and the FVF milk by Humana Milchwerke Westfalen. The study was supported by the Albert Påhlsson Foundation,Sweden, and Deutsche Forschungsgemeinschaft, Bonn, Germany.

Risk of bias table

Bias	Authors' judgement	Support for judgement

Random sequence generation (selection bias)	Unclear risk	Method not mentioned.
Allocation concealment (selection bias)	Unclear risk	Method not mentioned.
Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection)	Low risk	Total nutrient intakes and growth did not differ statistically among the four diet groups.
Similarity of baseline characteristics (checking for confounding, a potential consequence of selection bias)	Unclear risk	Not clear.
Incomplete outcome data (attrition bias)	Low risk	Sixteen children were excluded from the study. They had all participated less than 2 months. One child exhibited milk intolerance after gastroenteritis and was removed from the study by the investigator. Three children were removed by the investigator because of failure to adhere to protocol, and one child was withdrawn after injury in a car accident. Eleven children were removed from the study by the parents. Of these children, three did not like the milk, two were afraid of blood sampling, and two had loose stools without signs of milk intolerance. Of the remaining four, one had a urinary tract infection and vesicoureteral reflux, one had constipation, one had gastroenteritis followed by constipation, and one parent would not agree to feeding low-fat yogurt.
Blinding of participants and personnel (performance bias)	Low risk	The parents and investigators were blinded to the type of milk, and the investigators were blinded to the dairy products.
Blinding of outcome assessment (detection bias)	High risk	A dietitian controlled the children records and was aware of the results of the randomization. Staff was thoroughly instructed in the feeding regimen and taught parents how to record the child's diet.
Contamination (performance bias)	Unclear risk	Not clear.
Selective reporting (reporting bias)	High risk	The study was approved by the Ethics Committee, Lund University, Sweden. The study had a very small sample size.
Other bias	Unclear risk	Not clear.
Recruitment bias	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Baseline imbalance	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Loss of clusters	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Incorrect analysis	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Compatibility with individual RCTs	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.

Szymlek-Gay 2009

Methods	Randomized placebo-controlled trial with 3 arms.
Participants	Healthy non anaemic 12−20-mo-old children from urban centres (at sea level) in New Zealand.
Interventions	Children (n=225) were randomly assigned to 1 of 3 groups: 1. group 1 (n=90) received red meat (toddlers encouraged to consume 2.6 mg iron from red meat dishes daily). 2. group 2 (n=45) received fortified milk -toddlers' regular milk replaced with iron-fortified cow milk (1.5 mg iron/100 g prepared milk). 3. group 3 (n=90) control -toddlers' regular milk replaced with non-fortified cow milk (0.01 mg iron/100 g prepared milk). Participants in the milk groups were asked to replace their regular milk with either commercially available iron-fortified powdered cow milk, or non-fortified powdered cow milk. The control group was a non-treatment control for the red meat group, and a placebo control for the fortified milk group. Mothers continued to breastfeed at their discretion. The intervention lasted 20 weeks. For the purpose of this review only groups 2 and 3 were used.
Outcomes	Haemoglobin, serum ferritin, serum transferrin receptor, and C-reactive protein. The prevalence of sub optimal iron status (i.e. depleted iron stores, iron-deficient erythropoiesis, and iron deficiency anaemia) was determined, and body iron was calculated.
Notes	1. Fortified cow milk had iron as ferrous sulfate, calcium, magnesium, zinc, vitamin C, vitamin E, niacin, vitamin A, vitamin D, vitamin B-6, thiamin, and folate (Heinz Nurture Toddler Enriched Milk Drink; Heinz Wattie's Ltd, Hastings, New Zealand; fortified milk group). 2. Non fortified (Standard Instantized Whole Milk Powder with required A and D added; Fonterra, Auckland, New Zealand; control group) 3. The milks were packaged (Sutton Group Ltd, Auckland, New Zealand) into identical white 900-g cans (Canpac International, Hamilton, New Zealand) along with identical 17-mL scoops. Source of funding: Supported by the Health Research Council of New Zealand, Meat and Livestock Australia, Meat and Wool New Zealand, and the University of Otago. Heinz Wattie's New Zealand Ltd provided the iron-fortified milk; Fonterra New Zealand provided the non-fortified milk; Canpac International Ltd donated the cans and spoons;

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	Low risk	Computer-generated random assignment process based on the minimization method stratified by baseline C-reactive protein (<10 or ≥10 mg/L) and serum ferritin (<25 or ≥25 lg/L).
Allocation concealment (selection bias)	Low risk	Two investigators not involved in recruitment and data collection randomly assigned participants to each group.
Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection)	Low risk	No statistically significant differences were shown between variables at baseline.
Similarity of baseline characteristics (checking for confounding, a potential consequence of selection bias)	Low risk	Children´s age and sex were similar between groups.

Incomplete outcome data (attrition bias)	Low risk	A total of 215 completed the study: 10 children (4.4%) were lost to follow-up and a further 10 failed to provide the final blood sample because of unsuccessful blood sampling. In cow milk group 5, in read meat group 3, and in fortified milk group 2 children were lost as their caregivers desire them to withdraw,
Blinding of participants and personnel (performance bias)	Low risk	Milks were packaged into identical white 900-g cans along with 17-mL scoops. The cans were marked with only a concealed code number, which was known to only one research assistant who was not involved in data collection or analysis. The individuals who carried out the laboratory analyses were unaware of the participants' group assignments.
Blinding of outcome assessment (detection bias)	Low risk	Allocation code was broken when all data had been collected and analyzed.
Contamination (performance bias)	Low risk	There are no other clear signs of other bias.
Selective reporting (reporting bias)	Low risk	The study protocol was approved by the Human Ethics Committee of the University of Otago, Dunedin, New Zealand and was registered at actr.org.au as ACTRN12605000487617
Other bias	Low risk	There are no other clear signs of other bias.
Recruitment bias	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Baseline imbalance	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Loss of clusters	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Incorrect analysis	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Compatibility with individual RCTs	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.

Van der Gaag 2015

Methods	Nonrandomized retrospective case control study
Participants	Children aged between 1-6 years with recurrent respiratory tract symptoms
Interventions	Children (n= 99) were assigned 1 of 2 groups: 1. group 1 (n= 50) received dietary advice of daily whole milk/ yoghourt and natural butter 2. group 2 (n= 49) could continue their usual semi skimmed milk and low-fat margarine consumption as before (control group). The intervention duration was 2 months.
Outcomes	Respiratory symptoms
Notes	Source of funding: Not mentioned.

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Not used.
Allocation concealment (selection bias)	High risk	Not used.
Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection)	Unclear risk	Not mentioned.
Similarity of baseline characteristics (checking for confounding, a potential consequence of selection bias)	Unclear risk	Not mentioned.
Incomplete outcome data (attrition bias)	Unclear risk	Not clear.
Blinding of participants and personnel (performance bias)	High risk	Not used.
Blinding of outcome assessment (detection bias)	High risk	Not used.
Contamination (performance bias)	Unclear risk	Not clear.
Selective reporting (reporting bias)	Unclear risk	Not clear.
Other bias	Unclear risk	Not clear.
Recruitment bias	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Baseline imbalance	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Loss of clusters	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Incorrect analysis	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Compatibility with individual RCTs	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.

Villalpando 2006

Methods	Double-blind RCT trial with 2 arms
Participants	Healthy children 10-30 mo old selected from registry of children younger than 5 y of age in a poor peri urban community of 5000 inhabitants in the outskirts of Puebla, a city located 120 km east of Mexico City. Such a registry is maintained and periodically updated by the local health facility.

Interventions	Children (n=115) were randomly assigned to 1 of 2 groups: 1. group 1 (n=58) received 400 mL/d of fortified cow's whole milk. 2. group 2 (n=57) received 400 mL/d of unfortified cow´s whole milk. Milks were distributed in powder form and mothers were instructed how to reconstitute. Milk was delivered weekly to each family, and about the amount of milk intended for the infant to drink daily.
Outcomes	Hemoglobin, serum ferritin, soluble transferrin receptors (TfR), and C-reactive protein concentrations were measured, and prevalence of anaemia estimated.
Notes	1. Fortified cow´s milk contained 5.8 mg/400 mL of iron as ferrous gluconate, 5.28 mg/400mL of zinc as zinc oxide, and 48 mg/400mL of ascorbic acid. 2. Unfortified cow´s milk contained 0.2 mg iron/400 mL, 1.9 mg zinc/400 mL, and 6.8 mg ascorbic acid/400 mL. 3. Units of 220 g of the product were packed in metallic foil sachets. The packages of fortified milk and unfortified milks were undistinguishable, except for a colour-coded band in the upper corner of the sachet. Source of funding: Supported in part by The Ministry of Social Development of Mexico and Instituto Nacional de Salud Publica

Risk of bias table

Bias	Authors' judgement	Support for judgement
Random sequence generation (selection bias)	High risk	Method not reported. Quote: "The randomization procedure did not result in an even distribution of baseline anaemia in the 2 intervention groups (30.0% non-fortified milk, 41.4% fortified milk)." Quote: "Theoretically, such a difference made the fortified milk group more susceptible to a larger improvement."
Allocation concealment (selection bias)	Low risk	Milks allocation was blinded to researchers, field workers and participants.
Similarity of baseline outcome measurements (checking for confounding, a potential consequence of selection)	High risk	Groups did not differ at 6 mo and the changes between baseline and 6 mo did not differ between the groups. However, baseline outcomes showed that fortification milk group was more susceptible to a larger improvement.
Similarity of baseline characteristics (checking for confounding, a potential consequence of selection bias)	Low risk	Age, weight, length, energy intake, and distribution by gender did not differ between groups at baseline.

Incomplete outcome data (attrition bias)	Unclear risk	Not clear.
Blinding of participants and personnel (performance bias)	Unclear risk	The study was blinded to personnel.
Blinding of outcome assessment (detection bias)	Low risk	The colour code was unknown to researchers, field workers, and users and was disclosed after data analysis.
Contamination (performance bias)	Low risk	It is unlikely that contamination occurred as milk packages had a colour-coded band.
Selective reporting (reporting bias)	Low risk	The protocol was reviewed and approved by the Research, Ethics and Biohazards Committees from the National Public Health Institute, Cuernavaca, Mexico.
Other bias	Unclear risk	Not clear.
Recruitment bias	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Baseline imbalance	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Loss of clusters	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Incorrect analysis	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.
Compatibility with individual RCTs	Unclear risk	Not applicable. We only evaluated this domain in cluster-randomized studies.

Data and analyses

1 Animal milk (full-fat or lower-fat milk) versus no other milk

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate

2 Animal milk (full-fat or lower-fat milk) versus follow-on formula

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
2.1 Weight (kg) (All)	4	604	Mean Difference (IV, Random, 95% CI)	0.13 [-0.11, 0.36]
2.2 Weight (kg) (subgroup by stunting)	4	604	Mean Difference (IV, Random, 95% CI)	0.13 [-0.11, 0.36]
2.2.1 Stunting non-specified	4	604	Mean Difference (IV, Random, 95% CI)	0.13 [-0.11, 0.36]
2.2.2 Stunting very low prevalence (<2.5%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.2.3 Stunting low prevalence (2-5-9.9%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.2.4 Stunting medium prevalence (10-19.9%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.2.5 Stunting high prevalence (>20%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.3 Weight (kg) (subgroup by wasting)	4	604	Mean Difference (IV, Random, 95% CI)	0.13 [-0.11, 0.36]
2.3.1 Wasting non-specified	4	604	Mean Difference (IV, Random, 95% CI)	0.13 [-0.11, 0.36]
2.3.2 Wasting very low prevalence (<2.5%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.3.3 Wasting low prevalence (2.5-4.9%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.3.4 Wasting medium prevalence (5-10%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.3.5 Wasting high prevlaence (>10%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable

2.4 Weight (kg) (subgroup by type of feeding before 12 mo)	4	604	Mean Difference (IV, Random, 95% CI)	0.13 [-0.11, 0.36]
2.4.1 Non-specified	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.4.2 Breastmilk only	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.4.3 Breastfeeding + complementary feeding	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.4.4 Breastmilk substitute + complementary feeding	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.4.5 Breastfeeding + breastmilk substitute + complementary feeding	4	604	Mean Difference (IV, Random, 95% CI)	0.13 [-0.11, 0.36]
2.5 Weight (kg) (subgroup by anaemia)	4	604	Mean Difference (IV, Random, 95% CI)	0.13 [-0.11, 0.36]
2.5.1 Anaemia non-specified	2	176	Mean Difference (IV, Random, 95% CI)	0.34 [-0.13, 0.81]
2.5.2 Anaemia between 0-4.9%	2	428	Mean Difference (IV, Random, 95% CI)	0.05 [-0.22, 0.33]
2.5.3 Anaemia between 5-19.9%	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.5.4 Anaemia between 20-39.9%	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.5.6 40% or higher	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.6 Weight (kg) (subgroup by funding source)	4	604	Mean Difference (IV, Random, 95% CI)	0.13 [-0.11, 0.36]
2.6.1 Private industry/brand	3	571	Mean Difference (IV, Random, 95% CI)	0.13 [-0.10, 0.37]
2.6.2 Other	1	33	Mean Difference (IV, Random, 95% CI)	-0.20 [-1.72, 1.32]
2.7 Height (cm) (All)	4	604	Mean Difference (IV, Random, 95% CI)	0.20 [-0.31, 0.72]

2.8 Height (cm) (subgroup by stunting)	4	604	Mean Difference (IV, Random, 95% CI)	0.20 [-0.31, 0.72]
2.8.1 Stunting non-specified	4	604	Mean Difference (IV, Random, 95% CI)	0.20 [-0.31, 0.72]
2.8.2 Stunting very low-prevalence (<2.5%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.8.3 Stunting low prevalence (2.5-9.9%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.8.4 Stunting medium prevalence (10-19.9%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.8.5 Stunting high prevalence (>20%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.9 Height (cm) (subgroup by wasting)	4	604	Mean Difference (IV, Random, 95% CI)	0.20 [-0.31, 0.72]
2.9.1 Wasting non-specified	4	604	Mean Difference (IV, Random, 95% CI)	0.20 [-0.31, 0.72]
2.9.2 Wasting very low prevalence (<2.5%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.9.3 Wasting low prevalence (2.5-4.9%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.9.4 Wasting medium prevalence (5-10%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.9.5 Wasting high prevalence (>10%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.10 Height (cm) (subgroup by type of feeding before 12 mo)	4	604	Mean Difference (IV, Random, 95% CI)	0.20 [-0.31, 0.72]
2.10.1 Non-specified	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.10.2 Breastmilk only	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.10.3 Breastfeeding + complementary feeding	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.10.4 Breastmilk substitute + complementary feeding	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable

2.10.5 Breastfeeding + breastmilk substitute + complementary feeding	4	604	Mean Difference (IV, Random, 95% CI)	0.20 [-0.31, 0.72]
2.11 Height (cm) (subgroup by anaemia)	4	604	Mean Difference (IV, Random, 95% CI)	0.20 [-0.31, 0.72]
2.11.1 Anaemia non-specified	2	176	Mean Difference (IV, Random, 95% CI)	0.21 [-0.87, 1.30]
2.11.2 Anaemia between 0-4.9%	2	428	Mean Difference (IV, Random, 95% CI)	0.20 [-0.38, 0.78]
2.11.3 Anaemia between 5-19.9%	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.11.4 Anaemia between 20-39.9%	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.11.5 Anaemia 40% or higher	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.12 Height (cm) (subgroup by funding source)	4	604	Mean Difference (IV, Random, 95% CI)	0.20 [-0.31, 0.72]
2.12.1 Private industry/brand	3	571	Mean Difference (IV, Random, 95% CI)	0.22 [-0.30, 0.74]
2.12.2 Other	1	33	Mean Difference (IV, Random, 95% CI)	-1.20 [-5.70, 3.30]
2.13 Weight-for-height z score (WHZ)	1	143	Mean Difference (IV, Random, 95% CI)	0.30 [-0.01, 0.61]
2.14 Head circumference (cm)	2	425	Mean Difference (IV, Random, 95% CI)	-0.05 [-0.36, 0.26]
2.15 Body composition-Body Mass Index	2	176	Mean Difference (IV, Random, 95% CI)	0.28 [-0.15, 0.70]
2.16 Body composition-% Body fat	1	134	Mean Difference (IV, Random, 95% CI)	2.40 [-0.16, 4.96]
2.17 Nutritional status-Vit D as serum 25-hydroxyvitamin D [25(OH)D], nmol/L	2	455	Mean Difference (IV, Random, 95% CI)	-16.27 [-21.23, -11.31]
2.18 Nutritional status-Vit D deficiency	2	455	Risk Ratio (M-H, Random, 95% CI)	2.64 [1.57, 4.45]

2.19 Nutritional status-Iron as serum iron (µmol/l)	1	134	Mean Difference (IV, Random, 95% CI)	-0.70 [-2.63, 1.23]
2.20 Child development-Bayley psychomotor development index (PDI)	2	395	Mean Difference (IV, Random, 95% CI)	-1.15 [-3.07, 0.77]
2.21 Child development-Bayley mental development index (MDI)	2	395	Mean Difference (IV, Random, 95% CI)	1.55 [-0.64, 3.73]
2.22 Iron deficiency anaemia (IDA)	2	445	Risk Ratio (M-H, Random, 95% CI)	5.37 [0.94, 30.48]
2.23 Iron deficiency (ID, serum ferritin <12 µg/l)	2	452	Risk Ratio (M-H, Random, 95% CI)	2.33 [1.40, 3.86]
2.24 Haemoglobin (g/L) (All)	6	663	Mean Difference (IV, Random, 95% CI)	-2.61 [-4.86, -0.37]
2.25 Haemoglobin (g/L) (subgroup by stunting)	6	663	Mean Difference (IV, Random, 95% CI)	-2.61 [-4.86, -0.37]
2.25.1 Stunting non-specified	6	663	Mean Difference (IV, Random, 95% CI)	-2.61 [-4.86, -0.37]
2.25.2 Stunting very low prevalence (<2.5%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.25.3 Stunting low prevalence (2.5-9.9%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.25.4 Stunting medium prevalence (10-19.9%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.25.5 Stunting high prevalence (>20%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.26 Haemoglobin (g/L) (subgroup by wasting)	6	663	Mean Difference (IV, Random, 95% CI)	-2.61 [-4.86, -0.37]
2.26.1 Wasting non-specified	6	663	Mean Difference (IV, Random, 95% CI)	-2.61 [-4.86, -0.37]
2.26.2 Wasting very low prevalence (<2.5%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.26.3 Wasting low prevalence (2.5-4.9%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.26.4 Wasting medium prevalence (5-10%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable

2.26.5 Wasting high prevalence (>10%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.27 Haemoglobin (g/L) (subgroup by type of feeding before 12 mo)	6	663	Mean Difference (IV, Random, 95% CI)	-2.61 [-4.86, -0.37]
2.27.1 Non specified	1	78	Mean Difference (IV, Random, 95% CI)	-8.00 [-12.60, -3.40]
2.27.2 Breastmilk only	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.27.3 Breastfeeding + complementary feeding	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.27.4 Breastmilk substitute +complementary feeding	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.27.5 Breastfeeding + breastmilk substitute + complementary feeding	5	585	Mean Difference (IV, Random, 95% CI)	-1.46 [-3.13, 0.21]
2.28 Haemoglobin (g/L) (subgroup by anaemia)	6	663	Mean Difference (IV, Random, 95% CI)	-2.99 [-5.31, -0.66]
2.28.1 Anaemia non-specified	2	160	Mean Difference (IV, Random, 95% CI)	-0.89 [-3.24, 1.45]
2.28.2 Anaemia between 0-4.9%	2	107	Mean Difference (IV, Random, 95% CI)	-4.06 [-9.94, 1.82]
2.28.3 Anaemia between 5-19.9%	2	396	Mean Difference (IV, Random, 95% CI)	-4.63 [-10.46, 1.21]
2.28.4 Anaemia between 20-39.9%	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.28.5 Anaemia 40% or higher	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.29 Haemoglobin (g/L) (subgroup by funding source)	6	663	Mean Difference (IV, Random, 95% CI)	-2.54 [-4.66, -0.42]
2.29.1 Private industry/brand	5	630	Mean Difference (IV, Random, 95% CI)	-3.44 [-5.73, -1.16]
2.29.2 Other	1	33	Mean Difference (IV, Random, 95% CI)	0.00 [-1.07, 1.07]
2.30 Ferritin (µg/L) (All)	6	796	Mean Difference (IV, Random, 95% CI)	-9.87 [-15.02, -4.72]

2.31 Ferritin (µg/L) (subgroup by stunting)	6	796	Mean Difference (IV, Random, 95% CI)	-9.87 [-15.02, -4.72]
2.31.1 Stunting non-specified	6	796	Mean Difference (IV, Random, 95% CI)	-9.87 [-15.02, -4.72]
2.31.2 Stunting very low prevalence (<2.5%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.31.3 Stunting low prevalence (2.5-9.9%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.31.4 Stunting medium prevalence (10-19.9%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.31.5 Stunting high prevalence (>20%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.32 Ferritin (µg/L) (subgroup by wasting)	6	796	Mean Difference (IV, Random, 95% CI)	-9.87 [-15.02, -4.72]
2.32.1 Wasting non-specified	6	796	Mean Difference (IV, Random, 95% CI)	-9.87 [-15.02, -4.72]
2.32.2 Wasting very low prevalence (<2.5)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.32.3 Wasting low prevalence (2.5-4.9%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.32.4 Wasting medium prevalence (5-10%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.32.5 Wasting high prevalence (>10%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.33 Ferritin (µg/L) (subgroup by type of feeding before 12 mo)	6	796	Mean Difference (IV, Random, 95% CI)	-9.87 [-15.02, -4.72]
2.33.1 Non specified	1	78	Mean Difference (IV, Random, 95% CI)	-17.50 [-26.13, -8.87]
2.33.2 Breastmilk only	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.33.3 Breastfeeding + complementary feeding	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.33.4 Breastmilk substitute + complementary feeding	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable

2.33.5 Breastfeeding + breastmilk substitute + complementary feeding	5	718	Mean Difference (IV, Random, 95% CI)	-8.81 [-14.30, -3.32]
2.34 Ferritin (µg/L) (subgroup by anaemia)	6	796	Mean Difference (IV, Random, 95% CI)	-9.87 [-15.02, -4.72]
2.34.1 Anaemia non-specified	2	167	Mean Difference (IV, Random, 95% CI)	-16.52 [-19.11, -13.92]
2.34.2 Anaemia between 0-4.9%	2	261	Mean Difference (IV, Random, 95% CI)	-3.10 [-11.53, 5.33]
2.34.3 Anaemia between 5-19.9%	2	368	Mean Difference (IV, Random, 95% CI)	-11.05 [-22.34, 0.25]
2.34.4 Anaemia between 20-39.9%	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.34.5 Anaemia 40% or higher	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
2.35 Ferritin (µg/L) (subgroup by funding source)	6	796	Mean Difference (IV, Random, 95% CI)	-9.87 [-15.02, -4.72]
2.35.1 Private industry/brand	5	763	Mean Difference (IV, Random, 95% CI)	-8.47 [-14.12, -2.82]
2.35.2 Other	1	33	Mean Difference (IV, Random, 95% CI)	-16.60 [-20.01, -13.19]
2.36 Gut health-Stool frequency (per day)	1	306	Risk Ratio (M-H, Random, 95% CI)	2.00 [0.18, 21.83]
2.37 Gut health-Stool consistency (soft-formed)	1	306	Risk Ratio (M-H, Random, 95% CI)	1.22 [0.98, 1.52]

3 Full-fat animal milk versus lower-fat milk

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
3.8 Nutritional status-serum cholesterol (mmol/l)	1	17	Mean Difference (IV, Random, 95% CI)	-0.17 [-0.92, 0.58]
3.9 Nutritional status-serum low density lipoprotein (LDL) (mmol/l)	1	17	Mean Difference (IV, Random, 95% CI)	-0.25 [-0.94, 0.44]

3.10 Nutritional status-serum high density lipoprotein (HDL) (mmol/l)	1	17	Mean Difference (IV, Random, 95% CI)	-0.10 [-0.30, 0.10]
3.11 Nutritional status-serum tryglicerides (mmol/l)	1	17	Mean Difference (IV, Random, 95% CI)	0.34 [-0.12, 0.80]
3.12 Nutritional status-serum LDL/HDL	1	17	Mean Difference (IV, Random, 95% CI)	-0.05 [-0.83, 0.73]

4 Animal milk versus plant-based milk alternatives

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
4.7 Nutritional status-serum cholesterol (mmol/l)	1	21	Mean Difference (IV, Random, 95% CI)	-0.16 [-0.76, 0.44]
4.8 Nutritional status-serum low density lipoprotein (LDL) (mmol/l)	1	21	Mean Difference (IV, Random, 95% CI)	0.03 [-0.48, 0.54]
4.9 Nutritional status-serum high density lipoprotein (HDL) (mmol/l)	1	21	Mean Difference (IV, Random, 95% CI)	-0.18 [-0.85, 0.49]
4.10 Nutritional status-serum triglycerides (mmol/l)	1	21	Mean Difference (IV, Random, 95% CI)	-0.08 [-0.63, 0.47]
4.11 Nutritional status-serum LDL/HDL	1	21	Mean Difference (IV, Random, 95% CI)	0.33 [-0.36, 1.02]

5 Animal milk (full-fat or lower-fat) versus fortified milk (full-fat or lower-fat)

Outcome or Subgroup	Studies	Participants	Statistical Method	Effect Estimate
5.1 Weight (kg)	1	36	Mean Difference (IV, Random, 95% CI)	0.04 [-0.83, 0.91]
5.2 Undernutrition-Stunting	1	378	Risk Ratio (M-H, Random, 95% CI)	0.98 [0.74, 1.28]
5.3 Undernutrition-Wasting	1	378	Risk Ratio (M-H, Random, 95% CI)	1.06 [0.78, 1.44]

5.4 Undernutrition-Stunting and wasting	1	378	Risk Ratio (M-H, Random, 95% CI)	1.14 [0.86, 1.50]
5.5 Nutritional status-Iron as serum iron (µmol/l)	2	115	Mean Difference (IV, Random, 95% CI)	-0.46 [-4.38, 3.46]
5.6 Nutritional status-Zinc as plasma zinc (µmol/L)	2	493	Mean Difference (IV, Random, 95% CI)	0.43 [0.11, 0.76]
5.7 Anaemia	3	1324	Risk Ratio (M-H, Random, 95% CI)	2.29 [1.12, 4.69]
5.8 Iron deficiency anaemia (IDA)	1	465	Risk Ratio (M-H, Random, 95% CI)	4.15 [2.93, 5.87]
5.9 Iron deficiency (ID, serum ferritin <12 µg/l)	1	349	Risk Ratio (M-H, Random, 95% CI)	1.21 [0.57, 2.56]
5.10 Haemoglobin (g/L) (All)	6	1354	Mean Difference (IV, Random, 95% CI)	-5.91 [-9.84, -1.99]
5.11 Haemoglobin (g/L) (subgroup by stunting)	6	1354	Mean Difference (IV, Random, 95% CI)	-5.91 [-9.84, -1.99]
5.11.1 Stunting non-specified	2	158	Mean Difference (IV, Random, 95% CI)	-0.75 [-3.82, 2.32]
5.11.2 Stunting very low prevalence (<2.5%)	2	616	Mean Difference (IV, Random, 95% CI)	-8.08 [-12.00, -4.16]
5.11.3 Stunting low prevalence (2.5-9.9%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
5.11.4 Stunting medium prevalence (10-19.9%)	1	115	Mean Difference (IV, Random, 95% CI)	-3.30 [-3.50, -3.10]
5.11.5 Stunting high prevalence (>20%)	1	465	Mean Difference (IV, Random, 95% CI)	-13.60 [-16.06, -11.14]
5.12 Haemoglobin (g/L) (subgroup by wasting)	6	1354	Mean Difference (IV, Random, 95% CI)	-5.91 [-9.84, -1.99]
5.12.1 Wasting non-specified	2	158	Mean Difference (IV, Random, 95% CI)	-0.75 [-3.82, 2.32]
5.12.2 Wasting very low prevalence (<2.5%)	3	731	Mean Difference (IV, Random, 95% CI)	-6.36 [-10.80, -1.92]
5.12.3 Wasting low prevalence (2.5-4.9%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable

5.12.4 Wasting medium prevalence (5-10%)	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
5.12.5 Wasting high prevalence (>10%)	1	465	Mean Difference (IV, Random, 95% CI)	-13.60 [-16.06, -11.14]
5.13 Haemoglobin (g/L) (by type of feeding before 12 mo)	6	1354	Mean Difference (IV, Random, 95% CI)	-5.91 [-9.84, -1.99]
5.13.1 Non specified	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
5.13.2 Breastmilk only	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
5.13.3 Breastfeeding + complementary feeding	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
5.13.4 Breastmilk substitute + complementary feeding	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
5.13.5 Breastfeeding + breast milk substitute + complementary feeding	6	1354	Mean Difference (IV, Random, 95% CI)	-5.91 [-9.84, -1.99]
5.14 Haemoglobin (g/L) (subgroup by anaemia)	6	1354	Mean Difference (IV, Random, 95% CI)	-5.91 [-9.84, -1.99]
5.14.1 Anaemia non specified	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
5.14.2 Anaemia between 0-4.9%	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
5.14.3 Anaemia between 5-19.9%	3	623	Mean Difference (IV, Random, 95% CI)	-5.11 [-14.71, 4.49]
5.14.4 Anaemia between 20-39.9%	3	731	Mean Difference (IV, Random, 95% CI)	-6.36 [-10.80, -1.92]
5.14.5 Anaemia 40% or higher	0	0	Mean Difference (IV, Random, 95% CI)	Not estimable
5.15 Haemoglobin (g/L) (by funding source)	6	1354	Mean Difference (IV, Random, 95% CI)	-5.91 [-9.84, -1.99]
5.15.1 Private industry/brand	2	587	Mean Difference (IV, Random, 95% CI)	-7.56 [-19.61, 4.49]
5.15.2 Other	4	767	Mean Difference (IV, Random, 95% CI)	-5.09 [-8.91, -1.28]

5.16 Ferritin (µg/L)	3	852	Mean Difference (IV, Random, 95% CI)	-5.70 [-7.49, -3.92]
5.17 Oral health-mean dmfs (mean decayed, missing, and filled surfaces index in molars and canines)	1	186	Mean Difference (IV, Random, 95% CI)	1.30 [0.37, 2.23]
5.18 Oral health -caries free (dmfs in molars and canines= 0)	1	186	Risk Ratio (M-H, Random, 95% CI)	0.73 [0.59, 0.91]
5.19 Morbidity-Respiratory episodes per child per year	1	554	Mean Difference (IV, Random, 95% CI)	0.03 [-0.14, 0.20]
5.20 Morbidity-Diarrhea episodes per child per year	2	1187	Mean Difference (IV, Random, 95% CI)	0.80 [0.27, 1.33]