Conduct and reporting of formula milk trials: systematic review

Bartosz Helfer,1,2 Jo Leonardi-Bee,3 Alexandra Mundell,4 Callum Parr,1 Despo Ierodiakonou,5 Vanessa Garcia-Larsen,1,6 Cynthia M Kroeger,7 Zhaoli Dai,7,8 Amy Man,1 Jessica Jobson,1 Fatemah Dewji,1 Michelle Kunc,1 Lisa Bero,1 Robert J Boyle1,10

Abstract

Objective
To systematically review the conduct and reporting of formula trials.

Design
Systematic review.

Data Sources
Medline, Embase, and Cochrane Central Register of Controlled Trials (CENTRAL) were searched from 1 January 2006 to 31 December 2020.

Review Methods
Intervention trials comparing at least two formula products in children less than three years of age were included, but not trials of human breast milk or fortifiers of breast milk. Data were extracted in duplicate and primary outcome data were synthesised for meta-analysis with a random effects model weighted by the inverse variance method. Risk of bias was evaluated with Cochrane risk of bias version 2.0, and risk of undermining breastfeeding was evaluated according to published consensus guidance. Primary outcomes of the trials included in the systematic review were identified from clinical trial registries, protocols, or trial publications.

Results
22,201 titles were screened and 307 trials were identified that were published between 2006 and 2020, of which 73 (24%) trials in 13,197 children were prospectively registered. Another 111 unpublished but registered trials in 17,411 children were identified. Detailed analysis was undertaken for 125 trials. For 68 recently published superiority trials, a pooled standardised mean difference of 0.51 (range −0.43 to 3.29) was calculated with an asymmetrical funnel plot (Egger’s test P=0.001), which reduced to 0.19 after correction for asymmetry. Primary outcomes were reported by authors as favourable in 86 (69%) trials, and 115 (92%) abstract conclusions were favourable. One of 38 (3%) trials in partially breastfed infants reported adequate support for breastfeeding and 14 of 87 (16%) trials in non-breastfed infants confirmed the decision not to breastfeed was firmly established before enrolment in the trial.

Conclusions
The results show that formula trials lack independence or transparency, and published outcomes are biased by selective reporting.

Systematic Review Registration
PROSPERO 2018 CRD42018091928.

Introduction
Breast milk substitutes, commonly termed formula milks, are consumed by most infants in North America and Europe, and global sales rose from 3.5 kg per child in 2005 to 7.4 kg per child in 2019.1 Formula milk is very rarely medically indicated, but the social changes of industrialisation and aggressive marketing of formula products have contributed to the cultural normalisation of formula feeding in many regions.2 Use of formula in place of breastfeeding is associated with short and long term health risks, from burns, colic, and constipation to pneumonia, developmental effects, and fatal diarrhoea.3 Formula is sometimes a young infant’s only source of nutrition, and can be consumed at very high levels, up to 20% of an infant’s body weight per day, during a sensitive period of development.2 Clinical trials are therefore required by regulators to show the safety of new formula products, and trials are also used by manufacturers to support marketing claims.5 The widespread use of formula products, their associated health risks, and the high individual levels of intake during early development highlight the importance of the evidence generated from formula milk trials. Associations for the nutrition industry cite their investment in world class research and development, but several groups from academia and regulatory science have raised concerns about the conduct and reporting of formula milk trials.5 6 8 Some have suggested that trial procedures might contravene the International Code of Marketing of Breast-milk Substitutes, for example, by providing free formula to participants.7 8 We therefore conducted a systematic
review to evaluate the conduct and reporting of formula milk trials. We were specifically interested in understanding the risk of bias in published formula trials and if trial procedures could cause harm by undermining breastfeeding of participants.

**Methods**

This systematic review is reported according to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA).9 We searched Medline, Embase, and Cochrane Central Register of Controlled Trials (CENTRAL) for clinical trials comparing two or more different formula products in children aged <3 years, published between 1 January 2015 and 31 December 2020 (the supplementary material describes the search strategy). We also identified all trials registered on the International Clinical Trials Registry Platform from when registration of clinical trials was made mandatory, and the status of trial registration of all trials published since 1 January 2006, to evaluate the risk of selective reporting at the trial level.10 Prospective trial registration was defined as registration of the trial in a clinical trial registry approved by the World Health Organization before, or up to one month after, the reported start of the study. We excluded trials of supplements given separately to formula, breast milk or fortifiers of breast milk, and trials comparing the frequency or volume of consumption of formula without changing the ingredients or formulation. The primary outcome of the systematic review was overall risk of bias for the trial primary outcome, evaluated with Cochrane risk of bias version 2.0 (ROB2).11 Secondary outcomes were risk of bias in individual domains of ROB2, how favourable the primary outcomes and conclusions of the trial were, measures of selective reporting at the trial level, and risk of trial procedures undermining breastfeeding. Selection of studies, extraction of data, and coding of outcomes were extensively piloted and carried out by at least two researchers. Disagreements were resolved by rechecking the original articles and discussion leading to a consensus decision.

To identify the primary outcome of each trial, we used the first primary outcome listed in the record of a clinical trial registry approved by the World Health Organization or, if not registered, we used the methods section or first reported result in the main trial manuscript. Risk of undermining breastfeeding and clarity of describing the trial interventions were evaluated with criteria from recent Delphi consensus guidance on the conduct and reporting of formula trials.12 The extent to which the outcomes and conclusions were favourable was evaluated as previously described.13 For statistical analysis of the primary outcomes of the trial, effects were estimated by calculating standardised mean difference for continuous outcomes and odds ratios for dichotomous outcomes; odds ratios were then transformed into standardised mean difference, as previously described.14 Unreported means and standard deviations were computed from available raw data or estimated from interquartile range and median values.15

Publication bias was evaluated with contour enhanced funnel plots and the trim-and-fill procedure, and by summarising unpublished formula trials and abstract publications.16 We recorded the source of funding and conflicts of interest of the authors, as declared in published reports, and evaluated conflicts of interest as low where the trial funder did not have a commercial interest in the outcome of the trial and no study authors had a financial link to an organisation with a commercial interest in the outcome of the trial.

Statistical analyses with Review Manager 5.3 and R 4.0.2 were conducted separately for superiority trials and equivalence or non-inferiority trials. Superiority trials are randomised controlled trials designed to show that the intervention is better than the comparator, non-inferiority trials are designed to show that the intervention is no worse than the comparator, and equivalence trials are designed to show that no meaningful difference exists between the intervention and the comparator. Random effects meta-regression was used to evaluate associations between standardised mean difference and features of the trial. High heterogeneity was expected because of the diverse features of the trials, and therefore we used the Hartung-Knapp adjustment for random effects model weighted by the inverse variance method. Heterogeneity was assessed by $\chi^2$ with the Sidik-Jonkman estimator. The protocol of this systematic review was registered at PROSPERO after piloting the processes for selection of studies, extraction of data, and assessment of risk of bias but before piloting of the analysis of the data.17

**Patient and public involvement**

Patients or the public were not directly involved in this study, but contributed to development of the Delphi consensus guidance which we used to evaluate the risk of undermining breastfeeding and reporting the use of other nutritional intake in formula trials.12

**Results**

Our search results are summarised in the PRISMA flowchart (supplementary fig 1). After screening the title, abstract, and full text, we found 307 eligible main trial publications since 2006, the first calendar year after registration of clinical trials was made mandatory.10 We summarised the status of trial registration for all 307 trials published since 1 January 2006, and undertook a more detailed evaluation of 125 trials published since 1 January 2015. Of the 307 main trial publications identified since 2006, 149 (49%) had a record of trial registration, of which 73 (24%) were registered prospectively, evaluating 13 197 children (supplementary table 1). We identified another 111 unpublished formula trials in 17 411...
Characteristics of recently published formula milk trials
Between 2015 and 2020, we identified 125 published trials evaluating 23,757 randomised children (supplementary table 3). Six more trials were published as abstracts only and were not analysed further (supplementary table 4). The recently published full text formula trials were mainly conducted in infants (90%, supplementary table 3), and were mainly carried out in Europe (42%), Asia (28%), or North America (18%). Trials were small (median sample size 114, interquartile range 60–204) and usually had a superiority aim (76% of trials where the aim was clear). Formula interventions studied were commonly prebiotics or probiotics, alone or combined with other ingredients (42%), changes to protein source or content (29%), or fat content (12%). The most common primary outcome was weight gain (36%), measured in half of these trials according to regulatory guidance for establishing adequate growth in infants exclusively fed formula milk. Other common primary outcomes were intestinal health (26%), including faecal analysis and gastrointestinal symptoms, measures of nutrient absorption (10%), behaviour (6%), and allergy (6%).

Transparency of aims in recently published formula trials
Between 2015 and 2020, 57 (46%) of 125 recently published trials had a prospective registration record and two trials had a prospectively publicly available protocol (table 1 and supplementary table 5). Another 33 (26%) trials had a retrospective trial registration record, dated a median of 24 (interquartile range 7–48) months after the start of the study. A clear study aim was described in 79 (63%) trials, and this proportion was similar for registered and unregistered trials. A clear primary outcome was described in 59 (47%) trials and was more common in prospectively registered trials (60%) than unregistered trials (30%). Protocols dated before the start of the study were only available for 11 (9%) trials, of which eight had a clear study aim and clear primary outcome (table 1).

Evidence of publication and reporting bias in recently published formula trials
In formula trials published between 2015 and 2020, analysis of the primary outcome was favourable in 53 (42%) trials, the author’s interpretation of the outcome was favourable in 86 (69%) trials, and the abstract conclusions were favourable in 115 (92%) trials (supplementary table 6). For 68 published superiority trials, the standardised mean difference favoured the intervention formula in 57 (84%) trials, the control formula in eight (12%), and neither product in three (4%) trials. For superiority trials, we calculated the pooled standardised mean difference (0.51, 95% confidence interval 0.34 to 0.68, I²=84%), range −0.43 to 3.29; fig 1) with an asymmetrical funnel plot (fig 2 (top); Egger’s test P<0.001) and found a reduced effect size after correction for asymmetry (fig 2 (bottom); adjusted standardised mean difference 0.19, −0.03 to 0.41). For 34 published non-inferiority or equivalence trials, we calculated a pooled standardised mean difference of −0.06 (−0.15 to 0.02, I²=44%, range −1.04 to 0.42; supplementary fig 2). For 23 published trials designed according to regulatory guidance for demonstrating adequate growth, the authors interpreted the primary outcome as favourable in 22 (96%) trials. Only 12 of 23 (52%) trials showed equivalence with a two sided 95% confidence interval (or 15 (65%) trials with a 90% confidence interval), using the American Academy of Pediatrics recommended equivalence margin of a difference in weight gain of 3 g/day (fig 3).

Cochrane risk of bias assessment of recently published formula trials
For trials published between 2015 and 2020, risk of bias was low in five (4%) and high in 100 (80%) trials (table 2, supplementary table 7). In domain 1, the randomisation process, four (3%) trials were at high risk of bias because of inadequate concealment of allocation. In domain 2, deviations from the intended interventions, 52 (42%) trials were at high risk: 32 trials that evaluated adherence, mainly because the trials used a simple per protocol analysis which excluded ≥10% of participants (median 30%, interquartile range 15–45%) without accounting for missing data in the analysis; and 20 trials that evaluated assignment, mainly because of inappropriate exclusion of ≥10% of participants from the analysis (23%, 18–34%) for reasons such as adverse events or non-adherence. In domain 3, missing outcome data, 18 (14%) trials were at high risk of bias because of missing outcome data (35%, 23–42%) with no evidence that the estimate was stable with methods such as multiple imputation, and a high likelihood that missingness in the outcome data depended on their true value. In domain 4, outcome measurement, five (4%) trials were at high risk of bias because assessors of outcome were not blinded and the outcome was subjective. In domain 5, selective reporting, 16 (13%) trials were at high risk of bias, most commonly because of analysis of an inappropriate dataset for the stated aim, for example analysis of an intention to treat dataset for an equivalence aim without presenting the per protocol results. In 83 (66%) trials, there were some concerns about selective reporting because prospective trial registration was absent, multiple primary outcomes or evaluation methods were registered, methods for measurement or analysis of primary outcomes were poorly specified, or reporting was incomplete.

Conflict of interest in recently published formula trials
Eight (6%) trials published between 2015 and 2020 did not report a funding source. Of 117 trials that reported sources of funding, 98 (84%) received...
support from the formula milk industry, although in seven trials funding was limited to donating the study formula (table 3; supplementary table 8). Seventy five of 98 (77%) trials supported by the formula industry had at least one author affiliated to a formula company, and the formula industry was involved in the statistical analysis or writing of the trial report, or both, in 60 of 81 (76%) trials where their role was clearly described. Eighty seven of 112 (78%) trials that reported conflicts of interest of authors had at least one author who had a conflict of interest related to the formula industry. Overall, 17 (14%) published trials had a low level of conflict of interest related to the formula industry.

Support for breastfeeding evaluated with international guidance for formula trials

We evaluated reported support for breastfeeding in formula trials published between 2015 and 2020 according to criteria derived from consensus guidance for the conduct and reporting of formula trials. Eighty four (67%) trials clearly described if participants were receiving breast milk at enrolment (supplementary table 9). In 33 of 38 (87%) trials where some infants were likely to be receiving breast milk at enrolment, the International Code of Marketing of Breast-milk Substitutes seemed to be contravened by providing free formula to participants in the trial, which would not have been provided outside of the trial (supplementary table 9). In five of these, participants were required to consume a minimum or fixed daily volume of formula, which might have undermined breastfeeding.

In 73 of 87 (84%) trials in infants fully formula fed, whether the decision not to breastfeed was firmly established before enrolment was unclear, and 18 of these trials enrolled infants in the first week of life. Overall, in our evaluation, we found that 15 (12%) trials were at low risk of undermining breastfeeding, 45 (36%) trials were at high risk, and 65 (52%) had some concerns (table 4). Trials also failed to adhere to other recommendations for reporting nutritional intakes in formula trials. Insufficient information was available to link the products used in the trials with marketed formula products in 101 (81%) recent trials (supplementary table 10), and 53 (42%) trials did not provide sufficient information to confirm that the composition of the control formula conformed to international or local standards. In 93 (74%) trials, no information was reported about length of breastfeeding in the study participants, before or during the trial, and in 100 of 115 (87%) trials where non-milk feeds could be relevant, we found no information about timing of introduction of complementary food or intake of other relevant nutritional products.

Trial features associated with risk of bias and favourability of outcomes

Supplementary tables 11-16 and supplementary figs 3-6 show the relations between the characteristics of the trials and risk of bias, outcomes of the trials, or support for breastfeeding. All trials with a low risk of bias had a superiority aim, and a low risk of bias was associated with low conflicts of interest (odds ratios 0.03; 95% confidence interval 0.01 to 0.07). Consequences of findings lay outside the scope of this study.

Table 1 | Transparency of aims in formula trials published between 2015 and 2020

<table>
<thead>
<tr>
<th>Issues with definition of study aim:</th>
<th>Trials</th>
<th>Prospectively registered</th>
<th>Non-prospectively registered</th>
<th>Unregistered</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unclear which arm is control group, or whether aim is superiority, equivalence, or non-inferiority</td>
<td>12/57 (21)</td>
<td>6/35 (17)</td>
<td>8/33 (24)</td>
<td>26/125 (21)</td>
<td></td>
</tr>
<tr>
<td>Dataset inappropriate for stated aim of trial</td>
<td>9/57 (16)</td>
<td>6/35 (17)</td>
<td>5/33 (15)</td>
<td>20/125 (16)</td>
<td></td>
</tr>
<tr>
<td>Clear study aim</td>
<td>36/57 (63)</td>
<td>23/35 (66)</td>
<td>20/33 (61)</td>
<td>79/125 (63)</td>
<td></td>
</tr>
<tr>
<td>Issues with description of primary outcome:</td>
<td>12/57 (21)</td>
<td>7/35 (20)</td>
<td>16/33 (48)</td>
<td>35/125 (28)</td>
<td></td>
</tr>
<tr>
<td>Multiple primary outcomes or unclear description of primary outcome</td>
<td>11/57 (19)</td>
<td>13/35 (37)</td>
<td>7/33 (21)</td>
<td>31/125 (25)</td>
<td></td>
</tr>
<tr>
<td>Clear primary outcome</td>
<td>34/57 (60)</td>
<td>15/35 (43)</td>
<td>10/33 (30)</td>
<td>59/125 (47)</td>
<td></td>
</tr>
<tr>
<td>Both study aim and primary outcome clearly described</td>
<td>26/57 (46)</td>
<td>10/35 (29)</td>
<td>7/33 (21)</td>
<td>43/125 (34)</td>
<td></td>
</tr>
</tbody>
</table>

| Protocol availability: | Prospectively protocol publicly available, or provided by author or funder on request | 9/57 (16) | 2/35 (6) | 0/33 (0) | 11/125 (9) |
| Retrospective or undated protocol publicly available, or provided by author or funder on request | 10/57 (18) | 9/35 (26) | 0/33 (0) | 19/125 (15) |
| Author or funder declined to provide protocol | 20/57 (35) | 5/35 (14) | 2/33 (6) | 27/125 (22) |
| No response from author or funder and no publicly available protocol | 5/57 (9) | 8/35 (23) | 6/33 (18) | 19/125 (15) |
| Protocol availability: | 13/57 (23) | 11/35 (31) | 25/33 (76) | 49/125 (39) |
| Prospective protocol and clear study aim and clear primary outcome | 8/57 (14) | 0/35 (0) | 0/33 (0) | 8/125 (6) |

Values are No of trials/total No of trials (%).
**Fig 1** | Primary outcomes for formula trials with a superiority aim published between 2015 and 2020. Standardised mean difference and 95% confidence intervals for the primary outcome of the 68 formula trials designed with the aim of showing that the intervention was better than the comparator, where standardised mean difference could be calculated.
Adjusted standardised mean difference is 0.19 (−0.03 to 0.41). Tweedie’s trim-and-fill procedure until funnel plot symmetry is reached (bottom panel). $P<0.001$. White dots represent standardised mean difference imputed with Duval and mean difference is 0.51 (0.34 to 0.68), Egger’s test for funnel plot asymmetry=4.111, indicate regions of statistical significance (top panel). Unadjusted pooled standardised dot for each trial primary outcome, against the corresponding standard errors. Colours 2020. Effect size, expressed as standardised mean difference, is shown as a purple emphasises the importance of scientific rigour in this in children at a sensitive period of development. In this systematic evaluation of formula milk trials, we found an almost universal lack of transparency, and selective reporting delay. We did not find any relation between the characteristics of the trials and support for breastfeeding, or any changes in risk of bias, favourable conclusions, or support for breastfeeding over time.

**Discussion**

**Principal findings**

In this systematic evaluation of formula milk trials, we found an almost universal lack of transparency, and evidence of selective reporting between and within trials. The widespread and increasing use of formula in children at a sensitive period of development emphasises the importance of scientific rigour in this area of clinical investigation. Our findings suggest that recent formula trials lack that scientific rigour, and published outcomes are biased by selective reporting.

We found more registered unpublished trials than prospectively registered published trials. Few of the recent published trials had a clear a priori registered aim and primary outcome, or a publicly available protocol. Most formula trials had a high risk of bias for their primary outcome, usually because of inappropriate exclusions of participants from the analysis, and selective reporting. This lack of transparency was complemented by favourable conclusions in more than 90% of recent trials, and evidence of publication bias in recent superiority trials where a favourable standardised mean difference for 84% of primary outcomes was found. The favourable outcomes in equivalence and non-inferiority trials were also affected by selective reporting; almost half of the recent trials designed to show adequate weight gain based on regulatory guidance failed to show equivalence, but authors’ own interpretation of the data supported adequate weight gain in 95% of trials. In most recent formula trials, investigators were employed by, or had financial links to, the formula industry, who were often involved in the statistical analysis and writing. Conflicts of interest could explain some of the transparency issues identified in formula trials because work in other disciplines has shown a relation between commercial funding for clinical trials and favourable results.24

Low rates of breastfeeding are an important global public health issue, formula marketing can adversely affect breastfeeding rates, and the International Code of Marketing of Breast-milk Substitutes is a key international agreement used by policy makers to protect breastfeeding by limiting the marketing of formula.25 In recent formula trials, we found evidence that trial procedures might contravene the code, for example by providing free formula to parents of breastfed or mixed fed infants. Only 12% of trials reported adequate measures to ensure breastfeeding was not undermined in participants of the trial. The code states that “health workers should not give samples of infant formula to pregnant women, mothers of infants and young children, or members of their families”. Therefore, oversight of formula trials might be inadequate, and there might be a failure to adhere to the Declaration of Helsinki by protecting research participants.8 12 26 In a setting where trialists and regulators might require high consumption of formula to better answer scientific uncertainties, there is a risk of creating an incentive structure within a trial that promotes the use of formula in place of breast milk, with substantial health risks for participants in the trial.

**Policy implications**

Since the invention of industrially produced formula milk in the 19th century, ongoing scientific studies have been conducted aimed at improving formula products, for example by adding vitamins and modifying levels of protein. Clinical trials are potentially important in the development of improved formula products with reduced risks for health and nutrition. In this study,
we found the trials to be unreliable, suggesting that much of the recent information generated about formula products might be misleading.27 Given the lack of transparency about the aims of the trials and almost universally favourable conclusions, some trials might have a marketing aim and no robust scientific aim.27 Marketing of formula is cited by the World Health Organization as a major barrier to promoting and protecting breastfeeding. Claims arising from formula trials can contribute to formula marketing by narrowing the perceived benefits of breast milk over formula for consumers. Some have called for a change in the regulatory environment for formula products, including a ban on health and nutrition claims.3 28 Our findings suggest that such changes should include improved oversight, conduct, and reporting of formula trials to ensure they provide a rigorous evidence base to inform nutrition in infants and young children.

Limitations
Our study was a comprehensive evaluation of the conduct and reporting of formula trials. Because of lack of resources, our detailed evaluation was largely restricted to the analysis of the primary outcome for the 23 trials that were designed according to guidance of Food and Drug Administration for demonstrating adequate growth. Equivalence margins of 3 g/day weight gain are indicated by dotted lines. Authors interpreted the primary outcome as favourable in 22 (96%) trials.

Fig 3 | Primary outcomes for formula trials designed to show adequate infant growth, published between 2015 and 2020. Data are mean differences (95% confidence intervals) and mean differences (90% confidence intervals) in weight gain between intervention and control formula during the intervention period, for the 23 trials that were designed according to guidance of Food and Drug Administration for demonstrating adequate growth. Equivalence margins of 3 g/day weight gain are indicated by dotted lines. Authors interpreted the primary outcome as favourable in 22 (96%) trials.

Table 2 | Risk of bias in formula trials published between 2015 and 2020

<table>
<thead>
<tr>
<th>Risk of bias</th>
<th>Randomisation process</th>
<th>Deviation from interventions</th>
<th>Missing outcome data</th>
<th>Outcome measurement</th>
<th>Selective reporting</th>
<th>Overall risk of bias</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>103/125 (82)</td>
<td>36/125 (29)</td>
<td>67/125 (54)</td>
<td>107/125 (86)</td>
<td>26/125 (21)</td>
<td>5/125 (6)</td>
</tr>
<tr>
<td>Some concerns</td>
<td>18/125 (14)</td>
<td>37/125 (30)</td>
<td>40/125 (32)</td>
<td>13/125 (10)</td>
<td>83/125 (66)</td>
<td>20/125 (16)</td>
</tr>
<tr>
<td>High</td>
<td>4/125 (3)</td>
<td>52/125 (42)</td>
<td>18/125 (14)</td>
<td>5/125 (4)</td>
<td>16/125 (13)</td>
<td>100/125 (80)</td>
</tr>
</tbody>
</table>

Values are No of trials/total No of trials (%). Risk of bias was evaluated with Cochrane risk of bias version 2.0 (ROB2).9
Table 3 | Conflict of interest in formula trials published between 2015 and 2020

<table>
<thead>
<tr>
<th>Risk</th>
<th>Conflict of interest, formula industry or analysis or funding</th>
<th>Independent funding (n=19)*</th>
<th>Formula donated by formula industry but no other industry support reported (n=7)</th>
<th>Formula industry funding reported (n=91)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>At least one author reported a conflict of interest related to formula industry</td>
<td>2/16 (13)</td>
<td>4/7 (57)</td>
<td>81/89 (91)</td>
<td>87/112 (78)</td>
<td></td>
</tr>
<tr>
<td>At least one author affiliated to formula industry</td>
<td>0/19</td>
<td>2/7 (29)</td>
<td>73/91 (80)</td>
<td>75/117 (64)</td>
<td></td>
</tr>
<tr>
<td>Formula industry sponsor involved in statistical analysis or trial reporting</td>
<td>–</td>
<td>1/6 (17)</td>
<td>59/75 (79)</td>
<td>60/81 (74)</td>
<td></td>
</tr>
</tbody>
</table>

Values are No of trials/total No of trials (%). *Independent funding means a funding source not related to the formula industry (n=17) or authors reported that the study did not have any formal funding (n=2). Denominators vary because of lack of reporting for some variables. Source of funding was not reported for eight studies, author conflict of interest was not reported for nine studies, and role of formula sponsor in statistical analysis or trial reporting was not reported for 17 studies.

conducted, and the results of many formula trials are not in the public domain. Some groups have found that the Cochrane ROB2 tool is difficult to use and have criticised its reproducibility.29 We found that ROB2 had good reproducibility within our team after initial training and piloting, with κ scores of 0.61-0.87 for blinded assessments of ROB2 domains.

Our assessments of support for breastfeeding were made with criteria developed by an international group of experts in formula trials and their regulation.12 Analysis of pooled standardised mean difference from trials testing varied interventions in diverse populations for heterogeneous outcomes must be interpreted with caution. We are confident that findings were, in general, reported as favourable; but the relations identified between characteristics of the study and standardised mean difference should be regarded as exploratory because of the heterogeneous nature of the data contributing to the analysis of standardised mean difference and small sample sizes of some subgroups.

We did not include formal patient and public involvement in the conduct or reporting of this project but the Delphi consensus guidance which informed the design of this project was undertaken with the involvement of patients and the public.30 We did not formally compare the risk of bias in this area with clinical trials in other areas, but the effect sizes seen in this analysis suggest that the outcomes are at least as favourable as trials of pharmaceuticals licensed for treating general medical or psychiatric conditions.30 Our analyses of primary outcomes from published clinical trials are likely to represent conservative estimates of effect size (standardised mean difference) and selective reporting bias, compared with the outcome analyses selected by authors, because interpretation of primary outcomes by authors was more favourable than our own analysis of the raw data. Our analyses of primary outcomes are also likely to underestimate reporting biases compared with exploratory outcomes or post hoc subgroup analyses, which are often used by the formula industry as the basis of health and nutrition claims.3 31

Conclusion

Our study suggests that formula trials are not reliable and might not adequately protect participants in the trials. The formula industry is closely involved in formula trials, findings are almost always reported as favourable, and little transparency exists about the aims of the trial or reporting of results. Our findings support the need for a substantial change in the conduct and reporting of formula trials to adequately protect participants from harm and protect consumers from misleading information.

We are grateful to Ewa Szymlek-Gay, Huiven Zhang, Dirk Haller, and Johannes van Goudoever for sharing prospective protocols for their trials, to Manjiraj Rao from Nestlé Clinical Research for sharing four retrospective trial protocols and one prospective trial protocol, to Hania Szajewska and to Barbara Marriage from Abbott Laboratories for sharing retrospectively dated protocols, to Mark Corkins and Vanessa Scheefler for sharing statistical analysis plans, and to Antonia Monayo, Marc Rhoads, Rosella Turco, Tetsuhiro Sakihara, Flavia Indrio, Olle Hernel, Roberto Canani, and Hanna Petersen for sharing undated trial protocols. A summary of an earlier version of information reported in this manuscript was presented by BH at the Star Child Health Forum on 4 February 2021 https://www.starchildhealth.org/past-events.

Contributors: RJB, JL-B, DI, and BH designed the study, with advice and input from LB, VGL, CMK, and ZD. BH, RJB, AM, MK, FD, JJ, and CP developed and piloted the study search, data extraction, and coding tools. BH conducted the search. BH, RJB, AM, MK, FD, and JJ screened titles, extracted data, and coded outcomes. BH, JL-B, and DI undertook the statistical analyses. RJB wrote the first draft of the manuscript. All authors critically reviewed and revised the manuscript and approved the final manuscript. RJB acts as guarantor of the manuscript. AM and

Table 4 | Risk that breastfeeding was undermined in formula trials published between 2015 and 2020

<table>
<thead>
<tr>
<th>Risk</th>
<th>Receiving breast milk at enrolment</th>
<th>Most common reason</th>
<th>No breast milk at enrolment</th>
<th>Most common reason</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Low</td>
<td>1/38 (3)</td>
<td>Good breastfeeding support. Free formula during inpatient stay but likely consistent with local practice</td>
<td>14/87 (16)</td>
<td>No breastfeeding for a substantial length of time before enrolment*</td>
<td>15/125 (12)</td>
</tr>
<tr>
<td>Some concerns</td>
<td>11/38 (29)</td>
<td>Free formula which was not available outside of the trial context</td>
<td>55/87 (63)</td>
<td>No information to judge if the decision not to breastfeed was firmly established. Enrolment occurred after the first week of life</td>
<td>65/125 (52)</td>
</tr>
<tr>
<td>High</td>
<td>26/38 (68)</td>
<td>Free formula and no statement that breastfeeding was support</td>
<td>18/87 (21)</td>
<td>No information to judge if the decision not to breastfeed was firmly established. Enrolment occurred in the first week of life</td>
<td>45/125 (36)</td>
</tr>
</tbody>
</table>

Values are No of trials/total No of trials (%). *Three trials enrolled infants fed only formula or not breastfed for ≥3 days before enrolment; one trial ≥5 days, and eight trials between ≥2 weeks and ≥2 months. We judged risk of undermining breastfeeding as low in trials where children did not receive breast milk at enrolment, if the publication made it clear the decision not to breastfeed was firmly established before enrolment, for example an eligibility criterion of no breast milk for ≥3 days.16 For trials with some breastfeeding at enrolment, we judged risk of undermining breastfeeding as low if all Delphi consensus criteria were adhered to, and high if no criterion was adhered to.10
Not commissioned, externally peer reviewed. Provenance and peer review: All authors have completed the ICMJE uniform disclosure form for www.icmje.org/disclosure-of-interests/ and declare support from Imperial Health Charity. RJ B received personal fees from Cochrane, DBV Technologies, and Prota Therapeutics, and from expert witness work in cases of food anaphylaxis and class actions related to infant formula health claims, outside the submitted work, and received personal fees from Public Health England as a member of the UK Nutrition and Health Claims Committee and the Maternal and Child Nutrition Subgroup of the Scientific Advisory Committee on Nutrition. JL B received fees from Danone Nutricia Research and the Food Standards Agency, outside of the submitted work. The University of Colorado receives remuneration for LB’s work as senior editor, Cochrane, which is outside the submitted work. The authors report no other relationships or activities that could appear to have influenced the submitted work.

Ethical approval: Ethics approval and informed consent were not required for this study. All information used for the analyses was in the public domain, with the exception of the clinical trial protocols and statistical analysis plans which were kindly provided by some trialists and trial sponsors at our request, to confirm methods used in the relevant trials.

Data sharing: All data collected for this study, including detailed responses to ROB2, breastfeeding, and conflict of interest signalling questions, are in the data supplement (web appendix 2). All other data are in the public domain or publication databases, except for trial protocols and statistical analysis plans received from authors and funders. We have not been given permission to share trial protocols and statistical analysis plans outside of our study team, but they can be requested directly from the trial authors and funders, as indicated in supplementary table 5.

RJB affirms that this manuscript is an honest, accurate, and transparent account of the study being reported; that no important aspects of the study have been omitted; and that any discrepancies from the study as planned and registered have been explained.

Dissemination to participants and related patient and public communities: The findings of this study will be disseminated to regulators, industry clinical trials departments, and charities which fund formula and infant food research. The findings of this study will be disseminated to trial sponsors at our request, to confirm methods used in the relevant trials.

Provenance and peer review: Not commissioned, externally peer reviewed. This is an Open Access article distributed in accordance with the Creative Commons Attribution Non Commercial (CC BY-NC 4.0) license, which permits others to distribute, remix, adapt, build upon this work non-commercially, and license their derivative works on different terms, provided the original work is properly cited and the use is non-commercial. See: http://creativecommons.org/licenses/by-nc-4.0/

2 Tran HT, Nguyen TT, Mathisen R. The use of human donor milk. *BMJ* 2020;371:m4243. DOI: 10.1136/bmj.m4243
3 Mumblt D, Crawley H, Hyde R, Boyle RJ. Health and nutrition claims for infant formula are poorly substantiated and potentially harmful. *BMJ* 2020;369:m875. DOI: 10.1136/bmj.m875
4 Infant formula requirements pertaining to current good manufacturing practice, quality control procedures, quality factors, records and reports, and notifications. Electronic Code of Federal Regulations, Title 21, Chapter IB, Part 106.
13 Münz T, chimney BM, Gianola S, Gonzalez-Lorenzo M, Banzi R. The revised Cochrane risk of bias tool for randomized trials (RoB 2) showed lower interrater reliability and challenges in its application. *J Clin Epidemiol* 2020;126:37-44. DOI: 10.1016/j.jclinepi.2020.06.015