REGULAR ARTICLE

Intervention Fidelity Within Trials of Infant Feeding Behavioral Interventions to Prevent Childhood Obesity: A Systematic Review

Elaine Toomey, PhD[1] • Karen Matvienko-Sikar, PhD[2] • Caroline Heary, PhD[3] • Lisa Delaney, MSc[4] •
Michelle Queally, PhD[5] • Catherine B. Hayes, MD[6] • Patricia M. Kearney, PhD[2] • Molly Byrne, PhD[3];
On behalf of the Choosing Healthy Eating for Infant Health (CHErIsH) study team

Abstract

Background Intervention fidelity refers to whether an intervention has been implemented as intended. Trials of infant feeding behavioral interventions to prevent childhood obesity show inconsistent evidence of effectiveness. However, intervention fidelity has not been previously explored within these trials, limiting interpretation of findings.

Purpose To review the use and/or reporting of strategies to enhance and assess intervention fidelity within trials of infant feeding interventions to prevent childhood obesity, and their association with study quality, effectiveness, and publication year.

Methods Seven electronic databases were searched, with articles screened for inclusion by two reviewers. The National Institutes of Health Behaviour Change Consortium fidelity checklist was used to assess use and/or reporting of fidelity strategies across five domains (design, provider training, delivery, receipt, and enactment).

Results Ten trials (16 papers) were identified. Average use/reporting of fidelity strategies was moderate (54%), ranging from 28.9% to 76.7%. Levels of use/reporting ranged from 15.9% in the domain of provider training to 95% for enactment. No association was found between these levels and study quality, effectiveness, or publication year.

Conclusions The moderate use/reporting of fidelity strategies within trials of infant feeding interventions suggests that previous findings of inconsistent effectiveness may not fully reflect the intended interventions. The review highlights key considerations for improving future research, both in the area of behavioral infant feeding and wider behavior change literature. This includes improving reporting across all fidelity domains and ensuring an enhanced focus on provider training and control group content to optimize the translation of research into practice.

PROSPERO Registration number CRD42016033492.

✉ Elaine Toomey
Elaine.toomey@nuigalway.ie

1 Health Behaviour Change Research Group, School of Psychology, National University of Ireland Galway, Room 2058, Arts Millennium Building, Galway, Ireland

2 School of Public Health, University College Cork, Cork, Ireland

3 School of Psychology, National University of Ireland Galway, Galway, Ireland

4 Department of Epidemiology and Public Health, University College Cork, Cork, Ireland

5 Discipline of Economics, JE Cairnes School of Business and Economics, National University of Ireland Galway, Galway, Ireland

6 School of Medicine, Trinity College Dublin, Dublin, Ireland

Background

Childhood obesity is one of the most serious current global public health challenges according to the World Health Organization, with an estimated 40.6 million children under the age of 5 years affected by overweight or obesity in 2016 [1]. Overweight and obesity in children are associated with a number of adverse health outcomes; some of which include the development of cardiovascular disease, type 2 diabetes, musculoskeletal disorders, certain forms of cancer, and issues such as anxiety, depression, and lower self-esteem [2–7]. Moreover, overweight and obese children are likely

to become overweight adults, with such health issues tracking into later life [8]. As such, childhood obesity is an issue requiring urgent attention, necessitating an increased focus on prevention.

Observational research indicates that early infant feeding behaviors in the first 2 years of life are associated with the later development of childhood overweight and obesity [9, 10]. These include behaviors such as inappropriate weaning practices [11], early introduction of solid foods, and nonresponsive feeding [12]. Additionally, the protective effects of breastfeeding against the development of obesity have been consistently demonstrated [13–15]. As such, behavior change interventions to improve feeding practices in infants up to 2 years of age have an important role to play in terms of childhood obesity prevention [16]. However, to be able to fully interpret the findings of behavioral infant feeding interventions and to translate successful findings into practice, knowledge of how and why these interventions work or do not work is crucial. To achieve this, a thorough understanding and evaluation of their implementation is necessary.

Intervention fidelity is a key element of the implementation of behavior change interventions [17, 18] and refers to the extent to which an intervention is actually implemented as intended by the intervention developers. This includes not only aspects such as the delivery of the intervention by its providers, but also how the developers intend intervention participants or recipients to engage with the intervention [19, 20] (e.g., parents/carers will need to understand and acquire intervention-related skills in order to apply them and change infant feeding practices). In behavior change research, intervention fidelity has also been defined as "the methodological strategies used to enhance and monitor the reliability and validity of behavioural interventions" [21]. This includes using strategies or methods to enhance or promote fidelity (e.g., intervention manuals), as well as the methods used to assess it (e.g., using direct observations, self-reported checklists). Enhancing intervention fidelity may influence the effectiveness of interventions, and previous studies have shown positive correlations between fidelity and intervention effectiveness [22, 23]. Intervention fidelity is therefore of paramount importance for interpreting the effectiveness outcomes of a behavior change intervention, as without it we are only really *assuming* that the intended intervention is actually being evaluated. Greater transparency of the intervention process across the designer–provider–recipient pathway also facilitates scientific replication and improved translation of knowledge from research into policy and practice [24]. In particular, assessing intervention fidelity within interventions delivered by healthcare professionals is vitally important as it enables a greater understanding of the implementation challenges faced in real-life healthcare settings, of critical importance for informing

the development of relevant and realistic policy and practice-based guidance [17, 25]. Despite its importance, several reviews have previously demonstrated that intervention fidelity is often poorly addressed within trials of behavioral interventions [19, 20, 26–29]. Moreover, although there is evidence to suggest that fidelity may be improving over time in specific areas such as aspects of tobacco research [23], the same has not been found for childhood obesity research involving children over 2 years of age [30]. However, it is unknown how intervention fidelity in infant feeding intervention research (i.e., involving children aged 0–2) has been addressed to date or its progress over time.

Currently, evidence from reviews of clinical trials is inconclusive or shows minimal effectiveness of infant feeding behavioral interventions on outcomes [31–33]. However, little is known about the actual implementation of these interventions and the fidelity with which they have been put into practice in these trials. This limits our ability to interpret effectiveness outcomes or to successfully replicate interventions. Although they did not assess fidelity of included studies, in their recent review of prevention interventions in childhood obesity, Redsell et al. suggested that issues in relation to the reporting of participant adherence and fidelity of delivery may have contributed towards explaining the smaller effect sizes found [31]. Jaka et al. (2016) recently reviewed the reported use of fidelity strategies within childhood obesity prevention and intervention studies and found that fidelity was poorly addressed and reported within this area; however, studies of higher quality performed significantly better [30]. However, this review only involved studies with children above 2 years of age and did not include infant feeding interventions. Furthermore, only fidelity in one treatment/intervention arm selected by the authors was evaluated, regardless of whether multiple intervention arms were present, as is often the case in trials. Fidelity information pertaining to any control groups was also not included, which is considered a key component of intervention fidelity [26, 34] and will be explored in this review.

The National Institutes of Health Behaviour Change Consortium (NIHBCC) fidelity checklist was specifically developed in 2005 to assess how intervention fidelity is addressed within health behavior change research [21, 26] and has been previously used in several reviews of fidelity of behavioral interventions [23, 27, 28, 35–37]. Validity and reliability of the checklist have also been previously established [23, 26, 38]. The checklist was updated in 2011 [39] to address additional aspects such as cultural considerations and behavioral theory and now consists of 40 components or strategies to enhance and assess intervention fidelity across five domains, including the intended treatment of participants in both intervention and control groups [26, 39]. The five domains

are as follows: (i) Study Design relates to whether an intervention adequately assess its hypotheses in relation to the underlying theory and mechanisms of action of the study; (ii) Training of Providers relates to training procedures and whether providers are *able* to deliver the intervention as intended; (iii) Treatment Delivery refers to providers' actual delivery of the intended intervention in both control and treatment groups; (iv) Treatment Receipt relates to whether participants are *able* to perform intervention skills and behaviors; and (v) Treatment Enactment addresses whether participants actually apply intervention skills and behaviors in daily life. The structured checklist approach enables fidelity levels or scores to be generated for all included studies, facilitating comparison between studies, individual components/ domains, and other systematic reviews as well as examination of correlation with intervention effectiveness or other variables [26, 39].

The aim of this review is to systematically review the use and/or reporting of strategies to enhance and assess intervention fidelity within trials of behavioral infant feeding interventions delivered by healthcare professionals, involving infants aged 0–2 years, to prevent childhood obesity. The review objectives are to establish NIHBCC fidelity levels/scores for included studies and to explore associations between fidelity score and study quality, intervention effectiveness, and year of publication. As demonstrated in previous studies, it was hypothesized that intervention fidelity score would be positively associated with study quality [30], intervention effectiveness [23], and year of publication [23].

Methods

This review was conducted in conjunction with two corresponding reviews examining (i) effectiveness [40] and (ii) behavior change techniques (BCTs) and theory use of infant feeding interventions (Matvienko-Sikar et al., under review), registered on PROSPERO (registration number: CRD42016033492). One search was conducted for all three reviews (as detailed in the following), with separate data extraction and analyses conducted for each individual review.

Search Strategy and Study Selection

The electronic databases of CINAHL, the Cochrane Library, EMBASE, Medline, PubMed, PsycINFO, and Maternity and Infant Care were searched in May 2017 from the earliest date possible. The search combined extrapolated terms for "infant" AND "feeding" AND "trial" AND "weight" adapted for each database as needed (complete sample search strategy provided in Supplementary File 1). There were no restrictions on language of publication. Reference lists of included studies and recent systematic reviews in the area were hand-searched, and authors of included studies were contacted to identify gray literature. The criteria for inclusion of studies are described in Table 1. Two reviewers (KMS, LD) independently screened all study titles, abstracts, and full texts, with discrepancies resolved by consensus.

Data Extraction

Two reviewers (KMS, LD) independently extracted data from the included studies using a standardized form, including details of study design, participant and intervention characteristics, and outcome data. Study quality was assessed by the same two reviewers using the Cochrane Effective Practice and Organization of Care tool for assessing risk of bias [41] (seven potential sources of bias rated as high, low or unclear risk), with disagreement resolved through consensus.

Fidelity data were extracted for each of the eligible studies by one reviewer (ET) using the updated NIHBCC fidelity checklist [39]. Of the 40 components (48 components for four-armed trials), 17 fall within the domain of Study Design, seven within Training of Providers, nine within Treatment Delivery, and five and two in Treatment Receipt and Treatment Enactment respectively [39]. The checklist (detailing domains and subsequent components) is provided in full in Table 2. For all included

Table 1 Study inclusion criteria

"PICO" category	Description
Participants	Full-term infants ≤2 years at intervention commencement; infants considered full-term unless explicitly stated otherwise
Intervention	Randomized controlled trials, case–control, and quasiexperimental studies of any intervention aiming to promote healthy feeding practices to prevent overweight and obesity delivered in any healthcare setting (e.g., primary care, hospital) or by at least one healthcare professional with whom parents may have contact during infancy. A healthcare professional was defined as someone who has undergone professional training to provide any form of healthcare.
Comparator	Any active or "normal care" comparator
Outcome	Eligible studies must have included outcomes of early feeding practices (including, but not limited to, timing and type of weaning foods)

Table 2 Fidelity of NIHBCC[a] components and summary of findings

NIHBCC domain	NIHBCC component	Studies applicable, n (total = 10)	Component score (% of studies rating components as "present")	Fidelity methods: summary of findings
Study design	Provided information about treatment dose in the intervention condition:			Sufficient details provided on content and amount of contact for all interventions. No information provided regarding length of contact for half of the interventions
	• Length of contact (minutes)	10	50	
	• Number of contacts	10	100	
	• Content of treatment	10	100	
	• Duration of contact over time	10	100	
	Provided information about treatment dose in the comparison condition:			No information provided regarding length of contact for any comparison
	• Length of contact (minutes)	10	0	Often insufficient information regarding content or duration provided for comparison groups
	• Number of contacts	10	50	
	• Content of treatment	10	60	
	• Duration of contact over time	10	30	
	Provided information about treatment dose in the second comparison condition (three-armed trials)			
	• Length of contact (minutes)	3	0	
	• Number of contacts	3	66.7	
	• Content of treatment	3	100	
	• Duration of contact over time	3	66.7	
	Provided information about treatment dose in the third comparison condition (four-armed trials):			
	• Length of contact (minutes)	2	0	
	• Number of contacts	2	100	
	• Content of treatment	2	100	
	• Duration of contact over time	2	100	
	Method to ensure that dose is equivalent between conditions	2	100	Dose not designed to be equivalent between intervention and comparison in most studies. Methods used by two studies: • Treatment manual (n = 2; French, INSIGHT) • Reminder prompts/checklists (n = 1; French)

Table 2 Continued

NIHBCC domain	NIHBCC component	Studies applicable, n (total = 10)	Component score (% of studies rating components as "present")	Fidelity methods: summary of findings
	Method to ensure that dose is equivalent for participants within conditions	10	70	Methods used by seven studies: • Treatment manual/protocol ($n = 6$; Wen, French, NOURISH, INFANT, INSIGHT, Gross) • Reminder prompts/checklists ($n = 3$; French, Paul, Wen) • Provider support and contact/feedback ($n = 2$; INFANT, NOURISH)
	Specification of provider credentials that are needed	10	90	Provider credentials specified in nine studies, e.g., 'Providers were either GPs or pediatric nurse practitioners' (Adam et al., 1985)
Theoretical model upon which the intervention is based on is clearly articulated				
	• The active ingredients are specified and incorporated into the intervention	10	100	All studies specified the active ingredients or theory underpinning the intervention and incorporated this into the intervention, e.g. "…families educated and empowered through anticipatory guidance. Health promotion theories, such as the Health Belief Model, have been applied in informing the development of the intervention" (Wen)
	• Use of experts or protocol review group to determine whether the intervention protocol reflects the underlying theoretical model or clinical guidelines	10	30	Three studies reported use or involvement of an expert or review group in protocol development, e.g. "the institutional review board at Nationwide Children's Hospital approved the project protocol" (French)
	• Plan to ensure that the measures reflect the hypothesized theoretical constructs/mechanisms of action	10	80	Outcome measures that mapped onto specified theory/mechanisms of action were used in most studies, e.g. "assessment of maternal feeding style if intervention intended to change infant feeding outcomes through influencing maternal feeding style" (NOURISH)
	Potential confounders that limit the ability to make conclusions at the end of the trial are identified	10	90	Potential confounders identified in most studies, e.g. "outcome analysis controlled for mothers' education level (as a known predictor of the outcome)" (INFANT)

Table 2 *Continued*

NIHBCC domain	NIHBCC component	Studies applicable, n (total = 10)	Component score (% of studies rating components as "present")	Fidelity methods: summary of findings
	Plan to address possible setbacks in implementation (i.e., back-up systems or providers)	10	40	Methods/plan provided in four studies: • Training of extra providers (n = 3; Adam, NOURISH, Wen) • Establishing "buy-in" from providers (n = 1; French)
	If more than one intervention is described, all described equally well	3	30	Three applicable studies gave sufficient description of content of all intervention groups where there is more than one intervention
Training of providers	Description of how providers will be trained (manual of training procedures)	10	20	Insufficient description of provider training in most studies, basic description of training provided in two studies (French, INFANT). Training manuals specified in one study (INSIGHT), but not provided or made publicly available.
	Standardization of provider training (especially if multiple waves of training are needed for multiple groups of providers)	10	40	Insufficient description of training standardization in most studies, two studies trained all providers during same session (INFANT, French) and two described having a "standardised training procedure" (INSIGHT, NOURISH)
	Assessment of provider skill acquisition	10	0	No provider skill assessment detailed by any study.
	Assessment and monitoring of provider skill maintenance over time	10	50	Methods used in five studies: • Ongoing contact and feedback with providers (n = 3; INFANT, NOURISH, INSIGHT) • Refresher training provided (n = 2; French, Schroeder)
	Characteristics being sought in a treatment provider are articulated a priori. Characteristics that should be avoided in a treatment provider are articulated a priori	10	0	No specification of provider characteristics in any study
	At the hiring stage, assessment of whether or not there is a good fit between the provider and the intervention (e.g., ensure that providers find the intervention acceptable, credible, and potentially efficacious)	9	0	No specification of assessment of fit/acceptability of intervention to providers in any study
	There is a training plan that takes into account trainees' different education and experience and learning styles	10	10	Insufficient information provided by studies to determine any tailored training for providers

Table 2 *Continued*

NIHBCC domain	NIHBCC component	Studies applicable, n (total = 10)	Component score (% of studies rating components as "present")	Fidelity methods: summary of findings
Treatment delivery	Method to ensure that the content of the intervention is delivered as specified	10	70	Methods used by seven studies: • Treatment manual/protocol (n = 6; Wen, French, NOURISH, INFANT, INSIGHT, Gross) • Reminder prompts/checklists (n = 3; French, Paul, Wen) • Provider support and contact/feedback (n = 2; INFANT, NOURISH)
	Method to ensure that the dose of the intervention is delivered as specified	10	70	Methods used by seven studies: • Treatment manual/protocol (n = 6; Wen, French, NOURISH, INFANT, INSIGHT, Gross) • Reminder prompts/checklists (n = 3; French, Paul, Wen) • Provider support and contact/feedback (n = 2; INFANT, NOURISH)
	Mechanism to assess if the provider actually adhered to the intervention plan or in the case of computer delivered interventions, method to assess participants' contact with the information	10	70	Methods used by seven studies: • Self-reported assessment (n = 6; NOURISH, French, Paul, Wen, INSIGHT, Schroeder) • Attendance records (delivery of session dose/duration) (n = 3; INFANT, NOURISH, French) • Exit interviews/questionnaires with participants (n = 3; French, Schroeder, INSIGHT) • Direct observation assessment (n = 2; INFANT, NOURISH) • Chart review/treatment log (n = 2; French, INFANT) • Audio-recordings (n = 1; INSIGHT)
	Assessment of nonspecific treatment effects	10	0	No study assessed nonspecific treatment effects
	Used treatment manual	10	60	Use of treatment manual/protocol reported by six studies (INFANT, NOURISH, Wen, French, INSIGHT, Gross)

Table 2 *Continued*

NIHBCC domain	NIHBCC component	Studies applicable, n (total = 10)	Component score (% of studies rating components as "present")	Fidelity methods: summary of findings
	There is a plan for the assessment of whether or not the active ingredients were delivered	10	70	Methods used by seven studies: • Self-reported assessment (n = 6; NOURISH, French, Paul, Wen, INSIGHT, Schroeder) • Attendance records (delivery of session dose/duration) (n = 3; INFANT, NOURISH, French) • Exit interviews/questionnaires with participants (n = 3; French, Schroeder, INSIGHT) • Direct observation assessment (n = 2; INFANT, NOURISH) • Chart review/treatment log (n = 2; French, INFANT) • Audio-recordings (n = 1; INSIGHT)
	There is a plan for the assessment of whether or not proscribed components were delivered. (e.g., components that are unnecessary or unhelpful)	10	70	No explicit assessment of *proscribed* components, however methods used by seven studies to assess fidelity of delivery Methods used by seven studies: • Self-reported assessment (n = 6; NOURISH, French, Paul, Wen, INSIGHT, Schroeder) • Attendance records (delivery of session dose/duration) (n = 3; INFANT, NOURISH, French) • Exit interviews/questionnaires with participants (n = 3; French, Schroeder, INSIGHT) • Direct observation assessment (n = 2; INFANT, NOURISH) • Chart review/treatment log (n = 2; French, INFANT) • Audio-recordings (n = 1; INSIGHT)
	There is a plan for how will contamination between conditions be prevented	10	20	Method used in two studies to explicitly prevent contamination • Cluster randomization (n = 2; Schroeder, French)
	There is an a priori specification of treatment fidelity (e.g., providers adhere to delivering >80% of components)	10	0	No study specified a priori fidelity levels

Table 2 *Continued*

NIHBCC domain	NIHBCC component	Studies applicable, n (total = 10)	Component score (% of studies rating components as "present")	Fidelity methods: summary of findings
Treatment receipt	There is an assessment of the degree to which participants understood the intervention	10	20	Participant understanding of intervention assessed in two studies • Telephone exit interview (n = 1; French) • Written assessment of feeding knowledge (n = 1; Gross)
	There is specification of strategies that will be used to improve participant comprehension of the intervention	10	100	All studies detailed methods to improve participant comprehension of intervention • Participant handouts/workbooks (n = 9; Adam, INFANT, NOURISH, Fangupo, French, Paul, Schroeder, Gross, INSIGHT) • Audio–visual messages/information (n = 4; INFANT, Fangupo, Gross, INSIGHT) • Group discussions (n = 4; INFANT, NOURISH, INSIGHT, Gross) • Practical learning activities/skill rehearsal (n = 3; Wen, INSIGHT, Gross) • Unspecified "resources" (n = 1; Wen)
	The participants' ability to perform the intervention skills will be assessed during the intervention period	10	50	Methods used to assess participant performance of intervention skills used in five studies • Individual home visit—informal feedback and assessment (n = 4; Fangupo, Paul, Wen, INSIGHT) • Individual skill assessments during intervention (n = 1; Gross)
	A strategy will be used to improve subject performance of intervention skills during the intervention period	10	60	Methods used to improve participant performance of intervention skills used in six studies • Group skill rehearsals (n = 2; Fangupo, NOURISH) • Individual skills rehearsals (n = 4; Paul, Wen, Gross, INSIGHT)
	Multicultural factors considered in the development and delivery of the intervention (e.g., provided in native language; protocol is consistent with the values of the target group)	10	50	Multicultural factors explicitly considered in five studies • Participant materials provided in variety of languages (n = 3; Adam, French, Gross) • Participant feedback sought on intervention relevance and materials (n = 2; INFANT, Gross) • Intervention tailored for vulnerable/disadvantaged participants (n = 1; Wen)

Table 2 *Continued*

NIHBCC domain	NIHBCC component	Studies applicable, *n* (total = 10)	Component score (% of studies rating components as "present")	Fidelity methods: summary of findings
Treatment enactment	Participant performance of the intervention skills will be assessed in settings in which the intervention might be applied	10	100	All studies detailed methods to assess participant use of intervention skills in daily life • Questionnaire about use of intervention skills at home (e.g., current feeding practices) (*n* = 10; Adam, INFANT, NOURISH, Fangupo, French, Paul, Schroeder, Wen, Gross, Hohman) • Diaries of intervention skill use (*n* = 1; Paul)
	A strategy will be used to improve performance of the intervention skills in settings in which the intervention might be applied	10	90	Methods used to improve participant performance of intervention skills in daily life used in nine studies • Guidance, resources, and skills training provided during home visit (*n* = 4; Fangupo, Paul, Wen, INSIGHT) • Newsletter/reminders sent to participants outside of intervention (*n* = 3; INFANT, French, Schroeder) • Individual skills rehearsals (*n* = 4; Paul, Wen, Gross, INSIGHT) • Workbooks provided to participants (*n* = 1; NOURISH)

GP general practitioner.
[a]National Institutes of Health Behaviour Change Consortium treatment fidelity checklist [39].

studies, each component of the checklist was indicated as "present" (corresponding to a score of one), "absent but should be present" (score of zero), or "not applicable" [26]. A codebook was developed by one of the authors (ET) to provide further clarity on components in relation to this context, including definitions and specific examples for each component (Supplementary File 2). Authors were contacted for more details (including unreported/unpublished information), and any additional associated publications (e.g., protocols, intervention development) were reviewed. Thirty percent of studies were randomly selected for assessment by a second reviewer (KMS), and any disagreement between reviewers was resolved through discussion and consensus.

Data Analysis and Synthesis

Fidelity data were synthesized according to the individual studies, NIHBCC component, and NIHBCC domain as recommended by the checklist developers [26] and as conducted previously [23, 27, 28, 35–37]. For individual studies, a fidelity score was generated by calculating the number of components coded as "present" as a proportion of the total number of components deemed "applicable" for that study. Potentially nonapplicable components (e.g., information about treatment dose in a third intervention arm for two-armed studies) were specified a priori within the codebook (Supplementary File 2), which was used by both fidelity reviewers (ET, KMS) to guide this decision-making process. For NIHBCC component, the number of studies using the component ("present") was calculated as a proportion of the number of studies for which that component was applicable. For NIHBCC domain, this was conducted by calculating the total components present as a proportion of the total applicable components for each domain. Fidelity scores were interpreted as "low" (≤49%), "moderate" (50%–79%) or "high" (≥80%) [26, 42].

Effectiveness data on parental feeding behaviors (e.g., responsive and nonresponsive feeding behaviors), dietary outcomes (e.g., fruit and vegetable intake) and weight outcomes (i.e., body mass index z score) from the corresponding review of effectiveness were summarized using a vote-counting approach [43–46]. For each study, interventions were classified as: not effective (if 0% of the study's outcomes demonstrated positive effects in favor of the intervention); low effective (1%–34% of outcomes had positive effects in favor of the intervention); moderately effective (35%–69%); or generally effective (≥70%) [47]. Potential associations between fidelity data and outcomes of effectiveness or year of first publication were explored using Spearman's correlations and X–Y scatter plots. Potential associations between fidelity score and study quality were explored in SPSSv23 for each domain of the risk of bias tool separately using one-way ANOVAs.

Results

Ten trials or studies (detailed in 16 papers) were included (Fig. 1) [48–63]. Full details of the search results and included studies are available in the corresponding review of effectiveness [40]. Intervention characteristics for the included studies and effectiveness outcomes are summarized in Table 3. In general, the studies were found to be at a moderate to low risk of bias (Table 4). Agreement between reviewers (ET, KMS) regarding fidelity data extraction was 79.8%, with 100% consensus achieved following discussion. Two studies [56, 60] explicitly reported the result of a fidelity assessment for Treatment Delivery. French et al. reported fidelity of delivery to be 91%; however, Schroeder et al. reported fidelity of delivery as a range from 50% to 70%, making it difficult to synthesize these findings. Through informal contact with the author, fidelity of delivery was identified as 95% for the INFANT study [53, 54].

Use and/or Reporting of Fidelity Strategies Within Individual Studies

Fidelity scores ranged from 28.9% to 76.7% across all included studies (Table 5). The average use of fidelity strategies across the studies was found to be moderate at 54%. No study achieved "high" fidelity (≥80%); though, the majority ($n = 7$) scored above the "moderate" cutoff (>50%). In-depth fidelity data extracted for all studies were provided in Supplementary File 3.

Use and/or Reporting of Fidelity Strategies According to NIHBCC Domain and Component

Fidelity scores were variable across both NIHBCC domains and components. Domain scores ranged from 15.9% (Training of Providers) to 95% (Treatment Enactment) (Table 6). Individual component scores ranged from 0% to 100% (Table 2). All studies provided sufficient information about the intervention content and frequency of contact in the treatment group (Study Design); however, this information was often insufficient in the control group. Within the lowest scoring domain of Training of Providers, the majority of studies ($n = 8$) did not provide sufficient description of provider training, and no assessment of provider skill acquired from training was reported by any study. To enhance fidelity of Treatment Delivery (47.8%), six studies used treatment manuals or protocols, and three used reminder checklists or prompts. Self-reported assessment methods were most commonly used to assess fidelity within this domain ($n = 6$). For Treatment Receipt (56%), all studies detailed methods to improve participant comprehension of intervention, with the majority ($n = 9$) using participant handouts or workbooks to achieve this, but only two studies assessed participant understanding
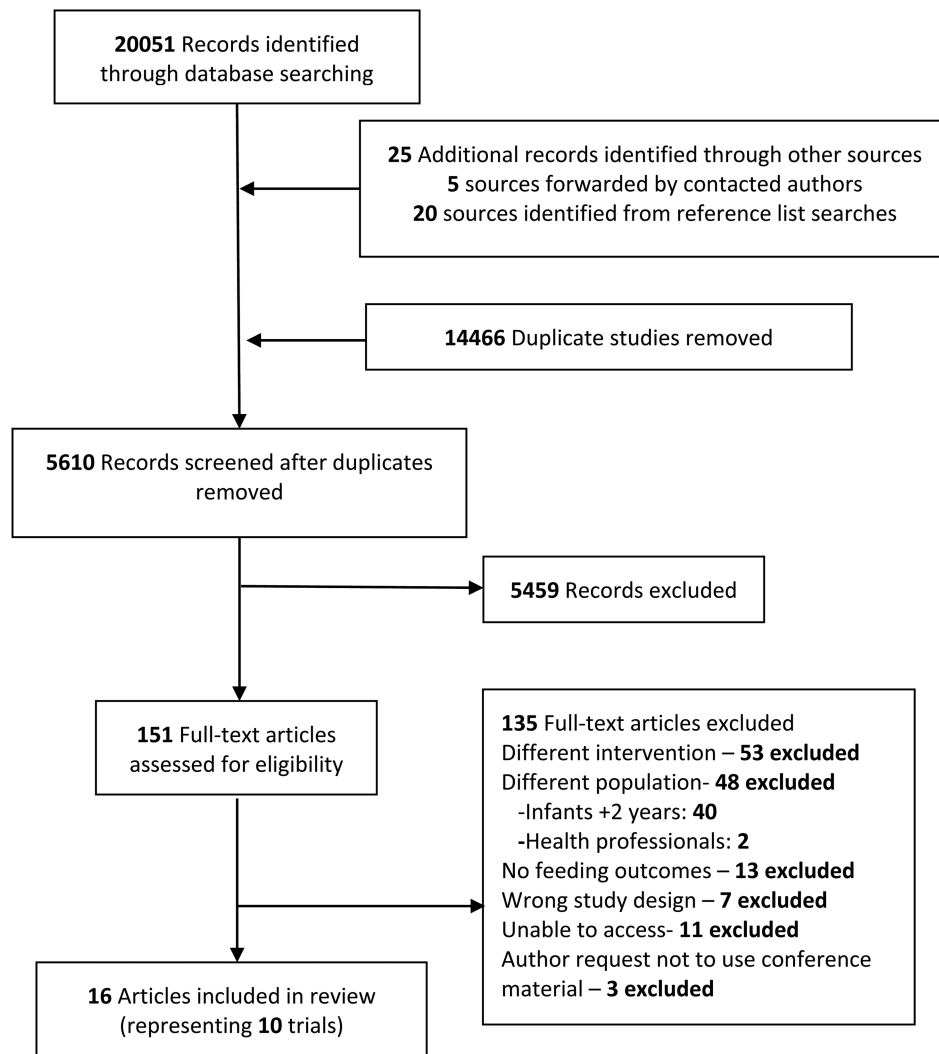
```
┌─────────────────────────────┐
│ 20051 Records identified    │
│ through database searching  │
└─────────────────────────────┘
          ┌──────────────────────────────────────────────────────┐
          │ 25 Additional records identified through other sources│
          │ 5 sources forwarded by contacted authors              │
          │ 20 sources identified from reference list searches    │
          └──────────────────────────────────────────────────────┘
          ┌──────────────────────────────────┐
          │ 14466 Duplicate studies removed  │
          └──────────────────────────────────┘
┌─────────────────────────────┐
│ 5610 Records screened after │
│ duplicates removed          │
└─────────────────────────────┘
                    ┌──────────────────────────┐
                    │ 5459 Records excluded    │
                    └──────────────────────────┘
┌─────────────────────────────┐
│ 151 Full-text articles      │
│ assessed for eligibility    │
└─────────────────────────────┘
┌─────────────────────────────┐
│ 16 Articles included in     │
│ review (representing 10     │
│ trials)                     │
└─────────────────────────────┘
```

135 Full-text articles excluded
Different intervention – **53 excluded**
Different population- **48 excluded**
  -Infants +2 years: **40**
  -Health professionals: **2**
No feeding outcomes – **13 excluded**
Wrong study design – **7 excluded**
Unable to access- **11 excluded**
Author request not to use conference material – **3 excluded**

Fig. 1. PRISMA flowchart of review search strategy.

or knowledge of the intervention [56, 57]. All studies described methods to assess participant use of intervention skills in daily life (Treatment Enactment), with most using a participant-reported questionnaire ($n = 9$). Further detail on strategies used by each study is provided in Table 2.

Association Between NIHBCC Fidelity Score and Study Quality/Effectiveness/Year of First Publication

No association was found between NIHBCC fidelity score and any domain of the Cochrane risk of bias tool (Table 4). No association was found between fidelity score and year of first publication ($p = .569$, Spearman's $\rho = 0.206$), or between fidelity score and effectiveness on parent feeding outcomes ($p = .747$, Spearman's $\rho = 0.200$), dietary outcomes ($p = 0.229$, Spearman's $\rho = –0.446$), or weight outcomes ($p = 0.511$, Spearman's $\rho = –0.339$). No observable trends were seen when data were graphed on X–Y scatter plots (Supplementary File 4).

Discussion

This is the first review examining intervention fidelity within trials of infant feeding behavioral interventions to prevent childhood obesity. Overall, the use and/or reporting of strategies to enhance and assess intervention fidelity was found to be moderate within the included studies, albeit only marginally above the "low" cutoff. Only two studies assessed and explicitly reported fidelity of delivery results. No patterns were observed between NIHBCC fidelity score and study quality, study effectiveness, or year of first publication. This review highlights several key findings of importance for both infant feeding research and implementation research in other behavior change topics.

This review emphasizes the critical importance of adequate reporting of intervention fidelity within trials of interventions to change infant feeding behaviors. Despite recent attempts to improve the quality and completeness of reporting of behavior change interventions,

Table 3 Intervention characteristics and summarized feeding, dietary, and weight outcomes

Study (trial)	Design	Intervention	Control	Parent feeding outcomes[a]	Dietary intake outcomes[b]	Weight outcomes[c]
Adam et al. (1985)	Controlled trial	$n = 49$ Group and individual sessions, on infant feeding. 4-month duration: 1 meeting during postpartum stay and 1 meeting at 1st well baby visit. Delivered by pediatrician.	$n = 54$ Usual care	N/A	+ve: 5; –ve: 0; NS: 0 100% (generally effective)	N/A
Campbell et al. (2013) Cameron et al. (2014) (INFANT)	Cluster RCT	$n = 271$ (2013); $n = 191$ (2014) Group sessions and peer support on infant feeding, physical activity, and sedentary behaviors. 15-month duration: 6 × 2 hr meetings delivered quarterly. Delivered by dietician.	$n = 271$ (2013); $n = 198$ (2014) Usual care from MCH nurse plus 6 general health newsletters, birthday and Christmas cards, and gifts also provided.	N/A	+ve: 6; –ve: 0; NS: 12 33% (low effectiveness)	+ve: 0; –ve: 0; NS: 3 0% (not effective)
Daniels et al. (2012; 2013; 2014; 2015) (NOURISH)	RCT	$n = 346$ Interactive group sessions on infant feeding. 12-month duration: 2 modules of 6 fortnightly sessions each, 1/1.5 hr. Delivered by dieticians and psychologists.	$n = 352$ Usual care	+ve: 34; –ve: 0; NS: 15 69% (moderately effective)	+ve: 3; –ve: 0; NS: 9 25% (moderately effective)	+ve: 3; –ve: 0; NS: 9 25% (moderately effective)
Fangupo et al. (2015)	RCT	$n = 205$ Three intervention arms: interactive family and individual sessions on breastfeeding (arm 1), sleep (arm 2) and both (arm 3) 1-month duration: 4 visits at 4, 7, 13, and 18 months. Delivered by nutritionists, pediatricians, and research staff.	$n = 209$ Usual care	+ve: 3; –ve: 0; NS: 18 14% (low effectiveness)	N/A	N/A

Table 3 *Continued*

Study (trial)	Design	Intervention	Control	Parent feeding outcomes[a]	Dietary intake outcomes[b]	Weight outcomes[c]
French et al. (2012)	Cluster RCT	$n = 101$ Two intervention arms: Group and individual sessions on infant feeding (arm 1) and maternal eating habits (arm 2). 10-month duration: 5 well child visits at 2, 4, 6, 9, and 12 months. Delivered by clinic physicians, nurses, and medical assistants.	$n = 104$ Usual care plus existing handouts on breast-feeding and infant feeding.	+ve: 1; –ve: 0; NS: 3 25% (low effectiveness)	+ve: 1; –ve: 0; NS: 1 50% (moderately effective)	+ve: 0; –ve: 0; NS: 1 0% (not effective)
Gross (2016) (starting early)	RCT	$n = 266$ Individual and group sessions on infant feeding and sleep. 33-month duration: Prenatal visit after 32 weeks; postnatal ward visit: 1, 2, 4, 6, 9, 12, 15, 18, 21, 24, 27, 30, 33 months. Delivered by dietitians certified as lactation counselors	$n = 267$ Usual care	N/A	+ve: 1; –ve: 0; NS: 0 100% (generally effective)	N/A
Hohman (2017) (INSIGHT)	RCT	$n = 129$ Individual sessions on infant feeding, sleep, emotional regulation, and active social play. 36- to 37-week duration: 5 visits: 3–4, 16, 28, and 40 weeks. Delivered by research nurses.	$n = 130$ Age-appropriate home safety information.	N/A	+ve: 1; –ve: 0; NS: 0 100% (generally effective)	N/A

Table 3 Continued

Study (trial)	Design	Intervention	Control	Parent feeding outcomes[a]	Dietary intake outcomes[b]	Weight outcomes[c]
Paul et al. (2011)	RCT	$n = 38$ Three intervention arms: individual sessions on introduction of solids (arm 1), sleep soothing (arm 2) and both (arm 3). Approximately 4- to 6-month duration: 1st visit at 2–3 weeks after birth. 2nd visit within 2 weeks of introduction of solids. Delivered by community nurses.	$n = 41$ Provided with parenting book and standard AAP handout on introduction of solids.	N/A	+ve: 3; −ve: 0; NS: 1 75% (generally effective)	+ve: 2; −ve: 0; NS: 0 100% (generally effective)
Schroeder et al. (2015)	Cluster RCT	$n = 134$ Individual sessions on infant feeding practices and physical activity. 24-month duration: 9 visits at 1, 2, 4, 6, 9, 12, 15, 18, and 24 months. Delivered by pediatricians.	$n = 144$ Usual care	+ve: 2; −ve: 0; NS: 1 67% (moderate effective)	+ve: 5; −ve: 0; NS: 0 100% (generally effective)	+ve: 0 −ve: 2; NS: 8 0% (not effective)
Wen et al. (2011; 2012; 2015) (healthy beginnings)	RCT	$n = 337$ (2011, 2012) $n = 236$ (2015) Individual sessions on infant feeding, physical activity, and social support. 24-month duration: 7 × 1- to 2-hr visits at 3- to 36-week gestation, and 1, 3, 5, 9, 12, 18, and 24 months. Delivered by community nurses.	$n = 330$ (2011, 2012) $n = 239$ (2015) Usual care	+ve: 6; −ve: 0; NS: 2 75% (generally effective)	+ve: 2; −ve: 0; NS: 16 11% (low effectiveness)	+ve: 1; −ve: 0; NS: 5 16% (low effectiveness)

AAP American Academy of Pediatrics; *MCH* Maternal and Child Health; *N/A* not applicable; *NS* not significant; *RCT* randomized controlled trial.

[a]Parental feeding outcomes included responsive feeding (awareness of hunger and satiety cues), nonresponsive feeding (pressuring to eat, controlling feeding), feeding practices or strategies (i.e., responses to food refusal).

[b]Dietary intake outcomes included timing of introduction of solids, fruit and vegetable consumption, beverage consumption.

[c]Weight outcomes included body mass index z score, triceps skinfold.

Table 4 Association between study quality (risk of bias domain) and NIHBCC[a] fidelity score

Study reference	NIHBCC fidelity score	Random sequence generation (selection bias)	Allocation concealment (selection bias)	Blinding of participants, personnel (performance bias)	Blinding of outcome assessment (detection bias)	Incomplete outcome data (attrition bias)	Selective outcome reporting (reporting bias)	Other bias
Adam et al. (1985)	28.9	High risk	High risk	High risk	High risk	High risk	Low risk	Low risk
INFANT Campbell et al. (2013) Cameron et al. (2014)	60.5	Low risk	Unclear risk	High risk	Low risk	Low risk	High risk	Low risk
NOURISH Daniels et al. (2012; 2013; 2014; 2015)	57.9	Low risk	High risk	High risk	Low risk	Low risk	Low risk	Low risk
Fangupo et al. (2015)	44.7	Low risk	Low risk	High risk	Low risk	Low risk	Low risk	Low risk
French et al. (2012)	76.7	High risk	Unclear risk	Low risk	Unclear risk	Low risk	Low risk	Low risk
Gross et al. (2016)	55.3	Low risk	Low risk	Unclear risk	Low risk	Low risk	Low risk	Low risk
INSIGHT Hohman et al. (2017)	64.1	Low risk	Unclear risk	Unclear risk	High risk	High risk	Low risk	Low risk
Paul et al. (2011)	53.2	Low risk	High risk	High risk	High risk	High risk	Unclear risk	Low risk
Schroeder et al. (2015)	36.8	Unclear risk	Unclear risk	High risk	Unclear risk	High risk	Low risk	Low risk
Wen et al. (2011)	60.5	Low risk	Low risk	Unclear risk	Low risk	Low risk	Low risk	Low risk
Test of between-subject effects (one-way ANOVA)		$p = .458$	$p = .534$	$p = .072$	$p = .788$	$p = .138$	$p = .904$	$p = .000$ (N/A)

N/A not applicable.

[a]National Institutes of Health Behaviour Change Consortium treatment fidelity checklist [39].

Table 5 Fidelity of individual studies

Study (trial)	NIHBCC[a] fidelity components used (%)	Applicable NIHBCC components[b] (*n*)
Adam et al. (1985)	28.9	38
INFANT (Campbell et al., 2013; Cameron et al., 2014)	60.5	38
NOURISH (Daniels et al., 2012; 2013; 2014; 2015)	57.9	38
Fangupo et al. (2015)	44.7	47
French et al. (2012)	76.7	43
Starting early (Gross, 2016)	55.3	38
INSIGHT (Hohman, 2017)	64.1	39
Paul et al. (2011)	53.2	47
Schroeder et al. (2015)	36.8	38
Healthy beginnings (Wen et al., 2011; 2012; 2015)	60.5	38
Total mean fidelity score	54	404

[a]National Institutes of Health Behaviour Change Consortium treatment fidelity checklist [39].

[b]Applicable components relate to the number of components deemed applicable to each study. Total possible applicable components = 48 (to include trials with four arms).

Table 6 Fidelity of NIHBCC[a] domain (average and per study)

Study (trial)	Study design %	Training of providers %	Treatment delivery %	Treatment receipt %	Treatment enactment %
Average components present per domain (domain score)	66.9	15.9	47.8	56	95
Adam et al. (1985)	53.3	0	0	40	50
INFANT (Campbell et al., 2013; Cameron et al., 2014)	71.4	42.9	66.7	40	100
NOURISH (Daniels et al., 2012; 2013; 2014; 2015)	71.4	28.6	66.7	40	100
Fangupo et al. (2015)	66.7	0	0	60	100
French et al. (2012)	85.7	50	77.8	60	100
Starting early (Gross, 2016)	73.3	0	33.3	100	100
INSIGHT (Hohman, 2017)	75	28.6	66.7	60	100
Paul et al. (2011)	62.5	0	55.6	60	100
Schroeder et al. (2015)	40	14.3	44.4	20	100
Healthy beginnings (Wen et al., 2011; 2012; 2015)	73.3	0	66.7	80	100

[a]National Institutes of Health Behaviour Change Consortium treatment fidelity checklist [39].

including aspects of intervention fidelity [24, 64, 65], this review documented several issues regarding reporting. For example, although six studies [49–54, 56, 58–63] reported using methods to assess fidelity of intervention delivery, only two studies [56, 60] explicitly reported the results of this assessment, and a third study [53, 54] provided this information only through informal contact with the author. Although ongoing fidelity assessment and monitoring may provide valuable information for the research team and identify potential issues with delivery [21, 39], knowledge of fidelity of delivery also has

key implications for the interpretation of intervention outcomes as discussed in the Introduction section. In an attempt to explain the lack of effectiveness within trials of UK-based breastfeeding interventions, Hoddinott et al. highlighted the importance of considering fidelity of intervention delivery of active, inactive, and/or detractive intervention components [66]. However, insufficient reporting of the use of intervention fidelity strategies and results precludes this from occurring. Therefore, the findings of this review highlight the need for further guidance and support for researchers on how to best

report and utilize the results of fidelity assessment, in order to realize the potential contribution of fidelity data to evaluation of infant feeding interventions.

Additionally, the review found that reporting of fidelity of treatment content and duration in the control group within trials of infant feeding interventions is insufficient. For trials of complex behavior change interventions, understanding what actually transpired in the comparison group is of paramount importance, in order to make accurate conclusions about the effectiveness of the intervention over the comparison [24]. Treatment differentiation, or the degree to which two or more trial arms (i.e., intervention and control groups) differ as intended in terms of aspects such as content and duration, is a key element of intervention fidelity [34], and the ability to meaningfully conduct and interpret intervention evaluations depends on a minimum degree of differentiation between intervention and control groups [67]. As such, without knowledge of the intervention fidelity across all study arms, accurate interpretation of the findings of infant feeding intervention research is compromised. Although poor reporting of control group characteristics has been previously highlighted by the WIDER recommendations to improve reporting of the content of behavior change interventions [65], to the best of our knowledge it has not been previously explored in infant feeding research and was not examined by Jaka et al. (2016) in their review of behavioral pediatric obesity interventions. Moreover, the review findings predominantly relate to reported fidelity, and the study authors may have used strategies to enhance or assess intervention fidelity, but not reported these. In an effort to address this, additional study publications (e.g., published protocols, process evaluation papers) were sourced, and authors were contacted for more information regarding low-scoring fidelity domains during data extraction to facilitate evaluation of actual use of intervention fidelity strategies, and not just reported information. However, the outcome of this contact often further highlighted the issue of insufficient reporting. For example, following contact with the authors of the INFANT study [53, 54], the score for Treatment Delivery improved from 33.3% to 66.6% due to the provision of additional information. Although journal restrictions on word count may have influenced the reporting of use of fidelity strategies within these studies [68], this could be addressed through the publication of additional papers such as process evaluations or specific fidelity reports [69–71]. Despite their potential value, previous research has suggested that process evaluations are not commonplace within childhood obesity prevention intervention research, which may have hampered the cumulative advancement of the evidence base in this area [72, 73].

Use of fidelity strategies within the domain of Training of Providers was found to be particularly low in the reviewed studies (15.9%), with little focus on how the healthcare professionals in these studies were trained or their skill levels in terms of delivering the infant feeding intervention. Previous research has shown that provider knowledge and experience significantly influence the fidelity of intervention delivery [70, 74, 75] and that appropriate evidence-based training is likely to contribute to higher intervention success rates [76]. It is possible that insufficient provider training, or a lack of provider knowledge or skill, may have hindered the delivery of the intervention as intended, potentially contributing to the inconsistent and minimal effects of infant feeding interventions [31, 40]. A limited focus on provider training and skills also has implications for the translation of successful and effective behavior change interventions into practice. Without knowledge about how providers were trained in the original research setting, it is extremely difficult to determine what might be necessary to enable providers to deliver the intervention successfully in a real-life setting [77]. This is relevant regardless of whether they are specifically trained research staff or existing healthcare professionals, as was the case in this review. Training received by intervention providers may also be considered a behavior change intervention in its own right, as behavior change interventions are defined as "coordinated sets of activities designed to change specified behaviour patterns" [78]. Recent work by Murphy et al. (2016) describes a "multi-level" behavior change intervention in cardiovascular rehabilitation, involving a provider- or staff-level intervention followed by a patient-level intervention [79], with equal attention devoted to the intervention fidelity and specification of the active intervention components designed to change behavior (i.e., BCTs) [80] in both. Considering provider training as a potential behavior change intervention in this way may facilitate better transparency and replicability of the training, as well as a more complete understanding of intervention fidelity across all domains, a finding relevant for both childhood obesity literature and broader health behavior change research.

Despite the aforementioned issues, the majority of studies in this review achieved a moderate level of fidelity in terms of strategies used to enhance and assess fidelity. The overall NIHBCC fidelity score of included studies was 54%, which is similar to the 55% found by Borrelli et al. [26] in the first review to use the NIHBCC checklist. This also compares favorably to average fidelity scores of 33%, 35%, 36%, and 47% found in previous reviews involving eating disorder prevention interventions [81], self-management physiotherapy interventions [27], behavioral childhood obesity interventions [30], and psychosocial childhood interventions [37] respectively. Several of the studies in this review explicitly mentioned

the term "fidelity," and six studies reported multiple different methods to enhance and assess fidelity of delivery, which is arguably the most commonly accepted fidelity domain [20, 21, 28]. Treatment Enactment was well addressed by the studies of this review, and participant handouts/workbooks were the most commonly used fidelity enhancement strategy (i.e., strategies to enhance fidelity to the intervention and improve participant performance of intervention skills and behaviors in daily life). However, this may more reflect the overlap between fidelity enhancement strategies and intervention components, and between fidelity assessment strategies (i.e., assessment of intervention skills in daily life) and study outcome measures, rather than a specific focus on fidelity within this domain. Additionally, this study found that self-report was the most commonly used strategy to assess fidelity across all domains, in keeping with the findings of a recent review of fidelity assessment measures used within complex health behavior change interventions [20]. Self-reported fidelity assessments have been consistently found to be less accurate than more objective methods such as direct observation or audio-recordings and may influence how the fidelity scores are interpreted [81, 82]. Furthermore, comparisons with average NIHBCC fidelity scores from other reviews in child psychology (73%) [28], psychosocial oncology (57%) [35], and in secondhand smoking interventions published between 2000 and 2008 (74%) [23] suggest that there remains further room for improvement in addressing intervention fidelity within the field of infant feeding research.

Strengths and Limitations

In addition to being the first review of intervention fidelity within trials of infant feeding interventions to prevent childhood obesity, this study has a number of strengths. Although several reviews of intervention fidelity exist in other previously mentioned research areas [27, 28, 35, 37, 83], few of these have specifically detailed the types of methods or strategies used to enhance and/or assess fidelity, across all domains of the NIHBCC framework [19]. This results in limited information to guide researchers in terms of developing such strategies for their own studies [84, 85] or to understand how intervention fidelity is being specifically addressed in behavior change research. Moreover, the use of an a priori codebook to guide the application of the NIHBCC checklist specifically to this review context ensures greater rigor of the findings. Provision of this codebook also provides greater transparency of the review process. Additionally, only a small amount of fidelity reviews have specifically explored the potential associations between use of fidelity strategies and study quality, effectiveness, or year of publication [23, 26, 30]. Such aspects are of crucial importance for fidelity reviews to increase the relevance of their findings

and help determine whether overall effectiveness findings are attributable to the reviewed intervention, or influenced by variability in implementation. Although this review found no patterns of association between any of these variables, the small sample size of ten included studies undoubtedly minimized the review's potential to detect any associations of significance. Moreover, despite previously identified associations between intervention effectiveness and NIHBCC fidelity scores [23], given that these scores pertain to the use and/or reporting of fidelity strategies and not actual quantitative fidelity results, this may also contribute somewhat toward explaining a lack of association.

There are some limitations of the review which should be taken into account. First, although the vote-counting approach and the fidelity cutoff categories [(e.g., "low" (≤49%), "moderate" (50%–79%) or "high" (≥80%)] used in this study are helpful for synopsizing and comparing findings with previous literature and are based on previously cited research [42, 44], it must be acknowledged that these are somewhat arbitrary. As such, these cutoffs should be interpreted with a degree of caution. Additionally, although the NIHBCC fidelity checklist ensures a standardized and structured approach to reviewing intervention fidelity [18], it has a number of issues such as the ambiguity and lack of weighting of components as previously identified [27, 86]. Moreover, although the checklist includes three components that specifically aim to address fidelity to the underlying intervention theory, this may not be sufficient to fully ensure that interventions were actually "theory-based" as opposed to just "theory-inspired" [87]. This is an integral component of behavior change interventions and as such has been explored in more depth in the corresponding review of BCTs and theory use (Matvienko-Sikar et al., under review). It also does not address the comprehensiveness or accuracy of fidelity assessments (e.g., assessing 5% of intervention delivery using self-report is likely to be less accurate than conducting direct observations of 100% of intervention sessions) [25]. More importantly, the checklist evaluates the use of strategies to enhance and assess fidelity, but does not include the results of a study's actual fidelity assessment (e.g., in three studies of this review that actually provided fidelity of delivery results). It is therefore unclear how best to incorporate reported fidelity assessment data, and as such the checklist in its current form may serve more as a reporting guideline for fidelity strategies rather than for quantifying and interpreting actual intervention fidelity. As such, there is a clear need for further methodological guidance, or revision of existing guidance, on conducting standardized assessments of intervention fidelity within systematic reviews, in relation to both the use of strategies and the actual assessment results themselves. Future research could also explicitly explore the most important components of this framework to be addressed within

behavior change research, that is via qualitative investigation of behavior change researchers' opinions, seeking expert consensus, or by quantitatively exploring potential associations between the presence of specific components and intervention effectiveness.

Implications of Research

This research has a number of implications for clinicians, researchers, and policymakers. The findings of this review enable a better interpretation and understanding of the effectiveness of infant feeding interventions [31, 40], in particular regarding those delivered by healthcare professionals to infants under 2 years. Specifically, the findings of this review suggest that the variability and inconsistency of intervention outcomes found in our corresponding review of effectiveness [40] may have been influenced by issues with intervention fidelity. For example, insufficiently trained providers may have adversely influenced intervention outcomes, such that the results may not provide an accurate reflection of the intended intervention.

In order to ensure that future behavior change infant feeding interventions are more accurately developed and evaluated, this review provides more information for researchers regarding the particular methods and strategies that can be used to enhance and assess fidelity across each fidelity domain, as the review has detailed at least one example for most components of the NIHBCC framework. Future work should also aim to determine which fidelity enhancement and assessment strategies are the most appropriate for certain types of interventions.

For clinicians and policymakers, improving intervention fidelity of behavioral infant feeding intervention research will enable effective infant feeding interventions to be more successfully and easily implemented in practice. For example, ensuring standardized training of providers (e.g., through development of training manuals and supervision protocols) and thoroughly reporting this training (e.g., through making these training manuals and protocols available) will enable effective interventions to be replicated and scaled up for other settings. Despite its aforementioned limitations, the NIHBCC checklist may facilitate this process better than existing reporting criteria (e.g., CONSORT, TIDieR, WIDER [64, 65, 88]) that focus more broadly on intervention and trial procedures and do not address provider training to the same extent.

Conclusions

This is the first review of intervention fidelity within the area of infant feeding and will enable an enhanced interpretation of the effectiveness of these interventions to change infant feeding behaviors, facilitate replication of effective interventions, and maximize knowledge translation for both policy and practice. The review shows that use and/or reporting of strategies to assess and enhance intervention fidelity within trials of behavioral infant feeding interventions to prevent childhood obesity was moderate. It also provides examples of strategies that have been previously used to enhance and assess fidelity within this area. The review highlights areas of key methodological importance where further progress can be made to improve the quality of behavioral infant feeding intervention research, as well as more general health behavior change research. Such areas include increasing attention regarding the training given to intervention providers and their subsequent levels of skill and knowledge; focusing more on the fidelity of treatment within control or comparison group; and ensuring better reporting across all fidelity domains, particularly regarding the results of fidelity assessments. These improvements will facilitate a more accurate test of the intended intervention, enabling more certainty in interpreting outcomes of effectiveness, and support the effective mobilization of successful interventions from research into practice.

Supplementary Material

Supplementary material is available at *Annals of Behavioral Medicine* online.

Compliance with Ethical Standards

References

1. UNICEF/WHO/World Bank Group. *Joint Child Malnutrition Estimates: Levels and Trends in Child Nutrition*. 2016.
2. Jeffreys M, Smith GD, Martin RM, Frankel S, Gunnell D. Childhood body mass index and later cancer risk: A 50-year follow-up of the Boyd Orr study. *Int J Cancer*. 2004;112:348–351.

3. Bridger T. Childhood obesity and cardiovascular disease. *Paediatr Child Health*. 2009;14:177–182.

4. Raj M. Obesity and cardiovascular risk in children and adolescents. *Indian J Endocrinol Metab*. 2012;16:13–19.

5. Morrison KM, Shin S, Tarnopolsky M, Taylor VH. Association of depression & health related quality of life with body composition in children and youth with obesity. *J Affect Disord*. 2015;172:18–23.

6. Halfon N, Larson K, Slusser W. Associations between obesity and comorbid mental health, developmental, and physical health conditions in a nationally representative sample of US children aged 10 to 17. *Acad Pediatr*. 2013;13:6–13.

7. Beck AR. Psychosocial aspects of obesity. *NASN Sch Nurse*. 2016;31:23–27.

8. Singh AS, Mulder C, Twisk JW, van Mechelen W, Chinapaw MJ. Tracking of childhood overweight into adulthood: A systematic review of the literature. *Obes Rev*. 2008;9:474–488.

9. Woo Baidal JA, Locks LM, Cheng ER, Blake-Lamb TL, Perkins ME, Taveras EM. Risk factors for childhood obesity in the first 1,000 days: A systematic review. *Am J Prev Med*. 2016;50:761–779.

10. Birch LL, Ventura AK. Preventing childhood obesity: What works? *Int J Obes (Lond)*. 2009;33(suppl 1):S74–S81.

11. Weng SF, Redsell SA, Swift JA, Yang M, Glazebrook CP. Systematic review and meta-analyses of risk factors for childhood overweight identifiable during infancy. *Arch Dis Child*. 2012;97:1019–1026.

12. Hurley KM, Cross MB, Hughes SO. A systematic review of responsive feeding and child obesity in high-income countries. *J Nutr*. 2011;141:495–501.

13. Dewey KG. Is breastfeeding protective against child obesity? *J Hum Lact*. 2003;19:9–18.

14. Arenz S, Rückerl R, Koletzko B, von Kries R. Breast-feeding and childhood obesity–a systematic review. *Int J Obes Relat Metab Disord*. 2004;28:1247–1256.

15. Owen CG, Martin RM, Whincup PH, Smith GD, Cook DG. Effect of infant feeding on the risk of obesity across the life course: A quantitative review of published evidence. *Pediatrics*. 2005;115:1367–1377.

16. Dattilo AM, Birch L, Krebs NF, Lake A, Taveras EM, Saavedra JM. Need for early interventions in the prevention of pediatric overweight: A review and upcoming directions. *J Obes*. 2012;2012:123023.

17. Moore GF, Audrey S, Barker M, et al. Process evaluation of complex interventions: Medical Research Council guidance. *BMJ*. 2015;350:h1258.

18. Resnick B, Bellg AJ, Borrelli B, et al. Examples of implementation and evaluation of treatment fidelity in the BCC studies: Where we are and where we need to go. *Ann Behav Med*. 2005;29(suppl):46–54.

19. Rixon L, Baron J, McGale N, Lorencatto F, Francis J, Davies A. Methods used to address fidelity of receipt in health intervention research: A citation analysis and systematic review. *BMC Health Serv Res*. 2016;16:663.

20. Walton H, Spector A, Tombor I, Michie S. Measures of fidelity of delivery of, and engagement with, complex, face-to-face health behaviour change interventions: A systematic review of measure quality. *Br J Health Psychol*. 2017;22:872–903.

21. Bellg AJ, Borrelli B, Resnick B, et al.; Treatment Fidelity Workgroup of the NIH Behavior Change Consortium. Enhancing treatment fidelity in health behavior change studies: Best practices and recommendations from the NIH Behavior Change Consortium. *Health Psychol*. 2004;23:443–451.

22. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. *Implement Sci*. 2007;2:40.

23. Johnson-Kozlow M, Hovell MF, Rovniak LS, Sirikulvadhana L, Wahlgren DR, Zakarian JM. Fidelity issues in second-hand smoking interventions for children. *Nicotine Tob Res*. 2008;10:1677–1690.

24. Michie S, Fixsen D, Grimshaw JM, Eccles MP. Specifying and reporting complex behaviour change interventions: The need for a scientific method. *Implement Sci*. 2009;4:40.

25. Toomey E, Matthews J, Guerin S, Hurley DA. Development of a feasible implementation fidelity protocol within a complex physical therapy-led self-management intervention. *Phys Ther*. 2016;96:1287–1298.

26. Borrelli B, Sepinwall D, Ernst D, et al. A new tool to assess treatment fidelity and evaluation of treatment fidelity across 10 years of health behavior research. *J Consult Clin Psychol*. 2005;73:852–860.

27. Toomey E, Currie-Murphy L, Matthews J, Hurley DA. Implementation fidelity of physiotherapist-delivered group education and exercise interventions to promote self-management in people with osteoarthritis and chronic low back pain: A rapid review part II. *Man Ther*. 2015;20:287–294.

28. Garbacz LL, Brown DM, Spee GA, Polo AJ, Budd KS. Establishing treatment fidelity in evidence-based parent training programs for externalizing disorders in children and adolescents. *Clin Child Fam Psychol Rev*. 2014;17:230–247.

29. O'Shea O, McCormick R, Bradley JM, O'Neill B. Fidelity review: A scoping review of the methods used to evaluate treatment fidelity in behavioural change interventions. *Phys Ther Rev*. 2016;21:207–214.

30. JaKa MM, Haapala JL, Trapl ES, et al. Reporting of treatment fidelity in behavioural paediatric obesity intervention trials: A systematic review. *Obes Rev*. 2016;17:1287–1300.

31. Redsell SA, Edmonds B, Swift JA, et al. Systematic review of randomised controlled trials of interventions that aim to reduce the risk, either directly or indirectly, of overweight and obesity in infancy and early childhood. *Matern Child Nutr*. 2016;12:24–38.

32. Blake-Lamb TL, Locks LM, Perkins ME, Woo Baidal JA, Cheng ER, Taveras EM. Interventions for childhood obesity in the first 1,000 days a systematic review. *Am J Prev Med*. 2016;50:780–789.

33. Reilly JJ, Martin A, Hughes AR. Early-life obesity prevention: Critique of intervention trials during the first one thousand days. *Curr Obes Rep*. 2017;6:127–133.

34. Moncher FJ, Prinz RJ. Treatment fidelity in outcome studies. *Clin Psychol Rev*. 1991;11:247–266.

35. Preyde M, Burnham PV. Intervention fidelity in psychosocial oncology. *J Evid Based Soc Work*. 2011;8:379–396.

36. Culloty T, Milne DL, Sheikh AI. Evaluating the training of clinical supervisors: A pilot study using the fidelity framework. *Cogn Behav Ther*. 2010;3:132–144.

37. McArthur BA, Riosa PB, Preyde M. Treatment fidelity in psychosocial intervention for children and adolescents with comorbid problems. *Child Adolesc Ment Health*. 2012;17:139–145.

38. Resnick B, Bellg AJ, Borrelli B, et al. Examples of implementation and evaluation of treatment fidelity in the BCC studies: Where we are and where we need to go. *Ann Behav Med*. 2005;29(suppl):46–54.

39. Borrelli B. The assessment, monitoring, and enhancement of treatment fidelity in public health clinical trials. *J Public Health Dent*. 2011;71(suppl 1):S52–S63.

40. Matvienko-Sikar K, Toomey E, Delaney L, Harrington J, Byrne M, Kearney PM; Choosing Healthy Eating for Infant Health (CHErIsH) study team. Effects of healthcare professional delivered early feeding interventions on feeding practices and dietary intake: A systematic review. *Appetite*. 2018;123:56–71.

41. Higgins JPT, Green S. *Cochrane Handbook for Systematic Reviews of Interventions Version 5.1.0*, 2011.

42. Perepletchikova F, Kazdin AE. Treatment integrity and therapeutic change: Issues and research recommendations. *Clin Psychol Sci Pract*. 2005;12:365–383.

43. Weir MC, Ryan R, Mayhew A, et al. The Rx for change database: A first-in-class tool for optimal prescribing and medicines use. *Implement Sci*. 2010;5:89.

44. Worswick J, Wayne SC, Bennett R, et al. Improving quality of care for persons with diabetes: An overview of systematic reviews—what does the evidence tell us? *Syst Rev*. 2013;2:26.

45. Ahmadzai N, Ansari M, Garritty C, et al. *Effects of Performing Complex Pediatric Intracavitary Surgical Procedures in Specialized vs Non-specialized Centers in High Risk Children: Cochrane Response Rapid Review*, 2013.

46. Toomey E, Currie-Murphy L, Matthews J, Hurley DA. The effectiveness of physiotherapist-delivered group education and exercise interventions to promote self-management for people with osteoarthritis and chronic low back pain: A rapid review part I. *Man Ther*. 2015;20:265–286.

47. Bushman B, Wang M. Vote counting procedures in meta-analysis. In: Cooper H, Hedges L, Valentine J, Cooper H, eds. *The Handbook of Research Synthesis and Meta-Analysis*. New York, NY: Russel Sage Foundation; 1994:207–220.

48. Adam HM, Stern EK, Stein RE. Anticipatory guidance: A modest intervention in the nursery. *Pediatrics*. 1985;76:781–786.

49. Daniels LA, Mallan KM, Battistutta D, et al. Child eating behavior outcomes of an early feeding intervention to reduce risk indicators for child obesity: The NOURISH RCT. *Obesity (Silver Spring)*. 2014;22:E104–E111.

50. Daniels LA, Mallan KM, Battistutta D, Nicholson JM, Perry R, Magarey A. Evaluation of an intervention to promote protective infant feeding practices to prevent childhood obesity: Outcomes of the NOURISH RCT at 14 months of age and 6 months post the first of two intervention modules. *Int J Obes (Lond)*. 2012;36:1292–1298.

51. Daniels LA, Mallan KM, Nicholson JM, Battistutta D, Magarey A. Outcomes of an early feeding practices intervention to prevent childhood obesity. *Pediatrics*. 2013;132:e109–e118.

52. Daniels LA, Mallan KM, Nicholson JM, et al. An early feeding practices intervention for obesity prevention. *Pediatrics*. 2015;136:e40–e49.

53. Campbell KJ, Lioret S, McNaughton SA, et al. A parent-focused intervention to reduce infant obesity risk behaviors: A randomized trial. *Pediatrics*. 2013;131:652–660.

54. Cameron AJ, Ball K, Hesketh KD, et al. Variation in outcomes of the Melbourne Infant, Feeding, Activity and Nutrition Trial (InFANT) Program according to maternal education and age. *Prev Med*. 2014;58:58–63.

55. Fangupo LJ, Heath AL, Williams SM, et al. Impact of an early-life intervention on the nutrition behaviors of 2-y-old children: A randomized controlled trial. *Am J Clin Nutr*. 2015;102:704–712.

56. French GM, Nicholson L, Skybo T, et al. An evaluation of mother-centered anticipatory guidance to reduce obesogenic infant feeding behaviors. *Pediatrics*. 2012;130:e507–e517.

57. Gross RS, Mendelsohn AL, Gross MB, Scheinmann R, Messito MJ. Randomized controlled trial of a primary care-based child obesity prevention intervention on infant feeding practices. *J Pediatr*. 2016;174:171–177.e2.

58. Hohman EE, Paul IM, Birch LL, Savage JS. INSIGHT responsive parenting intervention is associated with healthier patterns of dietary exposures in infants. *Obesity (Silver Spring)*. 2017;25:185–191.

59. Paul IM, Savage JS, Anzman SL, et al. Preventing obesity during infancy: A pilot study. *Obesity (Silver Spring)*. 2011;19:353–361.

60. Schroeder N, Rushovich B, Bartlett E, Sharma S, Gittelsohn J, Caballero B. Early obesity prevention: A randomized trial of a practice-based intervention in 0–24-month infants. *J Obes*. 2015;2015:795859.

61. Wen LM, Baur LA, Simpson JM, Rissel C, Flood VM. Effectiveness of an early intervention on infant feeding practices and "tummy time": A randomized controlled trial. *Arch Pediatr Adolesc Med*. 2011;165:701–707.

62. Wen LM, Baur LA, Simpson JM, et al. Effectiveness of home based early intervention on children's BMI at age 2: Randomised controlled trial. *BMJ*. 2012;344: e3732.

63. Wen LM, Baur LA, Simpson JM, et al. Sustainability of effects of an early childhood obesity prevention trial over time: A further 3-year follow-up of the healthy beginnings trial. *JAMA Pediatr*. 2015;169:543–551.

64. Hoffmann TC, Glasziou PP, Boutron I, et al. Better reporting of interventions: Template for intervention description and replication (TIDieR) checklist and guide. *BMJ*. 2014;348:g1687.

65. Albrecht L, Archibald M, Arseneau D, Scott SD. Development of a checklist to assess the quality of reporting of knowledge translation interventions using the Workgroup for Intervention Development and Evaluation Research (WIDER) recommendations. *Implement Sci*. 2013;8:52.

66. Hoddinott P, Seyara R, Marais D. Global evidence synthesis and UK idiosyncrasy: Why have recent UK trials had no significant effects on breastfeeding rates? *Matern Child Nutr*. 2011;7:221–227.

67. Lorencatto F, Gould NJ, McIntyre SA, et al.; AFFINITIE programme. A multidimensional approach to assessing intervention fidelity in a process evaluation of audit and feedback interventions to reduce unnecessary blood transfusions: A study protocol. *Implement Sci*. 2016;11:163.

68. Perepletchikova F, Hilt LM, Chereji E, Kazdin AE. Barriers to implementing treatment integrity procedures: Survey of treatment outcome researchers. *J Consult Clin Psychol*. 2009;77:212–218.

69. Hardeman W, Michie S, Fanshawe T, Prevost AT, Mcloughlin K, Kinmonth AL. Fidelity of delivery of a physical activity intervention: Predictors and consequences. *Psychol Health*. 2008;23:11–24.

70. Toomey E, Matthews J, Hurley DA. Using mixed methods to assess fidelity of delivery and its influencing factors in a complex self-management intervention for people with osteoarthritis and low back pain. *BMJ Open*. 2017;7:e015452.

71. French SD, Green SE, Francis JJ, et al. Evaluation of the fidelity of an interactive face-to-face educational intervention to improve general practitioner management of back pain. *BMJ Open*. 2015;5:e007886.

72. Hesketh KD, Campbell KJ. Interventions to prevent obesity in 0–5 year olds: An updated systematic review of the literature. *Obesity (Silver Spring)*. 2010;18(suppl 1):S27–S35.

73. Lunn PL, Roberts S, Spence A, Hesketh KD, Campbell KJ. Mothers' perceptions of Melbourne InFANT Program: Informing future practice. *Health Promot Int*. 2016;31:614–622.

74. Taylor CA, Shaw RL, Dale J, French DP. Enhancing delivery of health behaviour change interventions in primary care: A meta-synthesis of views and experiences of primary care nurses. *Patient Educ Couns*. 2011;85:315–322.

75. Wang B, Stanton B, Deveaux L, et al. Factors influencing implementation dose and fidelity thereof and related student outcomes of an evidence-based national HIV prevention program. *Implement Sci*. 2015;10:44.

76. Brose LS, McEwen A, Michie S, West R, Chew XY, Lorencatto F. Treatment manuals, training and successful provision of stop smoking behavioural support. *Behav Res Ther*. 2015;71:34–39.

77. Carroll C, Patterson M, Wood S, Booth A, Rick J, Balain S. A conceptual framework for implementation fidelity. *Implement Sci*. 2007;2:40.

78. Michie S, van Stralen MM, West R. The behaviour change wheel: A new method for characterising and designing behaviour change interventions. *Implement Sci*. 2011;6:42.

79. Murphy PJ, Mc Sharry J, Casey D, et al. Sexual counselling for patients with cardiovascular disease: Protocol for a pilot study of the CHARMS sexual counselling intervention. *BMJ Open*. 2016;6:e011219.

80. Michie S, Johnston M, Carey R. Behaviour change techniques. In Gellman M, Turner J, eds. *Encyclopaedia of Behavioural Medicine*. New York: Springer; 2016:1–8.

81. Carroll K, Nich C, Rounsaville B. Utility of therapist session checklists to monitor delivery of coping skills treatment for cocaine abusers. *Psychother Res*. 1998;8:307–320.

82. Breitenstein SM, Fogg L, Garvey C, Hill C, Resnick B, Gross D. Measuring implementation fidelity in a community-based parenting intervention. *Nurs Res*. 2010;59:158–165.

83. Schober I, Sharpe H, Schmidt U. The reporting of fidelity measures in primary prevention programmes for eating disorders in schools. *Eur Eat Disord Rev*. 2013;21:374–381.

84. Robb SL, Burns DS, Docherty SL, Haase JE. Ensuring treatment fidelity in a multi-site behavioral intervention study: Implementing NIH Behavior Change Consortium recommendations in the SMART trial. *Psychooncology*. 2011;20:1193–1201.

85. O'Donnell CL. Defining, conceptualizing, and measuring fidelity of implementation and its relationship to outcomes in k-12 curriculum intervention research. *Rev Educ Res*. 2008;78:33–84.

86. Leventhal H, Friedman MA. Does establishing fidelity of treatment help in understanding treatment efficacy? Comment on Bellg *et al*. (2004). *Health Psychol*. 2004;23:452–456.

87. Michie S, Carey RN, Johnston M, et al. From theory-inspired to theory-based interventions: A protocol for developing and testing a methodology for linking behaviour change techniques to theoretical mechanisms of action. *Ann Behav Med*. 2016. doi:10.1007/s12160-016-9816-6

88. Moher D, Schulz KF, Altman DG. The CONSORT statement: Revised recommendations for improving the quality of reports of parallel-group randomised trials. *Lancet*. 2001;357:1191–1194.