



Provided by the author(s) and NUI Galway in accordance with publisher policies. Please cite the published version when available.

Title	Health professional-delivered obesity prevention interventions during the first 1,000 days: A systematic review of external validity reporting
Author(s)	Hennessy, Marita; Heary, Caroline; Laws, Rachel; Van Rhoon, Luke; Toomey, Elaine; Wolstenholme, Hazel; Byrne, Molly
Publication Date	2019-10-21
Publication Information	Hennessy, M, Heary, C, Laws, R, Van Rhoon, L, Toomey, E, Wolstenholme, H, & Byrne, M. (2019). Health professional-delivered obesity prevention interventions during the first 1,000 days: A systematic review of external validity reporting [version 2; peer review: 2 approved]. HRB Open Research, 2(14). doi: 10.12688/hrbopenres.12924.2
Publisher	Health Research Board
Link to publisher's version	<a href="https://doi.org/10.12688/hrbopenres.12924.2">https://doi.org/10.12688/hrbopenres.12924.2</a>
Item record	<a href="http://hdl.handle.net/10379/15537">http://hdl.handle.net/10379/15537</a>
DOI	<a href="http://dx.doi.org/10.12688/hrbopenres.12924.2">http://dx.doi.org/10.12688/hrbopenres.12924.2</a>

Downloaded 2021-01-21T22:42:16Z

Some rights reserved. For more information, please see the item record link above.





## SYSTEMATIC REVIEW

**REVISED** **Health professional-delivered obesity prevention interventions during the first 1,000 days: A systematic review of external validity reporting [version 2; peer review: 2 approved]**

Marita Hennessy <sup>1</sup>, Caroline Heary <sup>2</sup>, Rachel Laws<sup>3</sup>, Luke Van Rhon<sup>1</sup>, Elaine Toomey <sup>1</sup>, Hazel Wolstenholme<sup>2</sup>, Molly Byrne <sup>1</sup>

<sup>1</sup>Health Behaviour Change Research Group, School of Psychology, National University of Ireland, Galway, Galway, Ireland

<sup>2</sup>School of Psychology, National University of Ireland, Galway, Galway, Ireland

<sup>3</sup>Institute for Physical Activity and Nutrition (IPAN), School of Exercise and Nutrition Sciences, Deakin University, Geelong, Victoria, Australia

**v2** **First published:** 19 Jul 2019, 2:14 (<https://doi.org/10.12688/hrbopenres.12924.1>)

**Latest published:** 21 Oct 2019, 2:14 (<https://doi.org/10.12688/hrbopenres.12924.2>)

**Abstract**

**Background:** Childhood obesity prevention interventions delivered by health professionals during the first 1,000 days of life show some evidence of effectiveness, particularly in relation to behavioural outcomes. External validity refers to how generalisable interventions are to populations or settings beyond those in the original study. The degree to which external validity elements are reported in such studies is unclear however. This systematic review aimed to determine the extent to which childhood obesity interventions delivered by health professionals during the first 1,000 days report on elements that can be used to inform generalizability across settings and populations.

**Methods:** Eligible studies meeting study inclusion and exclusion criteria were identified through a systematic review of 11 databases and three trial registers. An assessment tool based on the RE-AIM (Reach, Effectiveness, Adoption, Implementation, Maintenance) framework was used to assess the external validity of included studies. It comprised five dimensions: reach and representativeness of individuals, reach and representativeness of settings, implementation and adaptation, outcomes for decision making maintenance and/or institutionalisation. Two authors independently assessed the external validity reporting of 20% of included studies; discrepancies were resolved, and then one completed assessments of the remaining studies.

**Results:** In total, 39 trials involving 46 interventions published between 1999 and 2019 were identified. The majority of studies were randomized controlled trials (n=24). Reporting varied within and between dimensions. External validity elements that were poorly described included: representativeness of individuals and settings, treatment receipt, intervention mechanisms and moderators, cost effectiveness, and intervention sustainability and acceptability.

**Conclusions:** Our review suggests that more emphasis is needed on research designs that consider generalisability, and the reporting of external validity elements in early life childhood obesity prevention

**Open Peer Review**

**Reviewer Status**

Invited Reviewers

1 2

**REVISED**

**version 2**

published  
21 Oct 2019

**version 1**

published  
19 Jul 2019

report

report

- Luke Wolfenden**, Hunter New England Population Health, Wallsend, Australia
- Paula M. Watson** , Liverpool John Moores University, Liverpool, UK

Any reports and responses or comments on the article can be found at the end of the article.

interventions. Important gaps in external validity reporting were identified that could facilitate decisions around the translation and scale-up of interventions from research to practice.

### Keywords

External validity, childhood obesity, generalisability, intervention, implementation science, health professional, prevention, replication, systematic review

**Corresponding author:** Marita Hennessy ([m.hennessy11@nuigalway.ie](mailto:m.hennessy11@nuigalway.ie))

**Author roles:** **Hennessy M:** Conceptualization, Data Curation, Formal Analysis, Funding Acquisition, Investigation, Methodology, Project Administration, Writing – Original Draft Preparation, Writing – Review & Editing; **Heary C:** Conceptualization, Supervision, Writing – Review & Editing; **Laws R:** Conceptualization, Investigation, Supervision, Writing – Review & Editing; **Van Rhoon L:** Investigation, Writing – Review & Editing; **Toomey E:** Investigation, Writing – Review & Editing; **Wolstenholme H:** Investigation, Writing – Review & Editing; **Byrne M:** Conceptualization, Supervision, Writing – Review & Editing

**Competing interests:** No competing interests were disclosed.

**Grant information:** Health Research Board Ireland [SPHeRE/2013/1 to MH]

*The funders had no role in study design, data collection and analysis, decision to publish, or preparation of the manuscript.*

**Copyright:** © 2019 Hennessy M *et al.* This is an open access article distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**How to cite this article:** Hennessy M, Heary C, Laws R *et al.* **Health professional-delivered obesity prevention interventions during the first 1,000 days: A systematic review of external validity reporting [version 2; peer review: 2 approved]** HRB Open Research 2019, 2:14 (<https://doi.org/10.12688/hrbopenres.12924.2>)

**First published:** 19 Jul 2019, 2:14 (<https://doi.org/10.12688/hrbopenres.12924.1>)

**REVISED Amendments from Version 1**

We have clarified terminology to focus on “evaluating external validity reporting” and have defined what we mean by “first 1,000 days” in the abstract and introduction. More context to early childhood obesity prevention has been provided in the introduction, and we have clarified the definition of external validity, and trial generalisability. We have moved details of the external validity assessment tool used from the introduction to the methods section. We have added information on inter-rater reliability, using percent agreement, to the results section. In the discussion, we have noted that researchers could also use models such as RE-AIM to guide reporting of external validity elements, and that using RE-AIM as a planning tool when designing interventions and evaluation might also overcome some of the difficulties in reporting on RE-AIM components relevant to external validity. While we are not aware of any evidence regarding which of the elements of external validity are most important/useful, we have noted elements included in a recently published tool on assessment of scalability and have suggested these be prioritized by researchers in planning studies and reporting findings. We have also discussed whether efficacy trials should be held to the same standard of such reporting as effectiveness trials, given that the latter are intended to provide more externally valid findings. We argue that such information is important regardless of trial type, not only to inform generalisability, but also to enhance understanding of the active ingredients of interventions and core components to retain in effectiveness trials or the scale-up of interventions.

**Any further responses from the reviewers can be found at the end of the article**

**Introduction**

In 2018, approximately 41 million children under the age of five years were classified as having overweight or obesity<sup>1</sup>. Child obesity prevention is a public health priority, with early life intervention advocated<sup>2</sup>. Effective, scalable, and affordable strategies that do not widen health inequities are needed to address this issue<sup>3,4</sup>. In addition, interventions that can be embedded into ongoing practice and existing systems are required, rather than implementing interventions that are resource-intensive and cannot be maintained in the long-term<sup>5,6</sup>. This was echoed in a recent research prioritisation study in which ‘Implementation science’ and ‘How to integrate obesity prevention into existing service structures’ were the third and fourth ranked research priorities identified by researchers, policymakers and practitioners<sup>7</sup>. To date, there has been limited scale-up of childhood obesity prevention interventions.

Appraising scalability prior to investment is vital<sup>8,9</sup>. Scalability is defined by the World Health Organization as ‘the deliberate effort to increase the impact of successfully tested health interventions so as to benefit more people and to foster policy and program development on a lasting basis’<sup>10</sup>. An understanding of the external validity of the intervention is critical to determining scalability, in addition to understanding the fit between an intervention and the political and strategic context. External validity refers to the generalizability of the results of an intervention to or across target populations or settings, while applicability refers to generalizability to any populations or settings<sup>11</sup>. To understand the external validity of an intervention, decision makers need to have sufficient information on the reach and acceptability of the intervention, core intervention

components required for fidelity, any differential effects on the target population, unintended consequences, costs versus benefits, and the clinical or policy significance of the intervention effects to inform decisions about whether interventions should be scaled-up<sup>12–16</sup>. The poor reporting of external validity elements in childhood obesity prevention research also limits decision-makers’ ability to determine the likely scalability of these interventions, limiting the implementation of effective interventions into routine practice<sup>17,18</sup>.

As public health interventions are usually complex, and context dependent, it can be difficult to assess their generalisability to other contexts<sup>14</sup>. There are many tools for assessing generalisability, however, there is no consensus on which should be used, or when<sup>14</sup>. Indeed, Burchett and colleagues argue that such tools may not be the best method for generalisability assessments, instead advocating a focus on mechanisms of action through which an intervention exerts its effect – and which contextual elements underpin them, rather than solely on intervention characteristics<sup>14</sup>. To improve reporting across behavioural interventions and enhance the translation of research into practice, Glasgow and colleagues developed the RE-AIM (Reach, Effectiveness, Adoption, Implementation, Maintenance) framework<sup>19,20</sup>. Reach is the number, proportion of the intended target population, and the representativeness of participants compared with the intended target population. Effectiveness (or efficacy, depending on the study design) is the degree to which the intervention changes behavioural, quality of life, and participant satisfaction outcomes as well as physiologic endpoints, and includes attention to positive, unintended and negative results. Adoption is the number and proportion of settings and staff members that agree to initiate an intervention and how representative they are of the target setting and staff. Implementation is the degree to which settings and staff members deliver an intervention as intended, the adaptations made, and the related costs. Finally, maintenance is sustained effectiveness at the participant level and sustained (or adapted) delivery at the setting or staff level. At the individual level, it refers to the long-term results of intervention (defined as a minimum of six months following the last contact)<sup>20–22</sup>. RE-AIM is the most frequently applied framework in the translation of research evidence into policy and practice<sup>23</sup>. It has been used to assess reports of external validity factors across a variety of areas, including weight loss maintenance interventions<sup>24</sup>, behavioural interventions that target physical activity<sup>25</sup>, mobile health physical activity promotion interventions<sup>26</sup>, physical activity promotion in Latin American populations<sup>27</sup>, behavioural intervention studies conducted in community settings<sup>28</sup>, school health promotion studies<sup>29</sup>, behaviour change interventions in healthcare settings<sup>30</sup>, and housing improvement<sup>31</sup>.

Based on the RE-AIM framework, Green and Glasgow proposed a set of ratings to assess external validity<sup>21</sup>. These were further adapted by Laws and colleagues<sup>32</sup> and have been used to assess external validity in diabetes prevention research<sup>32</sup> and obesity prevention in children aged 0–5 years<sup>18</sup>.

Reviews of external validity reporting in childhood obesity interventions identify insufficient reporting of elements necessary to make decisions about generalisability<sup>17,18</sup>. A review of

external validity reporting in 19 long-term follow-up childhood obesity prevention trials (children aged 0–18 years) published between 1980 and 2004 found that all studies lacked full reporting on potential generalizability and dissemination elements; the most infrequent were reports of setting level inclusion and exclusion criteria and representativeness, characteristics regarding intervention staff, implementation of intervention content, costs, and program sustainability<sup>17</sup>. A more recent review of external validity reporting in 32 trials of interventions to prevent obesity or improve obesity related behaviours in children aged 0–5 years from socioeconomically disadvantaged or Indigenous families found similar issues with reporting<sup>18</sup>. Health professional-delivered interventions to prevent childhood obesity during the first 1,000 days of life (i.e. the period from conception to a child's second birthday) have limited impacts on adiposity/weight outcomes, but have more positive impacts on behavioural outcomes<sup>33</sup>. Despite the increasing numbers of trials to assess the impact of early life obesity prevention interventions, there is relatively little reporting on the potential for these interventions to be translated into routine practice. Furthermore, there is little evidence that interventions with demonstrated efficacy have been translated beyond the research setting and been broadly adopted. Given that it can take up to 17 years to translate evidence into practice<sup>34</sup>, it is important to assess the extent to which trials report on factors that can provide additional explanation for variability in intervention outcomes, insights into successful adaptations of interventions, inform generalizability across settings and populations, and help guide policy decisions.

This study aims to determine the extent to which childhood obesity interventions delivered by health professionals during the first 1,000 days report on factors that can be used to inform generalizability across settings and populations, and to provide recommendations for researchers planning to conduct similar studies.

## Methods

We conducted a systematic review of obesity prevention interventions delivered by health professionals targeting children in the first 1,000 days of life. A separate paper reports on the effectiveness of such interventions and what behaviour change theories and/or techniques are associated with more effective intervention outcomes<sup>33</sup>. The review protocol was registered with the International Prospective Register for Systematic Reviews (PROSPERO) [CRD42016050793](https://doi.org/10.1186/1745-6215-42016050793) on 3<sup>rd</sup> November 2016. This paper adheres to the PRISMA (Preferred Reporting Items for Systematic Reviews and Meta-Analyses) standardised reporting guidelines<sup>35</sup>; the PRISMA checklist is available on OSF<sup>36</sup>.

### Search strategy

Key word searches, using combinations of key words and Medical Subject Headings (or equivalent), were used across six concepts using the AND Boolean operator: (1) child; (2) mother/parent; (3) BMI/obesity; (4) nutrition/physical activity/sleep/parenting; (5) intervention/prevention; (6) randomised controlled trial (RCT)/quasi-randomised trials. Within each

of the categories, keywords were combined using the “OR” Boolean operator. The search strategy was purposefully broad enough to capture any study which might have assessed weight-related measures in children under the age of two. The search strategy was initially developed in Embase<sup>®</sup> (see extended data<sup>36</sup>), appropriately tailored for use within the other databases, and piloted before final searches were run.

One reviewer (MH) searched the following databases from inception to 04 April 2019 using pre-specified search strategies: [CINAHL Complete](#) (EBSCOhost; 1994); [Embase<sup>®</sup>](#) (Elsevier; 1980); [MEDLINE](#) (Ovid<sup>®</sup>; 1966); [PsycINFO](#) (Ovid<sup>®</sup>; 1978); [PubMed](#) (1996); [The Cochrane library](#) databases: The Central Register of Controlled Trials; Database of Systematic Reviews; Database of Abstracts of Reviews of Effect (Wiley; 1996). Conference proceedings and other grey literature were searched on: [Open Grey](#) (INIST-CNRS; 2011) and [Web of Science<sup>™</sup>](#) (Thomson Reuters). [ProQuest Dissertations & Theses Global](#), and [ProQuest Dissertations & Theses – UK and Ireland](#), were used to identify eligible dissertation and thesis studies internationally. We also searched the [International Clinical Trials Registry Platform Search Portal](#), [ClinicalTrials.gov](#), and the [ISRCTN registry](#) to identify any ongoing or unpublished research trials. Reference lists of previous systematic reviews on this topic were manually searched, as well as those of retrieved full texts.

Supplementary materials and trial registry protocols were also checked. No restrictions were applied to: language; date of publication; the length of follow-up of outcomes (given the diversity reported within systematic reviews to date); type of setting; mode of delivery. Records were de-duplicated in Endnote, imported into [COVIDENCE](#) and any remaining duplicates removed.

### Study inclusion and exclusion criteria

[Table 1](#) details the study inclusion and exclusion criteria. We included randomised controlled trials, including cluster-randomised controlled trials, or quasi-randomised trials comparing any behavioural intervention, delivered by health professionals, with ‘usual care’/active comparator which aimed to prevent overweight/obesity in children under the age of two that were born at term. Studies had to report at least one infant/child-related adiposity and/or weight outcome measure at follow-up, which could be immediately post-intervention, or at any time point thereafter; trials only reporting infant birth weight were excluded.

### Study selection

MH and LT independently screened titles and abstracts against the inclusion criteria, and following the retrieval of full-texts, MH and LvR independently reviewed them for inclusion. Disagreements were resolved through discussion, with a third author (MB / CH / RL) where necessary.

### Data extraction

All published papers and supplementary material related to the study (e.g. protocol papers and trial registry protocols,

**Table 1. Study inclusion criteria.**

<b>Design</b>	Randomised, and quasi-randomised, controlled trials, including individual cluster randomised trials
<b>Participants</b>	<p>Studies which targeted pregnant women and/or parents (including mothers/fathers/carers/guardians) of healthy infants less than two years old and/or infants born at term gestation (37 to 42 weeks of gestation) and up to two years of age.</p> <ul style="list-style-type: none"> <li>o No restrictions for sex, ethnicity, socio-economic group, or region, were applied.</li> <li>o Studies where children aged under two years were part of a family group receiving the intervention were included only if data could be extracted separately for these children</li> </ul>
<b>Intervention</b>	<ul style="list-style-type: none"> <li>o Behavioural interventions designed to prevent obesity (by directly/explicitly focusing on childhood obesity prevention, or by indirectly focusing on childhood obesity-related risk factors) in infants (e.g. individual counselling, face-to-face sessions, audio-visual packages, support groups, online interventions/forums) delivered by health professionals antenatally and/or up to a child's second birthday.</li> <li>o Behavioural interventions were defined as "those that require the active participation of a target group (e.g. patient, individual, health professional) in a programme delivered by a trained interventionist with the goal of changing health-related behaviour"<sup>37</sup>.</li> <li>o Interventions targeting key risk factors for childhood obesity<sup>38</sup>, including: early rapid weight gain, infant feeding method, timing of introduction of solid foods, and gestational weight gain</li> <li>o Health professionals were defined according to the International Standard Classification of Occupations (ISCO) ISCO-08<sup>39</sup>. For the purposes of this review, research nurses, lactation consultants, psychologists, and social workers were also classified as health professionals</li> </ul>
<b>Comparator</b>	Participants who were not exposed to an intervention/wait-list control, or an active comparator, or who received 'usual care'. 'Usual care' is defined as standard support and/or appointments without an obesity prevention focus
<b>Outcomes</b>	<p><u>Primary</u></p> <ul style="list-style-type: none"> <li>o Infant/child body mass index (BMI) z score</li> <li>o Additional anthropometric/growth-related: e.g. growth rates (weight gain, linear growth, and head growth, change in BMI z score), percent fat content, ponderal index, skin-fold thickness</li> </ul> <p><u>Secondary</u> (*intermediate behavioural outcomes)</p> <ul style="list-style-type: none"> <li>o Diet-related*: e.g. breastfeeding initiation and duration (total and exclusive); dietary intake and quality; timing of introduction of solid food(s)</li> <li>o Feeding/eating behaviour-related*: e.g. responsive feeding practices</li> <li>o Physical activity-related*: e.g. physical activity, tummy time, play, screen time</li> <li>o Sedentary time/behaviour-related*: e.g. frequency/time spent: being inactive, doing specific low-energy behaviours such as screen time</li> <li>o Sleep*</li> <li>o Environment-related*: e.g. outcomes related to the physical (e.g. food availability) and social environment</li> <li>o Cost effectiveness/costs of the intervention</li> </ul>
<b>Publications</b>	Trials reported only as abstracts were deemed eligible for inclusion if sufficient information was available from the report, or from contact with the authors, to fulfil the inclusion criteria

reference to websites with working hyperlinks, long-term follow-up studies) were used alongside the included article for data extraction. Data were extracted by one author (MH) using a pre-piloted data extraction tool (see extended data<sup>36</sup>), with 20% double-checked by a second reviewer (HCW). Intervention descriptions were extracted following the criteria outlined in the TIDieR reporting guidelines<sup>40</sup>. The external validity assessment tool previously developed by RL<sup>18,32</sup> was used to assess the extent to which included studies/trials reported on elements that would aid decision-making around whether the findings of such studies/trials could be generalised to populations or settings beyond those in the original study<sup>21</sup>. This tool includes five main dimensions (defined in Table 2): 1) reach and representativeness (individuals); 2) reach and representativeness (settings); 3) implementation and adaptation (of intervention), which includes fidelity considerations; 4) outcomes for decision makers; 5) maintenance and institutionalisation (i.e. the potential for implementation of the intervention in routine service delivery). Included studies

were coded according to whether they met each criterion (yes, no, or not applicable). Initially, two authors (MH and RL) independently assessed the external validity reporting of 20% of included studies. Any discrepancies were resolved through discussion, and then one author (MH) completed assessments of the remaining studies. We did not exclude any studies on the basis of the effectiveness and/or quality assessment.

## Results

Electronic and hand searches identified 27,609 references (see Figure 1). Following duplicate removal and title and abstract screening, 230 references were selected for full text review. We identified 39 eligible studies with 46 unique intervention arms and a total of 180 eligible papers<sup>41-79</sup>. Five trials had more than one eligible intervention arm<sup>60,65,68,75,77</sup>.

Studies were mostly published from 2011 onwards (n=34), conducted in high-income countries (n=33), and targeted the period from birth to 2 years only (n=26). They focused on a



**Table 2. Number and percentage of studies reporting external validity elements<sup>1</sup>.**

External validity dimension	Definition	Studies reporting	
		Yes/Total <sup>2</sup>	%
<b>Reach and representativeness of individuals</b>			
Target population for generalizability	Is the intended target population acknowledged/stated (at the individual level) for which the findings intend to be generalised to?	38/39	97
Method to recruit target population	Was information provided about how the target population was recruited/reached (e.g., radio, newspaper, TV, school meeting)?	30/39	77
Inclusion or exclusion criteria	Were individual inclusion and exclusion criteria stated?	38/39	97
Enrolment rate	Is the enrolment rate or data needed to calculate the enrolment rate among individuals reported? Proportion of people who are eligible for participation who actually enrol in the study	26/39	67
Recruitment rate	Is the recruitment rate or data needed to calculate the recruitment rate among individuals reported? Proportion of potential participants (those invited or expressing interest) who actually enrol in the study	26/39	67
Representativeness of individuals	Are there comparisons between individuals who participated versus either (1) those who declined to participate or (2) target population?	10/39	26
Participant characteristics	Are all of the following reported: <ul style="list-style-type: none"> <li>•Gender</li> <li>•Age</li> <li>•Any socioeconomic indicators (education, employment status, or income)</li> <li>•Participation by racial or ethnic minority groups</li> </ul>	21/39	54
<b>Reach and representativeness of settings</b>			
Target setting	Is the target setting for intervention delivery stated (such as workplace, general practice, outpatient facilities, churches, etc.)?	35/38	92
Method to recruit setting	Is information provided about how the site(s) within a given setting were recruited/reached to participate in delivering the intervention?	4/28	14
Inclusion or exclusion criteria	Were inclusion and exclusion criteria for selection of sites within a given setting stated? In the case of single sites, were the characteristics of the site described?	6/28	21
Participation rate	Is the participation level or data need to calculate the participation level among eligible sites reported (only applies to studies with more than one site)?	1/19	5
Representativeness of setting(s)	Are there comparisons between site(s) participating in the intervention and 1) those that decline to participate or 2) the target setting?	1/28	4
<b>Implementation and adaptation</b>			
Intervention characteristics	Were the intervention components described?	38/39	97
Intervention adaptation	Is information reported about how the study intervention is similar or different to original efficacy studies? Note: Only applicable to studies where an intervention is adapted from a previous trial	0/5	0
Time to deliver intervention described	Is the number and length of sessions or time required to deliver the intervention described?	24/37	65
Intervention delivery and exposure	Was the extent to which individuals were exposed to the intervention described? (e.g. proportion of planned intervention sessions actually attended (dose); content delivered as specified; provider adherence to intervention plan)	24/37	65
Delivery agents: characteristics and training	Is information provided on who delivered the intervention, such as the type of professional, or the amount of experience, skill or training required to deliver the intervention?	37/39	95
Methods to recruit delivery agents	Is information provided about how the delivery agents were identified/selected?	3/36	8
Delivery agents' participation	Is the participation level amongst delivery agents reported (% of delivery agents agreeing to participate)?	4/35	11
Fidelity assessment: treatment receipt	Is information reported about whether the program was received as intended? (e.g. degree to which the participants understood the intervention and/or ability to perform the intervention skills)	4/39	10
Mechanisms for intervention effects	Was retrospective analysis conducted to identify the mediating variables through which the intervention achieved its effect?	2/39	5

External validity dimension	Definition	Studies reporting	
		Yes/Total <sup>2</sup>	%
<b>Outcomes for decision making</b>			
Outcomes that can be compared to standards	Are outcomes (at least one) reported in a way that can be compared to either clinical targets or public health goals?	36/39	92
Adverse consequences	Does the article report whether they examined the occurrence of unintended consequences?	18/39	46
Effect moderators by participant characteristics	Are there any analyses of moderator effects by subgroups of participants	10/39	26
Effect moderator by delivery agent/ setting	Are there any analyses of moderator effects by delivery agents or settings	0/37	0
Dose response effect of intervention (sensitivity)	Are there sensitivity analyses to assess dose-response effects of the intervention?	1/39	3
Total costs of intervention	Are total costs of the intervention presented?	6/39	15
Cost of intervention components	If costs are presented, were the costs itemized by intervention components (e.g., personnel, equipment)?	4/6	67
Cost effectiveness	If costs are presented, was there any analysis done to assess cost-effectiveness or cost-benefit of the program or policy?	3/6	50
<b>Maintenance / institutionalisation</b>			
Long term effects (at least 12 months) <sup>3</sup>	Are data reported on longer term effects on health-related outcomes, at least 12 months following program implementation, or environmental or policy change?	19/39	49
Institutionalization: sustainability / plans for sustainability	Are data reported on the sustainability (or reinvention or evolution) or plans for sustainability of the intervention?	4/39	10
Attrition	Are data reported on the number of individuals dropping out and/or lost to follow up	38/39	97
Differential attrition (by condition or population sub-group)	Are data on attrition by condition or population sub-group reported?	35/39	90
Representativeness of completers/ dropouts	Did the study report statistically significant differences in those that dropped out of treatment and those that finished?	19/38	50
Acceptability of the intervention by stakeholders	Was information provided about acceptability of the intervention by stakeholders?	14/39	36

Notes:

<sup>1</sup>Laws *et al.* (adapted from Green *et al.*)

<sup>2</sup>Total = the no. of overall studies (n=39) minus the no. of studies reporting not applicable to the relevant element

<sup>3</sup>In RE-AIM (Reach, Effectiveness, Adoption, Implementation, Maintenance), long-term results of intervention are defined as a minimum of six months following the last contact; long-term is defined as a minimum of 12 months by Laws *et al.*

range of behaviours and outcomes, including: multiple infant behaviours (n=13); infant feeding: formula feeding / breast-feeding / introduction to solids (n=10); maternal diet/physical activity/gestational weight gain (n=9); infant feeding: breast-feeding only (n=8). Only 16 of the 46 interventions were clearly delivered as part of routine care, with a further two partly delivered as such. Details of intervention descriptions and outcomes are available as extended data<sup>36</sup>.

The assessment of the reporting of external validity elements of the 39 included studies is summarised in Table 2, with a summary by study available as extended data<sup>36</sup>. Inter-rater reliability, using percent agreement, was high (88.2%). Full details of these assessments, including supporting statements for each study, are available as extended data<sup>36</sup>. The number and percentage of studies reporting all elements of each dimension of external validity are outlined in Table 3.

### Reach and representativeness of participants

Only 15% of studies reported on all elements of this external validity dimension (Table 3). While almost all studies outlined the target population for generalizability (97%) and inclusion and exclusion criteria (97%), less reported the recruitment method (77%), enrolment rate (67%), and recruitment rate (67%) (Table 2). Just over half (54%) reported all of the specified participant characteristics - gender, age, any socioeconomic indicators (education, employment status, or income - and participation by racial or ethnic minority groups. Only one in four studies included comparisons between individuals who participated versus either (1) those who declined to participate or (2) target population.

### Reach and representativeness of settings

One in four studies reported on all elements of 'reach and representativeness of settings' (Table 3). Almost all studies



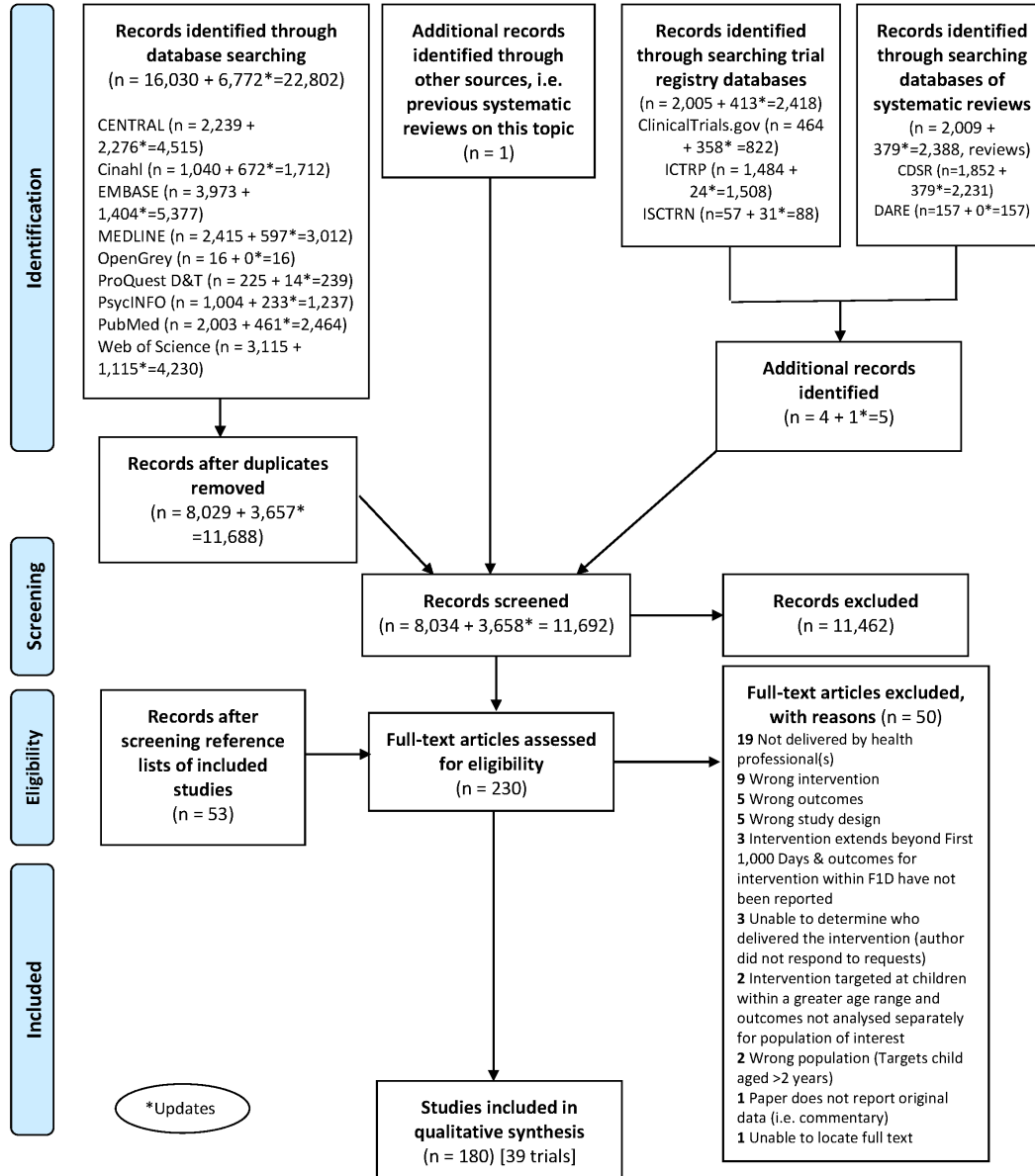


Figure 1. PRISMA flow diagram.

Table 3. Number and percentage of studies reporting all elements of each dimension of external validity<sup>1</sup>.

External validity dimension	No./Total <sup>2</sup>	%	Studies
Reach and representativeness of participants	6/39	15	NOURISH, PRIMROSE, Baby Milk Trial, INSIGHT, BLISS, POI
Reach and representativeness of settings	9/38	24	ProKind, Baby Milk Trial, Minding the Baby, SLIMTIME, INSIGHT, BLISS, POI, Healthy MOMS, Healthy Beginnings
Implementation and adaptation	0/39	0	None
Outcomes for decision making	0/39	0	None
Maintenance / institutionalisation	1/39	3	INSIGHT

Notes:

<sup>1</sup>Laws *et al.* (adapted from Green *et al.*)

<sup>2</sup>No. taken as sum of no. of studies reporting yes or not applicable to each of the element. Total excludes any studies for which the external validity criterion was not applicable (e.g. Grow2Gether was a social media intervention therefore 'research and representativeness of settings' criterion was not applicable).

provided details of the target setting for intervention delivery (92%); however, the remaining criteria were poorly described: inclusion and exclusion criteria (21%), how settings were recruited/reached to participate in delivering the intervention (14%) (Table 2). Only one study reported the participation level among eligible sites (5%); this was also the case for the representativeness of setting(s) (4%).

### Implementation and adaptation

No studies reported on all elements of this external validity dimension (Table 3). Most studies described the intervention characteristics (97%) and the characteristics and training of delivery agents (95%). Less described the time to deliver the intervention (65%), and intervention delivery and exposure (65%) (Table 2). Delivery agents' participation (11%), fidelity assessment: treatment receipt (10%), methods to recruit delivery agents (8%), and mechanisms for intervention effects (5%) were very poorly reported. Only five of the studies tested an intervention that was adapted from a previous trial - none reported on how the study intervention was similar or different to original efficacy studies.

### Outcomes for decision making

No studies reported on all elements of 'outcomes for decision making' (Table 3). Almost all studies reported outcomes in a way that could be compared to either clinical targets or public health goals (92%) (Table 2). Less than half of studies reported whether they examined the occurrence of unintended consequences (46%). Only six studies reported the total costs of the intervention (15%); of these, four studies reported the cost of intervention components (67%), and three examined cost effectiveness (50%). Ten studies (26%) examined effect moderators by participant characteristics; however, none reported effect moderators by delivery agent/setting. Only one study (3%) reported a sensitivity analyses to assess dose-response effects of the intervention.

### Maintenance / institutionalisation

Only one study – INSIGHT – reported on all elements of maintenance / institutionalisation (Table 3). Almost all studies (97%) reported on the number of individuals dropping out and/or lost to follow up (Table 2). Data on attrition by condition or population sub-group reported by 90% of studies (Note: we took condition to mean by intervention or control group). Only 50% of studies addressed the representativeness of completers/dropouts. Half of studies (49%) reported data on longer term effects on health-related outcomes (at least 12 months following program implementation, or environmental or policy change). Only 10% of studies reported on the sustainability (or reinvention or evolution) or plans for sustainability of the intervention. Only 36% reported on the acceptability of the intervention by stakeholders.

### Discussion/ Conclusion

Early life interventions delivered by health professionals have the potential to influence important health behaviours, in addition to child weight. Understanding the reporting of external validity elements of such interventions is vital to address their potential for translation and scalability, as well as replication efforts. In this systematic review we identified 39 studies,

representing 46 interventions. External validity elements that were generally well reported included target populations and settings, participant inclusion and exclusion criteria, intervention characteristics, delivery agents, outcomes, and attrition. Similar to other reviews of childhood obesity interventions<sup>17,18</sup>, however, we identified important gaps in the reporting of external validity elements within studies, and factors that could enhance translation and scale-up of interventions across all five external validity dimensions. External validity elements that were poorly reported included: representativeness of individuals and settings, treatment receipt, intervention mechanisms and moderators, cost effectiveness, and intervention sustainability and acceptability.

Key gaps in informing the translation and scalability of health professional-delivered early life obesity prevention interventions were identified in this review. These included understanding the representativeness of settings, and whether these settings and delivery agents could be engaged to deliver these types of interventions in a sustained way, in a way that is acceptable to those involved. This is especially important given that only 16 of the 46 interventions (35%) in this review were clearly delivered as part of routine care, with a further two partly delivered as such, i.e. contacts as part of routine care but additional contacts also (Starting Early<sup>78</sup> and STRIP<sup>79</sup>). The focus of the majority of studies was on establishing efficacy rather than effectiveness or how such interventions could be scaled up and translated into routine practice. This may account for the poor reporting of external validity in relation to settings and delivery agents. It could be argued that efficacy trials should not be held to the same level of accountability regarding reporting of external validity elements. Such information however is important regardless of trial type, not only to inform generalisability, but also to enhance understanding of the active ingredients of interventions and core components to retain in effectiveness trials or the scale-up of interventions.

Reporting of external validity elements considered important to inform decision makers was generally poor also. This included cost and cost-effectiveness measures, and an understanding of the intervention mechanisms and dose-response effects. While most interventions that are scaled up need to be adapted to fit the delivery context, knowing information about dose-response and the mechanism of intervention effects is essential in informing adaptations so that effectiveness of the intervention is not lost. The recent systematic review by McCrabb and colleagues highlights the decreased intervention effects when obesity interventions are scaled up – they found that effects on weight status, physical activity/sedentary behaviour, and nutrition reported in scaled-up interventions were typically 75% or less of the effects reported in pre-scale-up efficacy trials<sup>9</sup>. Reporting of fidelity components in our review was also varied – training (95%), delivery (65%), and receipt (10%). This has been noted in other childhood obesity-related reviews<sup>80,81</sup>, and has important implications for the interpretation, as well as the generalisability, of study findings.

Despite calls for greater attention to external validity for almost 40 years now<sup>16,82-84</sup>, we noted that problems with attention to generalisability persist. Only one trial within this review,

the INSIGHT trial<sup>56</sup>, reported on all elements of the external validity assessment tool developed by RL<sup>18,32</sup>. Earlier this year, Huebschmann and colleagues made a further call for increased attention to external validity<sup>82</sup>. For trialists, there is a tension between internal validity and external validity, with preference historically for ensuring the former and minimising the risk of bias, at the expense of generalisability and applicability to real-world settings. Standard reporting guidelines such as the CONSORT statement for the reporting of randomized controlled trials<sup>85</sup>, the CONSORT extension for cluster trials<sup>86</sup> and the CONSORT extension for pragmatic trials<sup>87</sup> traditionally focus on internal validity elements, with limited focus and guidance around external validity. The TIDieR reporting guidelines for intervention description and replication somewhat address this gap<sup>40</sup>.

We acknowledge the challenging context in which trialists work and that there are many positive activities in this area. We have a number of suggestions for moving work in this area forward nevertheless. Trialists could plan their interventions with scalability and sustainability in mind, giving due consideration to the type of trial conducted as well as the intervention characteristics. Few researchers plan for the sustainability of their interventions<sup>88</sup>. The aforementioned reporting guidelines can be used in combination to report on study findings, with additional materials published to enhance external validity assessment, including protocols and more detailed information made accessible via supplementary materials or open access repositories. Researchers could also use models such as RE-AIM to guide reporting of external validity elements. If researchers used RE-AIM as a planning tool when designing their intervention and evaluation, this might also overcome some of the difficulties in reporting on RE-AIM components relevant to external validity. Glasgow and Estabrooks note the challenges in comprehensively reporting on all RE-AIM dimensions within community and clinical settings with limited resources, however, highlighting that even well-funded NIH grants and published research studies, stating use of the RE-AIM framework, only employ it partially, and inconsistently when they do so<sup>89</sup>. Inconsistencies in the degree to which authors report each RE-AIM dimension in its entirety as well as inaccuracies in reporting elements within each dimension have been highlighted by other authors also<sup>22,90</sup>. Further work is needed with researchers to embed such frameworks appropriately. As mentioned above, increasing the availability of protocols and more detailed information via supplementary materials or open access repositories is one such step. Research to understand the facilitators and barriers to reporting elements of external validity, as well as work with stakeholders to prioritise the most important elements/dimensions of external validity reporting would be useful to enhance work in this area. A recently published tool on assessment of scalability contains several elements relevant to external validity assessment<sup>91</sup>. These include: costs and cost effectiveness, intervention characteristics, information on delivery agents, participation rate of settings, outcomes that can be compared to standards, effect moderators, adverse consequences, and acceptability. Such elements should be prioritised by researchers in planning studies and reporting findings.

Funding bodies, review panels, journals/journal editorial boards, and policymakers could also take action to promote the integration of external validity considerations into the funding, design, conduct, reporting, synthesis and translation of research<sup>19,82,84,92</sup>. This need not be at the expense of internal validity, and can help facilitate credible research and knowledge translation<sup>82,92</sup>. The inclusion of a PRECIS-2 graphic when proposing or reporting on a study can also be undertaken to enable the assessment of external validity<sup>82</sup>.

### Strengths and limitations

The strengths of this work are the use of a comprehensive and rigorous methodology, including a broad search strategy and range of databases, no language restrictions, and the screening of trials and extraction of data by two independent review authors. A number of limitations, however, must be noted. While we included journal articles, protocols, grey literature and supplementary materials, it is possible that researchers of the reviewed studies may have collected some of the information required to complete the external validity assessment but did not report it in the articles published to date. Furthermore, the external validity tool only codes items as present, absent, or not applicable. The extent, or quality, to which the studies report on the various external validity elements, e.g. fidelity, is not assessed; this may result in an over-estimation of the reporting quality of some studies. While it is not necessary for all studies to be strong on all of the external validity criterion, researchers, decision-makers and others could use this information, if provided, to make judgments as to the applicability or generalisability of a study or review<sup>16</sup>.

### Conclusion

This review examined the reporting of external validity elements within 39 studies encompassing 46 early-life health professional-delivered interventions. While such interventions have the potential to influence important health behaviours, in addition to child weight, we identified important gaps in the reporting of external validity elements within studies, and factors that could enhance translation and scale-up of interventions across all five external validity dimensions. External validity elements that were poorly described included: representativeness of individuals and settings, treatment receipt, intervention mechanisms and moderators, cost effectiveness, and intervention sustainability and acceptability. More emphasis is needed on research designs that consider generalisability, and the reporting of external validity elements in early life childhood obesity prevention interventions.

### Data availability

#### Underlying data

All data underlying the results are available as part of the article and no additional source data are required.

#### Extended data

Open Science Framework: Health professional-delivered obesity prevention interventions during the first 1,000 days: A systematic review of external validity reporting. <https://doi.org/10.17605/OSF.IO/G2ZMY><sup>36</sup>

This contains the following underlying data:

- SearchStrategy\_v1\_HealthProfessional-deliveredObesityPreventionFID.pdf (Search strategy)
- DataExtractionForm\_v3\_HealthProfessional-deliveredObesityPreventionFID.pdf (Data extraction form)
- InterventionDescriptions\_v7\_HealthProfessional-deliveredObesityPreventionFID.xlsx (Intervention descriptions)
- InterventionOutcomes\_v5\_HealthProfessional-deliveredObesityPreventionFID.xlsx (Intervention outcomes)
- FullDetailsOfExternalValidityAssessments\_v2\_HealthProfessional-deliveredObesityPreventionFID.pdf (Full details of external validity assessments)
- SummaryTable-ExternalValidity-ByStudy\_v3\_HealthProfessional-deliveredObesityPreventionFID.xlsx (Summary table of external validity assessments by study)

## Reporting guidelines

Open Science Framework: PRISMA checklist for 'Health professional-delivered obesity prevention interventions during the first 1,000 days: A systematic review of external validity reporting'. <https://doi.org/10.17605/OSF.IO/G2ZMY><sup>36</sup>

Data are available under the terms of the [Creative Commons Zero "No rights reserved" data waiver](#) (CC0 1.0 Public domain dedication).

## Acknowledgements

Ms Marita Hennessy is a PhD Scholar funded by the Health Research Board under SPHeRE/2013/1. The authors thank Ms Jane Mulligan and Ms Rosie Dunne (James Hardiman Library, NUI Galway) for their assistance in developing the search strategies, and Ms Louise Tully (Royal College of Surgeons in Ireland, Dublin) for assisting with the double screening of titles and abstracts for this review.

## References

1. World Health Organisation: **Obesity and Overweight Fact Sheet**. Geneva: World Health Organisation. 2018. [Reference Source](#)
2. World Health Organization: **Report of the Commission on Ending Childhood Obesity**. Geneva: World Health Organization. 2016. [Reference Source](#)
3. Wake M: **The failure of anti-obesity programmes in schools**. *BMJ*. 2018; **360**: k507. [PubMed Abstract](#) | [Publisher Full Text](#)
4. Blake-Lamb TL, Locks LM, Perkins ME, *et al.*: **Interventions for Childhood Obesity in the First 1,000 Days A Systematic Review**. *Am J Prev Med*. 2016; **50**(6): 780–789. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
5. World Health Organisation: **Consideration of the evidence on childhood obesity for the Commission on Ending Childhood Obesity: Report of the Ad hoc Working Group on Science and Evidence for Ending Childhood Obesity**. World Health Organisation: Geneva, Switzerland. 2016. [Reference Source](#)
6. Waters E, de Silva-Sanigorski A, Hall BJ, *et al.*: **Interventions for preventing obesity in children**. *Cochrane Database Syst Rev*. 2011; (12): CD001871. [PubMed Abstract](#) | [Publisher Full Text](#)
7. Hennessy M, Byrne M, Laws R, *et al.*: **Childhood obesity prevention: priority areas for future research and barriers and facilitators to knowledge translation, coproduced using the nominal group technique**. *Transl Behav Med*. 2019; **9**(4): 759–767. [PubMed Abstract](#) | [Publisher Full Text](#)
8. Indig D, Lee K, Grunseit A, *et al.*: **Pathways for scaling up public health interventions**. *BMC Public Health*. 2017; **18**(1): 68. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
9. McCrabb S, Lane C, Hall A, *et al.*: **Scaling-up evidence-based obesity interventions: A systematic review assessing intervention adaptations and effectiveness and quantifying the scale-up penalty**. *Obes Rev*. 2019; **20**(7): 964–982. [PubMed Abstract](#) | [Publisher Full Text](#)
10. World Health Organization: **Nine Steps for Developing a Scaling-Up Strategy**. Geneva: WHO ExpandNet. 2010. [Reference Source](#)
11. Murad MH, Katabi A, Benkhadra R, *et al.*: **External validity, generalisability, applicability and directness: a brief primer**. *BMJ Evid Based Med*. 2018; **23**(1): 17–19. [PubMed Abstract](#) | [Publisher Full Text](#)
12. Jamal F, Fletcher A, Shackleton N, *et al.*: **The three stages of building and testing mid-level theories in a realist RCT: a theoretical and methodological case-example**. *Trials*. 2015; **16**: 466. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
13. Bonell C, Oakley A, Hargreaves J, *et al.*: **Assessment of generalisability in trials of health interventions: suggested framework and systematic review**. *BMJ*. 2006; **333**(7563): 346–349. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
14. Burchett HED, Blanchard L, Kneale D, *et al.*: **Assessing the applicability of public health intervention evaluations from one setting to another: a methodological study of the usability and usefulness of assessment tools and frameworks**. *Health Res Policy Syst*. 2018; **16**(1): 88. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
15. Lavis JN, Oxman AD, Souza NM, *et al.*: **SUPPORT Tools for evidence-informed health Policymaking (STP) 9: Assessing the applicability of the findings of a systematic review**. *Health Res Policy Syst*. 2009; **7** Supp 1: S9. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
16. Glasgow RE, Green LW, Klesges LM, *et al.*: **External validity: we need to do more**. *Ann Behav Med*. 2006; **31**(2): 105–108. [PubMed Abstract](#) | [Publisher Full Text](#)
17. Glasgow RE, Dzewaltowski DA, Glasgow RE: **Review of external validity reporting in childhood obesity prevention research**. *Am J Prev Med*. 2008; **34**(3): 216–223. [PubMed Abstract](#) | [Publisher Full Text](#)
18. Laws R, Campbell KJ, van der Pligt P, *et al.*: **The impact of interventions to prevent obesity or improve obesity related behaviours in children (0-5 years) from socioeconomically disadvantaged and/or indigenous families: a systematic review**. *BMC Public Health*. 2014; **14**(1): 779. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
19. Glasgow RE, Klesges LM, Dzewaltowski DA, *et al.*: **The future of health behavior change research: what is needed to improve translation of research into health promotion practice?** *Ann Behav Med*. 2004; **27**(1): 3–12. [PubMed Abstract](#) | [Publisher Full Text](#)
20. Glasgow RE, Vogt TM, Boles SM: **Evaluating the public health impact of health promotion interventions: the RE-AIM framework**. *Am J Public Health*. 1999; **89**(9): 1322–1327. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
21. Green LW, Glasgow RE: **Evaluating the relevance, generalization, and applicability of research: issues in external validation and translation methodology**. *Eval Health Prof*. 2006; **29**(1): 126–153. [PubMed Abstract](#) | [Publisher Full Text](#)
22. Kessler RS, Purcell EP, Glasgow RE, *et al.*: **What does it mean to "employ" the RE-AIM model?** *Eval Health Prof*. 2013; **36**(1): 44–66. [PubMed Abstract](#) | [Publisher Full Text](#)
23. Milat AJ, Li B: **Narrative review of frameworks for translating research evidence into policy and practice**. *Public Health Res Pract*. 2017; **27**(1): pii: 2711704. [PubMed Abstract](#) | [Publisher Full Text](#)
24. Akers JD, Estabrooks PA, Davy BM: **Translational research: bridging the gap between long-term weight loss maintenance research and practice**. *J Am Diet Assoc*. 2010; **110**(10): 1511–1522, 1522.e1-3. [PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)



25. Estabrooks PA, Gyurcsik NC: **Evaluating the impact of behavioral interventions that target physical activity: issues of generalizability and public health.** *Psychol Sport Exerc.* 2003; 4(1): 41–55.  
[PubMed Abstract](#) | [Publisher Full Text](#)
26. Blackman KC, Zoellner J, Berrey LM, *et al.*: **Assessing the internal and external validity of mobile health physical activity promotion interventions: a systematic literature review using the RE-AIM framework.** *J Med Internet Res.* 2013; 15(10): e224.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
27. Galaviz KI, Harden SM, Smith E, *et al.*: **Physical activity promotion in Latin American populations: a systematic review on issues of internal and external validity.** *Int J Behav Nutr Phys Act.* 2014; 11: 77.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
28. Dzewaltowski DA, Estabrooks PA, Klesges LM, *et al.*: **Behavior change intervention research in community settings: how generalizable are the results?** *Health Promot Int.* 2004; 19(2): 235–245.  
[PubMed Abstract](#) | [Publisher Full Text](#)
29. Estabrooks P, Dzewaltowski DA, Glasgow RE, *et al.*: **Reporting of validity from school health promotion studies published in 12 leading journals, 1996-2000.** *J Sch Health.* 2003; 73(1): 21–28.  
[PubMed Abstract](#) | [Publisher Full Text](#)
30. Glasgow RE, Bull SS, Gillette C, *et al.*: **Behavior change intervention research in healthcare settings: a review of recent reports with emphasis on external validity.** *Am J Prev Med.* 2002; 23(1): 62–69.  
[PubMed Abstract](#) | [Publisher Full Text](#)
31. Thomson HJ, Thomas S: **External validity in healthy public policy: application of the RE-AIM tool to the field of housing improvement.** *BMC Public Health.* 2012; 12: 633.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
32. Laws RA, St George AB, Rychetnik L, *et al.*: **Diabetes prevention research: a systematic review of external validity in lifestyle interventions.** *Am J Prev Med.* 2012; 43(2): 205–214.  
[PubMed Abstract](#) | [Publisher Full Text](#)
33. Hennessy M, *et al.*: **The effectiveness of health professional-delivered interventions during the first 1,000 days to prevent overweight/obesity in children: A systematic review.** *Obesity Reviews.* 2019; In press.
34. Morris ZS, Wooding S, Grant J: **The answer is 17 years, what is the question: understanding time lags in translational research.** *J R Soc Med.* 2011; 104(12): 510–20.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
35. Moher D, Liberati A, Tetzlaff J, *et al.*: **Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement.** *BMJ.* 2009; 339: b2535.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
36. Hennessy M, *et al.*: **Extended Data: Health professional-delivered obesity prevention interventions during the first 1,000 days: A systematic review of external validity reporting.** O.S. Framework, Editor. 2019.
37. Bacon SL, Lavoie KL, Ninot G, *et al.*: **An International Perspective on Improving the Quality and Potential of Behavioral Clinical Trials.** *Curr Cardiovasc Risk Rep.* 2015; 9(1): 427.  
[Publisher Full Text](#)
38. Weng SF, Redsell SA, Swift JA, *et al.*: **Systematic review and meta-analyses of risk factors for childhood overweight identifiable during infancy.** *Arch Dis Child.* 2012; 97(12): 1019–1026.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
39. International Labour Office: **International Standard Classification of Occupations. Structure, group definitions and correspondence tables.** *ISCO – 08 Volume I*, International Labour Organisation: Geneva. 2012.  
[Reference Source](#)
40. 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.  
[PubMed Abstract](#) | [Publisher Full Text](#)
41. Albernaz E, Victora CG, Haisma H, *et al.*: **Lactation counseling increases breastfeeding duration but not breast milk intake as measured by isotopic methods.** *J Nutr.* 2003; 133(1): 205–210.  
[PubMed Abstract](#) | [Publisher Full Text](#)
42. Lakshman R, Sharp SJ, Whittle F, *et al.*: **Randomised controlled trial of a theory-based behavioural intervention to reduce formula milk intake.** *Arch Dis Child.* 2018; 103(11): 1054–1060.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
43. Eline V, *et al.*: **Effectiveness of a primary prevention program of overweight in 0-3 year old children, the BBOFT- a study; a cluster randomized trial.** *J Pediatr Gastroenterol Nutr.* 2016; 62(S1): 725.
44. Taylor RW, Williams SM, Fangupo LJ, *et al.*: **Effect of a Baby-Led Approach to Complementary Feeding on Infant Growth and Overweight: A Randomized Clinical Trial.** *JAMA Pediatr.* 2017; 171(9): 838–846.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
45. Carlsen EM, Kyhnaeb A, Renault KM, *et al.*: **Telephone-based support prolongs breastfeeding duration in obese women: a randomized trial.** *Am J Clin Nutr.* 2013; 98(5): 1226–1232.  
[PubMed Abstract](#) | [Publisher Full Text](#)
46. de Vries A, Huiting HG, van den Heuvel ER, *et al.*: **An activity stimulation programme during a child's first year reduces some indicators of adiposity at the age of two-and-a-half.** *Acta Paediatr.* 2015; 104(4): 414–421.  
[PubMed Abstract](#) | [Publisher Full Text](#)
47. Ekström A, Abrahamsson H, Eriksson RM, *et al.*: **Women's use of nipple shields- Their influence on breastfeeding duration after a process-oriented education for health professionals.** *Breastfeed Med.* 2014; 9(9): 458–466.  
[PubMed Abstract](#) | [Publisher Full Text](#)
48. Parat S, Nègre V, Baptiste A, *et al.*: **Prenatal education of overweight or obese pregnant women to prevent childhood overweight (the ETOIG study): an open-label, randomized controlled trial.** *Int J Obes (Lond).* 2019; 43(2): 362–73.  
[PubMed Abstract](#) | [Publisher Full Text](#)
49. Rauh K, Günther J, Kunath J, *et al.*: **Lifestyle intervention to prevent excessive maternal weight gain: mother and infant follow-up at 12 months postpartum.** *BMC Pregnancy Childbirth.* 2015; 15: 265.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
50. Froozani MD, Permezhadeh K, Motlagh AR, *et al.*: **Effect of breastfeeding education on the feeding pattern and health of infants in their first 4 months in the Islamic Republic of Iran.** *Bull World Health Organ.* 1999; 77(5): 381–385.  
[PubMed Abstract](#) | [Free Full Text](#)
51. Bonuck K, Avraham SB, Lo Y, *et al.*: **Bottle-weaning intervention and toddler overweight.** *J Pediatr.* 2014; 164(2): 306–312.e1-2.  
[PubMed Abstract](#) | [Publisher Full Text](#)
52. Fiks AG, Gruver RS, Bishop-Gilyard CT, *et al.*: **A Social Media Peer Group for Mothers To Prevent Obesity from Infancy: The Grow2Gether Randomized Trial.** *Child Obes.* 2017; 13(5): 356–368.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
53. 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(6): 543–551.  
[PubMed Abstract](#) | [Publisher Full Text](#)
54. Vesco KK, Leo MC, Karanja N, *et al.*: **One-year postpartum outcomes following a weight management intervention in pregnant women with obesity.** *Obesity (Silver Spring).* 2016; 24(10): 2042–2049.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
55. Campbell KJ, Lioret S, McNaughton SA, *et al.*: **A parent-focused intervention to reduce infant obesity risk behaviors: a randomized trial.** *Pediatrics.* 2013; 131(4): 652–660.  
[PubMed Abstract](#) | [Publisher Full Text](#)
56. Paul IM, Savage JS, Anzman-Frasca S, *et al.*: **Effect of a Responsive Parenting Educational Intervention on Childhood Weight Outcomes at 3 Years of Age: The INSIGHT Randomized Clinical Trial.** *JAMA.* 2018; 320(5): 461–468.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
57. Wake M, Price A, Clifford S, *et al.*: **Does an intervention that improves infant sleep also improve overweight at age 6? Follow-up of a randomised trial.** *Arch Dis Child.* 2011; 96(6): 526–532.  
[PubMed Abstract](#) | [Publisher Full Text](#)
58. Jonsdottir OH, Kleinman RE, Wells JC, *et al.*: **Exclusive breastfeeding for 4 versus 6 months and growth in early childhood.** *Acta Paediatr.* 2014; 103(1): 105–111.  
[PubMed Abstract](#) | [Publisher Full Text](#)
59. Kazemi H, Ranjkesh F: **Evaluating the problems of mothers in exclusive breastfeeding and educational intervention for improving nutrition status in Iran.** *HealthMED: Journal of Society for development in new net environment in B&H.* 2011; 5(6): 1517–1521.
60. French GM, Nicholson L, Skybo T, *et al.*: **An evaluation of mother-centered anticipatory guidance to reduce obesogenic infant feeding behaviors.** *Pediatrics.* 2012; 130(3): e507–e517.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
61. Kolu P, Raitanen J, Puhkala J, *et al.*: **Effectiveness and Cost-Effectiveness of a Cluster-Randomized Prenatal Lifestyle Counseling Trial: A Seven-Year Follow-Up.** *PLoS One.* 2016; 11(12): e0167759.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
62. Dodd JM, McPhee AJ, Deussen AR, *et al.*: **Effects of an antenatal dietary intervention in overweight and obese women on 6 month infant outcomes: follow-up from the LIMIT randomised trial.** *Int J Obes (Lond).* 2018; 42(7): 1326–1335.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
63. Tanvig M, Vinter CA, Jørgensen JS, *et al.*: **Effects of lifestyle intervention in pregnancy and anthropometrics at birth on offspring metabolic profile at 2.8 years: results from the Lifestyle in Pregnancy and Offspring (LIPO) study.** *J Clin Endocrinol Metab.* 2015; 100(1): 175–183.  
[PubMed Abstract](#) | [Publisher Full Text](#)
64. Ordway MR, Sadler LS, Holland ML, *et al.*: **A Home Visiting Parenting Program and Child Obesity: A Randomized Trial.** *Pediatrics.* 2018; 141(2): pii: e20171076.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
65. Mustila T, Raitanen J, Keskinen P, *et al.*: **Lifestyle counselling targeting infant's mother during the child's first year and offspring weight development until 4 years of age: a follow-up study of a cluster RCT.** *BMJ Open.* 2012; 2(1): e000624.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
66. Neyzi O, Olgun P, Kutluay T, *et al.*: **An educational intervention on promotion of breast feeding.** *Paediatr Perinat Epidemiol.* 1991; 5(3): 286–298.  
[PubMed Abstract](#) | [Publisher Full Text](#)
67. Daniels LA, Mallan KM, Nicholson JM, *et al.*: **An Early Feeding Practices Intervention for Obesity Prevention.** *Pediatrics.* 2015; 136(1): e40–49.  
[PubMed Abstract](#) | [Publisher Full Text](#)
68. Taylor RW, Gray AR, Heath AM, *et al.*: **Sleep, nutrition, and physical activity interventions to prevent obesity in infancy: follow-up of the Prevention of Overweight in Infancy (POI) randomized controlled trial at ages 3.5 and 5 y.** *Am*

- J Clin Nutr.* 2018; **108**(2): 228–236.  
[PubMed Abstract](#) | [Publisher Full Text](#)
69. Döring N, Ghaderi A, Bohman B, *et al.*: **Motivational Interviewing to Prevent Childhood Obesity: A Cluster RCT.** *Pediatrics.* 2016; **137**(5): pii: e20153104.  
[PubMed Abstract](#) | [Publisher Full Text](#)
70. Jungmann T, Kurtz V, Brand T, *et al.*: **[Child health as a prevention objective in the context of the pilot project “Pro Kind”. Preliminary results of a randomized controlled trial].** *Bundesgesundheitsblatt Gesundheitsforschung Gesundheitsschutz.* 2010; **53**(11): 1180–1187.  
[PubMed Abstract](#) | [Publisher Full Text](#)
71. Fornari E, *et al.*: **The Prevention of Obesity in Toddlers (PROBIT) trial.** *J Pediatr Gastroenterol Nutr.* 2018; **66**: 902.
72. Martin RM, Kramer MS, Patel R, *et al.*: **Effects of Promoting Long-term, Exclusive Breastfeeding on Adolescent Adiposity, Blood Pressure, and Growth Trajectories: A Secondary Analysis of a Randomized Clinical Trial.** *JAMA Pediatr.* 2017; **171**(7): e170698.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
73. Horan MK, McGowan CA, Gibney ER, *et al.*: **Maternal Nutrition and Glycaemic Index during Pregnancy Impacts on Offspring Adiposity at 6 Months of Age—Analysis from the ROLO Randomised Controlled Trial.** *Nutrients.* 2016; **8**(1): pii: E7.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
74. Schroeder N, Rushovich B, Bartlett E, *et al.*: **Early Obesity Prevention: A Randomized Trial of a Practice-Based Intervention in 0-24-Month Infants.** *J Obes.* 2015; **2015**: e795859.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
75. Schwartz R, Vigo Á, de Oliveira LD, *et al.*: **The Effect of a Pro-Breastfeeding and Healthy Complementary Feeding Intervention Targeting Adolescent Mothers and Grandmothers on Growth and Prevalence of Overweight of Preschool Children.** *PLoS One.* 2015; **10**(7): e0131884.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
76. Shah NB, Fenick AM, Rosenthal MS: **A Healthy Weight for Toddlers? Two-Year Follow-up of a Randomized Controlled Trial of Group Well-Child Care.** *Clin Pediatr (Phila).* 2016; **55**(14): 1354–1357.  
[PubMed Abstract](#) | [Publisher Full Text](#)
77. Paul IM, Savage JS, Anzman SL, *et al.*: **Preventing obesity during infancy: a pilot study.** *Obesity (Silver Spring).* 2011; **19**(2): 353–361.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
78. Gross RS, Mendelsohn AL, Gross MB, *et al.*: **Randomized Controlled Trial of a Primary Care-Based Child Obesity Prevention Intervention on Infant Feeding Practices.** *J Pediatr.* 2016; **174**: 171–177.e2.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
79. Niinikoski H, Lagström H, Jokinen E, *et al.*: **Impact of repeated dietary counseling between infancy and 14 years of age on dietary intakes and serum lipids and lipoproteins: the STRIP study.** *Circulation.* 2007; **116**(9): 1032–1040.  
[PubMed Abstract](#) | [Publisher Full Text](#)
80. 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**(12): 1287–1300.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
81. Toomey E, Matvienko-Sikar K, Heary C, *et al.*: **Intervention Fidelity Within Trials of Infant Feeding Behavioral Interventions to Prevent Childhood Obesity: A Systematic Review.** *Ann Behav Med.* 2019; **53**(1): 75–97.  
[PubMed Abstract](#) | [Publisher Full Text](#)
82. Huebschmann AG, Leavitt IM, Glasgow RE: **Making Health Research Matter: A Call to Increase Attention to External Validity.** *Annu Rev Public Health.* 2019; **40**(1): 45–63.  
[PubMed Abstract](#) | [Publisher Full Text](#)
83. Klesges LM, Dzawaltowski DA, Christensen AJ: **Are we creating relevant behavioral medicine research? Show me the evidence!** *Ann Behav Med.* 2006; **31**(1): 3–4.  
[PubMed Abstract](#) | [Publisher Full Text](#)
84. Rothwell PM: **External validity of randomised controlled trials: “to whom do the results of this trial apply?”.** *Lancet.* 2005; **365**(9453): 82–93.  
[PubMed Abstract](#) | [Publisher Full Text](#)
85. Schulz KF, Altman DG, Moher D: **CONSORT 2010 Statement: updated guidelines for reporting parallel group randomised trials.** *BMJ.* 2010; **340**: c332.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
86. Campbell MK, Piaggio G, Elbourne DR, *et al.*: **Consort 2010 statement: extension to cluster randomised trials.** *BMJ.* 2012; **345**: e6661.  
[PubMed Abstract](#) | [Publisher Full Text](#)
87. Zwarenstein M, Treweek S, Gagnier JJ, *et al.*: **Improving the reporting of pragmatic trials: an extension of the CONSORT statement.** *BMJ.* 2008; **337**: a2390.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
88. Johnson AM, Moore JE, Chambers DA, *et al.*: **How do researchers conceptualize and plan for the sustainability of their NIH R01 implementation projects?** *Implement Sci.* 2019; **14**(1): 50.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
89. Glasgow RE, Estabrooks PE: **Pragmatic Applications of RE-AIM for Health Care Initiatives in Community and Clinical Settings.** *Prev Chronic Dis.* 2018; **15**: E02.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
90. Harden SM, Gaglio B, Shoup JA, *et al.*: **Fidelity to and comparative results across behavioral interventions evaluated through the RE-AIM framework: a systematic review.** *Syst Rev.* 2015; **4**: 155.  
[PubMed Abstract](#) | [Publisher Full Text](#) | [Free Full Text](#)
91. Milat A, Lee K, Grunseit A, *et al.*: **Intervention Scalability Assessment Tool.** Sydney: The Australian Prevention Partnership Centre. 2019.
92. Green LW, Glasgow RE, Atkins D, *et al.*: **Making evidence from research more relevant, useful, and actionable in policy, program planning, and practice slips “twixt cup and lip”.** *Am J Prev Med.* 2009; **37**(6 Suppl 1): S187–191.  
[PubMed Abstract](#) | [Publisher Full Text](#)



# Open Peer Review

Current Peer Review Status:  

## Version 1

Reviewer Report 16 September 2019

<https://doi.org/10.21956/hrbopenres.14003.r26785>

© 2019 Watson P. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.



**Paula M. Watson** 

Research Institute for Sport and Exercise Sciences, Liverpool John Moores University, Liverpool, UK

### Overview

This is a well-conducted systematic review that assesses the extent to which early life childhood obesity intervention studies report external validity components. Methods and results are clear and transparent, and the discussion is good with some appropriate recommendations for researchers.

There is some ambiguity with terminology in places. As I understand it, the aim of this article is to assess the extent to which obesity prevention studies/trials *report* external validity components. However in places the authors refer to “evaluating external validity” and “evaluating external validity reporting” interchangeably. Whilst these processes are naturally linked, it needs to be clear they are distinct and the purpose of the article needs to be clearly outlined. I have highlighted some specific sections where this is unclear below, but it would be useful to check the whole article to ensure the aims are clear, and terminology is as intended throughout.

I also feel the introduction could benefit with some background information for non-specialist readers (see specific points below).

### Specific comments

#### Title/Throughout

It took me some time to understand what was meant by “the first 1000 days” (my initial interpretation was the first 1000 days of intervention implementation). For clarity, I think it would be useful to refer to “the first 1000 days of life” when using this phrase. It would also be useful to clarify for non-specialist readers that the first 1000 days of life includes the gestation period. Otherwise the inclusion criteria (which refers to 2 years, i.e. ~730 days) could appear to be misaligned with the aim/title.

#### Abstract

**Methods (3<sup>rd</sup> line from bottom)** – would it be more accurate to say “Two authors independently assessed the external validity *reporting* of 20% of included studies”?

## Introduction

For the non-specialist reader, it might be useful to insert a paragraph at the start of the introduction to outline the current state of play in early childhood obesity prevention, e.g. explain current obesity prevalence when starting school, highlight the importance of obesity prevention interventions within infancy (and during pregnancy), and outline what these interventions typically involve.

**3<sup>rd</sup> paragraph, lines 14-15** – suggest re-wording the final part of this sentence for accuracy. Whilst the RE-AIM model might provide a framework for *evaluating external validity reporting*, this was not the primary reason it was developed. In my understanding, the purpose of RE-AIM is to guide researchers in the planning, evaluation and reporting of interventions to maximise external validity (as correctly stated in the first part of the sentence, lines 10-13).

## Methods

**Data extraction paragraph, lines 9-13** – the authors indicate that the external validity assessment tool was used “to assess the extent to which findings of included studies could be generalised to populations or settings beyond those in the original study”. This implies that the data being extracted was about the *external validity* of interventions, rather than the *reporting* of external validity of interventions? Is this correct? Please clarify and amend wording as appropriate.

**Data extraction, lines 15-16** – as above, shouldn't this read external validity *reporting*?

## Results

**Paragraph 3, line 1** – the description of Table 2 suggests the table includes an assessment of the external validity of studies. This is a little misleading - the table reports on the extent to which external validity aspects are *reported*, but does not include an external validity assessment of the studies themselves. Suggest re-wording to make this clear.

**Figure 1 (Prisma flow diagram)** – in the “Full text articles excluded, with reasons” box, it would be useful to replace the word “wrong” with details of *why* the intervention/population/study design did not fit. E.g. “Intervention did not do XXX”, or “Study had no comparator group”. This would be more informative for readers.

## Discussion/Conclusion

**Paragraph 1, lines 3-5** – “an understanding of the external validity of such interventions is vital”. It is important to make clear this paper makes no attempt to improve our understanding of the external validity of interventions, the aim is to understand the extent to which process elements are reported (as reporting is crucial so that others can assess the external validity of interventions).

**Paragraph 3, final line** – it is worth being mindful that the TIDieR guidelines only address some aspects of “external validity”. Other elements such as reach and representativeness, adoption, and maintenance (settings) might require researchers to draw on RE-AIM also.

**Paragraph 4** – as RE-AIM is intended as a planning and evaluation tool, it would be useful to integrate RE-AIM into the discussion around planning for sustainability of interventions. If researchers use RE-AIM as a planning tool when designing their intervention and evaluation, this might also overcome some of the difficulties in reporting on RE-AIM components (as discussed later in the paragraph).

### Are the rationale for, and objectives of, the Systematic Review clearly stated?

Partly

### Are sufficient details of the methods and analysis provided to allow replication by others?

Yes

**Is the statistical analysis and its interpretation appropriate?**

Not applicable

**Are the conclusions drawn adequately supported by the results presented in the review?**

Yes

**Competing Interests:** No competing interests were disclosed.

**Reviewer Expertise:** Complex intervention development and evaluation, physical activity behaviour change (specific interests in child weight management and GP exercise referral).

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

Reviewer Report 29 August 2019

<https://doi.org/10.21956/hrbopenres.14003.r26701>

© 2019 Wolfenden L. This is an open access peer review report distributed under the terms of the [Creative Commons Attribution License](#), which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

**Luke Wolfenden**

Hunter New England Population Health, Wallsend, Australia

This systematic review examined the extent to which childhood obesity interventions delivered by health professionals during the first 1,000 days report on information to enable assessment of the generalisability of trial findings. Generalisability is an important attribute of public health intervention trials. While a number of studies have described reporting of trial information that may facilitate assessments of trial generalizability, there does not appear to have been previous work in this area on obesity in this important population group. The manuscript is well written, and the study methods are appropriate. I congratulate the authors on well undertaken review. I provide below points for clarification and suggestions for further improvement.

1. It may be helpful to provide an early definition of external validity, and trial generalisability in the introduction to set the scene. These terms are not necessarily inter-changeable but at times it appears as though they are used that way.
2. The introduction risks conflating external validity with scalability (among other things). It would help to provide a more nuanced description of how external validity is related to processes of scaling up, and also other processes suggested by the authors including replication and translation. A study (and its characteristics/findings) may not be particularly generalizable, but still meet local end-user decision making needs (particularly if it was specifically designed for or conducted with them). In the review the authors state that "An understanding of the external validity of such interventions is vital to address their potential for translation and scalability, as well as replication efforts" – I think what I am looking for is some text in the introduction stating how and why that's the case.

3. The manuscript presents a large list of factors important in assessment of trial generalisability. I wondered if there was any evidence regarding which of this comprehensive list, were most important/useful. The authors employed comprehensive systematic review methods in the identification of trials, including searching multiple data-bases, prospective registration and duplicate screening. While not expressed in as much detail in this manuscript, the authors provide references where such detail is available.
4. The external validity assessment tool used also appears comprehensive and appropriate for the study objectives. While the items in the tool are described in the tables, it would be nice if the authors could characterise the tool a little more in the methods text, to summarise the information it seeks to extract.
5. Data were extracted by one author (MH) using a pre-piloted data extraction tool (see extended data<sup>33</sup>), with 20% double-checked by a second reviewer. Are the authors able to report the inter-rater agreement where duplicate extraction was undertaken.
6. The results are clearly presented in table and text form.
7. The discussion does a good job of summarising and interpreting the key findings of the review, and appropriately acknowledges the study limitations.
8. As the authors state, the focus of the majority of studies was on establishing efficacy rather than effectiveness or how such interventions could be scaled up and translated into routine practice. I did wonder if efficacy trials which are explicitly and purposefully undertaken in ideal research conditions, and so are not designed with the intent of providing knowledge generalizable to routine clinical settings, should be held to the same standard of such reporting as studies (e.g effectiveness trials) which are intended to provide more externally valid findings. Perhaps this is worth discussing.

Similarly, at first glance the external validity tool used in this study appears a high bench mark for researchers to achieve given existing demands on research participants for data collection, word restrictions in peer reviewed journals, and research funding limits. I'm not discounting the importance of reporting information to allow an assessment of generalisability, but can you offer some recommendations as to how researchers can best manage this?

**Are the rationale for, and objectives of, the Systematic Review clearly stated?**

Yes

**Are sufficient details of the methods and analysis provided to allow replication by others?**

Yes

**Is the statistical analysis and its interpretation appropriate?**

Not applicable

**Are the conclusions drawn adequately supported by the results presented in the review?**

Yes

**Competing Interests:** No competing interests were disclosed.

**I confirm that I have read this submission and believe that I have an appropriate level of expertise to confirm that it is of an acceptable scientific standard.**

-----