McCrabb Sam (Orcid ID: 0000-0002-4216-0251)

Wolfenden Luke (Orcid ID: 0000-0002-6178-3868)

ABSTRACT

Maximising the benefits of investments in obesity research requires effective interventions to be adopted and disseminated broadly across populations (scaled-up). However, interventions often need considerable adaptation to enable implementation at scale, a process that can reduce the effects of interventions. A systematic review was undertaken for trials that sought to deliver an obesity intervention to populations on a larger scale than a preceding RCT that established its efficacy. Ten scaled-up obesity interventions (six prevention, four treatment) were included. All trials made adaptations to interventions as part of the scale-up process, with mode of delivery adaptations being most common. Meta-analysis of BMI/zBMI from three prevention RCTs found no significant benefit of scaled-up interventions relative to control (SMD=0.03; 95% CI: -0.06, 0.12, p=0.510 – I^2 = 0.0%). All four treatment interventions reported significant improvement on all measures of weight status. Pooled BMI/zBMI data from prevention trials found significantly lower effects among scaled-up intervention trials than those reported in pre-scale efficacy trials (SMD=-0.11; 95% CI: -0.17, -0.04, p=0.002 – I^2 = 0.0%). Across measures of weight status, physical activity/sedentary behaviour and nutrition, the effects reported in scale-up interventions were typically 75% or less of the effects reported in pre-scale-up efficacy trials. The findings underscore the challenge of scaling-up obesity interventions.

This is the author manuscript accepted for publication and has undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi: 10.1111/obr.12845

1| BACKGROUND

Since 1975 global rates of overweight have almost tripled, increasing the risk of a variety of preventable diseases, such as cardiovascular disease, cancer, and diabetes.[1-4] Interventions to reduce modifiable risk factors for obesity have been recommended to mitigate its adverse effects. Specifically, community-based approaches to reduce the risk of unhealthy weight gain have been suggested to be particularly beneficial, given their capacity to reach large numbers of those in the community who could benefit from obesity related services. As physical inactivity and diet are the primary drivers of unhealthy weight gain, they represent key targets for obesity prevention and treatment interventions.

Decades of research have identified a range of effective community-based interventions to reduce unhealthy weight gain and improve physical activity and diet.[5-7] However, maximising the benefits of investments in obesity research requires effective interventions to be adopted and disseminated broadly across populations, that is, scaled-up. Scale-up is defined by the World Health Organization (WHO) as *"deliberate efforts to increase the impact of health service innovations successfully tested in pilot or experimental projects to benefit more people and to foster policy and program development on a lasting basis."*[8] Whilst there have been calls from international agencies, including the WHO,[9] to scale-up the implementation of obesity interventions, and a range of frameworks available to do so,[8,10,11] the process of scaling-up interventions is complex. Many interventions as they were originally developed and tested are not suitable for delivery at scale and need considerable adaptation to enable population wide implementation.[12,13] Specifically, intervention adaptations are often required to accommodate broader population groups, to better align with existing delivery infrastructure, or improve reach. Such adaptations have been hypothesised to lead to a 'scale-

up penalty'[14,15] or an intervention 'voltage drop'[16] whereby the effects of interventions are reduced following scale-up.

Assessing the effect size of interventions delivered at scale is important to determine whether significant investments in their implementation are achieving worthwhile benefits to the community. Similarly, comparing the effects of interventions delivered at scale with those achieved during trials to establish their efficacy is useful to assess the extent to which any adaptations occurring as part of the scale-up process (or the broader context in which the intervention is being delivered), may impact on its effectiveness. That is, understanding the typical magnitude of any scale-up penalty can assist policy makers to appraise the likely impact of interventions delivered at scale, before significant investments in their population wide delivery occur. However, few studies have examined the impact of the scale-up of obesity interventions. A series of Cochrane reviews have reported the effect of strategies to improve the implementation of nutrition, physical activity, and obesity prevention interventions delivered at scale in community settings and found their impact on measures of health behaviour or adiposity to be equivocal.[17-19] However, these reviews did not identify whether interventions were scaled-up from earlier trials with established efficacy, and so did not describe any changes in the effects of interventions delivered at scale (relative to pre-scale assessments) or any program adaptations. In 2016, Reis et al.[20] conducted a review of physical activity interventions which had been scaled-up without the support of researchers. The effects of the 16 programs identified as relevant were not systematically synthesised and many did not originate from efficacy trials. Therefore comparison of their effect relative to pre-scale evaluations was not possible. In the field of diabetes, a community-based prevention program was effective in achieving weight loss of approximately -4.7kg.[21] However, the

effect of subsequent scale-ups of the program in Finland and Australia was reduced to an average of -4.5kg and -2.7kg at 12 months, respectively.[22,23]

Evidence regarding the effects of scaled-up obesity interventions in community settings has not been the subject of a systematic evidence synthesis. As such, uncertainty remains regarding: the type of adaptations typically made to an intervention as it transitions from a controlled research environment to a large scale, real-world enterprise; the potential real-world impact of such initiatives; and the magnitude of any scale-up penalty. In this context, the objectives of this review were to:

1) Describe adaptations to obesity interventions occurring as part of the scale-up process;
2) Assess the effects of evidence-based obesity interventions on comparable measures of weight status, physical activity (including sedentary behaviour), and nutrition following scale-up; and
3) Describe differences in effects of interventions established prior to and following scale-up (scale-up penalty) for comparable measures of weight status, physical activity (including sedentary behaviour), and nutrition.

2 | METHODS

To address the study aims, a systematic review was undertaken of peer reviewed and grey literature. The review methods were developed using guidance from the Cochrane handbook,[24] and were prospectively registered in PROSPERO (ID: CRD42018105785).

2.1 | Criteria for including and excluding studies

2.1.1 | Types of study designs

We included intervention studies of any design (including randomised, controlled before and after trials, and non-controlled before and after designs) examining the effects of a scaled-up obesity intervention on any measure of weight status (e.g., weight, BMI, percent body fat), physical activity (including sedentary behaviour), and/or nutrition. Interventions could focus on obesity prevention or weight treatment in children or adults and needed to include a measure of weight status, physical activity/sedentary behaviour, or nutrition as its primary trial outcome. To be included, trials must have reported the results of interventions which had been established as efficacious on any measure of weight status (as defined in the trial) from a previously published randomised controlled trial (RCT) ('efficacy' study). Studies were defined as efficacious if the results of the trial found a significant effect ($p<0.05$) on at least one measures of weight status between intervention and comparison trial arms. Included trials may, or may not have also been efficacious in improving other obesity related behaviours. As such, while trials of interventions that had been scaled-up could employ randomised, non-randomised or designs without a control group, a preceding trial of the intervention must have been undertaken demonstrating the efficacy of the intervention using a randomised design. Data was extracted from pairs of 'efficacy' trials and the 'scaled-up' trials in this review. There were no restrictions on the language of publication. Studies published in the peer reviewed and grey literature were eligible.

2.1.2 | Populations

We included interventions undertaken in community, non-clinical settings (e.g., schools, workplaces, or sport and recreation centres). Trials of interventions targeting populations with pre-existing medical conditions or obesity related co-morbidities were excluded. Intervention programs must be delivered direct to community populations, or within community settings

such as schools, childcare services, or recreational centres. Studies or strategies to scale interventions in medical settings such as hospitals or general practice were excluded (none were found).

2.1.3 | Types of interventions

Trials were categorised as either efficacy, effectiveness, or implementation or dissemination trials.[25] We excluded trials where the primary purpose was replication in the same translation stage, that is, the conduct of an efficacy trial to replicate the findings of a prior efficacy trial. We included trials that intentionally sought to deliver an intervention to populations on a larger scale (e.g., a greater number of individuals in the target population) than a preceding RCT that established its efficacy. This means both vertical and horizontal scaling interventions were included,[11] as well as interventions that had been 'scaled-out' (delivered to new populations and/or through different delivery systems from the efficacy trials).[12] There was no criteria regarding the absolute or relative increase in scale required of scale-up trials. Scale-up trials where the intervention was delivered to a greater number of the target population yet included a smaller number of participants in their evaluation, relative to corresponding efficacy trials (e.g. trials including nested evaluations), were also included.

2.1.4 | Types of outcome measures

Outcome measures were any measure of weight status (e.g., Body Mass Index [BMI], waist circumference or other measures), physical activity (including sedentary behaviour), and/or nutrition. Such measures could be derived from any data source including direct observation, questionnaire, or anthropometric or biochemical assessments.

2.2 | Search strategy

A comprehensive search strategy was undertaken to identify included studies. Firstly, MEDLINE, EMBASE, PsycINFO, Cochrane Central Register of Controlled Trials (CENTRAL), CINAHL, Education Resources Information Center (ERIC) electronic databases were searched. Search terms to identify scaled-up interventions were developed based on terminology used in reviews by Milat et al.,[26] Reis et al.,[20] and Charif et al.[27] and combined with published search filters for physical activity/sedentary behaviour,[28] nutrition,[17,29] and obesity.[17,30,31] Search terms are described in detail in Supplement file 1. Additionally:

1) Review authors contacted the corresponding authors of those interventions identified during full text screening which, if scaled-up, would have been eligible for inclusion to confirm: i) if their intervention had been subsequently scaled-up; and ii) whether the effects of the scale-up had been evaluated.

2) Corresponding authors of trials included in key systematic reviews on obesity prevention or treatment for children and adults[32-35] were contacted by email to assess if, to their knowledge: i) their intervention had been subsequently scaled-up; and ii) whether the effects of the intervention following scale-up were evaluated.

3) Studies included in reports on the processes or outcomes of health promotion interventions that have been scaled-up were checked for eligibility, with potentially eligible studies screened for inclusion.[20,25] Key individuals were contacted from: the WHO; the WHO Collaborating Centre for Physical Activity, Nutrition and Obesity; the New South Wales Ministry of Health; the National Cancer Institute, National Institute of Health; as well as general enquiries at Public Health England, and the Division of Nutrition, Physical Activity and Obesity at the Centers for Disease Control and Prevention (CDC) to request if they knew

of any other intervention which could be eligible for inclusion in the review. The review team assessed publications provided, or sought to locate via web searches for publications (peer reviewed or grey literature) of evaluations of obesity interventions identified as potentially eligible using details provided by contacted experts.

2.3 | Criteria for determination of independent findings

2.3.1 | Study selection

Two review authors, not blinded to the author or journal information, independently screened abstracts and titles. Google translate was used to assess the eligibility of abstracts not published in English. Full texts of manuscripts were obtained for all potentially eligible trials for further examination. Translation was sought for full texts of manuscripts not published in English. Decisions regarding study inclusion were made via consensus between review author pairs. Where consensus could not be reached, or where it was uncertain, decisions regarding eligibility were discussed with a third reviewer. Information regarding the primary reason for exclusion of full texts manuscripts were recorded.

2.3.2 | Data extraction and management

Data extraction occurred independently and in duplicate by review authors who were also unblinded to author or journal information. Discrepancies between review author pairs regarding data extraction were resolved via consensus or via a third reviewer if required. Data were extracted from both the scaled-up intervention trial, and the RCT on which the scaled-up intervention was based (and established as efficacious). We extracted data to describe the following: a) the characteristics of included studies, such as the country and year of publication, sample, study design, trials measures and outcomes; b) the translation stage of the trial using the pathways to scaling-up public health intervention used by Indig et al.[25]

(efficacy, effectiveness, implementation or dissemination); c) the nature of any scaled-up adaptations to the intervention or implementation strategy characterised using the Adaptome;[36] and d) information to enable assessment of study quality and meta-analysis. One reviewer transcribed information from data extraction templates into tables reported in the manuscript. One review author extracted, transcribed and conducted meta-analysis.

2.4 | Risk of Bias

Risk of bias was assessed independently by two reviewers using the risk of bias tool Effective Public Health Practice Project (EPHPP) Quality Assessment Tool for Quantitative Studies.[37] This tool covers any quantitative study design and includes components of intervention integrity. The tool provides an overall risk of bias ('strong', 'moderate' or 'weak') assessment for each included study based on consideration of selection bias, study design, confounders, blinding, data collection methods, and withdrawals and drop-outs. Discrepancies were resolved by consensus or a third reviewer was approached to adjudicate.

2.5 | Data synthesis

The study characteristics and key findings of scaled-up trials were described in a characteristics of included studies table. Scaled-up trials were also classified as either effectiveness or real world trials, or implementation or dissemination trials based on descriptions of research stage and scale-up pathways from Indig et al.[25]

2.5.1 | Adaptations to interventions or implementation

We narratively synthesised adaptations to the intervention as reported in the original efficacy and subsequent scaled-up trials. As descriptions of interventions in research manuscript often include poor descriptions, supplementary Google, and Google Scholar searches were undertaken to identify other reports or web-based material (e.g., additional studies, program

manuals, or protocols) that provided additional information regarding interventions or how they may have been modified.[38] We used the Adaptome model[36] as our evidence synthesis framework to describe adaptations to interventions.[36] Specifically we classified any adaptation as:

i) Service setting adaptations – Adaptations to any element of the environment where the intervention delivery takes place. This includes the physical location of delivery, who delivers the intervention (e.g., level of education, professional background etc.), the source of financing, and/or collaborative approaches to delivery.

ii) Target audience adaptations – Any adaptations to the target audience(s) and/or adaptations which make the intervention more appropriate for the population of interest. An example includes changing the target audience altogether.

iii) Mode of delivery adaptations – Adaptations to the channel(s) used for the delivery of the intervention. Examples of these types of adaptations include those made to the frequency or duration of group sessions, to the interventions technological format, to the delivery channel itself (internet-based versus in-person), and/or to the training of a population responsible for intervention delivery e.g., interventionists, parents or teachers.

iv) Cultural adaptations – Adaptations made to better reflect cultural appropriateness of the intervention. Cultural adaptations facilitate greater compatibility between the intervention and a particular cultural group through modifying cognitive and/or observable cultural components.[39]

Other adaptations that could not be classified into the above categories suggested by the Adaptome[36] were coded as 'Other adaptations'.

2.5.2 | The effects of evidence-based obesity interventions following scale-up

The effects of all interventions included in the review were narratively synthesised. Additionally, pooled quantitative syntheses methods using BMI or standardised BMI scores (i.e., BMI z-scores [zBMI]) were employed in instances when there was sufficient data available as this was the most commonly used measure of weight status across included studies. To assess the effectiveness of scaled-up interventions, we attempted to include all studies reporting on BMI or zBMI in a meta-analysis. We conducted random effects meta-analysis on BMI and zBMI scores, with standardised mean differences calculated using Hedges *g*. We performed meta-analysis separately for studies assessing obesity prevention and treatment. Furthermore, the effects of scaled-up interventions on other measures of weight status, physical activity/sedentary behaviour, and/or nutrition were synthesised narratively.

2.5.3 | Differences in effects of interventions established prior to and following scale-up

Similarly, to assess the magnitude of the scale-up penalty in trials reporting on BMI and zBMI scores, a random effects meta-analysis was conducted on the differences in effects of interventions assessed in the context of an efficacy trial relative to that achieved and following scale-up. Differences in effects were calculated by subtracting the between group effect size following scale-up, from the between group effect size estimated from the original efficacy trial (pre-scale RCT). For this meta-analysis we sought to include all trials where an effects size for the intervention on measures of BMI (or zBMI) could be calculated for both the scaled-up trial, and the corresponding efficacy trials on which it was based. For scaled-up trials using an uncontrolled before and after trial design, the effect sizes, calculated as a change from baseline was subtracted from an effect size calculated as the change from baseline within the intervention group in the original efficacy trial. Where multiple time-

periods were provided (i.e., DOiT efficacy trial)[40] data from the time-period that most closely matched the follow-up period used in the scaled-up trial was analysed.

Furthermore, to assess differences in effects of scaled-up interventions on other measures of weight status, physical activity/sedentary behaviour, and/or nutrition, we reported the findings of all measures that were included, assessed and reported in a standardised way across both efficacy and scaled-up interventions. For each included study, we calculated the percent of the effect size reported in the efficacy trial that was retained in the scaled-up intervention trial using the following formula: effect size reported in the scaled-up trial divided by the effect size reported in the pre-scale efficacy trial, multiplied by 100. From this equation, values higher than 100% indicate the intervention tested in the scale-up trial had a greater effect than it did in the efficacy trial; a value of 50% indicates the intervention tested in the scale-up trial was half as effective than it was in the efficacy trial, while values less than 0% (negative) indicate that the direction of effect of the intervention tested in the scale-up trial was opposite of the direction in the efficacy trial. Such analyses were described narratively.

2.5.4 | Dealing with missing data

Authors of included trials were also contacted to provide additional information if any outcome data were unclear or missing. Where necessary data was neither reported nor provided, we sought to calculate effect sizes from the available information. Specifically, the following calculations were performed: i) where BMI at follow-up for adult samples were not provided, mean BMI was estimated using mean weight at follow-up and mean height reported at baseline and standard deviation (SD) was imputed using the baseline BMI SD (HeLP-her trial); ii) where mean BMI values were either not reported but regression coefficients for change from baseline (i.e., Mend 7-13 scaled-up study),[41] or change from baseline values

were provided (i.e., PALs-ATLAS efficacy trial),[42] these values were used along with mean baseline BMI values to estimate mean BMI follow-up values and the SD imputed using the baseline BMI SD; iii) where mean BMI values were reported separately for sub-groups (i.e., by gender in DOiT)[43] estimates were combined following recommended formula outlined in the Cochrane Handbook;[24] and iv) where 95% CI were reported rather than SDs, the SD was calculated using recommended formula.[44]

2.5.5 | Assessment of heterogeneity

The I^2 statistic and Chi^2 tests for statistical heterogeneity were assessed for all meta-analyses,[45] with scored over 75% considered high.[46] Due to the small number of prevention and treatment interventions identified, we did not explore heterogeneity via subgroup analyses. Statistical analyses were programmed using Stata v13.0 (StataCorp Ltd, College Station, TX).

3| RESULTS

The systematic database searched identified a total of 3,320 titles to screen for inclusion in this review (see Figure 1).

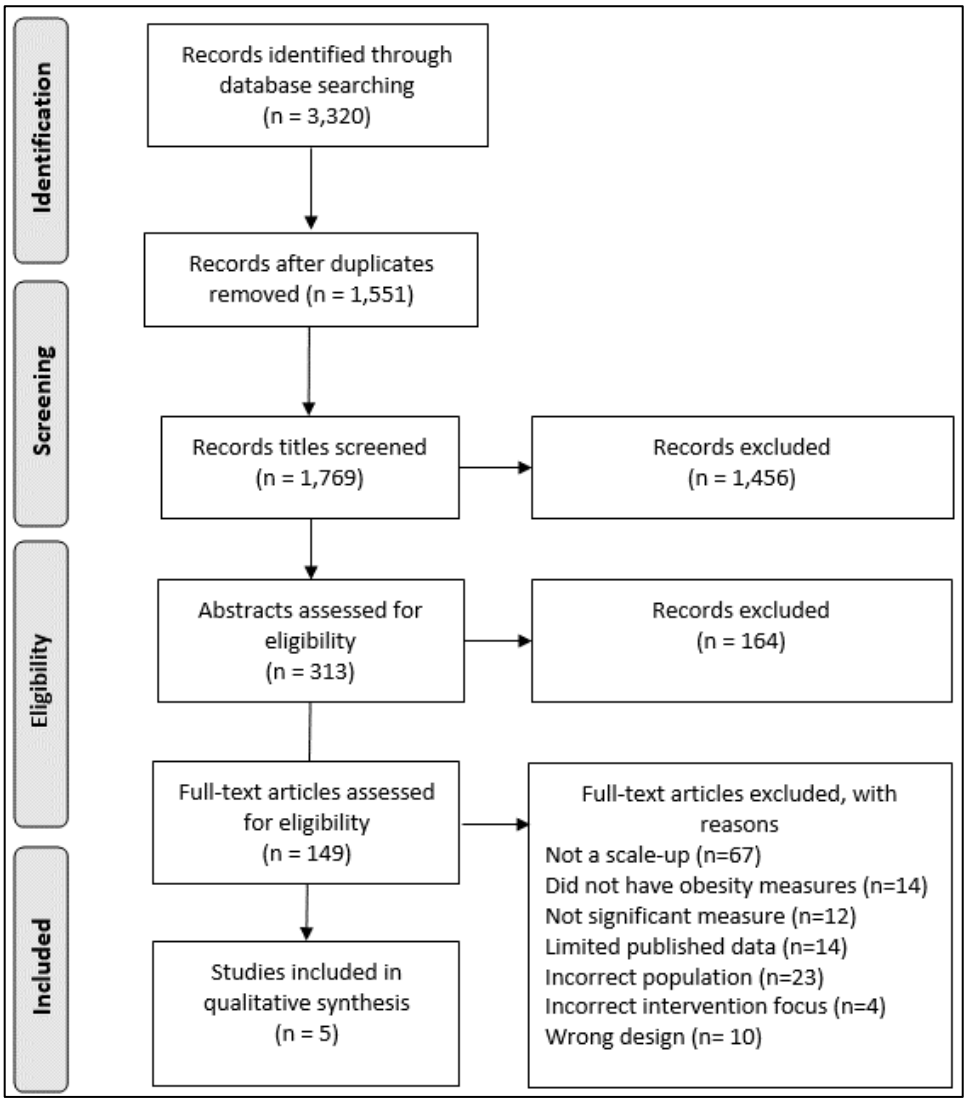

FIGURE 1. PRISMA diagram of included studies.

Additionally, corresponding authors of 301 trials and eight key experts or representatives from relevant international organisations were contacted. From these searches, initially 21 studies were identified as eligible, however 11 studies were subsequently excluded at the data-extraction phase for various reasons, details following. The efficacy trials for the PEACH QLD study[47] and the Active-For-Life study[48] reported within group improvement –

but no between group difference – on measures of weight status between the trial arms post-intervention;[49,50] Action Schools! BC,[51] TXT2BFiT,[52] and the YOG-Obesity[53] studies found no significant difference on a measure of weight status in the pre-scale RCT; Camp NERF[54] had no available data for the pre-scale RCT; CHAT[55] and PASOS had no follow-up outcome data (authors of these studies were contacted however no follow up data was able to be supplied);[54-56] the original pre-scale trial of EPODE[57] was a non-RCT design;[57] Healthy and Vital[58] did not include a measure of weight status, physical activity, or nutrition as its primary outcome measure; and finally, SHED-IT[59] tested an additional component in the larger trial where both control and intervention arms received SHED-IT[59] while the intervention arm also received a weight loss maintenance program.

In total, 10 scaled-up interventions were included in this review. Supplement file 2 lists the studies describing the scaled-up intervention and the pre-scale randomised efficacy trials.

3.1 | Characterises of included studies

Table 1 (characteristics of included studies) provides information on the 10 included scaled-up trials; four of which were conducted in Australia,[60-63] two each from the United Kingdom (UK)[41,64] and the United States of America (USA),[65,66] and one each from New Zealand (NZ)[67] and the Netherlands.[40]

Two of the scaled-up interventions included women only (≥40 years old and 18- 40 years old),[60,66] and one men only 30-65 years old.[64] The remaining trials focused on parent/child dyads,[41,62,65] preschool children (3-5 years old),[61] children (6-11 years old),[67] adolescents (aged 12-14 years),[40] and adolescent boys (in their first year of high school, aged approximately 11-12 years old).[63]

Four scaled-up trials used before and after design,[41,62,65,66] three cluster RCT,[61,63,68] and one study each employed an RCT,[64] historical comparison,[67] or a cluster controlled implementation trial.[40] Six studies[40,60,61,63,65,67] were classified as obesity prevention programs, the other four[41,62,64,66] were concerned with weight treatment (Table 1). Six[40,60,61,63-65] of the ten scaled-up interventions were classified as effectiveness trials, and four[41,62,66,67] as implementation or dissemination trials.

TABLE 1. Characteristics of included scaled-up interventions

Author, Year, Country Intervention name	Design Prevention/ Treatment	Setting	Population	Measures of weight status, physical activity/sedentary behaviour, and/or nutrition (What was measured and how)	Follow up time points	Key Findings	Translation stage
Obesity prevention interventions							
Smith, 2014, Australia[63] **ATLAS**	Cluster RCT	14 state-funded secondary schools within low-income communities in NSW, Australia	n = 361 Participants Eligibility: Adolescent boys in grade 7 (first year of secondary school)	Height (cm); weight (kg); BMI (kg/m²); zBMI; waist circumference (cm); % body fat using bioelectrical impedance; physical activity using accelerometers; self-reported screen-time (Adolescent Sedentary Activity Questionnaire); sugar sweetened beverage intake (2-item question); maximal strength (hand grip); muscular endurance (push-up); resistance training competency using video analysis	Intervention completion (8 months from baseline) and 18 months from baseline.	Weight status: No significant intervention effect for body composition (i.e., BMI, waist circumference and body fat percentage). Physical Activity/Sedentary Behaviour: Significant improvements in upper body muscular fitness in the intervention group. No significant intervention effect for overall physical activity. Nutrition: Significant reduction in sugar-sweetened beverage consumption in favour of the intervention.	Effectiveness
Van Nassau, 2014, the Netherlands[40] **DOiT**	Cluster-controlled trial	29 Prevocational Dutch Secondary schools	n = 1,486 participants (29 schools) Eligibility: Adolescents 12-14 years of age	Weight (kg); height (mm); BMI kg/m²; zBMI; waist circumference (cm); skinfold thickness measurements (mm); demographics, dietary and physical activity behaviour using adolescent questionnaires	Intervention completion (20- months)	Weight status: No significant intervention effect for measures of adiposity. Physical Activity/Sedentary Behaviour: Significant improvements in reported minutes of active transport in boys only (intervention subgroup). No significance overall, or in other measures of physical activity. Nutrition: Significant improvement in some intervention subgroups (i.e., reduced consumption of sugar-containing beverages in girls, and increased breakfast consumption in boys). No significance overall, or in other	Effectiveness

						measures of nutrition.	
Heerman, 2018, USA[65] **Healthier Families**	Pre-post study	Three Parks and Recreation Centres in three cities across the country (Michigan, Georgia and Nevada)	n= 26 families Eligibility: Parent (English speaking, ≥18 years of age) and child (3-5 years of age) dyads.	RE-AIM components- reach (community representativeness); effectiveness (pre-post qualitative surveys examining changes in healthy behaviours that are upstream mediators to BMI [i.e., fruits/veggies per day, days of child PA> 30 minutes]); adoption (number of facilities implementing the program & surveys with administrators and staff); implementation (fidelity); and maintenance (using key-informant interviews and participating parent questionnaires)	Intervention completion (12 weeks)	Physical Activity/Sedentary Behaviour: Significant increase in the percentage of families meeting physical activity goals. No significant change in parent-reports of their child's physical activity patterns. Nutrition: Significant increase in the percentage of families meeting healthy diet goals, and using healthy behaviour strategies (e.g., strategies to improve dietary intake). No significant change in parent-reports of their child's diet patterns.	Effectiveness
Lombard, 2016, Australia[60] **HeLP-her**	Cluster RCT	41 rural Australian towns	n= 649 participants Eligibility: Women 18-50 years of age residing in an eligible town	Height (cm); weight (kg); BMI; waist and hip circumference (cm); demographics, self-management, physical activity (international physical activity questionnaire; IPAQ long form), dietary intake (Cancer Council Australia Food Frequency Questionnaire) and health status (by Saelens et al.)	Intervention completion (12 months)	Weight status: Significant decrease in weight in favour of the intervention. No significant intervention effect for waist circumference or waist-hip ratio. Physical activity/Sedentary Behaviour: No significant intervention effect for physical activity. Nutrition: Significant increase in dietary self-management in the intervention group. No significant intervention effect for diet quality.	Effectiveness
Hardy, 2010, Australia[61] **Munch and Move**	Cluster Pre-post randomised controlled trial	29 Preschools in Sydney, NSW, Australia	n = 430 participants (29 preschools). Eligibility: Children attending preschool in Sydney, Western	Demographics using parent questionnaires; nutritional quality of children's lunches using lunch box audits; fundamental movement skills using Test of Gross Movement Development (TMGD-2)	Pre- (May/June 2008) and post-intervention (November	Nutrition: Significant decrease in sugar sweetened drinks provided in children's lunch boxes compared with the control group. No significant intervention effect on other factors of lunch box nutritional quality.	Effectiveness

			Sydney and South Western Sydney	checklist; preschool policies and practices using director interviews and staff questionnaires	2008)		
Rush, 2014, NZ[67] **Project Energize**	Compare results to historical comparison	193 Primary schools in Waikato, New Zealand	n = 4,804 participants (193 Primary schools) Eligibility: Younger Children ages 6-8 years (n = 2,474) Older Children ages 9-11 (n= 2,330)	Height (cm); weight (kg); BMI (kg/m²); cardiorespiratory endurance using 550m run tests; parent reported ethnicity; socioeconomic status using school deciles from a national register	2011 (up to six years of engagement)	Weight status: No significant difference in weight and BMI. Physical Activity/Sedentary Behaviour: Significant decrease in time taken to complete 550m run test	Dissemination/ implementation
Obesity treatment interventions							
Wyke, 2018, UK[64] **EuroFIT**	RCT	15 football clubs in the Netherlands, Norway, Portugal and the UK (England)	n = 1,113 participants Eligibility: Men 30-65 years of age, Gender: Male 100%; BMI ≥27 kg/m²	Height (mm); weight; BMI; waist circumference; SBP and DBP; cardio-metabolic blood biomarkers; food intake; self-reported physical activity, sedentary time, wellbeing, self-esteem, vitality, and quality of life	Post program and 12 months from baseline	Weight status: Significant improvement in self-reported weight in favour of the intervention. Physical Activity/Sedentary Behaviour: Significant increase in step count in favour of the intervention. No significant intervention effect for time spent sitting. Nutrition: Significant improvement in self-reported food intake in favour of the intervention.	Effectiveness
Hardy, 2015, Australia[62] **Go4Fun**	Pre-post non-controlled design	15 local health districts across NSW	n = 2,812 participants Eligibility: Children ages 6-15 years, BMI ≥ 85th percentile and a parent/carer	Height (m), weight (kg); BMI; zBMI; waist circumference (cm); waist to height ratios; lifestyle (physical activity, dietary habits, screen time) using parent questionnaires; cardiovascular fitness using validated 3-min step test with 1-min recovery heart rate;	Intervention completion (10 weeks)	Weight status: Significant improvement in BMI, zBMI, waist circumference and waist: height ratio. Physical Activity/Sedentary Behaviour: Significant improvement in measures of physical activity (i.e., increased moderate-intensity physical activity, reduced screen time	Dissemination/ implementation

				self-esteem using a modified Rosenberg Self-Esteem scale		and improved cardiovascular fitness). Nutrition: Significant improvement in daily servings of fruit and vegetables, and frequency of unhealthy food consumption. No significant change in understanding nutrition labels.	
Fagg, 2014, UK[41] **MEND 7-13**	Intervention evaluation using prospective data	All regions of England	n = 13,998 participants Eligibility: Overweight children 7-13 years of age and a parent/carer	Height; weight; BMI; zBMI; psychological distress using Strengths and Difficulties Questionnaire; self-esteem using a modified Rosenberg Self-Esteem scale; parent reported socioeconomic status and ethnicity	Intervention completion (10 weeks)	Weight status: Significant decrease in BMI and zBMI.	Dissemination/ implementation
Folta, 2015, USA[41] **Strong Women-Healthy Hearts**	Pre-post-test within-participants design	22 American States	n= 345 participants Eligibility: Women ≥40 years of age, BMI ≥24 kg/m², sedentary	Leader and site characteristics and self-efficacy. RE-AIM components- reach (and representativeness); effectiveness (change in body weight [kg], fruit and vegetable consumption [5-A Day for better health 7-item screener], physical activity [IPAQ Short form]); adoption; implementation (including cost); and maintenance	Intervention completion (12 weeks)	Weight status: Significant decrease in weight. Physical Activity/Sedentary Behaviour: Significant increase in mean metabolic equivalent (MET)-minutes per week. Nutrition: Significant increase in daily fruit and vegetable consumption.	Dissemination/ implementation

3.2 | Adaptation occurring as part of the scale-up process

The most common type of adaptation for prevention programs (five[40,60,61,63,65] of six)[40,60,61,63,65,67] and obesity treatment interventions (three[62,64,66] of four)[41,62,64,66] were mode of delivery adaptations (Table 2, also see Supplement File 3 for description of adaptations). Mode of delivery adaptations included increases (four prevention),[40,61,63,65] and reductions (three prevention,[60,63,65] two treatment interventions)[62,64] in the number of intervention element, in the frequency or duration of intervention contacts, or in the training and resources provided to those responsible for intervention delivery. Some trials included adaptations which increased some intervention elements while reducing the intensity of others (three prevention,[40,63,65] one treatment intervention);[64] and one treatment trial was judged to represent neither an increase nor reduction in intensity.[66] In a number of instances, mode of delivery adaptions were quite substantial. For example, the DOiT scaled-up intervention consisted of 16 mandatory and three optional in-school lessons over two years compared with just 11 sessions over one year in the efficacy trial;[40] and the HeLP-her scaled-up intervention, originally comprised of four group self-management sessions in the efficacy trial, reduced to one session.[60]

While target audience adaptations were not evident in scaled-up obesity treatment interventions, they were present in three of the prevention interventions. These included changes in the delivery of interventions to population groups that differed based on their demographic characteristics, as well as a shift in those targeted to support delivery of the intervention.[60,65,61] Cultural adaptations were made in one prevention[65] and two treatment interventions:[62,64] the language of delivery changed from Spanish (original) to English in Healthier Families,[65] the Go4Fun program was adapted to reflect an Australian context,[62] and

the content of EuroFIT was culturally-sensitised to reflect local physical activity and nutrition related norms for each country.[64] Service setting adaptations were evident in four[40,60,61,63] of six[40,60,61,63,65,67] prevention and two[62,66] of four[41,62,64,66] obesity treatment programs. A number of other adaptations were made to support scaling-up, including the establishment of infrastructure to support large scale dissemination of programs. For example, MEND Central (a social enterprise responsible for assisting with intervention delivery) was established and considered important for national implementation support.[41] The DOiT program established a support office with a dedicated recruitment employee, and promotional material and activities to facilitate dissemination.[40]

TABLE 2. Adaptations made to efficacy trials prior to scaled-up for each intervention.

Program	**Adaptation for scale-up**				
	Service setting	**Target audience**	**Mode of delivery**	**Cultural**	**Other**
Obesity prevention interventions					
ATLAS[63]	X		X		X
DOiT[40]	X		X		X
Healthier Families[65]		X	X	X	X
HeLP-her[60]	X	X	X		X
Munch and Move[61]	X	X	X		
Project Energize[69]					X
Obesity treatment interventions					
EuroFIT[64]			X	X	X
Go4Fun[62]	X		X	X	X
MEND 7-13[41]					X
Strong Women-Healthy Hearts[66]	X		X		X

3.3 | The effects of evidence-based obesity interventions following scale-up

3.3.1 | Weight status

Significant improvements in weight relative to control were reported in the HeLP-her randomised trial.[60] No other significant improvements were reported on any other measure of weight status for any of the other included prevention trials. Meta-analysis of BMI/zBMI from three prevention RCTs[40,60,63] including data from 2,339 participants, found no significant benefit of scaled-up interventions relative to control (SMD= 0.03; 95%CI: -0.06, 0.12; p= 0.510 - I^2 = 0%) (Figure 2).

All four treatment interventions,[41,62,64,66] including one randomised trial,[64] reported significant improvement on all measures of weight status included in these studies. Data could be combined for two trials only.[41,62] Meta-analysis of BMI/zBMI from two pre-post trials assessing obesity treatment interventions found a significant benefit of scaled-up interventions relative to the pre-intervention assessment (SMD= -0.20; 95% CI: -0.30, -0.10; p< 0.001 - I^2 = 87.6%) (Figure 2).

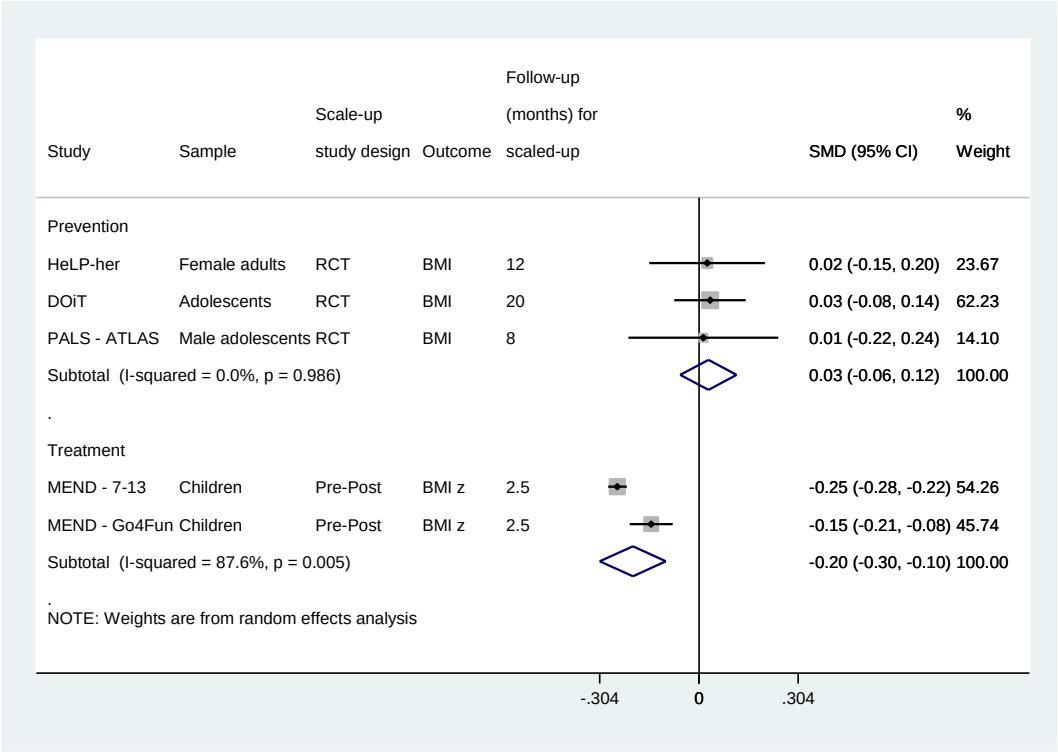

FIGURE 2. Forest plot of standardised mean differences in BMI and zBMI scores in scaled-up trials by intervention purpose and study design

3.3.2 | Physical Activity/Sedentary Behaviour

For prevention trials, the impact of scaled-up interventions was mixed. Of the three randomised trials reporting physical activity/sedentary behaviour measures, none reported significant improvements in overall measures. However, improvements were reported for minutes of active transport (in boys only).[40] Among the pre-post or historically controlled trials: the Healthier Families intervention found a significant increase in the percentage of families meeting physical activity goals set by the family while taking part in group sessions,[65] but not children's patterns of physical activity (measured in days of physical

activity greater than 30 minutes).Project Energize found significant reductions in time taken to run 550 metres[67] favouring participants in the intervention group (Table 1).

For three obesity treatment studies which measured physical activity/sedentary behaviour,[62,64,66] significant improvements in step count and metabolic equivalent (MET) minutes per week were found for EuroFIT[64] and StrongWomen-Healthy Hearts,[41] respectively. The third, Go4Fun,[62] reported significant improvements in self-reported physical activity. No improvements, however, were reported for measures of sitting time in the randomised trial of the EuroFIT intervention.[64] (Table 1).

3.3.3 | Nutrition

The effectiveness of scaled-up interventions in improving nutrition related outcomes were mixed. Among the five prevention trials reporting nutrition related outcomes,[40,60,61,63,65] three[40,61,63] controlled trials reported significant effects on sugar sweetened beverages: via decreased inclusion in child lunchboxes in Munch and Move,[61] reduced consumption in ATLAS,[63] and reduced consumption for girls only in DOiT[40]. With the exception of improved breakfast intake among boys in the DOiT[40] trial, no other measure of dietary intake improved significantly (Table 1). In the fourth RCT, the HeLP-her[60] study, no significant intervention effects were observed for measures of diet quality. In the only non-randomised trial to include nutrition related outcomes, the Healthier Families[65] intervention found a significant increase in the percentage of families meeting healthy diet goals, but not for parent reports of child's diet patterns.

Three obesity treatment studies measured nutrition.[62,64,66] The only randomised trial to do so, EuroFIT,[64] reported a significant improvement in self-reported food intake. Among trials of other designs, a significant increase in daily fruit and vegetable consumption was found for

Go4Fun[62] and StrongWomen-Healthy Hearts.[41] Go4Fun[62] also reported a significant change in frequency of unhealthy foods consumed (Table 1).

3.4 | Differences in effects established prior and following scale-up (scale-up penalty)

One intervention did not include measures of the same outcomes across pre-scale efficacy trials and trials of scaled-up interventions,[65] and one did not provide enough information to enable comparable assessment of effects.[64] For those interventions which provided sufficient data, the effect size reported in the scaled-up trials were typically less than 75% that of the reported in efficacy trials (Table 3).

TABLE 3. Effectiveness of intervention pre and post scale-up with proportion of original effect size

Post-scaled intervention name	Effectiveness measures common to both interventions for weight status, physical activity/sedentary behaviour and nutrition*		
	Pre-scale measure	**Post-scale measure**	**Scale-up penalty: proportion (%) of the efficacy trial effect size achieved in the scaled-up trial of the intervention (effectiveness of scaled-up/pre-scale)**
Obesity prevention interventions			
ATLAS[63]	RCT Mean between groups difference at 6 months follow-up <u>Weight status</u> Waist (Umb) (cm) 0.8, p= 0.23 BMI (kg/m^2) -0.8, p<0.001 BIA (body fat %) -1.8, p= 0.04 <u>Physical Activity/Sedentary Behaviour</u> Push-up test (reps) 1.7, p= 0.09 <u>Nutrition</u> SSB consumption (250ml) -1.17, p = 0.02	RCT Adjusted between groups difference at 8 months follow-up <u>Weight status</u> Waist circumference (cm) 0.5+/-0.45 p=0.16 BMI (kg/m^2) 0.0 +/-1.2 p= 0.84 Body fat % 0.0+/- 0.48 p=0.99 <u>Physical Activity/Sedentary Behaviour</u> Push-ups (reps) 0.9 +/- 0.49, p=0.04 <u>Nutrition</u> SSB consumption (glasses/day) -0.6 +/- 0.26,p =0.01	62.5% -^ -^ 52.9% 51.3%
DOiT[40]	RCT Between group difference at 20 months follow-up <u>Nutrition</u> SSB consumption for girls (mL/day) at 20 months: -88	RCT Between group difference at 20 months follow-up <u>Nutrition</u> SSB consumption for girls (mL/day) at 20 months: -188.2	213.9%
HeLP-her[60]	RCT Adjusted between group differences at 12 months follow-up <u>Weight status</u>	RCT Adjusted between group differences at 12 month follow-up <u>Weight status</u>	

	Weight (kg): -1.13, p<0.05 Waist circumference (cm): -1.51 <u>Physical Activity/Sedentary Behaviour</u> Activity self-management: 0.24, p<0.001 <u>Nutrition</u> Energy (kJ/day): -198 Nutrition self-management: 0.18, p=0.02	Weight (kg): -0.87, p=0.02 Waist circumference (cm): -0.96 <u>Physical Activity/Sedentary Behaviour</u> Activity self-management: 0.06 <u>Nutrition</u> Energy intake (kJ/day): -43 Nutrition self-management: 0.08, p<0.05	77% 63.6% 25% 21.7% 44.4%
Munch and Move[61]	Pre- post-intervention Adjusted between group differences at 12 months follow-up <u>Physical Activity/Sedentary Behaviour</u> FMS global score: 14.79, p<0.0001	Cluster pre-post RCT Adjusted group between group differences at 6 months follow-up <u>Physical Activity/Sedentary Behaviour</u> FMS global score: 5.33, p= 0.003	 36%
Project Energize[67]	RCT Adjusted between group differences at 2 years follow-up <u>Weight status</u> BMI SDS (younger): 0.00, p= 0.96 BMI SDS (older): 0.05, p= 0.29	Compare results to historical comparison Same population as the original RCT Adjusted between group differences at the time of evaluation, schools had been enrolled from as few as six school terms (1.5 years) to as many as twenty-six school terms (6.5 years) <u>Weight status</u> BMI SDS (younger): -7.0 BMI SDS (older): -4.4	 -^ -8,800%
Obesity treatment interventions			
Go4Fun[62]	RCT Within subjects change 0-6 months follow-up <u>Weight status</u> Waist circumference (cm): -4.2, p<0.0001	Pre-post non-controlled design Pre-post mean change in outcomes for all participants at 10 weeks follow-up <u>Weight status</u> Waist circumference (cm): -1.83, p<0.001	 43.6%

	BMI (kg/m^2): -1.0, p<0.0001 zBMI: -0.30, p<0.0001 <u>Physical Activity/Sedentary Behaviour</u> Recovery heart rate (beats/min): -17.9, p< 0.0001	BMI (kg/m^2): -0.65, p<0.0001 zBMI: -0.11, p<0.001 <u>Physical Activity/Sedentary Behaviour</u> Recovery heart rate (beats/min): -4.64, p=0.002	65% 36.7% 25.9%
MEND 7-13[41]	RCT Within subjects change 0-6 months follow-up <u>Weight status</u> BMI (kg/m^2): -1.0, p<0.0001 zBMI: -0.30, p<0.0001	Pre-Post Regression coefficients for change in BMI and zBMI post intervention completion <u>Weight status</u> BMI (kg/m^2): -0.76, p<0.0001 zBMI: -0.18, p<0.0001	 76% 60%
Strong Women - Healthy Hearts[66]	RCT Mean pre-post change for intervention group at 12 week follow-up <u>Weight status</u> Body weight (kg): -1.8 <u>Nutrition</u> Serves fruit and vegetables: 0.2	Pre-post-test within-participants design Post-intervention 12 week follow-up <u>Weight status</u> Body weight (kg): -0.5, p<0.001 <u>Nutrition</u> Serves fruit and vegetables: 2.1, p<0.001	 27.8% 1050%

* p-values not presented if not provided in the citing manuscript
^ Proportions cannot be reported as mathematical equation invalid

3.4.1 | Weight status

Three prevention interventions[60,63,67] provided comparable assessment of weight status measures across pre-scale efficacy and scaled-up trials. The proportion of the pre-scale effect size achieved in the scaled-up trials was highly variable across reported weight status measures, but were typically less than 80% of the effect achieved in pre-scale interventions. Pooled BMI/zBMI data from the three prevention trials found significantly lower effects among scaled-up intervention trials than those reported in pre-scale efficacy trials (SMD= -0.11; 95% CI: -0.17, -0.04;p= 0.002 - I^2 = 0%;) (Figure 3).

Three weight treatment trials[41,62,66] reported comparable weight status measures. The effects reported on these measures in scaled-up intervention trials were typically less than 75% of that reported in pre-scale efficacy trials. Pooled BMI/zBMI data from the two prevention trials found no significant differences between the effect of scaled-up and pre-scale efficacy trials (SMD= -0.11; 95% CI: -0.25, 0.04; p=0.147 - I^2 = 0%).

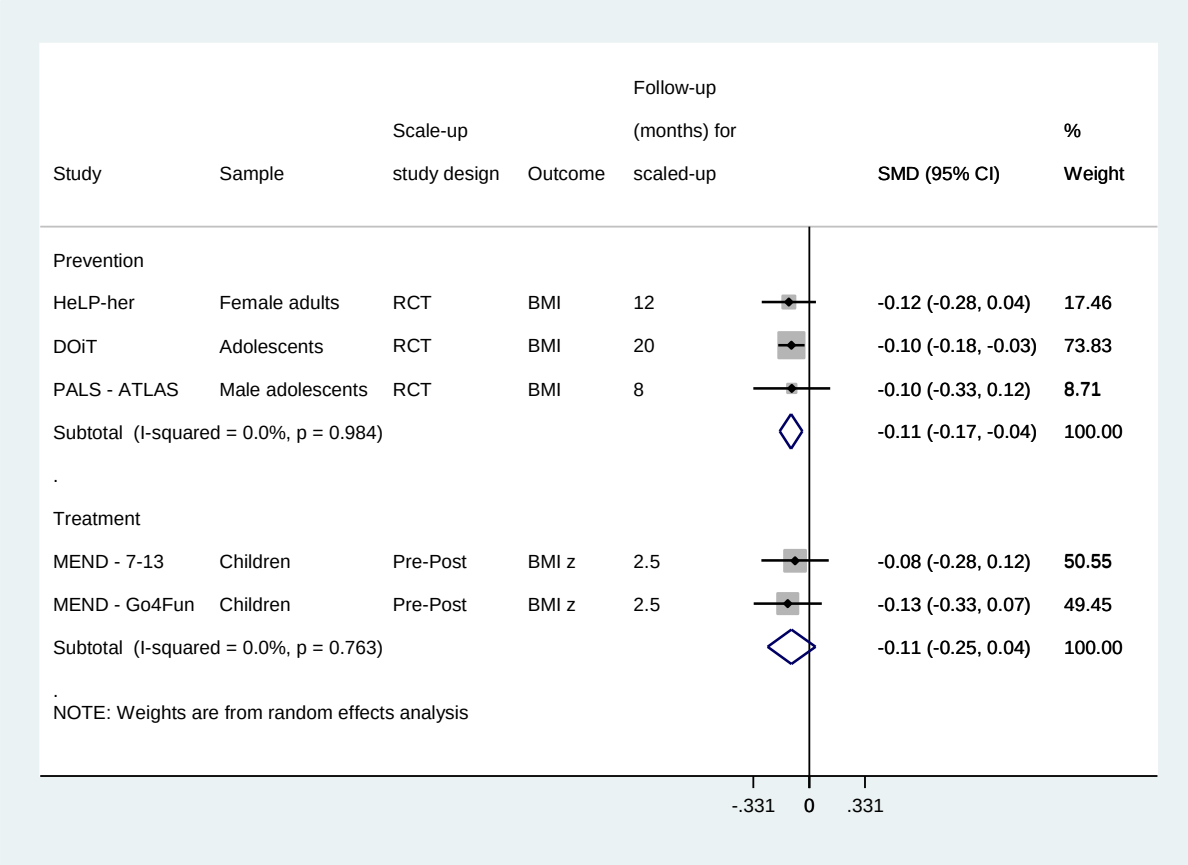

FIGURE 3. Forest plot of differences in effect sizes observed between scaled-up trials and their corresponding efficacy trials by intervention purpose and design of scaled-up study.

3.4.2 | Physical Activity/Sedentary Behaviour

Three prevention[60,61,63] and one weight treatment intervention[62] provided comparable measures of physical activity/sedentary behaviour across efficacy and scaled-up trials. In all cases the effect sizes reported in scaled-up trials of the intervention were lower: between 53% and 25% of that achieved in pre-scale efficacy trials.

3.4.3 | Nutrition

Three prevention[40,60,63] and one weight treatment intervention[66] provided comparable measures of nutrition related outcomes across efficacy and scaled-up trials. The effect sizes

reported in scaled-up trials of prevention interventions were variable but less than 55% of that achieved in pre-scale efficacy trials on all but one measure. The effect size on the nutrition measures in the scaled-up obesity treatment trial was higher than that in the pre-scale RCT.[66]

3.5 | Risk of bias

Using the EPHPP tool, [37] there were no studies assessed as being strong in quality, with two articles considered moderate in quality[61,63] and the remaining eight assessed as weak in quality (Table 4).[40,41,60,62,64-67] The quality ratings section related to *study design* and *confounders* achieved the most consistent number of strong ratings with four and five respectively. The quality ratings section related to *selection bias* and *blinding of assessments* resulted in the most consistent weak ratings, with eight and seven weak ratings respectively. Overall, the study quality was weak.

TABLE 4. Risk of bias rankings based on the Effective Public Health Practice Project (EPHPP) Quality Assessment Tool for Quantitative Studies

	Selection bias	Study design	Confounders	Blinding	Data collection methods	Withdrawal and drop out	Overall assessment
Obesity prevention interventions							
ATLAS[70]	Weak	Strong	Strong	Moderate	Strong	Strong	Moderate
DOiT[40]	Weak	Moderate	Strong	Weak	Strong	Moderate	Weak
Healthier Families[63]	Weak	Moderate	Weak	Weak	Weak	Moderate	Weak
HeLP-her[68]	Weak	Strong	Weak	Strong	Moderate	Weak	Weak
Munch and Move[61]	Weak	Strong	Strong	Moderate	Weak	Strong	Moderate
Project Energize[67]	Weak	Moderate	Strong	Weak	Moderate	N/A	Weak
Obesity treatment interventions							
EuroFIT[64]	Moderate	Strong	Weak	Weak	Moderate	Strong	Weak
Go4Fun[62]	Moderate	Moderate	Strong	Weak	Weak	Strong	Weak
MEND 7-13[41]	Weak	Moderate	Moderate	Weak	Weak	Moderate	Weak
Strong Women- Healthy Hearts	Weak	Moderate	Weak	Weak	Weak	Weak	Weak

4| DISCUSSION

To our knowledge, this is the first systematic review to examine the effects of obesity interventions that have been scaled-up. The review found a relatively small number of trials that varied in the types of interventions delivered, the reach of the scaled-up interventions, and the evaluation methods used to assess their effects. The review found that adaptations of interventions that had been trialled and found to be efficacious was common as part of the scale-up process, with all included studies reporting adaptations of some kind. Scaled-up, obesity treatment, but not prevention, interventions typically yielded significant improvements on measures of weight status. For both prevention and obesity treatment studies, relative to their preceding efficacy trials, there were substantial reductions in the effect size of interventions on measures of weight status, physical activity/sedentary behaviour, and nutrition outcomes reported for scaled-up trials. Specifically, the findings of scaled-up interventions typically represented less than 75% of the effect established in efficacy trials of the intervention.

The key findings of the review are comparable with previous research in related fields. Diabetes Prevention Programs achieved significant improvement in measures of weight status following scale-up, however the effects of these were less than achieved in preceding trials.[70] Trials of other physical activity and nutrition interventions that have been scaled-up have reported mixed findings regarding their effects at scale, but are consistent in reporting reductions in the magnitude of any effects relative to preceding efficacy trials[51,71] - findings similar to this review. While modest improvements are discernible for weight treatment interventions, evidence from this review suggest that current efforts to scale-up obesity prevention interventions may be unlikely to be having a beneficial impact on community

weight status. A number of frameworks exist to guide program scale-up,[11,26,72,73] however, given the importance of delivering interventions at scale in achieving population level improvements in health, further research to better understand and improve scale-up methods that preserve the effect of interventions is required.

All interventions were adapted before they were scaled-up. Interventions developed and trialled under optimal research conditions are often complex; intensive; delivered by experts; and require knowledge, skills, and resources that are not routinely available in more real-world settings. As such, policy makers and practitioners often have to pare back the intervention to ensure it is more amenable for implementation in real world contexts, a process hypothesised as responsible for reducing the effectiveness of interventions (scale-up penalty). Adaptations that both increased and reduced the intensity of interventions, however, were evident in this review. Further, reductions in the effect size of prevention interventions were evident across all trials including those where substantial enhancement to the intensity of the intervention were made as part of the scale-up process.[40,63] The existence of a scale-up penalty in these instances may be due to other factors. For example, difficulties in implementing the program, differences between study populations, or differences in trial methods between pre-scale efficacy trials and scaled-up trials of these interventions. Previous reviews of obesity research, for example, have smaller effect sizes among explanatory (i.e., efficacy) trials relative to trials employing more pragmatic methods.[74]

A number of limitations of the review methods are worth particular consideration. While the review used comprehensive search methods (e.g., searches of electronic databases; contacting study authors, institutions and experts in the field; and cross referencing with existing reviews in the field), the emerging and varied terminology used in the field of implementation

science,[75] including that related to scale-up, made the development of search terms difficult. It is possible, therefore, that not all eligible trials were identified. The identification of program adaptions was also a considerable challenge. Reports of interventions in published manuscripts are very brief and may often omit important details. While we conducted secondary searches of websites and related program publications for the interventions included in this review, it is likely that the reported adaptations do not represent a comprehensive description of adaptations that were undertaken. The use of existing tools[76] by authors of scaled-up programs to ensure that sufficient detail to capture adaptations is recorded and reported would benefit the field. Similarly, few trials reported the same measures between efficacy and scaled-up trials, thus reducing the number of measures for which effects could be compared. While heterogeneity was low for most pooled analyses (I^2 = 0%), there was evidence of considerable heterogeneity in pooled analyses of scaled-up treatment interventions (I^2= 87.6%). As such, the point estimate for this analysis should be interpreted with caution. Finally, to identify the scale-up of specific evidence-based programs we restricted our inclusion criteria to trials employing a randomised design that had been demonstrated to be efficacious on a measure of weight status following evaluation. Interventions that had been developed and delivered at scale based on the findings of prior non-randomised trials or systematic reviews were not included, limiting the generalisability of the review findings.

5| CONCLUSIONS

The findings underscore the challenges of delivering effective interventions at scale. Nonetheless, the review provides important information for policy makers and practitioners responsible for providing obesity prevention services. Specifically, the findings of the review

should enable a more realistic appraisal of the likely effects, and reduction in effects, of interventions delivered following scale-up – providing clearer eyes for decision making. Given the importance of implementation fidelity in achieving intervention effects, developing the science of implementation at scale[17-19] and appraising the 'scalability' of interventions before investments are made to scale-up interventions should be considered priority areas for future work in this field.[25]

REFERENCES

1. World Health Organisation. Obesity and overweight. 2017; http://www.who.int/news-room/fact-sheets/detail/obesity-and-overweight, 21/06/2018.
2. World Health Organisation. Obesity. 2017; http://www.who.int/news-room/facts-in-pictures/detail/6-facts-on-obesity. Accessed 26/06/2018.
3. World Health Organisation. Physical Activity. 2018; http://www.who.int/dietphysicalactivity/pa/en/. Accessed 21/06/2018.
4. World Health Organisation. Nutrition. 2018; http://www.who.int/topics/nutrition/en/.
5. Eakin EG, Lawler SP, Vandelanotte C, Owen N. Telephone interventions for physical activity and dietary behavior change: a systematic review. *Am J Prev Med.* 2007;32(5):419-434.
6. Bock C, Jarczok MN, Litaker D. Community-based efforts to promote physical activity: a systematic review of interventions considering mode of delivery, study quality and population subgroups. *J Sci Med Sport.* 2014;17(3):276-282.
7. O'Hara BJ, Phongsavan P, Venugopal K, et al. Effectiveness of Australia's Get Healthy Information and Coaching Service®: translational research with population wide impact. *Prev Med.* 2012;55(4):292-298.
8. World Health Organisation. *Practical Guidance for scaling up health service innovations.* 2009.
9. World Health Organisation. Global Action Plan for the Prevention and Control of NCDs 2013-2020 [Internet]. WHA66. Geneva: World Health Organization; 2013 In.

10. Barker PM, Reid A, Schall MW. A framework for scaling up health interventions: lessons from large-scale improvement initiatives in Africa. *Implementation Science.* 2016;11(1):12.
11. Milat A, Newson R, King L, et al. A guide to scaling up population health interventions. *Public Health Res Prac.* 2016.
12. Aarons GA, Sklar M, Mustanski B, Benbow N, Brown CH. "Scaling-out" evidence-based interventions to new populations or new health care delivery systems. *Implement Sci.* 2017;12(1):111.
13. Wolfenden L, Ziersch A, Robinson P, Lowe J, Wiggers J. Reducing research waste and improving research impact. *Aust N Z J Public Health.* 2015;39(4):303-304.
14. Welsh BC, Sullivan CJ, Olds DL. When Early Crime Prevention Goes to Scale: A New Look at the Evidence. *Prev Sci.* 2010;11(2):115-125.
15. Tommeraas T, Ogden T. Is There a Scale-up Penalty? Testing Behavioral Change in the Scaling up of Parent Management Training in Norway. *Adm Policy Ment Health.* 2017;44(2):203-216.
16. Kilbourne AM, Neumann MS, Pincus HA, Bauer MS, Stall R. Implementing evidence-based interventions in health care: application of the replicating effective programs framework. *Implement Sci.* 2007;2(1):42.
17. Wolfenden L, Jones J, Williams CM, et al. Strategies to improve the implementation of healthy eating, physical activity and obesity prevention policies, practices or programmes within childcare services. *Cochrane Database Syst Rev.* 2016;10:Cd011779.

18. Wolfenden L , Stacey FG, Grady A, et al. Strategies to improve the implementation of workplace-based policies or practices targeting tobacco, alcohol, diet, physical activity and obesity. *Cochrane Database Syst Rev.*. 2018;Issue 11. Art. No.: CD012439. .
19. Wolfenden L, Nathan NK, Sutherland R, et al. Strategies for enhancing the implementation of school-based policies or practices targeting risk factors for chronic disease. *Cochrane Database Syst Rev.* 2017;11:Cd011677.
20. Reis RS, Salvo D, Ogilvie D, Lambert EV, Goenka S, Brownson RC. Scaling up physical activity interventions worldwide: stepping up to larger and smarter approaches to get people moving. *The Lancet.* 2016;388(10051):1337-1348.
21. Eriksson J, Lindström J, Valle T, et al. Prevention of Type II diabetes in subjects with impaired glucose tolerance: the Diabetes Prevention Study (DPS) in Finland Study design and 1-year interim report on the feasibility of the lifestyle intervention programme. *Diabetologia.* 1999;42(7):793-801.
22. Lindström J, Louheranta A, Mannelin M, et al. The Finnish Diabetes Prevention Study (DPS). Lifestyle intervention and 3-year results on diet and physical activity. *Diabetes Care.* 2003;26(12):3230-3236.
23. Laatikainen T, Dunbar JA, Chapman A, et al. Prevention of Type 2 Diabetes by lifestyle intervention in an Australian primary health care setting: Greater Green Triangle (GGT) Diabetes Prevention Project. *BMC Public Health.* 2007;7(1):249.
24. *Cochrane handbook for systematic reviews of interventions.* Vol Version 5.1. 0. : The Cochrane Collaboration; 2011.
25. Indig D, Lee K, Grunseit A, Milat A, Bauman A. Pathways for scaling up public health interventions. *BMC Public Health.* 2017;18(1):68.

26. Milat AJ, Bauman A, Redman S. Narrative review of models and success factors for scaling up public health interventions. *Implement Sci.* 2015;10(1):113.

27. Ben Charif A, Zomahoun HTV, LeBlanc A, et al. Effective strategies for scaling up evidence-based practices in primary care: a systematic review. *Implement Sci.* 2017;12(1):139.

28. Dobbins M, Husson H, DeCorby K, LaRocca RL. School-based physical activity programs for promoting physical activity and fitness in children and adolescents aged 6 to 18. *Cochrane Database Syst Rev.* 2013(2).

29. U.S. National Library of Medicine, Medical Subject Headings. 2018; https://www.nlm.nih.gov/mesh/filelist.html.

30. Martin A, Booth JN, Laird Y, Sproule J, Reilly JJ, Saunders DH. Physical activity, diet and other behavioural interventions for improving cognition and school achievement in children and adolescents with obesity or overweight. *Cochrane Database Syst Rev.* 2018(3).

31. Waters E, de Silva-Sanigorski A, Burford BJ, et al. Interventions for preventing obesity in children. *Cochrane Database Syst Rev.* 2011(12).

32. Dombrowski SU, Knittle K, Avenell A, Araujo-Soares V, Sniehotta FF. Long term maintenance of weight loss with non-surgical interventions in obese adults: systematic review and meta-analyses of randomised controlled trials. *BMJ.* 2014;348:g2646.

33. Lemmens VEPP, Oenema A, Klepp K-I, Henriksen HB, Brug J. A systematic review of the evidence regarding efficacy of obesity prevention interventions among adults. *Obes Rev.* 2008;9(5):446-455.

34. Peirson L, Fitzpatrick-Lewis D, Morrison K, Warren R, Ali MU, Raina P. Treatment of overweight and obesity in children and youth: a systematic review and meta-analysis. *CMAJ Open.* 2015;3(1):E35.

35. Wang Y, Cai L, Wu Y, et al. What childhood obesity prevention programmes work? A systematic review and meta [illegible] *Obes Rev.* 2015;16(7):547-565.

36. Chambers DA, Norton WE. The Adaptome: Advancing the Science of Intervention Adaptation. *Am J Prev Med.* 2016;51(4 Suppl 2):S124-131.

37. Armijo [illegible] Olivo SA, Stiles CR, Hagen N [illegible] study quality for systematic reviews: a comparison of the Cochrane Collaboration Risk of Bias Tool and the Effective Public Health Practice Project Quality Assessment Tool: methodological research. *J Eval Clin Pract.* 2012;18(1):12-18.

38. Hill LG, Maucione K, Hood BK. A focused approach to assessing program fidelity. *Prev Sci.* 2007;8(1):25-34.

39. Barrera Jr M, Castro FG, Strycker LA, Toobert DJ. Cultural adaptations of behavioral health interventions: A progress report. *J Consult Clin Psychol.* 2013;81(2):196.

40. van Nassau F, Singh AS, Cerin E, et al. The Dutch Obesity Intervention in Teenagers (DOiT) cluster controlled implementation trial: intervention effects and mediators and moderators of adiposity and energy balance-related behaviours. *Int J Behav Nutr Phys Act.* 2014;11(1):158.

41. Fagg J, Chadwick P, Cole T, et al. From trial to population: a study of a family-based community intervention for childhood overweight implemented at scale. *Int J Obes.* 2014;38(10):1343.

42. Lubans DR, Morgan PJ, Aguiar EJ, Callister R. Randomized controlled trial of the Physical Activity Leaders (PALs) program for adolescent boys from disadvantaged secondary schools. *Prev Med.* 2011;52(3-4):239-246.

43. Singh AS, Paw MJCA, Brug J, van Mechelen W. Dutch obesity intervention in teenagers: effectiveness of a school-based program on body composition and behavior. *Arch Pediatr Adolesc Med.* 2009;163(4):309-317.

44. Fu R, Vandermeer BW, Shamliyan TA, et al. Handling continuous outcomes in quantitative synthesis. 2013.

45. Higgins JP, Green S. Cochrane handbook for systematic reviews of interventions. Wiley Online Library; 2008.

46. Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. *BMJ.* 2003;327(7414):557.

47. Moores CJ, Miller J, Daniels LA, Vidgen HA, Magarey AM. Pre–post evaluation of a weight management service for families with overweight and obese children, translated from the efficacious lifestyle intervention Parenting, Eating and Activity for Child Health (PEACH). *Br J Nutr.* 2018;119(12):1434-1445.

48. Wilcox S, Dowda M, Leviton LC, et al. Active for life: final results from the translation of two physical activity programs. *Am J Prev Med.*2008;35(4):340-351.

49. Magarey AM, Perry RA, Baur LA, et al. A parent-led family-focused treatment program for overweight children aged 5 to 9 years: the PEACH RCT. *Pediatrics.* 2011:peds. 2009-1432.

50. Dunn AL, Marcus BH, Kampert JB, Garcia ME, Kohl III HW, Blair SN. Comparison of lifestyle and structured interventions to increase physical activity and cardiorespiratory fitness: a randomized trial. *JAMA.* 1999;281(4):327-334.

51. McKay HA, Macdonald HM, Nettlefold L, Masse LC, Day M, Naylor P-J. Action Schools! BC implementation: from efficacy to effectiveness to scale-up. *Br J Sports Med.* 2015;49(4):210-218.

52. Allman-Farinelli M, Partridge SR, McGeechan K, et al. A mobile health lifestyle program for prevention of weight gain in young adults (TXT2BFiT): Nine-month outcomes of a randomized controlled trial. *JMIR Mhealth Uhealth.* 2016;4(2).

53. Wang Z, Xu F, Ye Q, et al. Childhood obesity prevention through a community-based cluster randomized controlled physical activity intervention among schools in china: the health legacy project of the 2nd world summer youth olympic Games (YOG-Obesity study). *Int J Obes.* 2018;42(4):625.

54. Hopkins LC, Webster A, Sharn A, Gunther C. Camp NERF: Caregiver Outcomes from a Theory-Based Nutrition Education Recreation and Fitness Program Aimed at Preventing Unhealthy Weight Gain in Underserved Children during Summer Months. *The FASEB Journal.* 2017;31(1_supplement):640.623-640.623.

55. Wen LM, Rissel C, Baur LA, et al. A 3-Arm randomised controlled trial of Communicating Healthy Beginnings Advice by Telephone (CHAT) to mothers with infants to prevent childhood obesity. *BMC Public Health.* 2017;17(1):79.

56. Borelli MR, Riden HE, Bang H, Schenker MB. Protocol for a cluster randomized controlled trial to study the effectiveness of an obesity and diabetes intervention

(PASOS) in an immigrant farmworker population. *BMC Public Health.* 2018;18(1):849.

57. Borys JM, Richard P, du Plessis HR, Harper P, Levy E. Tackling health inequities and reducing obesity prevalence: the EPODE community-based approach. *Ann Nutr Metab.* 2016;68(Suppl. 2):35-38.
58. Hopman-Rock M, Westhoff MH. Development and evaluation of "aging well and healthily": a health-education and exercise program for community-living older adults. *J Aging Phys Act.* 2002;10(4):364-381.
59. Young M, Collins C, Callister R, Plotnikoff R, Doran C, Morgan P. Testing the efficacy of a gender-tailored intervention to prevent weight regain in men: The SHED-IT Weight Loss Maintenance RCT. *J Sci Med Sport.* 2014;18:e66.
60. Lombard C, Harrison C, Kozica S, Zoungas S, Ranasinha S, Teede H. Preventing weight gain in women in rural communities: a cluster randomised controlled trial. *PLoS Med.* 2016;13(1):e1001941.
61. Hardy LL, King L, Kelly B, Farrell L, Howlett S. Munch and Move: evaluation of a preschool healthy eating and movement skill program. *Int J Behav NutrPhys Act.* 2010;7(1):80.
62. Hardy LL, Mihrshahi S, Gale J, Nguyen B, Baur LA, O'Hara BJ. Translational research: are community-based child obesity treatment programs scalable? *BMC Public Health.* 2015;15(1):652.
63. Smith JJ, Morgan PJ, Plotnikoff RC, et al. Smart-phone obesity prevention trial for adolescent boys in low-income communities: the ATLAS RCT. *Pediatrics.* 2014:peds. 2014-1012.

64. Wyke S, Bunn C, Andersen E, et al. OP80 Effectiveness of european fans in training (EuroFIT): randomised controlled trial in england, portugal, the netherlands and norway. *J Epidemiol Community Health*. 2018.

65. Heerman WJ, Schludnt D, Harris D, Teeters L, Apple R, Barkin SL. Scale-out of a community-based behavioral intervention for childhood obesity: pilot implementation evaluation. *BMC Public Health.* 2018;18(1):498.

66. Folta SC, Seguin RA, Chui KK, et al. National dissemination of StrongWomen–Healthy Hearts: a community-based program to reduce risk of cardiovascular disease among midlife and older women. *American J Public Health.* 2015;105(12):2578-2585.

67. Rush E, McLennan S, Obolonkin V, et al. Project Energize: whole-region primary school nutrition and physical activity programme; evaluation of body size and fitness 5 years after the randomised controlled trial. *Br J Nutr.* 2014;111(2):363-371.

68. Lombard C, Deeks A, Jolley D, Ball K, Teede H. A low intensity, community based lifestyle programme to prevent weight gain in women with young children: cluster randomised controlled trial. *BMJ.* 2010;341:c3215.

69. Rush E, Reed P, McLennan S, Coppinger T, Simmons D, Graham D. A school-based obesity control programme: Project Energize. Two-year outcomes. *Br J Nutr.* 2011;107(4):581-587.

70. Ackermann RT. From programs to policy and back again: the push and pull of realizing type 2 diabetes prevention on a national scale. *Diabetes Care.* 2017;40(10):1298-1301.

71. Sutherland RL, Nathan NK, Lubans DR, et al. An RCT to Facilitate Implementation of School Practices Known to Increase Physical Activity. *Am J Prev Med.* 2017;53(6):818-828.

72. Milat A, Newson R, King L. Increasing the scale of population health interventions: a guide. *North Sydney: NSW Ministry of Health.* 2014.

73. World Health organisation and ExpandNet. Nine steps for developing a scaling-up strategy. 2010.; http://www.expandnet.net/tools.htm

74. Yoong SL, Wolfenden L, Clinton-McHarg T, et al. Exploring the pragmatic and explanatory study design on outcomes of systematic reviews of public health interventions: a case study on obesity prevention trials. *J Public Health.* 2014;36(1):170-176.

75. Colquhoun H, Leeman J, Michie S, et al. Towards a common terminology: a simplified framework of interventions to promote and integrate evidence into health practices, systems, and policies. *Implement Sci.* 2014;9(1):781.

76. Stirman SW, Miller CJ, Toder K, Calloway A. Development of a framework and coding system for modifications and adaptations of evidence-based interventions. *Implement Sci.* 2013;8(1):65.

77. Kingsland M, Wiggers J, Wolfenden L. Interventions in sports settings to reduce alcohol consumption and alcohol-related harm: a systematic review protocol. *BMJ Open.* 2012;2(2):e000645.

78. Folta SC, Lichtenstein AH, Seguin RA, Goldberg JP, Kuder JF, Nelson ME. The StrongWomen–Healthy Hearts program: reducing cardiovascular disease risk factors

in rural sedentary, overweight, and obese midlife and older women. *Am J Public Health.* 2009;99(7):1271-1277.

79. Barkin SL, Gesell SB, Po'e EK, Escarfuller J, Tempesti T. Culturally tailored, family-centered, behavioral obesity intervention for Latino-American preschool-aged children. *Pediatrics.* 2012:peds. 2011-3762.
80. Sacher PM, Kolotourou M, Chadwick PM, et al. Randomized controlled trial of the MEND program: a family -based community i *Obesity.* 2010;18(S1):S62-S68.
81. Zask A, Adams JK, Brooks LO, Hughes DF. Tooty Fruity Vegie: an obesity prevention intervention evaluation in Australian preschools. *Health Promot J Austr.* 2012;23(1):10-15.
82. Hunt K, Wyke S, Gray CM, et al. A gender-sensitised weight loss and healthy living programme for overweight and obese men delivered by Scottish Premier League football clubs (FFIT): a pragmatic randomised controlled trial. *The Lancet.* 2014;383(9924):1211-1221.

SUPPLEMENT FILE 1. Search strategy.

Search topic	Included search terms
Scaled-up interventions	Scaling-up OR scalability[26] ("scaling-up" or "scaled-up" or "scale up" or "up-scaling" or "upscaling").ti,ab. (scalability or scalable or "at scale").ti,ab. (spread adj5 (innovation* OR intervention* OR technolog* OR practice OR care)).ti,ab. ((bring* or brought or taking or take* or increas* or going or implement*) adj5 scale)).ti,ab.[27] scaled-up OR scale-up OR scaling-up OR scalability OR scalable OR reach OR expanding OR expandable OR expandability OR institutionalization OR institutionalisation OR roll-out OR rolling-out OR dissemination OR disseminating[20]
Physical activity	21. exp exercise/ 22. physical inactivity.mp. 23. physical activity.mp. 24. exp motor activity/ 25. (physical education and training).mp. 26. exp "Physical Education and Training"/ 27. exp physical fitness/ 28. sedentary.ab. or sedentary.ti. 29. exp life style/ 30. exp leisure activities/ 31. exp walking/ 32. exp sports/ 33. exp dancing/ 34. dancing.mp. 35. exp exercise therapy/ 36. (exercise$ adj aerobic$).tw. 37. (physical$ adj5 (fit$ or train$ or activ$ or endur$)).tw. 38. (exercis$ adj5 (train$ or physical$ or activ$)).tw. 39. sport$.tw. 40. walk$.tw. 41. cycle$.tw. 42. (("lifestyle" or life-style) adj5 activ$).tw. 43. (("lifestyle" or life-style) adj5 physical$).tw. 44. or/21-43[28] 10. exp Exercise/ 11. physical inactivity.mp. 12. physical activity.mp. or Physical Activity/

	13. Motor Activity.mp. 14. (physical education and training).mp. 15. Physical Education/ 16. Physical Fitness/ 17. sedentary.mp. 18. exp Lifestyle/ 19. leisure time/ or recreation/ 20. exp Sports/ 21. Dance/ 22. (exercise* adj2 aerobic*).mp. 23. sport*.mp. 24. ((life style or lifestyle) adj5 activ*).mp. 25. (dance* or dancing).mp. 26. or 11 or 12 or 13 or 14 or 15 or 16 or 17 or 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25[17] physical activity OR physically active OR physical inactivity OR physically inactive OR fitness OR exercis* OR sport* OR walk OR walking OR sedentary OR sitting OR television OR TV OR screen time OR screen-time OR active transport* OR active transit OR active travel OR commut* OR active commuting OR bicycle OR bicycling OR bike OR biking OR active living OR active-living[20] MeSH terms Human Activities [I03] Exercise [I03.305][29]
Nutrition	27. exp Nutrition/ 28. nutrition*.mp. 29. (health* adj2 eat*).mp. 30. Child Nutrition Sciences/ 31. Fruit/ or fruit*.mp. 32. Vegetables/ or vegetable*.mp. 33. canteen*.mp. 34. Food Services/ 35. menu.mp. 36. (calorie or calories or kilojoule*).mp. 37. Energy Intake/ 38. energy density.mp. 39. Eating/ 40. Feeding Behavior/ or feeding behaviour.mp. 41. nutritionary intake.mp. 42. Food Habits/ 43. Food/ 44. Carbonated Beverages/ or soft drink*.mp.

	45. soda.mp. 46. sweetened drink*.mp. 47. Nutritionary Fats/ 48. confectionary.mp. 49. (school adj2 (lunch* or meal*)).mp. 50. Menu Planning/ 51. feeding program*.mp. 52. food program*.mp. 53. (nutrition* adj2 program*).mp. 54. cafeteria*.mp. 55. Nutritional Status/ 56. 27 or 28 or 29 or 30 or 31 or 32 or 33 or 34 or 35 or 36 or 37 or 38 or 39 or 40 or 41 or 42 or 43 or 44 or 45 or 46 or 47 or 48 or 49 or 50 or 51 or 52 or 53 or 54 or 55 57. 9 or 26 or 56[17] MeSH terms Healthy Nutrition [F01.829.458.205.500] Nutrition [G07.203.650.240][29]
Obesity	#1 MeSH descriptor Overweight explode all trees #2 MeSH descriptor Body Weight, this term only #3 (obes* or overweight or over-weight) #4 MeSH descriptor Body Weight Changes explode all trees #5 (weight near/2 (loss or lost or losing or reduc*)) #6 (weight near/2 (gain* or increas*)) #7 MeSH descriptor Body Fat Distribution explode all trees #8 MeSH descriptor Body Mass Index explode all trees #9 MeSH descriptor Skinfold Thickness explode all trees #10 MeSH descriptor Waist-Hip Ratio explode all trees #11 ("body weigh*" or bodyweigh* or "body mass*" or bodymass or "body fat*" or bodyfat*) #12 MeSH descriptor Overnutrition, this term only #13 (overeat* or over-eat* or overnourish* or over-nourish* or overnutrit* or over-nutrit*) #14 (#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7 OR #8 OR #9 OR #10 OR #11 OR #12 OR #13)[30] 1. exp Obesity/ 2. exp Weight Gain/ 3. exp Weight Loss/ 4. obes$.af. 5. (weight gain or weight loss).af. 6. (overweight or over weight or overeat$ or over eat$).af. 7. weight change$.af. 8. ((bmi or body mass index) adj2 (gain or loss or change)).af.

	9. or/1-8[31] 1. exp obesity/ 2. Weight Gain/ 3. exp Weight Loss/ 4. obes*.mp. 5. (weight gain or weight loss).mp. 6. (overweight or over weight or overeat* or over eat*).mp. 7. weight change*.mp. 8. ((bmi or body mass index) adj2 (gain or loss or change)).mp. 9. 1 or 2 or 3 or 4 or 5 or 6 or 7 or 8[17]
Study design	71. randomized controlled trial.pt. 72. controlled clinical trial.pt. 73. randomized.ab. 74. randomised.ab. 75. clinical trials as topic.sh. 76. randomly.ab. 77. trial.ti. 78. doubleblind.ab. 79. singleblind.ab. 80. experiment*.mp. 81. (pretest or pre test).mp. 82. (posttest or post test).mp. 83. (pre post or prepost).mp. 84. Before after.mp. 85. (Quasi-randomised or quasi-randomized or quasi-randomized or quazi-randomised).mp. 86. stepped wedge.mp. 87. Preference trial.mp. 88. Comprehensive cohort.mp. 89. Natural experiment.mp. 90. (Quasi experiment or quazi experiments).mp. 91. (Randomised encouragement trial or randomized encouragement trial).mp. 92. (Staggered enrolment trial or staggered enrollment trial).mp. 93. (Nonrandomised or non randomised or nonrandomized or non randomized).mp. 94. Interrupted time series.mp. 95. (Time series and trial).mp. 96. Multiple baseline.mp. 97. Regression discontinuity.mp. Or/17-97[77]

SUPPLEMENT FILE 2. List of included scale-up intervention trials and their pre-scale efficacy trials.

Pre-scale intervention name	Scaled-up intervention name
StrongWomen - Healthy Hearts[78]	StrongWomen - Healthy Hearts[66]
HeLP-her[68]	HeLP-her[60]
Salud Con La Familia (Health with the Family)[79]	Healthier Families[65]
MEND[80]	MEND 7-13[41]
MEND[80]	Go4Fun[62]
Project Energize[69]	Project Energize[67]
Tooty Fruity Vegie[81]	Munch and Move[61]
DOiT[43]	DOiT[40]
PALS[42]	ATLAS[63]
FFIT[82]	EuroFIT[64]

SUPPLEMENT FILE 3. Description of intervention adaptations.

Program	Pre-scale intervention description	Adaptation of intervention for scale-up
Obesity prevention interventions		
ATLAS[63]	A multi-component school-based intervention that aimed to help adolescent boys become physical activity leaders The six month intervention was delivered by male teaching staff at each school and consisted of: a) ten, 90-minute self- and teacher- directed physical activity sessions focused on resistance training; b) three, 30-minute interactive seminars addressing health behaviours and leadership delivered by research team members; c) eight, 30-minute self-directed lunch-time physical activities; d) physical activity and nutrition handbooks; e) six, 30-minute leadership sessions where participants mentored younger students in resistance training technique; and f) pedometers for goal setting and self-monitoring.	1. Addition of parent encouragement to manage their children's recreational screen-time (Service setting adaptation). 2. Physical activity sessions increased in frequency from 10 to 20; physical activity sessions provided a greater variety of activities and a greater focus on movement skill development; interactive seminars decreased in duration from 30 minutes to 20 minutes; and lunch time physical activity sessions decreased in frequency from eight to six sessions. Further, the scaled-up intervention included a smartphone application (app) and website designed to supplement the delivery of the enhanced school sport and interactive sessions (Mode of delivery adaptations). 3. Reducing screen time added as a key intervention focus, whereas the pre-scale key areas of focus were promoting physical activity, healthy nutrition and leadership (Other adaptation).
DOiT[40]	A multicomponent school-based health promotion intervention that aimed to improve body composition and health behaviours in Dutch adolescents. The intervention took place over one school year and consisted of an individual component (i.e., educational program coverage through 11 in-school lessons) and an environmental component (i.e., school-specific advice such as offering additional PE classes and healthy changes to school cafeterias). Target behaviours included consumption of high-energy snacks and beverages, physical activity, and screen-viewing behaviour.	1. Changed source of program funding from a national grant to funding facilitated through a national database. Also added parent engagement through an information book, information on the website, an optional parent meeting, and assignments to work on collaboratively with their adolescent (Service setting adaptations). 2. The scaled-up intervention consisted of 16 mandatory and three optional sessions over the course of two school years (compared with 11 sessions over the course of one school year in the pre-scale); and was supported by a five-minute instruction video, teacher manual and website resources to guide teacher implementation (Mode of delivery adaptations). 3. Addition of a support office with a dedicated recruitment employee, and promotional material and activities; contact with health promotion professionals added as strategies to

		facilitate school engagement and recruitment; and addition of another target behaviour: daily breakfast consumption (Other adaptations).
Healthier Families[65]	A culturally tailored, family-centered short-term behavioural intervention that aimed to reduce the BMI of Latino-American preschool-aged children. The 12 week intervention consisted of weekly, 90-minute small-group sessions, delivered in Spanish by a single trained facilitator in a community recreation centre. Both parents and children engaged in the skill-building activities to promote healthier family behaviours (i.e., improved nutrition, increased physical activity, and decreased sedentary behaviour).	1. Changed target population from Latino-American preschool children to English speaking preschool children (Target audience adaptation). 2. Group session duration decreased from 90 minutes to 60 minutes, and multiple, rather than a single, facilitator were used for intervention delivery. Further, training modules for facilitators were enhanced to include an online component and emphasis on facilitation and behaviour change techniques (Mode of delivery adaptation). 3. The language of delivery changed from Spanish (original) to English (scaled-up) (Cultural adaptation). 4. Community partnership were established prior to the scaled-up intervention and a pre-implementation assessment undertaken to identify feasible/interested parks and recreation centres to implement the program (Other adaptations).
HeLP-her[60]	A self-treatment intervention that aimed to prevent weight gain in women with young children. The one year program consisted of: a) one 60 minute group session during weeks one, two, three and 16 delivered by the same trained facilitator (a nutritionitian) at the local primary school, with a focus on behaviour change skills for nutrition and physical activity; b) provision of a pedometer to self-monitor physical activity; and d) follow up support via monthly text messages from week four to 52 personalised by name. Participants were encouraged to take part in low intensity voluntary physical activities and regular self-weighing.	1. Changed to multiple, rather than a single, trained facilitators to lead group intervention exercises. Facilitators were also trained and required to have a tertiary qualification (but not specifically in nutritionetics) in health sciences (Service setting adaptations). 2. Changed target population from women with a child attending primary school to women aged 18-50 years residing in or near a participating town (Target audience adaptation). 3. Changed from four group sessions, to only one group session combined with a manual and a 20 minute telephone coaching session (Mode of delivery adaptation). 4. Addition of theoretical influences of motivational interviewing, as well as a communication plan and engagement framework to ensure multilevel engagement across communities and program implementation (Other adaptations).
Munch and	A preschool-based intervention that aimed to improve Fundamental Movement Skills, increase fruit and vegetable intake, and reduce intake of unhealthy snack	1. Changed funding providers and stakeholders from a joint Australian, State and Territory government initiative (pre-scale intervention) to a state government funded program

Move[61]	items for preschool aged children. The 10 month intervention consisted of twice-weekly Fundamental Movement Skill development; small sports equipment grant; parent workshops, newsletters and a DVD; playground environment review and alterations; nutrition policies; posters to improve health of food, staff role-modelling healthy eating, and accessible drinking water. Health professionals employed by local health services, preschool staff or parent volunteers delivered these various program components.	(Service setting adaptation). 2. Changed the audience of focus from pre-school staff, parents and children, to solely pre-school staff in the scaled-up intervention. . (Target audience adaptation) 3. Addition of a one-day professional development workshop for preschool staff, parent workshops, newsletters and DVDs were removed (Mode of delivery adaptation).
Project Energize[69]	A school-based intervention that aimed to improve childhood obesity and cardiovascular health in primary school children. The intervention was delivered by an 'Energizer'; a trained employee of a Regional Sports Trust who worked with 8-10 schools as an agent of change to achieve school-specific health goals through facilitating within-school changes (e.g., modification to food provided by canteens', in class physical activities, newsletter health updates, parent information sessions, etc.) and accessing regional initiatives.	1. More Energizers were required to deliver the scaled-up intervention (27 compared with 11 in the pre-scale), and each were assigned a greater number of schools (8-12 compared with 8-10 in the pre-scale) (Other adaptation).
1. Obesity treatment interventions		
EuroFIT[64]	A gender-sensitised, group-based weight loss and healthy living intervention for male fans in Scottish professional football clubs. The 12 month intervention was conducted in two phases. The first, a 12 week active phase, involved weekly 90 minute classroom and physically active group sessions delivered by community coaching staff at the club's home stadium. Coaching staff were employed by the local football clubs and received two days of training by the research team. The second maintenance phase, involved a six month group reunion and six email prompts over the course of nine months.	1. Novel technologies were added for self-monitoring and to promote competition, and email prompts were removed (Mode of delivery adaptation). 2. Content was culturally-sensitised for each country to reflect local physical activity and nutrition norms (Cultural adaptation). 3. A consortium was established for implementation, sedentary behaviour was added as a focus of interest, and weight loss was removed as a core focus (Other adaptation).
Go4Fun[62]	As described above for MEND 7-13	1. The scaled-up intervention was managed and funded by New South Wales (Australia) Ministry of Health (Service setting adaptation). 2. Facilitator training was reduced from four days of in-person training, to two days of in-person plus online training. (Mode of delivery adaptation).

		3. MEND was translated to reflect an Australian context (Cultural adaptation). 4. Implementation took place progressively by local health services and intervention was marketed to assist with program enrolment (Other adaptation).
MEND 7-13[41]	A multicomponent family-based childhood obesity intervention. The six month intervention consisted of 18, two-hour group sessions delivered in a community setting twice a week over nine weeks, followed by a 12 week free family swim pass at a local community pool. Two MEND leaders and an assistant delivered group sessions which focused on nutrition education, behaviour change and exercise. Eight to 15 parent-child dyads and their siblings composed each group. MEND leaders received four days of training and identical materials including detailed methods for group sessions to ensure standardized delivery.	1. MEND Central was established and considered important for national implementation support (Other adaptation).
Strong Women-Healthy Hearts[66]	A community-based intervention with the aim to reduce the risk of cardiovascular disease in middle age overweight/obese women. The 12 week intervention consisted of 24 group sessions, delivered twice weekly by an educator trained by research staff. Each one hour session involved a 30 minute physical activity component and a 30 minute nutrition component. Study personnel at each site observed one class to assess intervention fidelity.	1. Educators recruited from/using a different organization (Service setting adaptation). 2. Educators were trained through a workshop that consisted of a series of seminars on all program aspects as opposed to training by the research staff in the pre-scale (Mode of delivery adaptation). 3. Nutrition guidelines used in the intervention changed to be consistent with the American Heart Association guidelines and the 2010 Nutrition Guidelines for Americans. The pre-scale intervention use general healthy eating practices without a specified information source (Other adaptation).