



A systematic review of adaptations and effectiveness of scaled-up nutrition interventions

Rachel L. Sutherland, Jacklyn K. Jackson , Cassandra Lane, Sam McCrabb , Nicole K. Nathan, Sze Lin Yoong, Melanie Lum, Judith Byaruhanga, Matthew McLaughlin, Alison Brown, Andrew J. Milat, Adrian E. Bauman, and Luke Wolfenden

Context: Public health nutrition interventions shown to be effective under optimal research conditions need to be scaled up and implemented in real-world settings.

Objectives: The primary aim for this review was to assess the effectiveness of scaled-up public health nutrition interventions with proven efficacy, as examined in a randomized controlled trial. Secondary objectives were to: 1) determine if the effect size of scaled-up interventions were comparable to the prescale effect, and; 2) identify any adaptations made during the scale-up process. **Data sources:** Six electronic databases were searched and field experts contacted. **Study selection:** An intervention was considered scaled up if it was delivered on a larger scale than a preceding randomized controlled trial (“prescale”) in which a significant intervention effect ($P \leq 0.05$) was reported on a measure of nutrition. **Data extraction:** Two reviewers independently performed screening and data extraction. Effect size differences between prescale and scaled-up interventions were quantified. Adaptations to scale-up studies were coded according to the Adaptome model.

Results: Ten scaled-up nutrition interventions were identified. The effect size difference between prescale trials and scaled-up studies ranged from -32.2% to 222% (median, 50%). All studies made adaptations between prescale to scaled-up interventions. **Conclusion:** The effects of nutrition interventions implemented at scale typically were half that achieved in prior efficacy trials. Identifying effective scale-up strategies and methods to support retainment of the original prescale effect size is urgently needed to inform public health policy. **Systematic Review Registration:** PROSPERO registration no.CRD42020149267.

Results: Ten scaled-up nutrition interventions were identified. The effect size difference between prescale trials and scaled-up studies ranged from -32.2% to 222% (median, 50%). All studies made adaptations between prescale to scaled-up interventions. **Conclusion:** The effects of nutrition interventions implemented at scale typically were half that achieved in prior efficacy trials. Identifying effective scale-up strategies and methods to support retainment of the original prescale effect size is urgently needed to inform public health policy. **Systematic Review Registration:** PROSPERO registration no.CRD42020149267.

Affiliation: R.L. Sutherland, J.K. Jackson, C. Lane, S. McCrabb, N.K. Nathan, S.L. Yoong, M. Lum, J. Byaruhanga, M. McLaughlin, A. Brown, A.E. Bauman, and L. Wolfenden are with the School of Medicine and Public Health, Faculty of Health and Medicine, University of Newcastle, Callaghan, New South Wales, Australia. R.L. Sutherland, J.K. Jackson, C. Lane, S. McCrabb, N.K. Nathan, S.L. Yoong, M. Lum, J. Byaruhanga, M. McLaughlin, A. Brown, and L. Wolfenden are with the Hunter Medical Research Institute, New Lambton Heights, New South Wales, Australia. R.L. Sutherland, C. Lane, N.K. Nathan, S.L. Yoong, M. Lum, J. Byaruhanga, M. McLaughlin, A. Brown, and L. Wolfenden are with the Hunter New England Population Health Unit, Hunter New England Local Health District, Wallsend, New South Wales, Australia. S.L. Yoong is with the School of Health Sciences, Swinburne University of Technology, Hawthorn, Victoria, Australia. A.J. Milat and A.E. Bauman are with the School of Public Health, University of Sydney, Sydney, New South Wales, Australia

Correspondence: R.L. Sutherland, Hunter New England Population Health, Locked Bag 10, Booth Building, Wallsend, New South Wales 2287, Australia. E-mail: Rachel.Sutherland@health.nsw.gov.au.

Key words: adaptation, public health nutrition, scale-up, systematic review

© The Author(s) 2021. Published by Oxford University Press on behalf of the International Life Sciences Institute.

This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial-NoDerivs licence (<https://creativecommons.org/licenses/by-nc-nd/4.0/>), which permits non-commercial reproduction and distribution of the work, in any medium, provided the original work is not altered or transformed in any way, and that the work is properly cited. For commercial re-use, please contact journals.permissions@oup.com

INTRODUCTION

Poor dietary intake, including the overconsumption of foods high in energy, saturated fat, salt, and sugar, and suboptimal intake of fruit, vegetables, and fiber, are leading causes of noncommunicable diseases internationally.^{1,2} Poor dietary intake accounts for > 11 million deaths globally per year and is linked to a variety of preventable diseases, such as cardiovascular disease, cancer, type 2 diabetes, and stroke.^{3–5} Given that 1 in 5 deaths could be averted by improving dietary intake,² public health nutrition interventions that align population-level consumption with dietary guidelines have been recommended across all age groups.² Specifically, community- and settings-based intervention approaches to improve dietary intake have been suggested to be particularly beneficial because they provide opportunities for repeated exposure reaching large numbers of people.⁶

Decades of research have identified a range of effective community- and settings-based interventions to improve dietary intake to prevent chronic disease.^{7,8} There are now numerous systematic reviews of school- and childcare-based nutrition interventions^{9,10} and reviews of public health nutrition interventions conducted in community settings, including workplaces,¹¹ sporting clubs,¹² and places of worship, that have demonstrated improvements in dietary intake aimed at preventing chronic diseases.¹³ Despite the plethora of evidence-based public health nutrition interventions targeting chronic disease prevention, interventions that are effective under ideal research conditions offer little benefit to population health unless they are scaled up.¹⁴ Scale-up is defined by the World Health Organization 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.”¹⁵

Although governments¹⁶ and international agencies recommend scaling-up evidence-based public health interventions at a population level,⁶ few effective interventions are ever delivered to large numbers in the population. Consequently, effects of dietary interventions targeting chronic disease prevention when delivered at scale are largely unknown.¹⁷ For example, a 2018 systematic review of scaled-up public health interventions targeting all chronic disease risk factors (eg, smoking, alcohol, nutrition, physical activity, weight) across all ages and settings identified just 40 scaled-up interventions globally.¹⁸ Of these, 55% followed a comprehensive scale-up pathway including phases involving efficacy and effectiveness testing¹⁸ despite not following the recommended scientific pathway to warrant public health investment.¹⁸

Of concern, findings of a recent review demonstrate that scale-up trials often fail to generate the effect size achieved in their prescale efficacy trial, which, in turn, often results in limited effect when delivered in the population of interest.¹⁹ This phenomenon has been termed as a *scale up penalty* or *voltage drop*.²⁰ One reason for this observed penalty may be due to the adaptations that are typically undertaken as part of the scale-up process to increase fit of evidence-based interventions to the needs of the users and the delivery context.²¹ For example the mode of delivery, dose delivered, target audience, or setting may all be adapted in the scale-up process. A 2019 systematic review of 10 obesity prevention and management trials, for example, demonstrated that adaptations occurred in all cases between the original, efficacious prescale trial and the scale-up of the intervention and resulted in a 25% “scale-up penalty or voltage drop”^{22,23} whereby the effects of interventions are reduced after scale-up.¹⁹

Given this, it is critical to assess the effect size of interventions delivered at scale and determine whether significant investments in their implementation are achieving the intended benefits to the community. In addition, comparing the effects of interventions delivered at scale with those achieved during trials to establish their efficacy (prescale) is useful to assess the extent to which adaptations as part of the scale-up process may influence effectiveness. This information is important for community-based nutrition research to allow policy makers to appraise the likely impact of interventions delivered at scale before significant investments in their population-wide delivery occur.

To our knowledge, the literature regarding the effects and/or adaptations of scaled-up nutrition interventions specifically targeting chronic disease prevention in community settings have not been subject to 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 systematic review, we address this evidence gap for community-based interventions directed at dietary behaviors. Such findings will be useful for public health practitioners and policy makers, and contribute relevant evidence to guide the delivery of nutrition interventions at scale.

AIMS AND OBJECTIVES

The primary aim for this review was to assess the effectiveness, when scaled up, of public health nutrition interventions with proven efficacy as examined in a

randomized controlled trial (RCT). Specifically, the objectives were to:

- assess the effects of evidence-based health promotion interventions targeting the prevention of chronic disease on measures of nutrition after scale-up;
- describe differences in effects of interventions targeting the prevention of chronic disease established before and after scale-up (scale-up penalty) for study pairs with directly comparable measures of nutrition; and
- describe the types of adaptations made to the nutrition intervention as part of the scale-up process.

METHODS

To address the study aims, a systematic search was undertaken of peer-reviewed and grey literature, on the basis of an existing review conducted by McCrabb et al.¹⁹ Review methods were developed in accordance to the *Cochrane Handbook for Systematic Reviews of Interventions* version 5.1.0,²⁴ and has been registered with PROSPERO, the international prospective register for systematic reviews (registration no. CRD42020149267). This systematic review was conducted and reported in accordance to the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) guidelines ([Supplementary Material](#) in the Supporting Information online).²⁵

Search strategy

To identify published peer-reviewed literature, a systematic search strategy was undertaken in September 2019 of the following electronic databases: MEDLINE, Embase, PsycINFO, Cochrane Central Register of Controlled Trials, CINAHL, and Education Resources Information Center. Search terms to identify scaled-up interventions were developed on the basis of terminology used in previous reviews,^{26–28} combined with published search filters for nutrition and study design.^{29,30} We did not impose any language or time restrictions on the searches. Because this review was embedded within a series of systematic reviews evaluating the effectiveness of scaling up obesity,¹⁹ physical activity,³¹ and nutrition interventions targeting the prevention of chronic disease, and because obesity and physical focused interventions also often include a dietary component, the search strategy also included obesity- and physical activity-related terms. Studies that included nutrition as a primary or secondary outcome were filtered during the full-text screening process and included in this review. Search terms are listed in Table S1 in the Supporting Information online.

In addition to electronic databases, we searched for relevant published, unpublished, and grey literature in the following ways. We contacted corresponding authors of studies about interventions who were identified during full-text screening and whose interventions, which, if scaled up, would have been eligible for inclusion to confirm 1) if their intervention had been subsequently scaled-up, and 2) whether the effects of the scale-up had been evaluated. We contacted corresponding authors of trials included in key systematic reviews by email to assess their knowledge of whether an intervention had been subsequently scaled up and if the effects of the intervention following scale-up had been evaluated. We also checked if studies included in process and outcome reports of health promotion interventions that were scaled up were eligible for inclusion in this review. Key individuals were contacted from the World Health Organization; the World Health Organization Collaborating Center for Physical Activity; Nutrition and Obesity; the New South Wales Ministry of Health; the National Cancer Institute, National Institutes of Health; and we made general enquires at Public Health England and the Division of Nutrition, Physical Activity, and Obesity at the Centers for Disease Control and Prevention to request if they were aware of any other trials that could be eligible for inclusion in this review. Trials identified as potentially eligible, using provided contacts, were assessed by the review team.

Criteria for including and excluding studies

Types of study designs We included pairs of studies (a prescale trial and a scaled-up study) that fit the following criteria: 1)

The prescale trial (ie, “efficacy trial”) was an RCT with established statistical significance (eg, $P \leq 0.05$) for at least 1 outcome measure of participants’ dietary intake (primary or secondary outcome); and 2)

the scale-up study was intentionally delivered on a larger scale (eg, to a larger number of participants) than the prescale trial, and included at least 1 outcome measure (primary or secondary) of participants’ dietary intake consistent with the prescale trial so outcome could be compared before and after scale-up (eg, fruit and/or vegetable intake [F&V]; discretionary foods or energy-dense nutrient-poor [EDNP] foods; sugar sweetened beverages [SSB]). The scaled-up study could be of any design (including randomized, controlled, before-and-after trials; and noncontrolled before-and-after designs).

Prescale trials and the scaled-up studies were linked using forward and/or backward searching from those identified in the search, using the citation search on

Medline. Data were extracted from paired efficacy trial and scaled-up studies in this review. Studies published in the peer-reviewed and grey literature were eligible (Table 1).

Population Eligible trials targeted children, adolescents, or adults in a community nonclinical setting that aimed to improve participant dietary intake (eg, preschool, childcare service, school, workplace, sport and recreation facilities, general community). Trials recruiting participants on the basis of preexisting medical diagnosis or obesity-related comorbidities (eg, high blood pressure, high cholesterol level, diabetes, allergies, eating disorders) were excluded. Studies conducted in clinical settings such as hospitals or general practices were excluded.

Types of interventions Trials were included if the researchers intentionally sought to deliver an intervention to a population on a larger scale (eg, a greater number of individuals or settings in the target population), than the preceding RCT that established its efficacy and target dietary intake (as a single or multicomponent study) of the individuals within the setting as a primary or secondary outcome. The scaled-up study needed to be a progress along translation pathway (eg, efficacy, effectiveness, implementation, dissemination, institutionalization) from its previous prescale trial that established its efficacy. As a result, we excluded scale-up trials where the primary purpose was replication. There were no criteria regarding the absolute or relative increase in scale required of scaled-up evaluations. Thus, scaled-up interventions that were delivered to more of the target population but included fewer participants in the evaluation relative to the prescale trial were included (eg, more schools but evaluation included fewer students). Vertically scaled (ie, introduced across a whole system at the same time, as with a mandated policy or practice), horizontally scaled (ie, gradually introduced across different sites or groups over time, as with a phased implementation), and scaled-out interventions were included.³² Scaled-out interventions included those delivered to new populations and/or were reached through a new delivery system from those in the efficacy trial.³²

Scaled-up studies were categorized as effectiveness (ie, evaluating the effectiveness of an program or intervention in a real-world setting), implementation (ie, evaluating strategies to enhance the uptake or adoption of an evidence-based program or intervention in a specific setting), or dissemination (ie, evaluating the targeted distribution of a program or materials to a specific public health or clinical practice audience) studies.¹⁸ Evaluations were excluded if their primary

purpose was to replicate interventions in the same translation phase (ie, an efficacy trial conducted to replicate findings of a prior efficacy trial).¹⁸ Although the prescale trial needed to demonstrate efficacy for at least 1 measure of dietary intake using an RCT design, as recommended for establishing intervention effectiveness,¹⁸ the scaled-up evaluation could use a randomized design, a nonrandomized design, or designs without a control group, because assessment of effects at scale using an RCT is often not feasible or appropriate.

Types of outcome measures Outcome measures of interest included any measure of dietary intake. Such measures could be derived from any data source, including objective measures (eg, biomarker assessments), self-reported measures (eg, food frequency questionnaire [FFQ], short diet questionnaire), proxy measures (eg, lunchbox audits, purchasing data), or direct observations.

Study selection, data extraction, and data analysis

Selection of studies Pairs of reviewers not blinded to the author or journal information independently screened titles and abstracts of all studies. Where required, Google Translate was used to assess the eligibility of abstracts not published in English. Full-text articles were obtained for eligible studies or studies that could not clearly be excluded on the basis of study title and abstract. Full-text article inclusion was decided via consensus between reviewer pairs. When consensus could not be reached (n = 3 instances), eligibility was discussed with a third reviewer to determine final inclusion. The primary reason for exclusion of full-text manuscripts was recorded.

Data extraction and management Pairs of authors of the present review independently extracted data in duplicate from included studies. Data extractors were unblinded to author and journal information. When discrepancies between reviewers could not be resolved by consensus, a third reviewer was consulted for final decision-making. Data extraction from each pair of included studies (ie, the RCT that established efficacy and the scaled-up evaluation) related to the following: 1) study characteristics (ie, country, year of publication, sample population and size, study design, trial measures and outcomes, including the reporting of any economic evaluation); 2) the translation stage of each intervention per criteria described by Indig et al¹⁸ (ie, efficacy, effectiveness, implementation, or dissemination); 3) the nature of any adaptations made for the scale-up trial using a modified Adaptome model³³; 4) any measure of dietary intake reported using the same measure across

Table 1 PICOS criteria for inclusion of studies

Parameter	Description
Population	Inclusion: Presumably health participants (including children, adolescents, or adults) in nonclinical community settings (inclusive of preschools, childcare services, schools, workplaces, sport and recreational facilities, and general community) Exclusion: Participants with a preexisting medical diagnosis or obesity-related comorbidity (including high blood pressure, cholesterol, diabetes, allergies, or eating disorders). Participants were recruited from clinical settings such as hospitals or general practices.
Intervention	Inclusion: Nutrition interventions targeting the prevention of chronic disease. Interventions were intentionally delivered to a population on a larger scale (eg, greater number of individuals or settings in the target population) than the preceding randomized controlled trial that established the intervention's efficacy for improving at least 1 dietary outcome. Exclusion: Single or repeated efficacy trials
Comparison	Inclusion: Prescale trials must have had a control group (defined as a true, nonintervention control, delayed intervention control, or alternative intervention control). Scaled-up trials: Eligibility criteria were not applicable.
Outcome	Inclusion: Prescale: Must have established statistical significances for least 1 measure of participant dietary intake. Scale-up: Included at least 1 outcome measure of dietary intake consistent with the prescale trial. Exclusion: None
Study design	Inclusion: Prescale trials were randomized controlled trials with established statistical significant results for at least 1 dietary outcome. Scaled-up trials could be of any study design (including randomized, controlled, before-and-after trials, and noncontrolled before-and-after designs) Exclusion: Nonexperimental studies. Efficacy randomized controlled trials that did not have a preceding scaled-up trial.

both trials to enable assessment of study quality and meta-analysis; and 5) study risk of bias.

Data synthesis

The characteristics of the included scaled-up studies and their classification as either effectiveness, implementation, or dissemination studies, based on the scale-up pathways described by Indig et al,¹⁸ are included in Table 3.

The effect of evidenced-based nutrition interventions after scale-up The effects of all interventions included in the review were narratively synthesized. Meta-analysis of nutrition outcomes were not undertaken, because of the large heterogeneity in reported dietary outcomes included across studies. Scaled-up studies were categorized as F&V, EDNP foods, SSBs, and other, according to the nutrition outcome targeted. Interventions were reported as effective if the scaled-up study indicated a statistically significant outcome ($P \leq 0.05$) between groups at follow-up or between baseline and follow-up, depending on study design.

Differences in intervention effect established before and after scale-up To assess the effect size retained at scale from the previous efficacious prescale trial and to identify any scale-up penalty as a result of scaling nutrition interventions, the differences in the between-group effect sizes reported from prescale to scaled-up evaluations were extracted. For data to be included, the

prescale trial had to have produced a statistically significant intervention effect on a measure of dietary intake, and outcome measures and methodology had to have been consistent across both studies. Studies in which the methodology to assess the same dietary outcome differed (eg, a food diary was used in the efficacy trial and then an FFQ was used in the scaled-up trial) were deemed ineligible for inclusion. Included trials that provided sufficient data to allow comparable assessment of effects were grouped by type of dietary outcome: 1) F&V (where increases in intake indicated improvements at prescale); 2) EDNP foods (where reductions in intake indicated improvements at prescale); and 3) SSBs (where reductions in intake indicated improvements at prescale). To calculate the percentage of the effect size (ie, the between-group difference between follow-up and baseline) reported in the efficacy trial that was achieved in the scale-up trial of the intervention, the following formula was used, similar to that used in a previous review¹⁹:

$$\frac{\text{Effect Size of Scaled – Up study}}{\text{Effect Size of Pre – Scale study}} \times 100$$

A calculation of a scale-up penalty of 100% indicates that the intervention tested in the scaled-up trial had an effect equal to that achieved in the efficacy trial; values > 100% indicate the intervention tested in the scaled-up trial had a greater effect than it did in the efficacy trial; and values < 100% indicate a scale-up penalty (eg, a score of 50% indicates the intervention tested in

the scale-up trial was half as effective than it was in the efficacy trial). Scale-up penalty values < 0% (ie, a negative value) indicate the direction of the effect of the intervention tested in the scaled-up trial was opposite that of the direction of the efficacy trial.

Adaptation to interventions or implementation Study adaptations made for the scaled-up study were narratively described by comparing the intervention described in the prescale trial with the intervention described in the scaled-up study. Where additional intervention descriptions were required, we searched Google and Google Scholar to identify key [supplementary materials](#) (eg, study protocols). Using the Adaptome model,³³ adaptations were classified as 1) service setting adaptations: adaptations made to the environment of the intervention, including intervention delivery setting, and may also include changes to intervention delivery personnel; 2) target audience adaptations: adaptations relating to the target population of intervention; 3) mode of delivery adaptations: included changes made to the channel used to deliver the intervention (eg, change in intervention dose or modality of delivery, such as in-person vs via the internet); or 4) cultural adaptations: adaptations made to the intervention to improve the cultural appropriateness. Other adaptations that could not be classified into these categories were coded as “other.”

Dealing with missing data If any outcome data were deemed missing, the authors of included trials were contacted to provide additional information or clarify information. If data were not provided, the effect size was calculated, where possible, using the available information.

RESULTS

The systematic database search identified a total of 6277 titles to screen for inclusion in this review (Figure 1). An additional 218 titles were identified from other sources. Of these, a total of 174 full-text articles were assessed. Overall, 10 scaled-up intervention pairs were deemed eligible and included in this review. Table 2^{34–53} outlines the initial efficacy RCTs and the corresponding scaled-up interventions.

Characteristics of included studies

The characteristics of the included studies are outlined in Table 3.^{34–43} Of the 10 included scaled-up studies, 3 each were conducted in Australia^{34–36} and the United States,^{37–39} and 1 each was conducted in Canada,⁴⁰ the Netherlands,⁴¹ Sweden,⁴² and the United Kingdom.⁴³

One of the scaled-up interventions included preschool children (aged 3–5 years),³⁵ 3 included primary-school children (aged 6–12 years),^{37,40,42} and another focused on parent-child dyads.³⁴ Two included trials focused on adolescents (aged 12–14 years),^{36,41} 1 of which specifically focused on adolescent boys.³⁶ One scaled-up intervention targeted women only (aged ≥ 40 years),³⁹ 1 focused on male adults (aged 30–65 years),⁴³ and another included adults (ages ≥ 18 years) who attended church.³⁸ Five of the scaled-up trials used a cluster RCT design^{35–37,40,42}; 2 used an RCT study design^{34,43} and 1 study each used a cluster controlled trial,⁴¹ a prospective group randomized trial,³⁸ or a pretest-posttest within-participant design.³⁹ Only 2 of the scale-up trials had conducted an economic evaluation.^{39,43} Researchers also used a variety of methods to assess dietary outcomes, such as FFQs, 24-hour recalls, direct observations, and validated surveys.

The effect of evidenced-based nutrition interventions after scale-up

In 9 of the 10 scaled-up studies, authors reported statistically significant improvements ($P < 0.05$) in at least 1 diet related outcome due to the intervention^{34–37,39–43} (Table 3). These included studies across a variety of settings targeting preschools, primary and secondary schools, families, and sporting clubs and communities, and ranging in intervention length from 12 weeks to 2 years. Only 1 study, which used a pre-post within-participant design and targeted F&V consumption in adults within the church setting, was deemed not effective.³⁸ Across the 10 scaled-up studies, 20 different diet-related measures were reported, with 16 of the 20 measures indicating significantly improved outcomes after scale-up. Fruit and/or vegetable consumption was measured across 8 of the intervention pairs,^{34,35,37–40,42,43} reduction in EDNP foods^{35,41–43} and SSBs^{35,36,41,42} was measured across 4 studies, and other dietary outcomes (ie, portion size, frequency of breakfast consumption) measured in 2 pairs of studies.^{34,41}

Increased F&V consumption Four of the 7 interventions measuring fruit and/or vegetable intakes were found to improve F&V intakes.^{37,39,40,43} Two cluster RCTs conducted with primary school-aged children described improvements in child F&V intakes. Authors reported increased number of fruit servings, F&V servings, and increased variety of F&V intakes in Action Schools! British Columbia (AS! BC),⁴⁰ whereas the 5-a-Day study³⁷ reported increased intakes of F&V servings (with and without potato), and fruit servings (with or without juice). The StrongWomen–Healthy Hearts³⁹ intervention led to increases in daily F&V intakes in adult

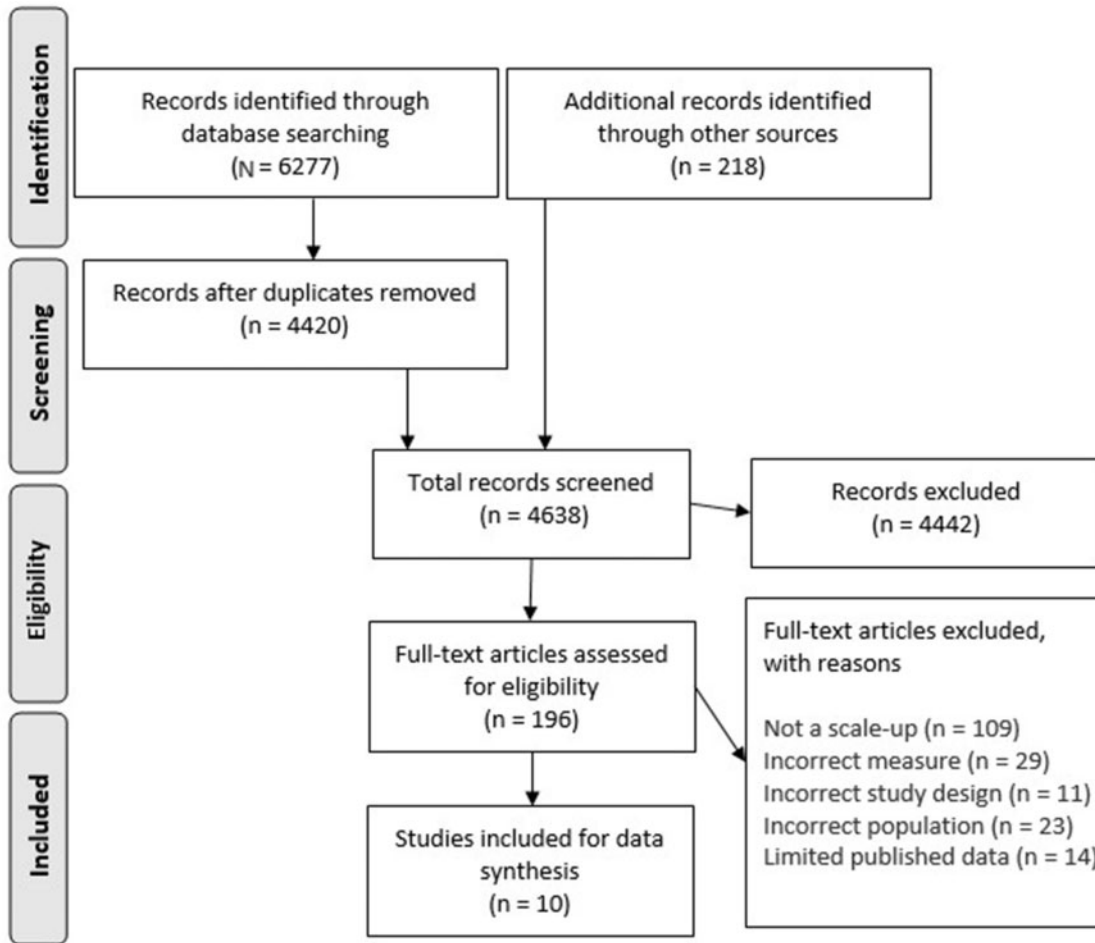


Figure 1 PRISMA flow diagram of included studies

women. European Fans in Training (EuroFIT),⁴³ an intervention targeting adult males, resulted in improved F&V intakes at intervention completion (12 weeks) and long-term follow-up (12 months).

Reduction in intake of EDNP foods Two of the 4 interventions that measured intakes of EDNP foods resulted in reductions in consumption. In Healthy School Start,⁴² a cluster RCT conducted in Swedish preschools, the intervention group had lower intakes of EDNP foods (including snacks, ice cream, cookies, and other sweets) at 6 months' follow-up; however, the effect was no longer significant at 10 months. In an RCT conducted with adult men,⁴³ reductions in fatty food scores and sugary food scores were found at intervention completion (12 weeks), an effect that was retained at 12 months from baseline.

Reduction in SSB consumption All 4 interventions that measured SSB intakes reported reductions in SSB consumption as a result of the scale-up intervention. In

Healthy School Start,⁴² reductions in the consumption of SSBs (eg, soft drinks, flavored milk, juice) were reported at 6-month follow-up; however this effect was not sustained at 10 months. For the Munch & Move program,³⁵ authors reported an improved intervention effect on SSB consumptions, indicated by a reduction in the number of SSBs provided in preschool lunchboxes. The Active Teen Leaders Avoiding Screen Time (ATLAS) study,³⁶ a cluster RCT conducted with adolescent boys, led to reduced SSB consumption, whereas a reduction in SSB consumption was reported in adolescent girls only in the Dutch Obesity Intervention in Teenagers (DOiT) study.⁴¹

Other dietary outcomes Healthy Dads, Healthy Kids (HDHK)³⁴ was the only trial to include parent-child dyads. The intervention reduced the usual portion size (measured via portion-size factor) reported by fathers, and significantly reduce energy intakes by children. Frequency of breakfast consumption as a result of

Table 2 List of included interventions evaluated in the prescale efficacy trial and corresponding scaled-up study, including the target population and general intervention focus

Prescale RCT intervention name	Scaled-up intervention name (population, focus)
AS! BC ⁴⁴	AS! BC (primary school children, primarily physical activity with nutrition outcomes) ⁴⁰
PALs ⁴⁵	ATLAS (adolescent boys, obesity) ³⁶
DoiT ⁴⁶	DOiT (adolescents, obesity) ⁴¹
Healthy School Start ⁴⁷	Healthy School Start II (primary school children, lifestyle) ⁴²
StrongWomen–Healthy Hearts ⁴⁸	StrongWomen–Healthy Hearts (women, lifestyle) ³⁹
Body & Soul ⁴⁹	Body & Soul (adults; fruit and vegetable intake) ³⁸
Healthy Dads, Healthy Kids ⁵⁰	Healthy Dads, Healthy Kids (parent/child dyads, obesity) ³⁴
Football Fans in Training ⁵¹	EuroFIT (men, lifestyle) ⁴³
Tooty Fruity ⁵²	Munch & Move (preschool children, obesity) ³⁵
5-a-Day ⁵³	5-a-Day Cafeteria Power Plus (primary school children; fruit and vegetable intake) ³⁷

Abbreviations: AS! BC, Action Schools! British Columbia; ATLAS, Active Teen Leaders Avoiding Screen Time; DOiT, Dutch Obesity Intervention in Teenagers; EuroFIT, European Fans in Training; PALs, Physical Activity Leaders; RCT, randomized controlled trial.

DOiT⁴¹ significantly improved in adolescent boys but not girls.

Differences in effects established before and after scale-up (scale-up penalty)

Four of the 10 included pairs of interventions did not provide sufficient information to enable an assessment of the scale-up effect on dietary outcomes.^{36,39,41,42} One did not have the same measure of F&V,³⁹ another did not have a common measure for SSBs,³⁶ and the other 2 (although they had a common measure of SSB⁴¹ or F&V⁴² intake at prescale and scale-up) did not have a significant effect on those measures at prescale. Of the 6 pairs of interventions that provided sufficient data for comparison for at least 1 standardized measure of diet, the scale-up effect was highly varied, ranging from –32.2% (Munch & Move³⁵) to 222.2% (change in F&V for AS! BC⁴⁰) (Table 4)^{34,35,37,38,40,43} and a scale-up penalty was observed for at least 1 dietary outcome in 6 pairs of scaled-up studies. Overall, the effect size reported in the scaled-up trials was typically 50% of the effect reported in the prescale efficacy trials (Table 4).

Five of the 6 pairs of studies included a measured F&V consumption common across both trials from prescale to scale-up.^{35,37,38,40,43} The effect size retained in scale-up interventions measuring F&V consumption ranged from –3.3% to 222.2% (median, 50.0% effect retained). Only 1 trial (Munch & Move) did not have any effect on F&V once scaled.³⁵ The remaining 5 were all effective at scale-up, with researchers on 2 trials (AS! BC⁴⁰ and EuroFIT⁴³) reporting a higher effect for F&V consumption at 12-month follow-up in the scale-up study compared with the original effect from the prescale trial.^{40,43}

In 2 of the 6 pairs of studies, researchers included consistent measures of consumption of EDNP foods common across both trials from prescale to scale-up.^{35,43} Of these studies, those that had been scaled up

retained –32.2% to 77.0% of the intervention effect size achieved at prescale. The Munch & Move study³⁵ did not retain any effect size at scale-up (the intervention resulted in a negative effect) for EDNP foods packed in lunchboxes. EuroFIT⁴³ had an effect on EDNP foods at scale-up, but it did not retain the effect of the prescale intervention for fatty foods or sugary food scores at 12 weeks or 12 months.

There were no trials that had a significant effect on SSB intake at prescale or that had a common measure of SSB intake; thus, the scale-up effect was not quantified.

HDHK³⁴ was the only intervention that measured child energy intakes at prescale and scale-up. Only 40% of the prescale effect for energy intakes was retained at scale-up.

Adaptation occurring as part of the scale-up process A full description of each prescale intervention and details of all reported adaptations made for the scaled-up variation are provided in Table S2 in the Supporting Information online. Table 5^{34–43} outlines the broad categories of these reported adaptations. For all the included interventions, researchers made adaptations to the intervention related to mode of delivery. For example, the frequency of sessions increased by 10 minutes in ATLAS,³⁶ and the parent workshops, newsletters, and DVDs were removed in Munch & Move.³⁵ Other common adaptations related to service setting (n = 7 of 10; eg, workshops were conducted at local schools by local physical education teachers instead of research staff from HDHK³⁴) and “other” (n = 6 of 10; eg, reducing screen time was added as a program component in ATLAS³⁶). Three of the included intervention trials reported adaptations related to the target audience (eg, an additional 5 school grades were included in AS! BC⁴⁰) and only 2 interventions made cultural adaptations (eg, the parent brochure was translated into

Table 3 Characteristics of included scaled-up studies, by outcome

Reference; country; INT name	Study design	Setting	Population	Measure of diet	INT length and follow-up time points	Key dietary findings	Translation stage
McKay et al (2015) ⁴⁰ ; Canada; ASI BC	Cluster RCT	10 primary schools within British Columbia province, Canada	<ul style="list-style-type: none"> N = 527 participants Eligibility: students in grades 4 and 5, 8–12 y old 	24 h recalls and food frequency questionnaire	<ul style="list-style-type: none"> INT length: 2 school y Follow-up: INT completion (20 mo from baseline) 	<p>Significant effect on:</p> <ul style="list-style-type: none"> Number of fruit (INT: +0.2 servings/d vs CON: –0.5 servings/d; $P \leq 0.05$) Number of F&V (INT: +0.4 servings/d vs CON: –0.7 servings/d; $P \leq 0.05$) Variety of F&V (INT: +0.3 servings/d vs CON: –0.2 servings/d; $P \leq 0.05$) 	Dissemination
Nyberg et al (2016) ⁴² ; Sweden; Health School Start II	Cluster RCT	31 preschool classes from 13 low-income schools in Stockholm County, Sweden	<ul style="list-style-type: none"> N = 378 participants Eligibility: 6 y old children from families of low socioeconomic status schools (below the average income level) (preschool class) 	Validated parent-proxy questionnaire, the EPAQ	<ul style="list-style-type: none"> INT length: 6 mo. Follow-up: INT completion (6 mo from baseline: [T1]), and 10–11 mo from baseline (T2) 	<p>NS effect post-INT or 10–11 mo after INT follow-up:</p> <ul style="list-style-type: none"> Difference in F&V: –0.02 servings/d, $P = 0.79$ (T1); –0.03 servings/d, $P = 0.68$ (T2). 	Effectiveness
Folta et al (2015) ³⁹ ; United States; StrongWomen–Healthy Hearts	Pre–post-test within-participant design	22 US states	<ul style="list-style-type: none"> N = 345 participants Eligibility: women ≥ 40 y old, BMI ≥ 24 kg m^{–2}, sedentary 	Validated “5-a-Day for Better Health 7-item screener”	<ul style="list-style-type: none"> INT length: 12 wk Follow-up: INT completion 	<p>Significant increase in mean daily servings of F&V of 2.1 (SE = 0.3; $P < 0.001$).</p>	Effectiveness
Allcock et al (2012) ³⁸ ; United States; Body & Soul	Prospective group randomized trial	16 churches located within diverse urban areas with large populations of African Americans	<ul style="list-style-type: none"> N = 1033 participants Eligibility: adults (≥ 18 y old) who attend a participating church 	2 validated measures: <ul style="list-style-type: none"> 10-item F&V questionnaire, measuring frequency of F&V intake over the last month 2-item measure assessing usual F&V intake 	<ul style="list-style-type: none"> INT length: 6 mo Follow-up: INT completion 	<p>NS effect on daily F&V intakes post-INT (INT: 4.7 servings/d vs CON: 4.4 servings/d). NS P-values not reported.</p>	Dissemination and
implementation Burrows et al (2012) ³⁴ ; Australia; Healthy Dads, Healthy Kids	RCT	The Hunter region of NSW (Singleton and Maitland)	<ul style="list-style-type: none"> n = 93 fathers, n = 132 children Eligibility: overweight and obese fathers of primary-school aged children 	<ul style="list-style-type: none"> Father intake: 74-item Dietary Questionnaire for Epidemiological Studies, version 2 Child intake: Australian Child and 	<ul style="list-style-type: none"> INT length: 3 mo Follow-up: 6 mo from baseline 	<ul style="list-style-type: none"> NS effect on father or child consumption of F&V post-INT Father: mean difference post INT: vegetables: –0.2 servings/ 	Dissemination

(continued)

Table 3 Continued

Reference; country; INT name	Study design	Setting	Population	Measure of diet	INT length and follow-up time points	Key dietary findings	Translation stage
Wyke et al, (2019) ⁴³ ; United Kingdom; EuroFIT	RCT	15 football clubs in the Netherlands, Norway, Portugal, and the United Kingdom (England)	<ul style="list-style-type: none"> N = 1113 participants Eligibility: men, 30–65 y old, BMI \geq 27 kg/m² 	<p>Adolescent Eating Survey FFQ (reported by mothers)</p> <ul style="list-style-type: none"> Dietary instrument for Nutrition Education Used to calculate F&V score (range: 1–12) 	<ul style="list-style-type: none"> INT length: 12 wk Follow-up: INT completion and 12 mo from baseline 	<ul style="list-style-type: none"> d; fruit: +0.4 servings/d Child: mean difference post-INT: vegetables: -0.8 servings/d; fruit: -0.4 servings/d NS <i>P</i>-values not reported Significant improvements in F&V scores: Post-program: +1.26; <i>P</i> < 0.001 12-mo post: + -0.96; <i>P</i> < 0.001 	Effectiveness
Hardy et al (2010) ³⁵ ; Australia; Munch & Move	Cluster RCT	29 preschools in Sydney, NSW, Australia	<ul style="list-style-type: none"> N = 430 participants Eligibility: children attending preschool in Sydney, Western Sydney, and South Western Sydney 	Nutritional Quality of lunchboxes were assessed using lunchbox audits as a proxy for quality of dietary intake.	<ul style="list-style-type: none"> INT length: ~4 mo. Follow-up: INT completion 	<ul style="list-style-type: none"> NS effect on the number of F&V servings packed in lunchboxes Fruit servings: -0.05 servings/d, <i>P</i> = 0.75 Vegetable servings: +0.12, <i>P</i> = 0.16 Significant effect on: <ul style="list-style-type: none"> F&V without potato: +0.14 servings, <i>P</i> = 0.03 F&V without potato or juice: +0.15 servings, <i>P</i> = 0.02 Fruit: +0.16 servings, <i>P</i> = 0.01 Fruit without juice: +0.17 servings, <i>P</i> < 0.001 NS effect on: <ul style="list-style-type: none"> F&V: +0.09 servings, <i>P</i> = 0.33 Vegetables: -0.06 servings, <i>P</i> = 0.32 	Effectiveness
Perry et al (2004) ³⁷ ; United States; 5-a-Day Cafeteria Power Plus	Cluster RCT	26 schools in the Minneapolis–St. Paul metropolitan area of Minnesota	<ul style="list-style-type: none"> N = 1820 participants Eligibility: elementary school students in the first and third grades 	Cafeteria lunch observations conducted by trained observers to record all items eaten at lunch and their portion sizes.	<ul style="list-style-type: none"> INT length: 2 school y. Follow-up: INT completion 	<ul style="list-style-type: none"> Significant effect on: <ul style="list-style-type: none"> F&V without potato: +0.14 servings, <i>P</i> = 0.03 F&V without potato or juice: +0.15 servings, <i>P</i> = 0.02 Fruit: +0.16 servings, <i>P</i> = 0.01 Fruit without juice: +0.17 servings, <i>P</i> < 0.001 NS effect on: <ul style="list-style-type: none"> F&V: +0.09 servings, <i>P</i> = 0.33 Vegetables: -0.06 servings, <i>P</i> = 0.32 	Dissemination

(continued)

Table 3 Continued

Reference; country; INT name	Study design	Setting	Population	Measure of diet	INT length and follow-up time points	Key dietary findings	Translation stage
EDNP foods: reduced intakes represent INT improvements van Nassau et al (2014) ⁴¹ ; The Netherlands; DOIIT	Cluster-CON-led trial	29 prevocational Dutch secondary schools	<ul style="list-style-type: none"> N = 1486 participants Eligibility: adolescents 12–14 y old 	Questionnaire that included items assessing the consumption of high-energy snacks and sweets	<ul style="list-style-type: none"> INT length: 20 mo. Follow-up: INT completion 	<ul style="list-style-type: none"> Vegetables without potato: -0.02 servings, $P = 0.55$ NS effect on intakes of EDNP foods. NS P-values not reported. 	Effectiveness
Nyberg et al (2016) ⁴² ; Sweden; Healthy School Start II	Cluster RCT	31 pre-school classes from 13 low income schools in Stockholm country, Sweden	<ul style="list-style-type: none"> N = 378 participants Eligibility: 6-y-old children in families with low socioeconomic status schools (preschool class) 	Validated parent proxy questionnaire, the Eating and Physical Activity Questionnaire (EPAQ).	<ul style="list-style-type: none"> INT length: 6 mo Follow-up: INT completion (6 mo from baseline: [T1]), and 10–11 mo from baseline (T2) 	<ul style="list-style-type: none"> Significantly lower intake of EDNP foods (eg, snacks, ice-cream, cookies, sweets; -0.32 servings/d, $P = 0.01$), at T1 NS effect at T2 (-0.15 servings/d, $P = 0.42$) 	Effectiveness
Wyke et al (2019) ⁴³ ; United Kingdom; EuroFIT	RCT	15 football clubs in the Netherlands, Norway, Portugal, and the United Kingdom (England)	<ul style="list-style-type: none"> N = 1113 participants Eligibility: men, 30–65 y old, BMI ≥ 27 kg/m² 	<ul style="list-style-type: none"> Self-reported diet reported using an adapted DINE. Used to calculate: sugary food score (range, 3–18) Fatty food score (range, 6.5–66.5) 	<ul style="list-style-type: none"> INT length: 12 wk Follow-up: INT completion and 12 mo from baseline 	<ul style="list-style-type: none"> Significant INT effects on fatty food score and sugary food scores post program and at 12 mo after INT. Fatty food score: Post-program: -1.65, $P < 0.001$; 12 mo after: -1.4, $P < 0.001$ Sugary food score: after program: -0.94, $P < 0.001$; 12-mo after: -0.67, $P < 0.001$ 	Effectiveness
Hardy et al (2010) ³⁵ ; Australia; Munch & Move	Cluster RCT	29 preschools in Sydney, NSW, Australia	<ul style="list-style-type: none"> N = 430 participants Eligibility: children attending preschool in Sydney, Western Sydney, and South Western Sydney 	Nutritional quality of lunchboxes was assessed using lunchbox audits as a proxy for quality of dietary intake.	<ul style="list-style-type: none"> INT length: ~ 4 mo Follow-up: INT completion 	<ul style="list-style-type: none"> NS effect on the number of EDNP snack servings packed in lunchboxes Snack servings: $+0.06$, $P = 0.75$ 	Effectiveness
SSB: reduced intakes represent INT improvements	Cluster RCT		<ul style="list-style-type: none"> N = 361 participants 		<ul style="list-style-type: none"> INT length: 20 wk. 		Effectiveness

(continued)

Table 3 Continued

Reference; country; INT name	Study design	Setting	Population	Measure of diet	INT length and follow-up time points	Key dietary findings	Translation stage
Smith et al (2014) ³⁶ ; Australia; ATLAS		14 state-funded secondary schools in low-income communities in NSW, Australia	<ul style="list-style-type: none"> Eligibility: adolescent boys in grade 7 (first year of secondary school), aged 12–14 y old 	NSW SPANS: 2 items related to SSB consumption	<ul style="list-style-type: none"> Follow-up: INT completion 	Significantly reduced intakes of SSB: –0.6 servings/d, $P = 0.01$ in INT group.	
van Nassau et al (2014) ⁴¹ ; the Netherlands; DOIIT	Cluster-CON-led trial	29 prevocational Dutch secondary schools	<ul style="list-style-type: none"> $N = 1486$ participants Eligibility: adolescents 12–14 y old 	Questionnaire that included items assessing consumption of SSB	<ul style="list-style-type: none"> INT length: 20 mo Follow-up: INT completion 	<ul style="list-style-type: none"> Significantly lower consumption of SSB by girls (INT vs CON: 974.4 mL/d vs 1259.1 mL/d; $P \leq 0.05$) This effect was NS in boys (INT vs CON: 1464.2 mL/d vs 1338.0 mL/d). NS P values not reported. 	Effectiveness
Nyberg et al. (2016) ⁴² ; Sweden; Healthy School Start II	Cluster RCT	31 preschool classes from 13 low-income schools in Stockholm County, Sweden	<ul style="list-style-type: none"> $N = 378$ participants Eligibility: 6-y-old children of families who attend schools in low-socio economic-status areas (pre-school class) 	Validated parent-proxy questionnaire, the EPAQ	<ul style="list-style-type: none"> INT length: 6 mo Follow-up: INT completion (6 mo from baseline: [T1]), and 10–11 mo from baseline 	<ul style="list-style-type: none"> Significantly lower intakes of SSBs at 5–6 mo in the INT group (–0.51 servings/d) compared with CON ($P = 0.01$) NS at 10–11 mo (–0.05 servings/d), $P = 0.83$ 	Effectiveness
Hardy et al (2010) ³⁵ ; Australia; Munch & Move	Cluster RCT	29 preschools in Sydney, NSW, Australia	<ul style="list-style-type: none"> $N = 430$ participants Eligibility: children attending preschool in Sydney, Western Sydney, and South Western Sydney 	Nutritional quality of lunchboxes was assessed using lunchbox audits as a proxy for quality of dietary intake	<ul style="list-style-type: none"> INT length: ~4 mo. Follow-up: INT completion 	Significant decrease in SSBs packed in lunchboxes (–0.13 servings/d, $P = 0.05$).	Effectiveness
Other dietary outcomes Burrows et al (2012) ³⁴ ; Australia; Healthy Dads, Healthy Kids	RCT	The Hunter region of NSW (Singleton and Maitland), Australia	<ul style="list-style-type: none"> $n = 93$ fathers, $n = 132$ children Eligibility: overweight and obese fathers of primary-school aged children 	<ul style="list-style-type: none"> Fathers' dietary intake: measured using a 74-item dietary questionnaire for Epidemiological Studies version 2, assessing usual eating habits during the last 12 mo 	<ul style="list-style-type: none"> INT length: 3 mo. Follow-up: 6 mo from baseline 	<ul style="list-style-type: none"> INT significant for: <ul style="list-style-type: none"> Father's usual portion size (–0.3, $P = 0.05$) Child's total energy intake (–1809 kJ/d, $P = 0.05$); Child energy intake/kg (–106 kJ/kg, $P = 0.05$). 	Dissemination

(continued)

Table 3 Continued

Reference; country; INT name	Study design	Setting	Population	Measure of diet	INT length and follow-up time points	Key dietary findings	Translation stage
van Nassau et al (2014) ⁴¹ ; the Netherlands; DOIIT	Cluster-CON-led trial	29 prevocational Dutch secondary schools	<ul style="list-style-type: none"> N = 1486 participants Eligibility: adolescents 12–14 y old 	<ul style="list-style-type: none"> Child's dietary intake: measured using Australian Child and Adolescent Eating Survey FFQ to estimate usual child intake (reported by mothers) Questionnaire included items assessing frequency of breakfast consumption 	<ul style="list-style-type: none"> INT length: 20 mo Follow-up: INT completion 	<ul style="list-style-type: none"> Significant effect on frequency of breakfast consumption by boys (INT vs CON: 6 d/wk vs 5.7 d/wk; $P \leq 0.05$). This effect was NS in girls (INT vs CON: 5.7 d/wk vs 5.7 d/wk). NS P values not reported. 	Effectiveness

Abbreviations: ASI, BC, Action Schools; BMI, body mass index; CON, control; DINE, Dietary Instrument for Nutrition Education; EDNP, energy-dense nutrient-poor; EPAC, Eating and Physical Activity Questionnaire; EuroFIT, Europeans Fans in Training; F&V, fruit and vegetables; INT, intervention; NS, not significant; NSW, New South Wales; RCT, randomized controlled trial; SE, standard error; SPANS, schools physical activity and nutrition survey; SSB, sugar-sweetened beverage.

Arabic and Somali—the most common non-English languages in the region for Health School Start⁴²).

DISCUSSION

This review provides the first evaluation, to our knowledge, of the impact of public health nutrition interventions delivered at scale. Across all ages and settings, we found just 10 public health nutrition studies that have been scaled-up after an effective RCT to establish the study intervention's efficacy. These scaled-up interventions varied considerably in their length as well as the reach, dietary assessment measures, and evaluation methods used to assess their effect on nutrition outcomes. Even so, most scaled-up community-based nutrition interventions in this review appear to have had a significant effect on at least 1 dietary outcome measure. Overall, however, relative to their preceding efficacy trial, there appeared to be considerable reductions in the effect size—of approximately 50% (from 17% to 22%)—reported in evaluations of scaled-up interventions.

Although current data are limited, the reduction in effect size after the scale-up process reported in this review appears consistent but slightly larger than findings from other effects of scaling up public health interventions.¹⁹ Authors of a recent systematic review evaluating the effectiveness of obesity management and prevention interventions reported that scaled-up obesity interventions typically represent < 75% of the effect established in the efficacy trials of the intervention.¹⁹ Of the 10 obesity interventions included in that review, 4 studies also reported on nutrition-related outcomes comparable between the efficacy and scaled-up trials. Although scaled-up obesity interventions appear to also report statistically significant nutrition outcomes, a scale-up penalty was similarly observed, with the effect size retained ranging from 22% to 76% of the prescale effect. Because the purpose of scaling up public health interventions is to achieve population health benefits, understanding the potential scale-up penalty is an important consideration for policy makers to ensure such interventions achieve the intended therapeutic health outcomes.

Because the effect size of public health nutrition interventions appears to attenuate quite substantially once interventions are no longer implemented under tightly controlled conditions,^{19,20} more research is warranted to better understand why this occurs. There is a need to remain critical about whether the apparent scale-up penalty is due to the methods used in the scaling-up process or the scalability of the original intervention,⁵⁴ or due to other research or contextual factors, such as the varying sample sizes, evaluation processes, and measures associated with delivering

Table 4 Effect size difference calculated using measures of dietary intake common to both prescale trial and scaled-up study

Study pair	Prescale RCT	Scaled-up study	Proportion of the efficacy trial effect size achieved in the scaled-up study
F&V intake (increases indicate improvements)			
ASI BC ⁴⁰	RCT F&V assessed using: 24-h food recall and FFQ Change in intervention group at 3–6 mo' follow-up: Servings of fruit: +0.24, $P < 0.05$ Servings of F&V: +0.18, $P < 0.05$ Variety of F&V: +0.47, $P < 0.05$	Cluster RCT F&V assessed using: 24 h food recall and FFQ Change in intervention group at 18-mo follow-up: Servings of fruit: +0.2 Servings of F&V: +0.4 Variety of F&V: +0.3	83.3 222.2 63.8
Body & Soul ³⁸	RCT F&V assessed using 2-item measure (frequency and portions) of F&V Post-test mean differences at 6-mo follow-up, adjusted for baseline values F&V servings/d: 2-item measure: +0.7, $P < 0.05$ Fruit servings/d: 1 item: +0.4, $P < 0.05$ Vegetables/d: 1 item: +0.2, $P < 0.05$	Prospective group randomized trial F&V assessed using 2-item measure of F&V Post-test mean difference at 6-mo follow-up, adjusted for baseline values F&V servings/d: 2-item measure: +0.3, $P = 0.16$ Fruit servings/d: 1 item: +0.1, $P = 0.34$ Vegetables servings/d: 1 item: +0.1, $P = 0.11$	42.9 25.0 50.0
EuroFIT ⁴³	RCT F&V scores measured using an adapted version of the DINE Adjusted b/n group differences for 12-wk and 12-mo follow-up F&V score: 12 wk: 1.32 (95%CI, 1.07–1.57), $P < 0.0001$ 12 mo: 0.54 (95%CI, 0.29–0.79), $P < 0.0001$	RCT F&V scores measured using an adapted version of the DINE Adjusted between group differences for 12-wk and 12-mo follow-up F&V score: 12 wk: 1.26 (95%CI, 0.94–1.58), $P < 0.001$ 12 mo: 0.96 (95%CI, 0.63–1.28), $P < 0.001$	95.5 177.8
Tooty fruity; Munch & Move ³⁵	RCT Servings of F&V packed in lunchboxes assessed via lunchbox audits Adjusted b/n group difference at 10-mo follow-up F&V servings in lunchboxes: +0.61, $P = 0.0013$	Cluster RCT Servings of F&V packed in lunchboxes assessed via lunchbox audits Adjusted b/n group difference at 6-mo follow-up F&V servings in lunchboxes: –0.02	–3.3 ^b
5-a-Day ³⁷	RCT F&V intakes assessed via lunchtime observation Post-test b/n group differences at 12-mo follow-up. F&V servings: +0.47, $P < 0.001$ Fruit servings: +0.30, $P < 0.001$ F&V servings/1000 kcal: +0.83, $P < 0.001$ Fruit servings/1000 kcal: +0.72, $P < 0.001$	Cluster RCT F&V intakes assessed via lunchtime observation Post-test b/n group differences at 24-mo follow-up F&V servings: +0.09, $P = 0.33$ Fruit servings: +0.16, $P = 0.01$ F&V servings/1000 kcal: +0.14, $P = 0.27$ Fruit servings/1000 kcal: +0.22, $P = 0.3$	19.1 53.3 16.9 30.6
EDNP intake (reductions indicate improvements) ^a			
EuroFIT ⁴³	RCT Diet scores measured using an adapted version of DINE. Alcohol measured using 7-d recall. Adjusted b/n group differences for 12-wk and 12-mo follow-up Fatty food score: 12 wk: –4.39 (95%CI, –5.16 to –3.61), $P < 0.0001$	RCT Diet scores measured using an adapted version of the DINE. Alcohol intake measured using 7-d recall. Adjusted b/n group differences for 12-wk and 12-mo follow-up Fatty food score: 12 wk: –1.65 (95%CI, –2.26 to –1.04), $P < 0.001$	37.6

(continued)

Table 4 Continued

Study pair	Prescale RCT	Scaled-up study	Proportion of the efficacy trial effect size achieved in the scaled-up study
	12 mo: $-2.74(95\%CI, -2.74 (-3.52 \text{ to } -1.96), P < 0.0001$	12 mo: $-1.40 (95\%CI, -1.97 \text{ to } -0.84), P < 0.001$	51.1
	Sugary food score: 12 wk: $-1.52 (95\%CI, -1.83 \text{ to } -1.21), P < 0.0001$	Sugary food score: 12 wk: $-0.94 (95\%CI, -1.23 \text{ to } -0.66), P < 0.001$	61.8
	12 mo: $-0.87 (95\%CI, -1.18 \text{ to } -0.56), P < 0.0001$	12 mo: $-0.67 (95\%CI, -0.97 \text{ to } -0.38), P < 0.001$	77.0
Tooty fruity-Munch & Move ³⁵	RCT Servings of EDNP items packed in lunchboxes assessed via lunchbox audits. Adjusted b/n group difference at 10-mo follow-up Children with EDNP items in lunch box, %: 0 EDNP items: $+29.1\%,^a P < 0.0001$ 2+ EDNP items: $-24.5\%, P < 0.0001$	Cluster-RCT Servings of EDNP items packed in lunchboxes assessed via lunchbox audits Adjusted b/n group difference at 6-mo follow-up. Children with EDNP items in lunchbox, %: 0 EDNP items: $-1\%,^a P = NS$ 2+ EDNP items: $+7.9\%, P = NS$	 $-3.4^{a,b}$ -32.2^b
Other dietary outcomes			
Healthy Dads, Healthy Kids ³⁴	RCT Mothers of children completed the 137-item ACAES FFQ Post-test b/n group differences a 6-mo follow-up. Child total energy intake (kJ/kg): mean differences between group, 87 (95%CI, 32 – 143), $P = 0.01$	RCT Mothers of children completed the 120-item ACAES FFQ Post-test b/n group differences at 14-wk follow-up Child total energy intake (kJ/kg); Mean differences between groups, 35 (95%CI, $-15 \text{ to } 85$), $P = 0.17$	 40.2

Abbreviations: ACAES, Australian Child and Adolescent Eating Survey; EDNP, energy-dense nutrient-poor; DINE, Dietary Instrument for Nutrition Education; F&V, fruits and vegetables; FFQ, food frequency questionnaire; NS, not significant.

^aA greater intervention effect if percentage of children with 0 EDNP items packed in lunchboxes increased.

^bNo proportion of the effect size was retained by scaled-up intervention (negative effect).

interventions in real-world settings. It may be reasonable to hypothesize that the effect size of an intervention would be variable and ultimately lower, once scaled up to more heterogeneous populations, in comparison with tightly controlled efficacy trials with high internal validity. Furthermore, many interventions tested under ideal research conditions may not be amenable to scale up, because they require expertise and resources not readily available outside of the research environment.⁵⁵ As such, careful consideration regarding intervention scalability, including the potential reach, cost, delivery infrastructure, and fit with the local context,¹⁸ in the development phases is crucial. Failing to do so may result in the development of interventions that require substantial adaptations to enable delivery at a population level,^{54,56} which invariably may reduce intervention effectiveness.⁵⁷

In our review, each of the scaled-up nutrition interventions included adaptations from the original trial that established its efficacy, indicating that considerations to the interventions' scalability within the efficacy phase may have been limited and potentially resulted in the associated scale-up penalty. We recognize that scale-up efforts may have been directed to achieve non-

nutrition-related outcomes where the intervention addressed multiple health risk behaviors. As a result, the adaptations to interventions may have occurred to preserve or enhance the effects on other health outcomes at the expense of those targeting nutrition. Therefore, the scale-up penalty reported in the review for some nutrition outcomes may represent an overestimation. Although a growing number of frameworks exist to guide policy makers and practitioners to make planned adaptations,⁵⁸ assess the scalability of proven effective interventions,⁵⁹ and develop detailed scale-up plans,^{60–63} more empirical data are needed to evaluate the effectiveness of such frameworks.^{54,64}

Despite the observed scale-up penalty, given the significant proportion of both adults and children not meeting current dietary recommendations and the significant savings in health care expenditure that could be achieved by improving dietary intake, even modest improvements in dietary intake may be beneficial when achieved at a population level. For example, in Australia, only 7% of adults and 4% of children⁶⁵ meet current vegetable intake recommendations. Research indicates improving daily vegetable intake at a population level by as little as 10% is estimated to save almost

Table 5 Adaptations made to nutrition interventions for scale up on the basis of the Adaptome model

Trial	Mode of delivery	Service setting	Target audience	Cultural	Other
ASI BC ⁴⁰	X	X	X		X
ATLAS ³⁶	X	X			X
DOiT ⁴¹	X	X			X
Healthy School Start ⁴²	X			X	
StrongWomen-Healthy Hearts ³⁹	X	X			X
Body & Soul ³⁸	X	X			X
HDHK ³⁴	X	X			
EuroFIT ⁴³	X			X	X
Munch & Move ³⁵	X	X	X		
5-a-Day ³⁷	X		X		

Abbreviations: ASI BC, Action Schools! British Columbia; ATLAS, Active Teen Leaders Avoiding Screen Time; DoiT, Dutch Obesity Intervention in Teenagers; EuroFIT, European Fans in Training; HDHK, Healthy Dads, Healthy Kids.

AU\$1 billion in health care expenditure annually.⁶⁶ Our review indicates modest changes in dietary consumption can be achieved when public health nutrition interventions are scaled up, with up to a 0.4-serving increase in F&V⁴⁰ and a reduction of 0.3⁴² and 0.6³⁶ servings of EDNP foods and SSBs, respectively. There is robust evidence from meta-analyses indicating a dose-response relationship between improvements in dietary intake in line with dietary guidelines and a reduction in morbidity associated with chronic disease⁶⁷⁻⁶⁹; thus, even these small improvements in dietary intake associated with scaled-up interventions could have public health benefits.

The results of this review should be interpreted in light of several limitations. First, we only evaluated the effectiveness of scaled-up interventions that resulted after a statistically significant RCT to demonstrate efficacy. To be included, we required the direct scale-up and evaluation of an intervention from a prior RCT. This, no doubt, precluded studies of nutrition intervention that, although not directly originating from a specific RCT, have been prioritized in government policy and delivered at scale. For example, child feeding programs such as those at schools have been delivered at scale across high-, middle-, and low-income countries for many years.⁷⁰ Examination of such literature may provide additional and important insights into the process of scale-up of nutrition interventions. Indig et al¹⁸ demonstrated a variety of pathways to scale up, and 45% of scale-up studies occurred despite the absence of an antecedent and effective RCT or effectiveness trial. Despite this, it is recommended that policy makers and practitioners prioritize the allocation of scarce health resources to the scale-up of interventions with an established evidence base.⁵⁵

Second, although we used a comprehensive search, including searching electronic databases; contacting study authors, institutions, and experts in the field; and cross-referencing with existing reviews in the field, the variability in terminology used in the field represents

challenges,⁷¹ and not all eligible trials may have been identified. For example, some interventions may only assess implementation once scaled up, with no reassessment of effectiveness. Our review was also embedded in a broader systematic review of effective obesity prevention, physical activity, and nutrition scale-up studies and, therefore, included a broad range of search terms. Despite the inclusion of diet-related terms within the exploded terms and Medical Subject headings, our search did not specifically include the term *diet*, which possible would have identified additional relevant studies.

Furthermore, the diet assessment methods used in different studies also differed in their risk of bias and measurement error.⁷² Studies included in the present review assessed diet using self-reported methods (eg, 24-hour recalls, FFQs, screeners) that are prone to socially desirable responding, whereby respondents provide information aligned with expected social norms (eg, overestimation of F&V intakes⁷³). It is also consistently documented in the literature that 24-hour recalls tend to underestimate habitual energy intakes,⁷⁴ whereas FFQs tend to overestimate intakes.^{75,76} However, the studies included in our review primarily used validated methods in an attempt to minimize such biases. Observational methods such as lunchbox audits or cafeteria observations used by Hardy et al³⁵ and Perry et al³⁷ are more objective; however, they assume that the foods purchased (or packed) are consumed. As such, these measurement limitations need to be considered when interpreting the findings of the review and when comparing the effect sizes between included studies. Although it is possible that this heterogeneity in dietary assessment methods used across included studies could explain why greater effect sizes were found in some studies than in others at scale-up, we only explored the differences in effect sizes from prescale to scale-up if the dietary assessment methods were consistent between the 2 studies. Last, accurate coding of adaptations made to scale-up intervention was complex,

because limited details are often included in published reports, due to word limits. The use of standardized reporting methods⁷⁷ may overcome such variability.

CONCLUSION

To improve the nutritional intake of the population, interventions with proven efficacy must be scaled up. The findings of this review demonstrate that current efforts to scale up public health nutrition interventions can be effective, although their effects seem to be considerably attenuated compared with effects reported from efficacy trials. Policy makers and practitioners should anticipate a reduction in the effects of nutrition interventions as they are scaled up. More research is warranted to identify the factors that may help identify approaches to scale-up that are more resilient to reductions in intervention effects.

ACKNOWLEDGEMENTS

Funding. This review was supported by a New South Wales Cancer Council Program Grant (PG 16–05) and The Australian Prevention Partnership Centre. R.L.S. is supported by a National Health and Medical Research Council (NHMRC) Translating Research into Practice fellowship (no. APP1150661). L.W. is supported by an NHMRC Career Development Fellowship (no. APP1128348) and a Heart Foundation Future Leaders Fellowship (no. 101175), and is a Hunter New England Clinical Research Fellow. S.L.Y. is supported by an ARC Discovery Early Career Researcher Award (no. DE170100382). N.K.N. is supported as a Hunter New England Clinical Research Fellow.

Author contributions. R.S., J.J., and L.W. conceived and designed the systematic review. R.S., C.L., and L.W. developed the search strategy. R.S., J.J., C.L., S.M., S.Y. and N.N. screened studies for inclusion. S.M. and C.L. completed the risk of bias assessment. J.J., A.B., C.L., M.L., J.B. and M.M. completed the data extraction. R.S. and J.J. drafted the manuscript, with all coauthors contributing to drafts of the paper. All authors approved the final manuscript.

Declaration of interest. The authors have no relevant interests to declare.

SUPPORTING INFORMATION

The following Supporting Information is available through the online version of this article at the publisher's website.

[Table S1 Electronic search strategy](#)

Table S2 Adaptation table for scaled-up nutrition review

PRISMA 2009 Checklist

REFERENCES

1. Forouzanfar MH, Afshin A, Alexander LT, et al. Global, regional, and national comparative risk assessment of 79 behavioural, environmental and occupational, and metabolic risks or clusters of risks, 1990–2015: a systematic analysis for the Global Burden of Disease Study 2015. *Lancet*. 2016;388:1659–1724.
2. Afshin A, Sur PJ, Fay KA, et al. Health effects of dietary risks in 195 countries, 1990–2017: a systematic analysis for the Global Burden of Disease Study 2017. *Lancet*. 2019;393:1958–1972.
3. World Health Organization. Obesity. Available at: <http://www.who.int/news-room/facts-in-pictures/detail/6-facts-on-obesity>. Accessed June 26, 2018.
4. World Health Organization. Physical activity. Available at: <http://www.who.int/diet-physicalactivity/pa/en/>. Accessed June 26, 2018.
5. World Health Organization. Nutrition. 2020. Available at: <http://www.who.int/topics/nutrition/en/>. Accessed December 2020.
6. World Health Organization. Global action plan for the prevention and control of noncommunicable disease, 2013–2020. 2013. Accessed March 2020.
7. Eakin EG, Lawler SP, Vandelanotte C, et al. Telephone interventions for physical activity and dietary behavior change: a systematic review. *Am J Prev Med*. 2007;32:419–434.
8. Gill TP. Obesity: prevention. In: Caballero B, ed. *Encyclopedia of Human Nutrition*. 3rd ed. Oxford, United Kingdom: Academic Press; 2013:367–373.
9. Racey M, O'Brien C, Douglas S, et al. Systematic review of school-based interventions to modify dietary behavior: does intervention intensity impact effectiveness? *J Sch Health*. 2016;86:452–463.
10. Evans CE, Christian MS, Cleghorn CL, et al. Systematic review and meta-analysis of school-based interventions to improve daily fruit and vegetable intake in children aged 5 to 12 y. *Am J Clin Nutr*. 2012;96:889–901.
11. Schliemann D, Woodside JV. The effectiveness of dietary workplace interventions: a systematic review of systematic reviews. *Public Health Nutr*. 2019;22:942–955.
12. McFadyen T, Chai LK, Wyse R, et al. Strategies to improve the implementation of policies, practices or programmes in sporting organisations targeting poor diet, physical inactivity, obesity, risky alcohol use or tobacco use: a systematic review. *BMJ Open*. 2018;8:e019151.
13. Campbell MK, Hudson MA, Resnicow K, et al. Church-based health promotion interventions: evidence and lessons learned. *Annu Rev Public Health*. 2007;28:213–234.
14. Owen N, Glanz K, Sallis JF, et al. Evidence-based approaches to dissemination and diffusion of physical activity interventions. *Am J Prev Med*. 2006;31:535–44.
15. World Health Organization. Practical guidance for scaling up health service innovations. 2009. http://apps.who.int/iris/bitstream/10665/44180/1/9789241598521_eng.pdf. Accessed April 2020.
16. National Strategic Framework for Chronic Conditions. 2020. Available at: <https://www.health.gov.au/resources/publications/national-strategic-framework-for-chronic-conditions>. Accessed September 2020.
17. Gaziano TA, Galea G, Reddy KS. Scaling up interventions for chronic disease prevention: the evidence. *Lancet*. 2007;370:1939–1946.
18. Indig D, Lee K, Grunseit A, et al. Pathways for scaling up public health interventions. *BMC Public Health*. 2017;18:68.
19. McCrabb S, Lane C, Hall A, et al. Scaling-up evidence-based obesity interventions: a systematic review assessing intervention adaptations and effectiveness and quantifying the scale-up penalty. *Obes Rev*. 2019;20:964–982.
20. Yohros A, Welsh BC. Understanding and quantifying the scale-up penalty: a systematic review of early developmental preventive interventions with criminological outcomes. *J Dev Life Course Criminol*. 2019;5:481–497.
21. Milat AJ, King L, Newson R, et al. Increasing the scale and adoption of population health interventions: experiences and perspectives of policy makers, practitioners, and researchers. *Health Res Policy Syst*. 2014;12:18.
22. Welsh BC, Sullivan CJ, Olds DL. When early crime prevention goes to scale: a new look at the evidence. *Prev Sci*. 2010;11:115–125.
23. Tommeras 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;44:203–216.
24. Higgins JP, Thomas J, Chandler J, et al. *Cochrane Handbook for Systematic Reviews of Interventions*. Chichester, United Kingdom: John Wiley & Sons; 2019.
25. Moher D, Liberati A, Tetzlaff J, et al.; PRISMA Group. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *PLoS Med*. 2009;6:e1000097.
26. Reis RS, Salvo D, Ogilvie D, et al. Scaling up physical activity interventions worldwide: stepping up to larger and smarter approaches to get people moving. *Lancet*. 2016;388:1337–1348.

27. Milat AJ, Bauman A, Redman S. Narrative review of models and success factors for scaling up public health interventions. *Implement Sci.* 2015;10:113.
28. Charif AB, Zomahoun HTV, LeBlanc A, et al. Effective strategies for scaling up evidence-based practices in primary care: a systematic review. *Implement Sci.* 2017;12:139.
29. 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.
30. US National Library of Medicine. Medical subject headings. 2019. Available at: <https://www.nlm.nih.gov/mesh/filelist.html>. Accessed January 2020.
31. Lane C, McCrabb S, Nathan N, et al. How effective are physical activity interventions when they are scaled-up: a systematic review. *Int J Behav Nutr Phys Act.* 2021;18:1–11.
32. Milat AJ, King L, Bauman AE, et al. The concept of scalability: increasing the scale and potential adoption of health promotion interventions into policy and practice. *Health Promot Int.* 2013;28:285–298.
33. Chambers DA, Norton WE. The adaptome: advancing the science of intervention adaptation. *Am J Prev Med.* 2016;51:S124–S131.
34. Burrows T, Morgan PJ, Lubans DR, et al. Dietary outcomes of the Healthy Dads Healthy Kids randomised controlled trial. *J Pediatr Gastroenterol Nutr.* 2012;55:408–411.
35. Hardy LL, King L, Kelly B, et al. Munch and Move: evaluation of a preschool healthy eating and movement skill program. *Int J Behav Nutr Phys Act.* 2010;7:80.
36. 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;134:e723–e731.
37. Perry CL, Bishop DB, Taylor GL, et al. A randomized school trial of environmental strategies to encourage fruit and vegetable consumption among children. *Health Educ Behav.* 2004;31:65–76.
38. Allicock M, Campbell MK, Valle CG, et al. Evaluating the dissemination of Body & Soul, an evidence-based fruit and vegetable intake intervention: challenges for dissemination and implementation research. *J Nutr Educ Behav.* 2012;44:530–538.
39. 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. *Am J Public Health.* 2015;105:2578–2585.
40. McKay HA, Macdonald HM, Nettlefold L, et al. Action Schools! BC implementation: from efficacy to effectiveness to scale-up. *Br J Sports Med.* 2015;49:210–218.
41. 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;11:158.
42. Nyberg G, Norman Å, Sundblom E, et al. Effectiveness of a universal parental support programme to promote health behaviours and prevent overweight and obesity in 6-year-old children in disadvantaged areas, the Healthy School Start Study II, a cluster-randomised controlled trial. *Int J Behav Nutr Phys Act.* 2016;13:4.
43. Wyke S, Bunn C, Andersen E, et al. The effect of a programme to improve men's sedentary time and physical activity: the European Fans in Training (EuroFIT) randomised controlled trial. *PLoS Med.* 2019;16:e1002736.
44. Day ME, Strange KS, McKay HA, et al. Action Schools! BC—healthy eating. *Can J Public Health.* 2008;99:328–331.
45. Lubans DR, Morgan PJ, Aguiar EJ, et al. Randomized controlled trial of the Physical Activity Leaders (PALs) program for adolescent boys from disadvantaged secondary schools. *Prev Med.* 2011;52:239–246.
46. Singh AS, Paw MJCA, Brug J, et al. Dutch obesity intervention in teenagers: effectiveness of a school-based program on body composition and behavior. *Arch Pediatr Adolesc Med.* 2009;163:309–317.
47. Nyberg G, Sundblom E, Norman Å, et al. Effectiveness of a universal parental support programme to promote healthy dietary habits and physical activity and to prevent overweight and obesity in 6-year-old children: the Healthy School Start study, a cluster-randomised controlled trial. *PLoS One.* 2015;10:e0116876.
48. Folta SC, Lichtenstein AH, Seguin RA, et al. 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;99:1271–1277.
49. Resnicow K, Campbell MK, Carr C, et al. Body and Soul: a dietary intervention conducted through African-American churches. *Am J Prev Med.* 2004;27:97–105.
50. Morgan PJ, Lubans DR, Callister R, et al. The 'Healthy Dads, Healthy Kids' randomized controlled trial: efficacy of a healthy lifestyle program for overweight fathers and their children. *Int J Obes.* 2011;35:436–447.
51. 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. *Lancet.* 2014;383:1211–1221.
52. Zask A, Adams JK, Brooks LO, et al. Tooty Fruity Veggie: an obesity prevention intervention evaluation in Australian preschools. *Health Promot J Aust.* 2012;23:10–15.
53. Perry CL, Bishop DB, Taylor G, et al. Changing fruit and vegetable consumption among children: the 5-a-Day Power Plus program in St. Paul, Minnesota. *Am J Public Health.* 1998;88:603–609.
54. Evans RE, Craig P, Hoddinott P, et al. When and how do 'effective' interventions need to be adapted and/or re-evaluated in new contexts? The need for guidance. *J Epidemiol Community Health* 2019;73:481–482.
55. Milat A, Lee K, Conte K, et al. Intervention Scalability Assessment Tool: a decision support tool for health policy makers and implementers. *Health Res Policy Syst.* 2020;18:1.
56. Power J, Gilmore B, Vallières F, et al. Adapting health interventions for local fit when scaling-up: a realist review protocol. *BMJ Open.* 2019;9:e022084.
57. Alvidrez J, Nápoles AM, Bernal G, et al. Building the evidence base to inform planned intervention adaptations by practitioners serving health disparity populations. *Am J Public Health.* 2019;109:S94–S101.
58. Stirman SW, Miller CJ, Toder K, et al. Development of a framework and coding system for modifications and adaptations of evidence-based interventions. *Implement Sci.* 2013;8:65.
59. Zamboni K, Schellenberg J, Hanson C, et al. Assessing scalability of an intervention: why, how and who? *Health Policy Plan.* 2019;34:544–552.
60. Milat AJ, Newson R, King L, et al. A guide to scaling up population health interventions. *Public Health Res Pract.* 2016;26:e2611604.
61. World Health Organization. *Nine steps for developing a scaling-up strategy.* France: World Health Organization; 2010.
62. Barker PM, Reid A, Schall MW. A framework for scaling up health interventions: lessons from large-scale improvement initiatives in Africa. *Implement Sci.* 2016;11:12.
63. Yamey G. Scaling up global health interventions: a proposed framework for success. *PLoS Med.* 2011;8:e1001049.
64. Fagan AA, Bumbarger BK, Barth RP, et al. Scaling up evidence-based interventions in US public systems to prevent behavioral health problems: challenges and opportunities. *Prev Sci.* 2019;20:1147–1168.
65. Australian Institute of Health Welfare. Australia's health 2018. 2018. <https://www.aihw.gov.au/reports/australias-health/australias-health-2018>. Accessed March 2020.
66. Deloitte Access Economics. The impact of increasing vegetable consumption on health expenditure. 2019. <https://www2.deloitte.com/content/dam/Deloitte/au/Documents/Economics/deloitte-au-economics-increasing-vegetable-consumption-health-expenditure-impact-040716.pdf>. Accessed April 2020.
67. Aune D, Giovannucci E, Boffetta P, et al. Fruit and vegetable intake and the risk of cardiovascular disease, total cancer and all-cause mortality—a systematic review and dose-response meta-analysis of prospective studies. *Int J Epidemiol.* 2017;46:1029–1056.
68. Wu Y, Zhang D, Jiang X, et al. Fruit and vegetable consumption and risk of type 2 diabetes mellitus: a dose-response meta-analysis of prospective cohort studies. *Nutr Metab Cardiovasc Dis.* 2015;25:140–147.
69. Bechthold A, Boeing H, Schwedhelm C, et al. Food groups and risk of coronary heart disease, stroke and heart failure: a systematic review and dose-response meta-analysis of prospective studies. *Crit Rev Food Sci Nutr.* 2019;59:1071–1090.
70. Wang D, Fawzi WW. Impacts of school feeding on educational and health outcomes of school-age children and adolescents in low-and middle-income countries: protocol for a systematic review and meta-analysis. *Syst Rev.* 2020;9:1–8.
71. 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;9:781.
72. National Institutes of Health National Cancer Institute. Dietary assessment primer. Available at: <https://dietassessmentprimer.cancer.gov/>. Accessed February 2, 2021.
73. Miller TM, Abdel-Maksoud MF, Crane LA, et al. Effects of social approval bias on self-reported fruit and vegetable consumption: a randomized controlled trial. *Nutr J.* 2008;7:1–7.
74. Lopes T, Luiz R, Hoffman D, et al. Misreport of energy intake assessed with food records and 24-h recalls compared with total energy expenditure estimated with DLW. *Eur J Clin Nutr.* 2016;70:1259–1264.
75. Noor Hafizah Y, Ang LC, Yap F, et al. Validity and reliability of a food frequency questionnaire (FFQ) to assess dietary intake of preschool children. *Int J Environ Res Public Health.* 2019;16:4722.
76. Fumagalli F, Monteiro JP, Sartorelli DS, et al. Validation of a food frequency questionnaire for assessing dietary nutrients in Brazilian children 5 to 10 years of age. *Nutrition.* 2008;24:427–432.
77. Plint AC, Moher D, Morrison A, et al. Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med J Aust.* 2006;185:263–267.