Review Article

Strengthening the rigour of population-wide, community-based obesity prevention evaluations

Luke Wolfenden[1],[*],[†] and John Wiggers[1],[2]

[1]School of Medicine and Public Health, Faculty of Health, The University of Newcastle, Newcastle, New South Wales, Australia: [2]Hunter New England Population Health, Newcastle, New South Wales, Australia

Abstract

Objective: The aim of the present study was to review the methodological literature regarding evaluation methods for complex public health interventions broadly and, based on such methods, to critically reflect on the evaluation of contemporary community-based obesity prevention programmes.

Design: A systematic review of the methods and community-based literature was performed by one reviewer.

Results: The review identified that there is considerable scope to improve the rigour of community-based obesity prevention programmes through: prospective trial registration; the use of more rigorous research designs, particularly where routine databases including an objective measure of adiposity are available; implementing strategies to quantify and reduce the risk of selective non-participation bias; the development and use of validated instruments to assess intervention impacts; reporting of intervention process and context information; and more comprehensive analyses of trial outcomes.

Conclusions: To maximise the quality and utility of community-based obesity prevention evaluations, programme implementers and evaluators need to carefully examine the strengths and pitfalls of evaluation decisions and seek to maximise evaluation rigour in the context of political, resource and practical constraints.

Excessive weight substantially increases the risk of a variety of chronic diseases including cancer, CVD and diabetes[1], and accounts for over 2 million deaths each year[2] and more than 30 million disability-adjusted life years[3]. While evidence regarding the effectiveness of obesity prevention initiatives is equivocal[4], there have been calls from international organisations[5], professional associations[6], academic experts[7,8] and the community[9] for governments to take decisive action to mitigate the adverse impacts of obesity. Indeed, across the globe, governments have developed policy, set obesity prevalence targets[10] and are implementing obesity prevention initiatives[11].

Similar to efforts to reduce other chronic disease risk factors in the 1980s and 1990s[12,13], community-based prevention programmes that attempt to address multiple determinants of obesity through multi-component population-wide strategies are frequently recommended[14]. Comprehensive evaluation of such initiatives is recommended

to ensure that conclusions regarding programme effectiveness are valid, to provide an understanding of why and how a programme may have succeeded or failed, and to facilitate programme improvement, policy development and service investment[15,16]. As such, programme evaluation is recognised as a key component of best-practice community-based obesity prevention practice[17].

In order to improve the rigour of future community-based obesity prevention programmes to better inform policy makers and practitioners, the aim of the present paper was to review the methodological literature regarding evaluation methods recommended for complex public health interventions broadly, and community-based interventions specifically, and to critically reflect on the evaluations of contemporary community-based obesity prevention programmes in the context of such recommendations. In doing so, we sought to outline a number of issues we believe represent significant impediments to the quality of community-based obesity prevention programme evaluation and offer some possible ways in which such impediments may be addressed.

† Address for correspondence: c/o Hunter New England Population Health, Locked Bag 10, Wallsend, Newcastle, NSW 2298, Australia.

Corresponding author: Email luke.wolfenden@hnehealth.nsw.gov.au

Method

A review of the literature was performed in January 2011 to identify: (i) key elements of rigorous evaluations of community-based interventions from the research methods literature and from methodological reviews of previous community-based chronic disease prevention trials; and (ii) the evaluation methods used in community-based obesity prevention programmes.

Guidance regarding rigorous evaluation of complex interventions

To identify recommended evaluation methods for community-based interventions we reviewed: (i) the UK Medical Research Council's guidelines for developing and evaluating complex interventions[18]; (ii) recommendations from an expert group on research designs for complex, multilevel health interventions sponsored by the US National Institutes of Health and the Centers for Disease Control and Prevention[19]; and (iii) selected research design and evaluation texts[20–22]. We also examined reviews of methodological issues of past community-based health intervention trials[12,13,23–25].

Community-based obesity prevention programmes

We employed systematic methods to identify previous community-based interventions. While there have been various definitions of what constitutes a community-based intervention[13,26], those which characterised a 'community' along geographical boundaries such as cities, villages or regions are most common[23] and are the subject of discussion in the present paper. As such, to be eligible for inclusion in the review, an obesity prevention intervention must have: (i) been implemented on a population basis across a defined geographic region (such as a town or city); (ii) included a measure of weight status; and (iii) been published in English and in a peer-reviewed journal. Published reports of evaluation designs or study protocols of incomplete trials were also included. Community-based interventions primarily focusing on reducing chronic diseases such as CVD or diabetes where obesity prevention was one of a number of risk factors targeted were excluded.

To identify such programmes, the electronic databases MEDLINE, EMBASE, Cochrane Central Register of Controlled Trials and Google Scholar were searched by one reviewer (L.W.) for articles published between 1990 and 2010. The search strategy for MEDLINE is described in the Appendix and was modified for other databases (a full search strategy for each database is available on request from the corresponding author). Reference lists of systematic reviews including the Cochrane review of interventions for the prevention of obesity[27] and narrative reviews or editorials on the issue of community-based obesity prevention interventions[28–30] were also searched.

Research synthesis

Information from community-based obesity prevention trials and texts providing recommendations regarding their evaluation was synthesised narratively in terms of issues considered to be important in the conduct of rigorous evaluations of community-based interventions, including[13,18,22,23,25,31,32]:

1. ethical and scientific conduct and reporting;
2. research design;
3. data collection;
4. measures; and
5. analysis.

Results

Articles retrieved and included

The electronic database search yielded 1197 citations. Following screening of titles and abstracts, full texts of thirty-three articles were reviewed. Of these, twelve were not population-wide community-based interventions, five targeted a broad range of chronic disease risks, two did not include a measure of weight status, one targeted weight loss rather than prevention, and the intervention and evaluation methods of two trials were still under development or yet to be reported. These trials were deemed ineligible. The remaining eleven articles[33–43] described ten eligible community-based obesity prevention initiatives (Table 1).

Ethical and scientific conduct and reporting

Research ethics

While impropriety has not been raised as an issue in the evaluation and reporting of previous community-based obesity interventions, all programme evaluations are recommended to begin with consideration of standards of ethical and scientific conduct and reporting. Ethics approval and monitoring of study procedures by independent ethics committees (or Institutional Review Boards) and adherence to the World Medical Association Declaration of Helsinki Ethical Principles for Medical Research (http://www.wma.net/en/30publications/10policies/b3/index.html) are an essential part of programme evaluations and a prerequisite for publication in most peer-reviewed journals. The American Evaluation Association's Guiding Principles for Evaluators[31] also provides guidance on proper professional conduct for evaluators and covers issues of competence, integrity, honesty, respect for people and responsibilities for general public welfare.

Trial registration

The UK Medical Research Council recommends that trials are reported regardless of the 'success' of the intervention[18]. Given this, evaluators of community-based obesity prevention programmes should be aware of the potential

Table 1 Evaluation characteristics of the population-wide community-based obesity prevention programmes included in the present review

Reference	Study design	Outcome measures (assessments of community health status)	Impact measures (assessments of community environments or settings)	Process measures (assessments of intervention implementation)	Context assessment	Research participation and attrition
APPLE Project[33,34]	Quasi-experimental: pre–post with comparison Intervention implemented in one community in New Zealand A second geographically separate community served as a comparison (broadly comparable on sociodemographic variables)	BMI (height and weight), WC, pulse rate and BP of children recruited through schools Diet assessed using a validated short FFQ PA assessed using accelerometers and a 7 d recall questionnaire Psychometric properties of the questionnaire not reported	Not reported	Not reported	Not reported	Response rate: 85–92 % and 81–87 % for intervention and comparison community, respectively Attrition rate: 27 % for both intervention and comparison communities
The California Endowment's Healthy Eating, Active Communities Program[35]	Quasi-experimental: pre–post with comparison Intervention delivered in 6 low-income California communities selected as they met criteria including high rates of obesity and capacity to implement a prevention programme Selection method and number of comparison communities not stated	BMI (height and weight) and physical fitness of children using existing quantitative data sets Diet and PA behaviour of children recruited through schools via survey Psychometric properties not reported	Foods and beverages sold in schools and health-care institutions; PA programming and equipment in schools, after-school programmes and parks; neighbourhood food retail and food advertising Psychometric properties of some assessment tools not reported	Not reported	Surveys and focus groups assessing the extent to which the community stakeholders and policy makers were aware of, engaged in and supported the programme	Not reported
The Kaiser Permanente Community Health Initiative[36]	Quasi-experimental: pre–post, repeat cross-sectional Intervention implemented in 3 ethnic minority communities in California, USA	Overweight (measure not specified) among children Fitness among children (measure not specified) PA and nutrition behaviour via automated telephone survey (for adults) or school-based survey Specific psychometric properties not reported	Not reported	A database documenting implementation status and number of people reached (exposed to and affected by the intervention); no other detail provided	Not reported	Not applicable (protocol/methods overview)
Romp & Chomp[37]	Quasi experimental: pre–post, repeated cross-sectional with comparison Intervention implemented in one Australian community selected based on existing collaborative links with the university and health services A sample of LGA from across the rest of the State served as a comparison	BMI (height and weight) collected by nursing staff during routine Maternal and Child Health Key Age and Stage Health Checks for 2–3·5-year-olds PA and nutrition behaviour using validated survey instrument completed by parents attending child health checks	Environmental audits of the physical, policy, socio-cultural and economic environments of Children's Services; active play survey to assess staff and child-care service activities promoting active play; and community capacity building assessment Psychometric properties of survey not reported	Implementation of the project plan (interviews, project records)	Documented other initiatives occurring at the time of the programme	Attendance at 2 years and 3·5 years child health checks: ~60 % and 50 %, respectively Anthropometric data received from 68 % of eligible LGA

Table 1 *Continued*

Reference	Study design	Outcome measures (assessments of community health status)	Impact measures (assessments of community environments or settings)	Process measures (assessments of intervention implementation)	Context assessment	Research participation and attrition
Shape Up Somerville[38]	Quasi-experimental: pre–post with comparison Intervention delivered in one community in the USA, selected due to existing relationships between key community organisations and the research entity Two communities with similar demographic characteristics served as comparisons	Child BMI (height and weight) based on survey of children recruited through schools Child PA and nutrition assessed via parent-completed questionnaire Psychometric properties not reported	Not reported	Project records used to document the extent of implementation and exposure to the intervention	Not reported	Participation rate: 28 % in intervention and comparison communities combined Attrition rate: 39 % in intervention community and 25 % in comparison communities
Be Active Eat Well[39]	Quasi-experimental: pre–post with comparison Intervention delivered in one community in Australia, selected based on existing community infrastructure and supportive networks Schools randomly selected from communities in the broader region served as a comparison	BMI (height and weight) and WC based on survey of children recruited through schools Nutrition and PA assessed using a survey based on a validated questionnaire completed by children and a parent-completed telephone survey A number of potential harms assessed including prevalence of underweight, unhappiness, attempts to lose weight and teasing Specific items of survey instruments or psychometric properties not reported	Home, community and school environments assessed via parent-completed items in a telephone survey and school staff questionnaire Psychometric properties not reported	Implementation of the project plan (interviews, project records)	Not reported	Response rate: 58 % for intervention community and 44 % for comparison communities Attrition rate: 16 % in intervention community and 17 % in comparison communities
Healthy Living Cambridge Kids[40]	Quasi-experimental: pre–post Intervention implemented in one community in the USA	Child BMI (height and weight) routinely collected by the Cambridge Public School system Child fitness test score routinely collected by the Cambridge Public School system	Not reported	Implementation of key components of the intervention reported Method of assessment not stated	Not reported	Participation rate not reported Attrition rate 45 %
Health Promoting Communities: Being Active Eating Well[41]	Quasi-experimental: pre–post repeat cross-sectional with comparison Intervention implemented in 5 disadvantaged communities in Australia Comparison sample from matched schools and workplaces from non-intervention communities across the state of Victoria	BMI (height and weight) and WC Nutrition and PA behaviours assessed via survey methods Psychometric properties not reported Quality of life using the Assessment of Quality of Life mark2 (AQoL) assessment tool adapted from a validated tool developed for use in adults	School environment assessed via questionnaire with school staff Psychometric properties not reported	Assessments of intervention implementation and reach through document analysis, key informant interview, participant feedback and information provided by research programme staff	Not reported	Not applicable (protocol/methods overview)

Table 1 *Continued*

Reference	Study design	Outcome measures (assessments of community health status)	Impact measures (assessments of community environments or settings)	Process measures (assessments of intervention implementation)	Context assessment	Research participation and attrition
The Pacific OPIC Project[42]	Quasi-experimental: pre–post with comparison Interventions implemented in communities in Australia, Tonga, Fiji and New Zealand Intervention communities selected based on criteria including amenable community capacity characteristics (such as existing project champions), geography and ease of access to research staff and other organisations Comparison communities selected to be as comparable as possible (ethnicity, SES, trajectory of weight gain) and geographically separate to minimise contamination	BMI (height and weight) and %BF Healthy eating and PA assessed via questionnaires using 'standard questions' where possible Validity of such assessments not reported Quality of life assessed using established instruments Knowledge assessed (no other information provided) In a sample subset, a validated assessment of perceived body image and body change questionnaire administered	Audit of school environment Validity of assessment tool or collection method not described Community readiness questionnaire (validity not reported) and stakeholder interviews used to assess community capacity	Coordinator reports on implementation activities, action plans, meeting minutes and project records will be used to assess implementation (reach, uptake, barriers, etc.) and implementation costs	Not reported	Not applicable (protocol/methods overview)
Fleurbaix-Laventie Ville Sante study (EPODE)[43]	Quasi-experimental: pre–post repeated cross-sectional Comparison communities assessed at post-intervention only Intervention implemented in two towns in France Two (no intervention) towns served as a comparison	Child BMI (height and weight) assessed during a survey of children recruited through schools	Not stated	Not stated	Not stated	95–98 % participation rate across all schoolchildren from both intervention and comparison towns

LGA, local government areas; SES, socio-economic status; WC, waist circumference; BP, blood pressure; PA, physical activity; %BF, percentage body fat.

for political and programme stakeholder interests to influence the proper conduct and reporting of community programme evaluations. Research has found that stakeholder interests can influence evaluation design decisions of programme evaluators[44] and can suppress[45] or pressure evaluators to misrepresent study findings[46]. The interests or inherent biases of obesity researchers have also been suggested to lead to misleading interpretations of research studies[47]. Prospective trial registration is a scientific convention that can help protect against such influence by requiring evaluators to document specific details of the evaluation design, intended sample, measures and planned analysis on a publically accessible database prior to the research evaluation taking place (Clinical Trial Registration). The requirement for registration of intervention trials was adopted by the International Committee of Medical Journal Editors in 2005 following the establishment of the WHO International Clinical Trials Registry Platform[48]. While a number of current and planned community-based obesity prevention programmes have been registered with an internationally recognised trial registry[33,34,39,41], the extent to which unregistered evaluations adhered to planned evaluation protocols or selectively reported trial outcomes is unable to be assessed. For trials which have utilised a community-based participatory approach, such as the Pacific OPIC Project (Obesity Prevention In Communities)[42], prospectively describing trial methods on a register may represent a considerable challenge given evolving research methods owing to the shared decision making and collective engagement of community members and organisational representatives in each research phase[49]. Nevertheless, documenting changes in research methods in trial registers over the course of the study as they occur and providing justifications for amendments in trial procedures may reduce the risk of selective reporting or other reporting biases.

Scientific reporting
The reporting of trial findings consistent with agreed standards, such as the Consolidated Standards of Reporting Trials (CONSORT) statement for randomised evaluation designs or the Transparent Reporting of Evaluations with Nonrandomized Designs (TREND) statement for non-randomised evaluation designs, is a requirement of many of the leading public health and general medical journals. Reporting of trial outcomes consistent with such standards has been a feature of some past published community-based obesity prevention evaluations[33,34,39]. Greater adherence to such standards improves transparent and consistent reporting of research trials[50] and may facilitate subsequent synthesis of trial outcomes by systematic reviews.

Protocol publication
A reported limitation of previous community-based health promotion programmes has been that a lack of sufficient documentation of intervention and evaluation strategies and procedures[25] has limited the ability to interpret study outcomes. Given word limit restrictions placed by non-electronic journals on submitted papers, the capacity to report the detail required to sufficiently describe evaluations of complex interventions such as community trials has been constrained. Two recent developments represent opportunities to redress this limitation. First, publication of evaluation protocols represents a mechanism whereby such detailed information is able to be provided. Publication of such protocols also serves as a further mechanism to address scientific integrity of subsequent reporting of results through a public statement of evaluation intent. Second, publication of protocols and/or results in electronic open access journals (which do not have space restrictions) or housing additional intervention or evaluation information in other publically accessible websites has been recommended[51] and employed successfully by some existing evaluations of community-based obesity prevention programmes[37,41].

Selection of research design
Previous evaluations of community-based obesity prevention interventions have employed uncontrolled pre–post[40], post-test only comparison[43] or quasi-experimental designs with[33–35,37–39,41,42] and without comparison conditions[36,40]. Such designs are vulnerable to a variety of biases, particularly the influence of secular trends or confounding[20]. The use of more internally valid research designs, such as those employing methods of random allocation (experimental designs), or efforts to strengthen the internal validity of more pragmatic quasi-experimental evaluation designs will provide greater confidence in the extent to which any observed effects can be attributed to a programme[18,20,22].

Randomised experimental designs
Randomised and cluster randomised controlled trials. Randomised (cluster) controlled trials are considered to represent the most internally valid evaluation design and have previously been used to assess the impact of cardiovascular and cancer risk factor community-based interventions[24,52]. A randomised controlled trial may be particularly valuable in innovation testing phased evaluations[22]. Ethical considerations, costs and logistical difficulties in randomising and intervening in large geographically separated communities, however, are well documented as impediments to the use of such experimental trial designs in community-based prevention programmes[53]. Nevertheless, both the UK Medical Research Council[18] and the recommendations from an expert group convened by the National Institutes of Health and the Centers for Disease Control and Prevention recommend the use of randomised controlled trials for programme evaluation whenever feasible[19]. Generally, a sample of at least ten intervention and ten control

	Randomly allocated groups (communities)			
	1	2	3	4
a				
b	■			
c	■	■		
d	■	■	■	
e	■	■	■	■
f	■	■	■	■

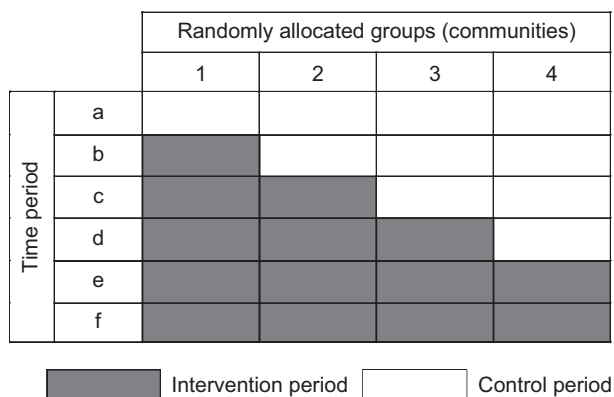■ Intervention period　　□ Control period

Fig. 1 Illustration of a stepped wedge cluster randomised trial

communities is suggested as a minimum number of sites for the conduct of experimental trials of community-based interventions[13].

Stepped wedge cluster randomised trial. A stepped wedge cluster randomised trial may overcome a number of constraints to the conduct of traditional randomised and cluster randomised community-based trials[54]. A stepped wedge design involves measurement of trial outcomes (i.e. weight status) being repeatedly undertaken simultaneously across a number of clusters (i.e. communities) over a period of time. In the context of such measurement, the delivery of the intervention occurs sequentially across communities, the order being randomly determined[55]. Comparisons are made between groups at the intervention section of the wedge (see Fig. 1) as well as within groups before and after the intervention. Advantages of the stepped wedge design for community-based prevention interventions include: the receipt of the intervention by all communities (overcoming ethical concerns regarding withholding intervention); the congruence of the design with intervention capacity constraints and typical schedules of programme rollout; the requirement of fewer clusters than a conventional randomised cluster trial; and the capacity of the design to detect and control for underlying trends and time effects. While the design requires repeated assessment of outcomes (e.g. weight status) which may be expensive and pragmatically challenging, the use of routinely collected weight status data – such as that utilised to assess the impact of Romp & Chomp and the Healthy Living Cambridge Kids obesity prevention programmes[37,40] – represents one approach to achieving these benefits at relatively low cost.

Quasi-experimental designs
Regression discontinuity. Where programme funding permits the use of large numbers of communities but ethical or political considerations deem randomisation not acceptable, regression discontinuity designs have been recommended as a potentially attractive alternative to experimental randomised designs[19,20]. Together with

randomised trials, regression discontinuity designs represent the only other trial design which can provide an unbiased estimate of intervention effect[19,20]. Rather than random assignment, in regression discontinuity designs researchers assign communities (or participants) to intervention or comparison conditions based on their exceeding a designated cut-off point on a pre-intervention assignment variable (e.g. BMI). A regression line through such an outcome variable which discontinues at the point of assignment for control group communities is taken as evidence of an intervention effect. The design may be particularly useful for community-based obesity programmes targeting socio-economically disadvantaged groups, where communities are assigned to receive an intervention based on their score on a measure of disadvantage. Owing to co-linearity between the assignment and outcome variables, however, regression discontinuity designs require substantially more research participants (or communities) compared with a similarly powered randomised trial[19].

Interrupted time series. In a time series evaluation of a community-based obesity intervention, repeated measures of weight status are required for a period prior to, during and following intervention delivery in one or more community. Comparisons are made between the change in level or slope of the outcome prior to and following the intervention to assess intervention effect[19]. The addition of comparison communities which do not receive the intervention improves the internal validity of the design. Like the stepped wedge cluster randomised trial design, the repeated measures allows for control of trends and time effects and is typically a more rigorous evaluation design than before-and-after trial designs[19]. While the use of routinely collected weight status data from institutions such as health services would permit a relatively low-cost evaluation using this design, the lack of such obesity surveillance data in most countries and communities has been previously identified as a significant impediment to its use[29].

Controlled and uncontrolled pre–post trials. Pre–post designs, where outcome assessments are taken at a single point in time prior to and following the delivery of an intervention, represent the most commonly utilised evaluation design in community-based interventions. Such designs, however, face a number of threats to internal validity[19]. Where the inclusion of controlled designs is feasible, programme evaluators should seek to maximise the number of intervention and comparison communities to help account for secular trends[19,20]. Stratifying or matching communities in such designs based on factors known to be closely associated with population weight status is also recommended to reduce the risk of selection bias and to maximise statistical power by increasing precision[20,23]. Identifying communities with comparable characteristics, however, represents a considerable challenge[13]. Even with large numbers of

communities, non-randomised pre–post designs should not be relied upon alone to infer casual inference[19].

With the exception of the Health Promoting Communities: Being Active Eating Well initiative[41], community-based obesity prevention interventions reviewed in the present paper that have employed a controlled design to date have not employed formal stratification or matching techniques. Further, a number of such obesity prevention initiatives[37–39,42] selected intervention communities on the basis that they had existing infrastructure to facilitate intervention delivery, unlike comparison communities. Such known non-equivalence between groups on factors such as community capacity or infrastructure represents a significant threat to the internal validity of trial findings in principle, as such factors are difficult to measure reliably and adjust for in analyses of programme outcomes[56].

Strengthening quasi-experimental designs. As a possible addition to all quasi-experimental designs, evaluators can maximise the internal validity of programme evaluations through the use of non-equivalent dependent variables[19,20]. For example, if a community-based intervention is delivered to all students attending government schools only, assessments of weight status of students in government and non-government schools in the intervention communities pre and post intervention delivery could be conducted. As students of both schools are likely to be subject to non-intervention factors in the community which may impact on weight status (such food marketing or macro-economic policies or events) but only those attending government schools are exposed to the intervention, a differential impact in the change in weight status over time would provide greater evidence of an intervention effect. A second strategy to enhance the internal validity of quasi-experimental designs is via assessment of a dose–response relationship between exposure to intervention activities and trial outcome (e.g. weight status). Dose–response analyses have been found to be a valuable feature of past community-based CVD risk factor intervention[25]. Neither strategy, however, appears to have been employed as an adjunct to evaluations by the community-based obesity prevention programme evaluations employing quasi-experimental designs included in the present review.

Data collection

Survey methods

Obesity prevention interventions have typically utilised survey methods to assess intervention outcomes. Instances where survey participation rates are low, attrition is high or either rate is different between groups limit the validity of inferences regarding intervention effects.

Research participation. Despite high anthropometric survey participation rates among New Zealand[33,34] and French[43] obesity prevention programmes, participation rates for programmes in the USA and Australia have typically ranged between 29% and 58%[38,39]. Research

suggests that obese persons may be less likely to participate in research requiring anthropometric assessments[57]. Examination of the potential for selective non-participation bias by comparing the characteristics of participants with non-participants[48] has not, however, been reported in any obesity prevention programme included in the present review. Identifying non-response bias may be particularly important in repeat cross-sectional evaluations of intervention effect (where independent cross-sectional samples are drawn from communities pre and post intervention), as small changes in the likelihood of participation in overweight participants at follow-up could result in artefactual post-intervention reductions in estimated obesity prevalence. As a crude illustration; assuming a pre-intervention prevalence of child overweight of 25% (250/1000), a relative 20% reduction in the likelihood of participation of overweight persons at post-intervention assessments would reduce post-intervention prevalence to 21% (200/950). Hypothetically, such changes in the propensity of overweight child participation may be an unintended consequence of the intervention itself, if intervention exposure increases perceived obesity stigma and reduces children's interest in participation. Such reductions could be misinterpreted as a public health intervention success.

Evaluation texts and methodological reviews recommend that evaluators seek to reduce the risk of non-response bias by employing intensive recruitment strategies to maximise study participation, such as the pre-notification and promotion of study participation, endorsement of the research by credible individuals or organisations, participation incentives and multiple follow-up reminders[58,59]. For trials that include a comparison condition, reporting participation rates among both intervention communities and comparison communities would allow an assessment of the potential risk of differences in non-participation between groups. More rigorous strategies, such as the recording and analysis of information regarding the weight status of non-participants, would provide more compelling evidence regarding the possible influence of non-participation, if permitted by ethics committees or Institutional Review Boards. For example, Booth and colleagues[60] describe a feasible, inexpensive and valid means of collecting weight status information of children not participating in school-based assessments of height and weight by having classroom teachers match the morphology of such children to those of participants.

Research attrition. Research attrition can threaten the internal and external validity of a community-based intervention evaluation if the characteristics of participants not providing follow-up data differ from those who do, and if the characteristics of participants are dissimilar between intervention and comparison groups. Such differences can lead to over- or underestimates of intervention effect[61]. With the exception of the Be Active Eat Well programme, participant attrition (loss to follow-up)

among child obesity prevention programmes with a cohort design in the present review ranged from 25 % to 45 %[33,35,38,40] and differed substantially between groups in the case of the Shape up Somerville programme (intervention 39 %; comparison 25 %). Even when loss to follow-up is small and not differing significantly between groups, as was the case for the Be Active Eat Well initiative (approximately 16 % loss to follow-up per group), the potential for bias to influence study findings remains present. For example, in that trial 34 % of study withdrawals in the intervention group were due to parental concern regarding their child's self-esteem following weight assessment; a reason for withdrawal which was not cited at all for children in comparison communities. Two of the reviewed cohort studies formally examined the potential for bias due to differences between participants and those lost to follow-up as recommended[48,61]. The Healthy Living Cambridge Kids initiative reported that participants who did not provide follow-up data were more likely to be older, Asian and less likely to pass fitness tests[40]. Project APPLE however found no meaningful differences between participants and those lost to follow-up[33,34].

Similar to those strategies to maximise study participation, evaluators should seek to employ intensive strategies to minimise study attrition such as through offering incentives, alternative locations for data collection and repeat reminders via mail and telephone[62,63], reporting attrition rates by group[48,61], and attempting to formally assess and control for bias due to study attrition in analyses.

Routinely collected institutional databases

Given the cost and feasibility challenges of collecting survey data, the use of routinely collected data such as health services records has been suggested as an inexpensive alternative for programme evaluation that does not require the development of new data collection systems and methodologies, often allows for comparisons with other regions or jurisdictions, can be used retrospectively and may be less likely to be subject to bias due to non-consent[64,65]. Despite these potential benefits, a number of limitations exist regarding the use of routinely collected data for evaluating public health interventions, including reliability between recording personnel, changes in data recording classifications, recording practices and systems, and limited population representativeness[64]. While evaluation of obesity programmes using such data needs to be mindful of these limitations, routinely collected information has a capacity to provide a valuable source of data for community-based obesity programme evaluations, particularly where they are accessed by a large proportion of the target population. For example, routinely collected BMI data were available from maternal and child health services in Victoria, Australia for approximately 60 % of all children of pre-school age and were used to evaluate the Romp & Chomp programme[37].

Similarly, height and weight data routinely collected annually as part of the physical activity curriculum were used to evaluate the effectiveness of the Healthy Living Cambridge Kids initiative[40].

Outcome, impact and process measures

While a number of programme evaluation frameworks exist, the inclusion of measures which assess the outcome of the intervention (as specified by the programme aim), programme impacts on intermediary factors (typically described as part of programme objectives), intervention processes and context are most frequently recommended by evaluation texts and reviews of past cardiovascular community-based interventions[18,22,24,25,66]. Importantly, the selection of evaluation measures should also be guided by programme logic models of how an intervention is intended to produce the intended intervention effect.

Outcome

Weight status. Given the limitations of self-reported measures of weight status[67] and the potential for such reports to be reactive to assessment[23], objective measures of would appear important in programme evaluations looking to examine the effect of a community-based intervention on population adiposity. BMI is the most widely utilised measure of weight status in previous studies[33,34,37–43] and represents a relatively inexpensive and feasible measure of adiposity[4]. To provide information regarding changes in both body and fat composition, additionally assessments of waist circumference and skinfold thickness have been recommended, particularly as interventions including physical activity promotion can actually increase lean body mass and reduce adiposity without any change in BMI[4]. Such assessments should be conducted and reported in accordance with standard measurement protocols to reduce measurement error and to facilitate comparison across trials and data pooling in meta-analyses.

Adverse events. In order to evaluate the merit of a community-based intervention to prevent obesity, both the benefits and adverse effects of the intervention need to be considered. As such, potential harms which may arise from community-based interventions should be hypothesised and assessed as part of programme evaluations[6,18,68]. Despite a few exceptions, assessment of harms has largely been overlooked in previous or planned community-based obesity prevention intervention evaluations included in the present review. The Be Active Eat Well child obesity programme was the only programme which explicitly included measures of harm such as the prevalence of underweight, weight loss attempts, weight-based teasing and unhappiness[39]. Similarly, assessments of the impact of the Pacific OPIC Project community interventions will include assessments of perceived body image (at least in a sub-sample of participants)[42]. While individual measures of intervention

harm are important, the potential adverse impacts at the setting or community level should also be considered by evaluators. For community safety, the use of committees (which include representatives from the community) to monitor adverse intervention effects and enforce 'stopping rules' in instances where adverse events exceed an acceptable level has been suggested as important[66].

Impact
Healthy eating, physical activity and sedentary behaviour. As weight status changes are mediated by dietary improvements to reduce excessive energy intake and/or increased physical activity to increase energy expenditure, assessments of these behaviours can enhance the internal validity of programme evaluations and provide evidence of the mechanism behind any intervention effect. For example, greater confidence of an intervention effect would be characterised by an inverse association between weight status and healthy eating and physical activity.

In their assessment of healthy eating and physical activity, most obesity prevention programme evaluations rely on brief self-reported measures which are vulnerable to social desirable responding[4]. Only a few trials included in the present review indicated that such measures had been previously demonstrated to be valid or reliable[33,34,39]. While short item questionnaires may represent the most feasible method to assess changes in broad dietary or activity patterns, such items represent crude measures and often are not sufficiently sensitive to detect small but meaningful changes at a population level. Indeed, in community-based programmes, these self-reported measures have been found to contradict objective measures. In the APPLE Project, for example, self-reported physical activity among children in intervention communities decreased significantly relative to control community children at the 1-year follow-up, in contrast to accelerometry assessments which found significantly increased counts among intervention community children over the same period[33].

The use of valid tools is key to robust programme evaluations. A number of papers have been published providing guidance to programme evaluators regarding the selection of measures to assess physical activity and nutrition in the context of purpose, capacity, skill and resource constraints[69–71]. In principle, obesity intervention programmes should seek to employ the most rigorous behavioural assessments which such constraints allow. The use of objective measures of physical activity such as accelerometry or fitness tests employed by previous interventions[33,35,40] or more rigorous dietary assessments such as 24 h food recall methods may be feasible for interventions with a small sample. For larger trials, the inclusion of such assessments on a sub-sample of the population could be considered in addition to the broader use of validated self-reported physical activity or food frequency questionnaires as a means of measurement triangulation.

Community environment. Community-based interventions often seek to reduce population adiposity through modifying community environments, such as the availability or accessibility of healthy foods or physical activity opportunities in key community settings (i.e. schools). For these interventions, community environmental characteristics which are the target of intervention represent important mediating variables and programme impacts in their own right[30]. While a number of trials in the present review have assessed community environments as part of programme evaluations, most have not reported the validity of the methods used to do so[35–37,39,41,42]. The psychometric properties of a number of tools to assess obesogenic characteristics of the home (e.g. family eating routines), child care (staff prompting child activity) and neighbourhood environments (availability of physical activity facilities)[72–74] have recently been published and their use would greatly strengthen the quality of future evaluations[75]. To our knowledge, however, there remains a lack of validated instruments for assessing other key community settings such as schools. The development of valid tools suitable for population-wide assessments of community settings and organisations would be particularly beneficial for programme evaluators and should be considered a research priority for the field.

Given the cost and complexity of evaluating individual health behaviour risks or chronic disease conditions noted in previous community-based health promotion programmes, the use of community-level environmental indicators has been proposed as a more feasible and cost-effective alternative when significant resource constraints exist[36,76]. The use of community environment measures in this way would require the availability and selection of valid measures of environmental characteristics known to be associated with population overweight and obesity, healthy eating or physical activity. Some simple and objective environmental proxy measures already exist. For example, grocery store shelf space has previously been found to reflect changes in some community dietary indicators with similar power to individual surveys at one-tenth of the cost[77].

Process
Measures of the extent to which the intervention has been delivered, its reach and variability, often termed process evaluation, are recommended by the UK Medical Research Council as an important means of assessing intervention fidelity and exposure in complex interventions[18]. Process evaluation can also provide insight into why and how an intervention worked, or failed, and how it may be improved[18,78]. Process evaluation should be conducted to the same methodological standard and reported as thoroughly as assessments of intervention outcomes[18]. Process evaluations reported as part of planned or previous community-based obesity prevention evaluations have varied in the extent to which they

have assessed intervention delivery, reach or dose[34,36–40]. Among the more comprehensive process evaluations have been the Romp & Chomp and Be Active Eat Well initiatives that were guided by explicit programme logic models and assessed exposure of the community to the intervention through changes in community capacity and organisational policies, practices or environments[79,80]. The lack of quality process information, particularly that related to intervention reach and dose, has previously been criticised as an impediment to the interpretation of community-based obesity prevention interventions[81].

Context

The effectiveness of community-based interventions is undoubtedly influenced by the social, political and organisational contextual factors occurring at the time of implementation. A common criticism of process evaluation is that it presupposes and mechanistically assesses specified intervention components[66]. While this has merit, it ignores other contextual factors which could operate as effect modifiers. As such, the UK Medical Research Council recommends that, in addition to measurement of intervention process, the evaluation of complex interventions assesses, monitors and documents important changes in community context over the life of the project[18]. The California Endowment's Healthy Eating, Active Communities Program[35] explicitly states an intention to collect and synthesise context data via interviews, surveys and focus groups with community members, programme stakeholders and policy makers. While not stated as a context evaluation, other initiatives included in the present review collected and reported context information, such as other interventions occurring in study communities prior to or during intervention implementation, or changes in media activity and resources contributed by local agencies[37]. As context evaluation is a developing science, little explicit guidance is available regarding what and how such information should be collected and utilised.

Analysis

Selection of the analytical procedure to assess the impact of a community-based approach to obesity prevention should, of course, be guided by the evaluation design[18]. As communities are typically the unit of analysis when assessing the effectiveness of community-based interventions, the use of generalised linear mixed models (also referred to as multilevel models, random-effects models, hierarchical linear models or covariance component models) is often most appropriate[23]. Such analytical techniques account for intra-class correlation and allow for individual- and community-level influences to be examined simultaneously. Encouragingly, trials included in the present review employed such techniques and avoided unit of analysis errors previously documented in the obesity literature[4].

Methodological reviews of previous community-based chronic disease prevention interventions have attributed modest intervention impacts, in part, to high rates of population migration into and out of intervention and comparison communities, or to inconsistent intervention exposure, diluting the intervention effect[13,32]. Such factors also represent a risk to evaluations of community-based obesity prevention programmes. The Shape Up Somerville child obesity prevention programme, for example, reported that 27% of children had moved out of intervention or comparison communities within the first 12 months of the initiative[38]. Variable intervention implementation has also been reported[37]. Two supplementary strategies may be useful in managing issues of intervention exposure when describing intervention effects in a community. First is a plausibility analyses, whereby comparisons are made between those receiving the intervention and those who did not after adjusting for confounders[82]. The second is a dose–response analyses where the effect of the intervention is examined according to the level of exposure to the intervention[12,81]. Despite the usefulness of such analyses, they have not been utilised in any of the trials included in the present review. Considering issues of intervention exposure, however, children residing within intervention communities but attending school outside the region will be excluded in the analysis of intervention effects of the Fiji OPIC initiative[42].

Discussion

The findings of the present review illustrate that there is opportunity for greater cross-disciplinary learning from the evaluation experiences of past community-based chronic disease risk factor interventions[12,13,25] and scope to improve the rigour of community-based obesity prevention interventions. Important methodological limitations apparent in a number of trials included in the present review, particularly in trial design and measurement, represent considerable impediments to inferences of causal attribution. Efforts to improve the rigour of future community-based interventions are therefore warranted.

Internationally, government and non-government organisations invest considerable sums in community-based initiatives to prevent excessive weight gain, often in the absence of – or with insufficient funding for – rigorous programme evaluation[81]. Quality evidence regarding the effects of community-based interventions is required to assess the community benefit of such expenditure, to help maximise future investment in obesity prevention and to facilitate more timely improvements in the health of populations. Under-resourced evaluations of community-based interventions offer little quality evidence to inform public health policy and practice. Selective investment in critical-mass funding of large, rigorously evaluated community-based interventions may represent a more efficient strategy to yield

robust evidence of intervention effects[83]. Funding comprehensive evaluations of targeted community-based obesity prevention interventions should therefore represent a priority for governments and other health-promoting funding agencies.

While the review provides useful guidance for the conduct of community-based obesity prevention intervention evaluations, there are a number of opportunities for further methods development to advance the field, particularly for initiatives utilising community-based participatory approaches. Community-based participatory research emphasises reciprocal transfer of expertise between community and researchers, sharing of decision making power and mutual ownership of the research process and products, and is increasingly being used to address a variety of public health issues in the community[49,84]. While such approaches can facilitate conceptualisation of the problem and culturally appropriate intervention development and delivery, and improve data collection, analysis and interpretation[84], participatory approaches introduce a number of unique challenges for evaluators. For example, evolving study procedures and processes may preclude the development of study protocols a priori and increasing the specificity of research to a community can reduce the generalisability of the research outcomes[84]. While broad guidelines exist to assist with the conduct of quality participatory research[85], further development of measures to assess participatory processes and constructs and more sophisticated analytical techniques to deal with such complexity are required[86].

Greater examination of factors which may mediate or moderate an intervention effect may also present an opportunity to further an understanding of the casual pathways in which interventions operate and to identify particular groups in the community for which an intervention may be beneficial[87,88]. Despite the benefits of such research, few trials have examined such relationships in obesity prevention research generally[89]. Encouragingly, within community-based obesity prevention research Johnson and colleagues recently published a multilevel analysis of the Be Active Eat Well initiative, in which a moderating effect of the intervention was found for the relationship between the frequency of watching television during meals and BMI[90]. A greater understanding of such relationships will aid future efforts to design and evaluate community-based obesity prevention initiatives.

For complex community-based interventions, the conduct of rigorous evaluations undoubtedly represents a considerable challenge. Evaluators need to carefully examine and the strengths and pitfalls of decisions regarding evaluation design, data collection, measurement and analysis, and seek to maximise evaluation rigour in the context of political, resource and practical constraints. The present paper attempts to provide guidance for evaluators to do so.

References

1. Stewart S, Tikellis G, Carrington C et al. (2008) *Australia's Future 'Fat Bomb'. A Report on the Long-Term Consequences of Australia's Expanding Waistline on Cardiovascular Disease.* Melbourne: Baker IDI Heart and Diabetes Institute.
2. Lopez AD, Mathers CD, Ezzati M et al. (2006) *Global Burden of Disease and Risk Factors.* Washington, DC: The World Bank and Oxford University Press.
3. World Health Organization (2004) *Comparative Quantification of Health Risks: Global and Regional Burden of Disease Attributable to Selected Major Risk Factors.* Geneva: WHO.
4. Livingstone MBE, McCaffrey TA & Rennie KL (2006) Child obesity prevention studies: lessons learned and to be learned. *Public Health Nutr* **9**, 1121–1129.
5. World Health Organization (2004) *Global Strategy on Diet, Physical Activity and Health.* Geneva: WHO; available at http://www.who.int/dietphysicalactivity/strategy/eb11344/strategy_english_web.pdf
6. Public Health Association of Australia (2008) Inquiry into Obesity in Australia. Submission from the Public Health Association of Australia to the House of Representatives Standing Committee on Health and Ageing. http://phaa.net.au/documents/ObesityintoAustraliaSubmission.pdf (accessed May 2011).
7. McCormick B & Stone I (2007) Economic costs of obesity and the case for government intervention. *Obes Rev* **8**, Suppl. 1, 161–164.
8. Gill TP, Baur LA, Bauman AE et al. (2009) Childhood obesity in Australia remains a widespread health concern that warrants population wide prevention programs. *Med J Aust* **190**, 146–148.
9. Lavery NS (2008) Stop all further research – and act. *BMJ* **336**, 7.
10. Crombie IK, Irvine L, Elliott L et al. (2009) Targets to tackle the obesity epidemic: a review of twelve developed countries. *Public Health Nutr* **12**, 406–413.
11. Fussenegger D, Pietrobelli A & Widhalm K (2008) Childhood obesity: political developments in Europe and related perspectives for future action on prevention. *Obes Rev* **9**, 76–82.
12. Nissinen A, Berrios X & Puska P (2001) Community-based non-communicable disease interventions: lessons from developed countries for developing ones. *Bull World Health Organ* **79**, 963–970.
13. Merzel C & D'Afflitti JD (2003) Reconsidering community-based health promotion: promise, performance and potential. *Am J Public Health* **93**, 557–574.

14. World Health Organization (2010) *Population-based Prevention Strategies for Childhood Obesity: Report of a WHO Forum and Technical Meeting, Geneva, 15–17 December 2009.* Geneva: WHO; available at http://www.who.int/dietphysicalactivity/childhood/child-obesity-eng.pdf

15. Ovretveit J (2002) Action evaluation: what is it and why evaluate? *Action Evaluation of Health Programmes and Changes: A Handbook for A User-focused Approach* pp. 1–7. Abingdon: Radcliffe Medical Press.

16. de Zoysa I, Habicht JP, Pelto G *et al.* (1998) Research steps in the development and evaluation of public health interventions. *Bull World Health Organ* **76**, 127–133.

17. King L, Gill T, Allender S *et al.* (2011) Best practice principles for community-based obesity prevention: development, content and application. *Obes Rev* **12**, 329–338.

18. Medical Research Council (2008) Developing and evaluating complex interventions: new guidance. http://www.mrc.ac.uk/complexinterventionsguidance (accessed July 2011).

19. Mercer SL, DeVinney BJ, Fine LJ *et al.* (2007) Study designs for effectiveness and translation research. *Am J Prev Med* **33**, 139–154.

20. Shadish WR, Cook TD & Campbell DT (2002) *Experimental and Quasi-Experimental Designs for Generalized Causal Inference.* New York: Houghton Mifflin Company.

21. Scott D & Weston R (editors) (1998) *Evaluating Health Promotion.* Cheltenham: Stanley Thornes.

22. Nutbeam D & Bauman AE (2006) *Evaluation in a Nutshell: A Practical Guide to the Evaluation of Health Promotion Programs.* New York: McGraw-Hill.

23. Atienza AA & King AC (2002) Community-based health intervention trials: an overview of methodological issues. *Epidemiol Rev* **24**, 72–79.

24. Hancock L, Sanson-Fisher R, Redman S *et al.* (1996) Community action for cancer prevention: overview of the Cancer Action in Rural Towns (CART) project, Australia. *Health Promot Int* **11**, 277–290.

25. Pirie PL, Stone EJ, Assaf AR *et al.* (1994) Program evaluation strategies for community-based health promotion programs: perspectives from the cardiovascular disease community research and demonstration studies. *Health Educ Res* **9**, 23–36.

26. McLeory KR, Norton BL, Kegler MC *et al.* (2003) Community-based interventions. *Am J Public Health* **93**, 529–533.

27. Summerbell CD, Waters E, Edmunds L *et al.* (2005) Interventions for preventing obesity in children. *Cochrane Database Syst Rev* issue 3, CD001871.

28. Friedrich MJ (2007) Researchers address childhood obesity through community-based programs. *JAMA* **298**, 2728–2730.

29. Swinburn BA & de Silva-Sanigorski AM (2010) Where to from here for preventing childhood obesity: an international perspective. *Obesity (Silver Spring)* **18**, Suppl. 1, S4–S7.

30. Huang TT & Story M (2010) A journey just started: renewing efforts to address childhood obesity. *Obesity (Silver Spring)* **18**, Suppl. 1, S1–S3.

31. American Evaluation Association (2005) American Evaluation Association: guiding principles for evaluators. *Am J Eval* **26**, 297–298.

32. Gilligan C, Sanson-Fisher R & Shakeshaft A (2010) Appropriate research designs for evaluating community-level alcohol interventions: what next? *Alcohol Alcohol* **45**, 481–487.

33. Taylor RW, Mcauley KA, Williams SM *et al.* (2006) Reducing weight gain in children through enhancing physical activity and nutrition: the APPLE project. *Int J Pediatr Obes* **1**, 146–152.

34. Taylor RW, McAuley KA, Barbezat W *et al.* (2007) APPLE Project: 2-y findings of a community-based obesity prevention program in primary school-age children. *Am J Clin Nutr* **86**, 735–742.

35. Samuels S, Craypo L, Boyle M *et al.* (2010) The California Endowment's Healthy Eating, Active Communities Program: a midpoint review. *Am J Public Health* **100**, 2114–2123.

36. Cheadle A, Schwartz P, Rauzon S *et al.* (2010) The Kaiser Permanente Community Health Initiative: overview and evaluation design. *Am J Public Health* **100**, 2111–2113.

37. de Silva-Sanigorski AM, Bell CA, Kremmer P *et al.* (2010) Reducing obesity in early childhood: results from Romp & Chomp, an Australian community-wide intervention program. *Am J Clin Nutr* **91**, 831–840.

38. Economos CD, Hyatt RR, Goldberg JP *et al.* (2007) A community intervention reduces BMI z-score in children: Shape Up Somerville first year results. *Obesity (Silver Spring)* **15**, 1325–1336.

39. Sanigorski AM, Bell AC, Kremer PJ *et al.* (2008) Reducing unhealthy weight gain in children through community capacity-building: results of a quasi-experimental intervention program, Be Active Eat Well. *Int J Obes (Lond)* **32**, 1060–1067.

40. Chomitz VR, McGowan RJ, Wendel JM *et al.* (2010) Healthy Living Cambridge Kids: a community based participatory effort to promote healthy weight and fitness. *Obesity (Silver Spring)* **18**, Suppl. 1, S45–S53.

41. de Silva-Sanigorski A, Bolton K, Haby M *et al.* (2010) Scaling up community-based obesity prevention in Australia: background and evaluation design of the Health Promoting Communities: Being Active Eating Well initiative. *BMC Public Health* **10**, 65.

42. Swinburn B, Pryor J, McCabe M *et al.* (2007) The Pacific OPIC Project (Obesity Prevention In Communities) – objectives and designs. *Pac Health Dialog* **14**, 139–153.

43. Romon M, Lommez A, Tafflet M *et al.* (2009) Downward trends in the prevalence of childhood overweight in the setting of 12-year school- and community-based programmes. *Public Health Nutr* **12**, 1735–1742.

44. Azzam T (2010) Evaluator responsiveness to stakeholders. *Am J Eval* **31**, 45–65.

45. Yazahmeidi B & Holman CDJ (2007) A survey of suppression of public health information by Australian governments. *Aust N Z J Public Health* **31**, 551–557.

46. Morris M (2010) The good, the bad, and the evaluator: 25 years of AJE ethics. *Am J Eval* **32**, 134–151.

47. Cope MB & Allison DB (2010) White hat bias: examples of its presence in obesity research and a call for renewed commitment to faithfulness in research reporting. *Int J Obes (Lond)* **34**, 84–88.

48. DeAngelis CD, Drazen JM, Frizelle FA *et al.* (2004) Clinical trial registration: a statement from the International Committee of Medical Journal Editors. *JAMA* **292**, 1363–1364.

49. Viswanathan M, Ammerman A, Eng A *et al.* (2004) *Community-Based Participatory Research: Assessing the Evidence. Evidence Reports/Technology Assessments* no. 99. Rockville, MD: Agency for Healthcare Research and Quality.

50. Plint AC, Moher D, Morrison A *et al.* (2006) Does the CONSORT checklist improve the quality of reports of randomised controlled trials? A systematic review. *Med J Aust* **185**, 263–267.

51. Wolfenden L, Wiggers J, Tursan d'Espaignet E *et al.* (2010) How useful are systematic reviews of child obesity interventions? *Obes Rev* **11**, 159–165.

52. COMMIT Research Group (1995) Community Intervention Trial for Smoking Cessation (COMMIT), I: cohort results from a four year community intervention. *Am J Public Health* **85**, 183–192.

53. Sanson-Fisher RW, Bonevski B, Green LW *et al.* (2007) Limitations of the randomized controlled trial in evaluating population-based health interventions. *Am J Prev Med* **33**, 155–161.

54. Mdege ND, Man MS, Taylor CA *et al.* (2011) Systematic review of stepped wedge cluster randomized trials shows that design is particularly used to evaluate interventions during routine implementation. *J Clin Epidemiol* **64**, 936–948.

55. Hussey MA & Hughes JP (2007) Design and analysis of stepped wedge cluster randomized trials. *Contemp Clin Trials* **28**, 182–191.

56. Simmons A, Reynolds RC & Swinburn B (2011) Defining community capacity building: is it possible? *Prev Med* **52**, 193–199.

57. Mellor JM, Rapoport RB & Maliniak D (2008) The impact of child obesity on active parental consent in school-based survey research on healthy eating and physical activity. *Eval Rev* **32**, 298–312.

58. Treweek S, Pitkethly M, Cook J *et al.* (2010) Strategies to improve recruitment to randomised controlled trials. *Cochrane Database Syst Rev* issue 4, MR000013.

59. Wolfenden L, Kypri K, Hodder R *et al.* (2009) Obtaining active parental consent for school based research: a guide for researchers. *Aust N Z J Public Health* **33**, 270–275.

60. Booth ML, Okely AD & Denney-Wilson E (2011) Validation and application of a method of measuring non-response bias in school-based surveys of pediatric overweight and obesity. *Int J Pediatr Obes* **6**, e87–e93.

61. Barry AE (2005) How attrition impacts the internal and external validity of longitudinal research. *J Sch Health* **75**, 267–270.

62. Davis LL, Broome ME & Cox R (2002) Maximizing retention in community-based clinical trials. *J Nurs Scholarsh* **34**, 47–53.

63. Booker C, Harding S & Benzeval M (2011) A systematic review of the effect of retention methods on response rates in population-based cohort studies. *BMC Public Health* **11**, 249.

64. Kane R, Wellings K, Free C *et al.* (2000) Uses of routine data sets in the evaluation of health promotion interventions: opportunities and limitations. *Health Educ* **100**, 33–41.

65. Breen C, Shakeshaft A, Slade T *et al.* (2011) Assessing the reliability of measures using routinely collected data. *Alcohol Alcohol* **46**, 501–502.

66. Hawe P, Shiell A, Riley T *et al.* (2004) Methods for exploring implementation variation and local context within a cluster randomised community intervention. *J Epidemiol Community Health* **58**, 788–793.

67. Gorber CS, Tremblay M, Moher D *et al.* (2007) A comparison of direct vs. self-report measures for assessing height, weight and body mass index: a systematic review. *Obes Rev* **8**, 307–326.

68. Roberts H (2004) Intervening in communities: challenges for public health. *J Epidemiol Community Health* **58**, 729–730.

69. Magarey A, Watson J, Golley RK *et al.* (2010) Assessing dietary intake in children and adolescents: considerations and recommendations for obesity research. *Int J Pediatr Obes* (Epublication ahead of print version).

70. Vanhees L, Lefevre J, Philippaerts R *et al.* (2005) How to assess physical activity? How to assess physical fitness? *Eur J Cardiovasc Prev Rehabil* **12**, 102–114.

71. Sirard JR & Pate RR (2001) Physical activity assessment in children and adolescents. *Sports Med* **31**, 439–454.

72. Benjamin SE, Neelon B, Ball SC *et al.* (2007) Reliability and validity of a nutrition and physical activity environment self-assessment for childcare. *Int J Behav Nutr Phys Act* **4**, 29.

73. Bryant MJ, Ward DS, Hales D *et al.* (2008) Reliability and validity of the Healthy Home Survey: a tool to measure factors within homes hypothesized to relate to overweight in children. *Int J Behav Nutr Phys Act* **5**, 23.

74. Cerin E, Saelens B, Sallis J *et al.* (2006) Neighborhood environment walkability scale: validity and development of a short form. *Med Sci Sports Exerc* **38**, 1682–1691.

75. Gattshall ML, Shoup JA, Marshall JA *et al.* (2008) Validation of a survey instrument to assess home environments for physical activity and healthy eating in overweight children. *Int J Behav Nutr Phys Act* **5**, 3.

76. Cheadle A, Sterling TD, Schmid TL *et al.* (2000) Promising community-level indicators for evaluation cardiovascular health-promotion programs. *Health Educ Res* **15**, 109–116.

77. Cheadle A, Psaty B, Wagner E *et al.* (1990) Evaluating community-based nutrition programs: assessing the reliability of a survey of grocery store product displays. *Am J Public Health* **80**, 709–711.

78. Oakley A, Strange V, Bonell C *et al.* (2006) Process evaluation in randomised controlled trials of complex interventions. *BMJ* **332**, 413–416.

79. de Silva-Sanigorski A, Robertson N & Nichols M (2009) Romp & Chomp. Health Eating + Active Play for Geelong Under 5s. Process report for objective 3: To evaluate the process, impact and outcomes of the project. Implementation strategies, process evaluation, lessons learned and recommendations for future practice. http://www.goforyourlife.vic.gov.au/hav/admin.nsf/Images/Process_report_Objective_3_Final.pdf/$File/Process_report_Objective_3_Final.pdf (accessed February 2011).

80. Simmons A, Sanigorski AM, Cuttler R *et al.* (2008) Nutrition and Physical Activity in Children and Adolescents, Barwon-South Western Region. Sentinel Site Series. Report 6: Lessons learned from Colac's Be Active Eat Well project (2002–2006). http://www.goforyourlife.vic.gov.au/hav/admin.nsf/Images/ssop6_report_6_baew_final.pdf/$File/ssop6_report_6_baew_final.pdf (accessed February 2011).

81. Swinburn B, Bell C, King L *et al.* (2007) Obesity prevention programs demand high-quality evaluations. *Aust N Z J Public Health* **31**, 305–307.

82. Victoria CG, Habicht JP & Bryce J (2004) Evidence-based public health: moving beyond randomized trials. *Am J Public Health* **94**, 400–405.

83. Hennekens CH & DeMets D (2009) The need for large-scale randomized evidence without undue emphasis on small trials, meta-analyses or subgrouop analyses. *JAMA* **302**, 2361–2362.

84. Cargo M & Mercer SL (2003) The value and challenges of participatory research: strengthening its practice. *Annu Rev Public Health* **29**, 325–340.

85. Green LW, George MA, Daniel M *et al.* (1995) *Study of Participatory Research in Health Promotion: Review and Recommendations for the Development of Participatory Research in Health Promotion in Canada*. Ottawa: Royal Society of Canada.

86. Sandoval JA, Lucero J, Oetzel J *et al.* (2012) Process and outcome constructs for evaluating community-based participatory research projects: a matrix of existing measures. *Health Educ Res* **27**, 680–690.

87. MacKinnon DP & Luecken LJ (2008) How and for whom? Mediation and moderation in health psychology. *Health Psychol* **27**, 2 Suppl., S99–S100.

88. Cerin E & Mackinnon DP (2009) A commentary on current practice in mediating variable analyses in behavioural nutrition and physical activity. *Public Health Nutr* **12**, 1182–1188.

89. Cerin E, Barnett A & Baranowski T (2009) Testing theories of dietary behavior change in youth using the mediating variable model with intervention programs. *J Nutr Educ Behav* **41**, 309–318.

90. Johnson BA, Kremer PJ, Swinburn BA *et al.* (2012) Multilevel analysis of the Be Active Eat Well intervention: environmental and behavioural influences on reductions in child obesity risk. *Int J Obes (Lond)* **26**, 901–907.

Appendix

Search strategy: MEDLINE

1. obesity.mp.
2. prevention.mp.
3. 1 and 2
4. nation$.mp.
5. state.mp.
6. count$.mp.
7. district.mp.
8. regio$.mp.
9. communit$.mp.
10. area.mp.
11. town.mp.
12. village.mp.
13. borough.mp.
14. municip$.mp.
15. province.mp.
16. shire.mp.
17. 4 or 5 or 6 or 7 or 8 or 9 or 10 or 11 or 12 or 13 or 14 or 15 or 16
18. randomized controlled trial.pt.
19. controlled clinical trial.pt.
20. randomized.ab.
21. randomised.ab.
22. clinical trials as topic.sh.
23. randomly.ab.
24. trial.ti.
25. double blind.ab.
26. single blind.ab.
27. (pretest or pre test).mp.
28. (posttest or post test).mp.
29. (pre post or prepost).mp.
30. Before after.mp.
31. (Quasi-randomised or quasi-randomized or quasi-randomized or quazi-randomised).mp.
32. stepped wedge.mp.
33. Preference trial.mp.
34. Comprehensive cohort.mp.
35. Natural experiment.mp.
36. (Quasi experiment or quazi experiments).mp.
37. (Randomised encouragement trial or randomized encouragement trial).mp.
38. (Staggered enrolment trial or staggered enrollment trial).mp.
39. (Nonrandomised ornon randomised or nonrandomized or non randomized).mp.
40. Interrupted time series.mp.
41. (Time series and trial).mp.
42. Multiple baseline.mp.
43. Regression discontinuity.mp.
44. 18 or 19 or 20 or 21 or 22 or 23 or 24 or 25 or 26 or 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
45. 3 and 17 and 44
46. 45
47. limit 45 to year = '1991-201'