Effectiveness of interventions to promote healthy weight in general populations of children and adults: a meta-analysis

Helene Luckner1,2, John R. Moss1, Christian A. Gericke3

1 School of Population Health and Clinical Practice, University of Adelaide, Australia
2 Department of Health Care Management, Berlin University of Technology, Germany
3 Peninsula CLAHRC, National Institute for Health Research, Peninsula Medical School, Universities of Exeter & Plymouth, Plymouth, UK

Correspondence: Christian A. Gericke, Deputy Director, Peninsula CLAHRC, National Institute for Health Research, Peninsula Medical School, Universities of Exeter & Plymouth, Plymouth, UK

Received 26 November 2010, accepted 9 September 2011

Background: Responding to the obesity epidemic requires robust evidence to help prioritize the allocation of scarce resources to preventive interventions. The aim of this study was to evaluate interventions that promote healthy weight [defined as reduction in body mass index (BMI) or percentage body fat] in general populations (unselected by weight) using a comprehensive meta-analysis. Interventions with both single and multiple components were considered. Methods: Studies were first identified through well-conducted systematic reviews complemented by a search for single studies in five large medical databases up to 6 November 2008. Sixty-eight controlled studies were included. For each intervention type and age group, all relevant studies were pooled in a random effects meta-analysis. Results: In children, the highest reductions in mean BMI were achieved through promoting reduced television viewing [−0.27 kg/m² (95% CI −0.4 to −0.13 kg/m²)]. Programmes combining physical activity, specifically themed or general health education and nutrition achieved a lower reduction [−0.1 kg/m² (95% CI −0.17 to −0.04 kg/m²)]. Other interventions had high heterogeneity or showed no statistically significant reduction in outcomes. In adults, single component interventions were found to reduce both outcome measures. Their mean percentage body fat was reduced through education by −1.22% (95% CI −1.92 to −0.52). Conclusion: The evidence for the effectiveness of promoting healthy weight in general populations is limited, though multi-component interventions in schools and encouraging reduced children’s television viewing are promising strategies. Improving the reporting of outcomes is vital, as imputation of inadequately reported measures may have contributed to the observed heterogeneity. Longer follow-up is essential for understanding policy relevance.

Introduction

Overweight and obesity represent a major public health burden by increasing the risk of type 2 diabetes, cardiovascular disease, osteoarthritis, various cancers and other severe chronic conditions. Given the rising prevalence of overweight and obesity in most western countries, policy-makers are seeking to allocate the limited resources available for implementing public health policy into effective preventative strategies. These strategies are being directed towards the general population and robust evidence of successful prevention is therefore vital.

The best evidence available can generally be found in a well-conducted systematic review of all relevant randomized controlled trials; however, evidence synthesis restricted to such studies may be too narrow, because this study design is not always feasible in public health. Body composition measured as the body mass index (BMI) or percentage of body fat (%BF) is well recognized as a valid outcome for population studies on prevention of overweight and obesity, but to date only a few meta-analyses have addressed studies with these outcome measures, and this was mostly done through pooling standardized effects. Two of these meta-analyses were limited to physical activity interventions in the school setting, while three pooled all studies regardless of intervention type. Only two distinguished between interventions with multiple components; however, one of these was restricted to the school setting and the other focused on minority children in the USA. Useful though all these studies have been, more still needs to be done to address the information requirements for evidence-based policy. This includes in particular an understanding of the performance of single interventions compared with multi-component strategies. The objective of this study was to evaluate interventions that promote healthy weight (defined as a reduction in BMI or %BF) in general populations (unselected by weight) from Western countries in a comprehensive meta-analysis.

Methods

Literature search

Articles were identified through well-conducted systematic reviews on overlapping body weight reduction-related topics. These were obtained by searching the Cochrane Database of Systematic Reviews and the Database of Abstracts and Reviews of Effects (DARE). The MeSH terms ‘obesity’, ‘body composition’, ‘nutrition’, ‘physical activity’ and ‘health promotion’ were used in combination with ‘prevention’. An additional search in PubMed was performed with the search strategy ‘Overweight (MeSH) and prevention and systematic and review’.

Once the relevant systematic reviews were identified, all full-text articles on interventions were retrieved from them and assessed for eligibility for meta-analysis. A second search was conducted in the databases PubMed, Embase, Scopus, CINAHL and the Cochrane Central Register of Controlled Trials to cross-check whether there were single studies that were unidentified by the systematic reviews (see Appendix 1 for search strategies in Supplementary Data). Eligible studies so identified were also included in the meta-analysis. All searches were run up to 6 November 2008.
Study selection

Records identified in the search for systematic reviews were screened. Those that clearly focused on the promotion of healthy weight in general populations (unselected by weight) were included, but not those with a sole focus on treatment. Primary studies were included if they reported on controlled, randomized or quasi-experimental projects to promote healthy weight in general populations together with sufficient outcome statistics measured in BMI and/or %BF. To be considered a general population, the mean BMI at baseline had to be within either the normal or overweight ranges, but not the obese (BMI<30 kg/m²; adjusted for age and sex in children). Studies were excluded if the objective included treatment of eating disorders or differed substantially from healthy weight promotion in a general population, or if they were conducted in a non-Western country.

Data extraction

Detailed information on the characteristics of all studies reviewed was extracted and summarized by one investigator. The results from the most recent date of data collection were used. An attempt was made to contact the authors if follow-up values or measures of variation were missing if outcomes were only presented graphically or if only a BMI z-score was reported.

Assessment of quality

The criteria for assessing study quality in Harris et al.⁶ which were based on two established assessment forms,¹⁴,¹⁵ were considered to be the most appropriate for studies to promote healthy weight. Studies were assessed for a clear description of inclusion/exclusion criteria and intervention content, an a priori power calculation, blinding of outcome assessment, a method for randomization (where applicable), reporting of the attrition rate, a reproducible description of statistical methods and whether baseline characteristics were similar between intervention and control groups. The criterion concerning the 'theoretical basis for the intervention' was not assessed as it was not relevant for this review.

Data synthesis and analysis

Studies were clustered into groups according to the outcome measured (BMI or %BF), the type of intervention they referred to and the target age. Meta-analysis was performed on those groups that contained more than one study. Four different types of intervention were identified. Regular exercise was classified as 'physical activity' (PA). Dissemination of information or teaching on either general healthy behaviour or specifically related to nutrition, physical activity or sedentary behaviour was classified as 'education' (EDU). If an intervention consisted of a change in at least one major daily meal, it was classified as 'nutrition' (NUTR). Combinations of components were also possible (e.g. EDU + PA). In addition, a second classification was formed for interventions that aimed at reducing TV viewing (TV) regardless of other components involved. One single study could contribute to more than one intervention group if it consisted of several study arms with different components. Age was categorized into two major groups: children and adolescents (0–18 years) or adults (>18 years).

In the meta-analysis, the generic inverse variance approach for continuous outcomes in RevMan 5¹⁶ was used with a random effects model as the base case. An $I^2 > 50\%$ was understood to indicate substantial heterogeneity. Funnel plots were used to assess the possible risk of publication bias. No adjustments were made for multiple statistical comparisons, it being understood that the reader would make due allowance for this.

The outcome of interest for each intervention group was the mean difference (MD) and its 95% confidence interval (95% CI). This required each study to report the MD in change scores compared with baseline between the intervention ($\Delta_i$) and control group ($\Delta_c$), namely ($\Delta_i - \Delta_c$). Where studies only reported point estimates of $\Delta_i$ and $\Delta_c$, the continuous outcome comparison in RevMan 5 was used to derive the CI of ($\Delta_i - \Delta_c$). Moreover, if only mean outcome values at baseline and follow-up were reported and not the MD, point estimates of $\Delta_i$ and $\Delta_c$ were calculated and the respective standard deviations were derived using single imputation with a correlation coefficient of 0.97. This value was derived from those studies within this systematic review having sufficient data.¹⁷–²²

Subgroup results within one study were combined into one estimate for males and females together or for different study arms of the same intervention (see Appendix 2 for the formula used in Supplementary Data).

To assess the extent of uncertainty, one-way sensitivity analyses were performed. First, a fixed-effects model was applied to those intervention groups where the base case $P < 50\%$. Second, an adjustment for cluster design was performed where this had not already been accounted for in the original study. The intra-class correlation coefficient for BMI and %BF needed for adjustment was taken from one study in this systematic review, which provided both measures.²³ Third, the correlation coefficient needed to impute standard deviations was replaced by 0.9 and 0.999.

Interaction in subgroups was tested with a two-tailed z-test based on all studies for the same outcome and age range, regardless of intervention. Studies were grouped according to whether the following applied: randomization; school-based setting; family involvement; active intervention for control group; participants described as being at risk or disadvantaged; final measurement >4 weeks after the intervention ended; data combined for study groups or outcome statistics imputed. In addition, the following were also taken into account when grouping studies: the length of the follow-up (<3 month, 3–6 months, >6 months to 1 year, >1–2 years and >2 years) and differences within age groups (younger children vs. children vs. adolescents). Whether the year of publication had an impact was also tested (after 1999; after 2004). Furthermore, the impact of removing each single study in turn from the intervention group was assessed.

Results

Literature search

Figure 1 depicts the process of the literature search according to the PRISMA statement.²⁴ Twenty-three potentially relevant systematic reviews were identified (see Appendix 3 in Supplementary Data). Sixty-nine out of 227 full-text articles met the inclusion criteria. These reported on 51 separate studies. Seventeen additional studies were identified in the second search. Thus, 68 studies were included in the meta-analysis. Additional information was obtained from the authors of four studies.²⁵–²⁸

Interventions

Details of the 68 studies included are provided in Appendices 4 and 5. They reported on 84 programme across 7 intervention types: 16 as PA, 25 as EDU, 3 as NUTR, 22 as PA + EDU, 3 as EDU + NUTR, 6 as PA + EDU + NUTR and 9 as TV. Forty-seven studies used a cluster as the unit of analysis and 54 were randomized. The length of the intervention varied from 1 month to >7 years. Forty-seven studies measured only BMI as an outcome, while 7 measured only %BF and 14 measured both outcomes. Assessment of %BF was performed by skin-fold thickness in seven studies; another seven studies used bioelectrical impedance. Dual-energy X-ray absorptiometry was used in five studies and two studies used underwater weighing. Three studies were about younger children (<6 years), 34 were about children (6–12 years), 18 about adolescents (12–18 years) and 13 involved adults (>18 years).
Most studies scored well for providing a clear description of their intervention, the attrition rate and the statistical analysis. Conversely, only a minority of studies reported information on an \textit{a priori} power calculation, a blinded outcome assessment or the method of randomization.

**Meta-analysis by type of intervention**

For the children, meta-analysis was undertaken for the intervention groups PA, EDU, PA + EDU and PA + EDU + NUTR regarding both BMI and %BF, but for EDU + NUTR and TV regarding BMI only. Three intervention groups contained only one study: NUTR measured in BMI as well as EDU + NUTR and TV measured in %BF. In adults, meta-analysis was feasible for PA and EDU in both outcome measures, but for NUTR only for BMI. An imputation of the CI around the outcome estimates was needed in 22 studies.

MD for each intervention group in children is summarized in table 1. In each situation where meta-analysis could be performed, there was a reduction (negative MD) in BMI. Except for PA and EDU + NUTR, these MD were statistically significant. The reduction was highest in TV followed by PA + EDU and EDU. The children’s intervention group for PA also showed a statistically significant reduction in %BF. In EDU and PA for adults, the reduction was statistically significant for both outcomes (table 2).
Furthermore, lowered the level of heterogeneity in all intervention groups. 

### Table 1: Meta-analyses of interventions to promote healthy weight in children (0–18 years) from general populations measured as MD in change from baseline in either BMI or %BF

<table>
<thead>
<tr>
<th>Intervention group by component(s) and outcome measure (change from baseline)</th>
<th>Sample size</th>
<th>Number of studies included (number of controlled but non-randomized studies)</th>
<th>Mean difference (MD) (95% CI)</th>
<th>$I^2$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PA (BMI)$^a$</td>
<td>2927</td>
<td>10 (1)</td>
<td>$-0.15$ (–0.33 to 0.03)</td>
<td>87</td>
</tr>
<tr>
<td>EDU (BMI)$^b$</td>
<td>5667</td>
<td>15 (6)</td>
<td>$-0.15$ (–0.24 to –0.07)</td>
<td>65</td>
</tr>
<tr>
<td>NUTR (BMI)$^c$</td>
<td>103</td>
<td>1 (0)</td>
<td>$-0.14$ (–0.55 to 0.27)</td>
<td>–</td>
</tr>
<tr>
<td>PA + EDU (BMI)$^d$</td>
<td>8399</td>
<td>21 (3)</td>
<td>$-0.19$ (–0.37 to –0.02)</td>
<td>94</td>
</tr>
<tr>
<td>EDU + NUTR (BMI)</td>
<td>1695</td>
<td>2 (1)</td>
<td>$-0.05$ (–0.2 to 0.1)</td>
<td>0</td>
</tr>
<tr>
<td>PA + EDU + NUTR (BMI)$^e$</td>
<td>10257</td>
<td>6 (2)</td>
<td>$-0.1$ (–0.17 to –0.04)</td>
<td>5</td>
</tr>
<tr>
<td>TV and other (BMI)</td>
<td>3962</td>
<td>8 (2)</td>
<td>$-0.27$ (–0.4 to –0.13)</td>
<td>20</td>
</tr>
<tr>
<td>PA (%BF)$^a$</td>
<td>1989</td>
<td>6 (1)</td>
<td>$-0.7$ (–1.05 to –0.31)</td>
<td>59</td>
</tr>
<tr>
<td>EDU (%BF)$^b$</td>
<td>1110</td>
<td>2 (0)</td>
<td>$+0.13$ (–0.04 to 0.3)</td>
<td>0</td>
</tr>
<tr>
<td>PA + EDU (%BF)$^d$</td>
<td>1915</td>
<td>6 (2)</td>
<td>$-1.07$ (–2.27 to 0.13)</td>
<td>97</td>
</tr>
<tr>
<td>EDU + NUTR (%BF)$^c$</td>
<td>1419</td>
<td>1 (0)</td>
<td>$+0.18$ (–1.75 to 2.11)</td>
<td>–</td>
</tr>
<tr>
<td>PA + EDU + NUTR (%BF)$^e$</td>
<td>1517</td>
<td>2 (1)</td>
<td>$-0.78$ (–0.29 to 1.86)</td>
<td>56</td>
</tr>
<tr>
<td>TV and other (%BF)$^c$</td>
<td>459</td>
<td>1 (0)</td>
<td>$+0.08$ (–0.68 to 0.84)</td>
<td>–</td>
</tr>
</tbody>
</table>

BMI in terms of kg/m². TV and other, promotion of reduced TV viewing in combination with other strategies

#### Table 2: Meta-analyses of interventions to promote healthy weight in adults (19–65 years) from general populations measured as mean difference (MD) in either BMI or %BF

<table>
<thead>
<tr>
<th>Intervention group by component and outcome measure (change from baseline)</th>
<th>Sample size</th>
<th>Number of studies included</th>
<th>Mean difference (MD) (95% CI)</th>
<th>$I^2$ (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>PA (BMI)$^b$</td>
<td>71</td>
<td>2</td>
<td>$-1.24$ kg/m² (–1.62 to –0.85)</td>
<td>0</td>
</tr>
<tr>
<td>EDU (BMI)$^c$</td>
<td>42567</td>
<td>8</td>
<td>$-0.41$ kg/m² (–0.63 to –0.19)</td>
<td>73</td>
</tr>
<tr>
<td>NUTR (BMI)</td>
<td>113</td>
<td>2</td>
<td>$-1.40$ kg/m² (–3.66 to 0.87)</td>
<td>94</td>
</tr>
<tr>
<td>PA (%BF)$^a$</td>
<td>130</td>
<td>2</td>
<td>$-1.97$ (–2.66 to –1.29)</td>
<td>0</td>
</tr>
<tr>
<td>EDU (%BF)$^b$</td>
<td>769</td>
<td>4</td>
<td>$-1.22$ (–1.92 to –0.52)</td>
<td>38</td>
</tr>
</tbody>
</table>

BMI: Body Mass Index in terms of kg/m²; Percentage of body fat; CI: Confidence interval; EDU: Education; MD: Mean difference; NUTR: Nutrition; PA: Physical activity.

a: All studies were randomized

b: One study contributes to both sets of meta-analyses (BMI and %BF)

c: Three studies contribute to both sets of meta-analyses (BMI and %BF)

d: Five studies contribute to both sets of meta-analyses (BMI and %BF)

e: Two studies contribute to both sets of meta-analyses (BMI and %BF)

The level of heterogeneity was substantial in eight of the intervention groups. The respective forest plots of the seven intervention groups with low heterogeneity can be found in Appendix 6 in Supplementary Data. Due to small numbers of studies being included or high levels of heterogeneity, the results of funnel plot analyses were not interpretable and thus not reported.

### Sensitivity analyses

Applying a fixed-effects model narrowed the CI in the group TV [MD $-0.27$ kg/m² (95% CI $-0.39$ to $-0.16$ kg/m²)]. There was a higher MD for PA + EDU + NUTR [MD: $-0.11$ kg/m² (95% CI: $-0.16$ to $-0.05$ kg/m²)], but a lower MD measured in %BF in EDU for adults [MD: $-1.17$% (95% CI: $-1.62$ to $-0.71$)].

Table 3 shows how additional cluster adjustment in 18 studies lowered the level of heterogeneity in all intervention groups. Furthermore, $I^2$ varied considerably when the correlation coefficient for imputation of standard deviations was replaced by 0.9 or 0.999. In adults, the $I^2$ for these two replacements were 70% and 75% in the group EDU measured in BMI, while these were 92% and 95% in NUTR.

### Subgroup analyses

Only statistically significant interactions are reported here. For studies on children regardless of intervention type, the MD was significantly larger in those aged younger than 6 years compared with those between 6 and 12 years ($P = 0.03$).

The test for interaction regarding studies reporting %BF was significant for studies with a longer intervention duration showed paradoxically a higher increase in children’s %BF ($P = 0.05$ for length >6 months), although why this should be so is unclear. Heterogeneity was high throughout the respective subgroups.

Across the adult interventions, there was a significantly lower MD in BMI for studies not requiring imputation ($P < 0.0001$).

### Discussion

#### Summary of findings

Although the two literature searches when combined revealed 453 (=227 + 298 – 72 duplicates) potentially relevant articles on the promotion of healthy weight in general populations, only 103 articles met the criteria for inclusion in the synthesis. One major reason for this was that many studies reported only process-related measures rather than BMI or %BF. The two separate searches proved necessary, because neither was exhaustive, with only 33 studies being found by both.

The relevant meta-analysis suggested that interventions with multiple components and those that aimed to reduce TV viewing in children led to a significant reduction in BMI. Two parallel meta-analyses suggested that interventions where adults received lifestyle education led to a statistically significant decrease in BMI or %BF, although the interpretation of this is complicated by heterogeneity. However, PA directed at adults had a considerable effect on both BMI and %BF and hence may be a promising strategy. There were also statistically significant reductions for five of the other intervention groups, but substantial heterogeneity limits the ability to interpret these results.
It cannot be ruled out that promoting a reduction in TV viewing may have led concurrently to reducing inactivity as well as lowering the energy intake derived from eating while watching TV or reducing food advertisement exposure. Moreover, only one of these eight studies aimed solely at reducing TV viewing, while the others also incorporated physical activity or other components, which may have had a separate impact on BMI. Evidence for the success of this intervention is promising, but warrants further research.

PA + EDU + NUTR measured in BMI also appear promising. The six studies included in this multi-component intervention group were all school based. Five of them lasted for at least 2 years and four also reported involvement of the family. Thus, this finding supports the current opinion that these features characterize good practice for obesity prevention. Nevertheless, the reduction in BMI appeared rather small.

The heterogeneity commonly observed reflects differences in study design within each intervention group. Despite the effort to pool similar studies, the programmes in each study varied by content, intensity and length and so did the characteristics of the study population by age in years, gender, ethnicity, socioeconomic status and initial weight at baseline. In addition, the necessity to impute outcome statistics may have had a considerable impact on the heterogeneity found. This was shown in the sensitivity analysis, where $I^2$ decreased with lower values for the correlation coefficient used for imputation. The value employed in the base case (0.97) may have been high, but was calculated from the studies included in this systematic review.

Even though more statistically significant reductions in BMI or %BF were found for different intervention types than in the previous meta-analyses, the effectiveness of the interventions studied remains relatively modest in both age groups. Since these studies mostly promoted individual behaviour change and targeted children in the school setting, it may be reasonable to question the impact of this type of intervention and to seek to compare it to what might be achievable by more comprehensive population-based approaches.

Had there been less heterogeneity in eight of the intervention groups where a meta-analysis was performed, potentially meaningful observations might have been derived from this review regarding the generally modest reductions in outcomes and a possible interaction of effects in multi-component interventions.

**Limitations of the present study**

There is a potential source of bias due to including controlled but non-randomized studies. In addition, the degree of heterogeneity made the interpretation of funnel plots difficult. Another limitation was that deriving missing CI around outcome estimates was approached by single instead of multiple imputations. The latter would have increased the standard errors to account for the uncertainty of imputation. Hence, the outcome estimates of those 22 studies that needed imputation may be overly precise in the present study. It should be noted that MD for different intervention types should be compared with caution since this is only an informative indirect comparison; had there been less heterogeneity, a more complex multiple-treatment analysis could have been conducted. Finally, there are known shortcomings in the outcomes analysed. BMI does not fully capture changes in body composition. It is widely held that %BF may be a better indicator of this. However, %BF was measured differently across the studies that reported it, which may have compromised the comparability of their outcomes.

**Recommendations for the evaluation of future interventions to promote healthy weight**

Krogsby et al. listed the need for improvements such as control for covariates of the outcome, appropriate adjustment for cluster design, subgroup analysis by gender and special assessment of the impact on overweight and obese participants. In addition, our findings demonstrate that all studies should report both BMI and %BF and present outcome estimates with a measure of variation derived from ANCOVA or equivalent approaches so imputation can be avoided. These outcome estimates should also be reported separately for each sex.

Measures of distribution should be reported because a reduction in mean BMI (or %BF) does not indicate whether this was general or predominantly in a subgroup. Improvements in the future design and reporting of overweight and obesity prevention trials are required to ensure that changes in body composition remain comparable across studies and compatible with a rigorous synthesis. In view of the modest weight reductions so far achieved, long-term follow-up is needed to gain greater insight into their maintenance and their relevance for healthy public policy.
Supplementary Data

Supplementary Data are available at Eurpub online.

Acknowledgements

The views expressed in this publication are those of the author(s) and not necessarily those of the South Australian Department of Health, the NHS, the NIHR or the Department of Health in England. An earlier draft of this work was presented as a poster at the International Conference of Obesity in Stockholm, 11–15 July 2010.

Funding

A Strategic Health Research Program Grant by the South Australian Department of Health (SHRP 9881) and support from the National Institute for Health Research (NIHR) in England for CAG’s contribution is also gratefully acknowledged.

Conflicts of interest: None declared.

Key points

- A considerable number of studies on promoting healthy weight in general populations could be identified and included in the meta-analysis by type of intervention. This number could have been even larger, had more studies reported comparable outcome measures.
- The reductions in body composition measured in the body mass index and/or percentage of body fat resulting from preventive interventions of one or multiple components are at best modest.
- Aiming to reduce TV viewing appeared to be the most promising single intervention despite low statistical power and the uncertainty resulting from multiple strategies being used across the eight studies included. Life-style education for adults also appeared promising although the interpretation of this is complicated by heterogeneity.
- The high level of heterogeneity restricted the drawing of clear conclusions in 8 out of the 15 intervention groups where meta-analyses could be performed. In many studies, imputation to overcome the absence of completely reported estimates with a measure of variation produced uncertainty and may have contributed to the heterogeneity.
- Future trials to promote healthy weight in general populations should evaluate the effect on both body mass index and percentage of body fat and should report confidence intervals around all outcome estimates.

References

14 Cochrane Effective Practice and Organisation of Care Group. Data collection checklist, Available at: http://epoc.cochrane.org/epoc-resources-review-authors (18 June 2010, date last accessed).

