Chapter 2

An evidence-update on the prospective relationship between childhood sedentary behaviour and biomedical health indicators: a systematic review and meta-analysis

Evi van Ekris, Teatske T.M. Altenburg, Amika S. Singh, Karin I. Proper, Martijn W. Heymans, Mai J.M. Chinapaw

Published in Obesity Reviews (2016) 17: 833-849
Chapter 2

Abstract

Evidence for adverse health effects of excessive sedentary behaviour in children is predominantly based on cross-sectional studies, measuring TV viewing as proxy for sedentary behaviour. This systematic review and meta-analysis summarizes the evidence on the prospective relationship between childhood sedentary behaviour and biomedical health indicators, overall and stratified by type of sedentary behaviour (TV viewing, computer use/games, screen time and objective sedentary time). PubMed, EMBASE, PsycINFO and Cochrane were systematically searched till January 2015. Methodological quality of all included studies was scored and a best evidence synthesis was applied. We included 109 studies of which 19 were of high quality. We found moderate-to-strong evidence for a relationship of overall sedentary time with some anthropometrics (overweight/obesity, weight-for-height), one cardiometabolic biomarker (HDL-cholesterol) and some fitness indicators (fitness, being unfit). For other health indicators, we found no convincing evidence due to inconsistent or non-significant findings. The evidence varied by type of sedentary behaviour. The meta-analysis indicated that each additional baseline hour of TV viewing ($\beta=0.01$, 95% confidence interval (CI)= [-0.002; 0.02]) or computer use ($\beta=0.00$, 95% CI= [-0.004; 0.01]) per day was not significantly related with BMI at follow-up. We conclude that the evidence for a prospective relationship between childhood sedentary behaviour and biomedical health is in general unconvincing.
Introduction

Children are nowadays exposed to a large diversity of sedentary activities, such as watching TV, playing video games, using the computer and sitting at school, resulting in high levels of sedentary time (1;2). Even children who are sufficiently physically active can still be highly sedentary (3-5). The potential health consequences of excessive sedentary time have received increased attention, resulting in a large number of publications in the last few years. Rezende et al. (6) recently evaluated the evidence from existing reviews and found strong evidence for an association of sedentary behaviour with obesity and moderate evidence for an association with increased blood pressure and increased cholesterol levels in children.

However, the evidence from previously published reviews linking sedentary behaviour to negative health consequences has some major limitations. First, the evidence is predominantly based on cross-sectional studies, which makes inference about causality of the association inappropriate. Prospective studies provide more insight in a potential causal relationship. Thus far, one review (Chinapaw et al. (7)) included prospective studies only and concluded moderate evidence for a negative relationship with aerobic fitness and insufficient evidence for relationships with all other biomedical health indicators. The authors emphasized the need for more high-quality prospective studies and inclusion of a broader range of health indicators besides the predominantly investigated anthropometric indicators. As a large number of prospective studies have been published in the past few years, an update of the review of Chinapaw et al. (7) is in place.

A second limitation of the evidence is that it is mainly based on studies including self- or parent-reported TV viewing time or screen time as markers for sedentary time. Besides that TV and screen time may not be representative of total sedentary time in children (8), the strength of the relationship with health indicators may depend on the type of sedentary behaviour. Especially TV viewing has been related to a low energy expenditure (9) and health effects may be mediated by a less healthy diet accompanied with TV viewing (10). Only one of the previously published reviews distinguished between different measures of sedentary behaviour in the evidence synthesis (11).

A last limitation of the evidence is that only one meta-analysis of observational studies has been performed previously (12). Other systematic reviews only applied qualitative evidence synthesis, because studies were often too heterogeneous in applied methods and statistical analysis to conduct a meta-analysis.

The present systematic review elaborates on the review of Chinapaw et al. (7), by updating the evidence on the prospective relationship between young people's sedentary behaviour and biomedical health indicators, for sedentary behaviour overall and stratified by type of sedentary behaviour (i.e.
TV viewing, computer use/games, screen time and objectively assessed total sedentary time). Moreover, when possible, we conduct meta-analyses to quantitatively synthesise the results.

Methods

The methods used in this review were replicated from the review by Chinapaw et al. (7), except where noted, to allow for an update of the evidence over the last five years. This review is registered in PROSPERO (CRD42013005020).

Literature search

We systematically searched for relevant studies in PubMed, EMBASE, PsycINFO and the Cochrane Library until January 2015. Because of the rapid development in sedentary behaviour research, the search strategy performed by Chinapaw et al. (7) was adapted for terms related to sedentary behaviour (e.g. terms for physical activity were removed and terms for sitting were added). Moreover, we limited the search to children (aged ≤18 years). The search strategy focussed on terms referring to (i) sedentary behaviour (e.g. television OR gaming OR screen time OR sitting time), (ii) prospective study design (e.g. longitudinal OR cohort OR intervention) and (iii) aged ≤18 years (e.g. preschool child OR adolescent OR youth OR infant). The three elements were linked by AND combinations and terms were used as controlled terms (e.g. MeSH in PubMed and Emtree in EMBASE) and as title-abstract words. We verified whether our search yielded all included studies by Chinapaw et al. (7). The full search strategy can be obtained upon request.

Inclusion criteria

Studies were included when the following criteria were met: (i) prospective study design (observational cohort or intervention study), (ii) sedentary behaviour (e.g. TV viewing, screen time, total sedentary time) was assessed during youth (mean age of study population ≤18 years) and examined for a prospective relation with at least one biomedical health indicator (e.g. anthropometrics, cardiometabolic biomarkers, blood pressure, fitness, asthma, bone mass and other biomedical health indicators) assessed during youth or adulthood, (iii) article was published in English, in a peer-reviewed journal.

Selection process

First, two reviewers (EE and TA) independently screened all titles and abstracts of the articles retrieved through the search for relevance based on the inclusion criteria. Discrepancies were discussed and full-text papers were obtained for articles meeting initial screening. Second, two reviewers (EE and TA)
Evidence for health effects

independently screened the full text papers to determine whether inclusion criteria were met. A third reviewer (MC) checked inconsistencies and made a final decision about inclusion.

Data extraction

Two independent reviewers (EE and TA) extracted data from the included studies by using a standardized table. Information was extracted regarding: (i) study sample at baseline; (ii) follow-up duration; (iii) type of sedentary behaviour; (iv) type and measurement (objective or self-reported) of biomedical health indicator; (v) statistical analysis and (vi) main results. Disagreement between the reviewers with regard to the extracted data was discussed until consensus was reached.

Quality assessment

The methodological quality of included studies was scored with the criteria list for prospective studies (table 1) as used by Chinapaw et al. (7). This list was originally adapted from other existing quality assessment lists (13-15). Studies were assessed on 13 different criteria divided over four dimensions: (i) study population and baseline participation (three items); (ii) study attrition (three items); (iii) data collection (three items) and (iv) data analysis (three items). Of these criteria, five criteria corresponded to informativeness (I) and eight criteria corresponded to validity/precision (V/P) of the study. Criteria have a ‘yes’ (+), ‘no’ (-) or don’t know (?) answer format. Two independent reviewers (EE and MC) performed the quality assessment. A positive score (+) was given to each criterion that was sufficiently described and met in the publication. A negative score (-) was given if a criterion was sufficiently described but lacked performance. Criteria with an unclear or incomplete description received a question mark (?). If a study referred to another publication for relevant information regarding the study design, we retrieved the additional publication to score the criterion of concern. Disagreement between the reviewers concerning scoring of each criterion was discussed until consensus was reached. A total quality score for each study was calculated by counting the number of positively scored criteria related to validity/precision and divided this by the total amount of criteria related to validity/precision (i.e. eight criteria). When a validity/precision criterion was scored with a question mark, we contacted the authors by email for additional information to ensure appropriate scoring of the study. A study was considered ‘high quality’ when the quality score was at least 0.75 (i.e. 75%). Studies scoring lower than 0.75 were considered ‘low quality’.
Chapter 2

Table 1: Criteria List for Assessment of the Methodological Quality of Prospective Studies (based on Chinapaw et al. (7))

<table>
<thead>
<tr>
<th>Criteria (rating of criteria: + = yes, - = no, ? = not or insufficiently described)</th>
<th>I</th>
<th>% of studies scoring +</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study population and participation (baseline):</td>
<td></td>
<td></td>
</tr>
<tr>
<td>The study sample represents the population of interest on key characteristics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Adequate† description of sampling frame, recruitment methods, period of recruitment, and place of recruitment (setting and geographical location)‡</td>
<td>I</td>
<td>73%</td>
</tr>
<tr>
<td>2. Participation rate at baseline at least 80%, or if the non-response was not selective (show that baseline study sample does not significantly differ from population of eligible subjects)</td>
<td>V</td>
<td>30%</td>
</tr>
<tr>
<td>3. Adequate description of baseline study sample (i.e. individuals entering the study) for key characteristics (number of participants, age, gender, sedentary behaviour, and health indicator)‡</td>
<td>I</td>
<td>73%</td>
</tr>
<tr>
<td>Study attrition:</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Loss to follow-up is not associated with key characteristics (i.e. the study data adequately represent the sample)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Provision of the exact number of participants at each follow-up measurement</td>
<td>I</td>
<td>75%</td>
</tr>
<tr>
<td>5. Provision of exact information on follow-up duration</td>
<td>I</td>
<td>95%</td>
</tr>
<tr>
<td>6. Response at short-term follow-up (up to 12 months) was at least 80% of the number of participants at baseline and response at long-term follow-up was at least 70% of the number of participants at baseline</td>
<td>V</td>
<td>49%</td>
</tr>
<tr>
<td>7. Not selective non-response during follow-up measurement(s)§</td>
<td>V/P</td>
<td>25%</td>
</tr>
<tr>
<td>Data collection</td>
<td></td>
<td></td>
</tr>
<tr>
<td>8. Adequate measurement of sedentary behaviour: done by objective measures (i.e. accelerometry, heart rate monitoring, observation) and not by self-report (self-report = -, no/insufficient information = ?)</td>
<td>V</td>
<td>12%</td>
</tr>
<tr>
<td>9. Sedentary behaviour was assessed at a time point prior to the measurement of the health indicator</td>
<td>V</td>
<td>100%</td>
</tr>
<tr>
<td>10. Adequate measurement of the health indicator: objective measurement of the health indicator done and not by self-report</td>
<td>V</td>
<td>83%</td>
</tr>
<tr>
<td>Data analyses</td>
<td></td>
<td></td>
</tr>
<tr>
<td>11. The statistical model used was appropriate¶</td>
<td>V/P</td>
<td>55%</td>
</tr>
<tr>
<td>12. The number of cases was at least 10 times the number of the independent variables</td>
<td>V/P</td>
<td>97%</td>
</tr>
<tr>
<td>13. Presentation of point estimates and measures of variability (confidence interval or standard error)</td>
<td>I</td>
<td>73%</td>
</tr>
</tbody>
</table>

*I = criterion on informativeness, V/P = criterion on validity/precision.
† Adequate = sufficient information to be able to repeat the study.
‡+§+¶+ is given only if adequate information is given on all items.
§+ is given only if non-selective dropout on key characteristics (age, gender, sedentary behaviour, health indicators) is reported in the text or tables or when the response rate was at least 95% of the number of participants at baseline.
¶+ is given only if a multivariate regression model was used adjusting for PA or for PA or diet in case the dependent variable was an indicator of body composition. When diet and PA were measured both, only a ‘+’ was given when the model adjusted for both variables.
Levels of scientific evidence

To draw conclusions regarding the evidence for a prospective relationship between sedentary behaviour and biomedical health indicators, we applied a best evidence synthesis. The evidence synthesis as performed by Chinapaw et al. (7;13-15) was repeated with the newly identified studies, taking the number of studies, the methodological quality of included studies and consistency of findings into account. We performed the evidence synthesis for sedentary time overall (i.e. independent of type of sedentary behaviour), and also stratified the evidence synthesis by type of sedentary behaviour, including TV viewing, computer use/games, total screen time and objectively assessed total sedentary time. The evidence synthesis for overall sedentary time was performed to allow for comparison of the evidence with the previous review and other sedentary behaviour reviews.

The level of evidence was rated as follows:

- **Strong evidence:** consistent findings in multiple (≥2) high-quality studies.

- **Moderate evidence:** consistent findings in one high-quality study and at least one low-quality study, or consistent findings in multiple (≥2) low-quality studies.

- **Insufficient evidence:** only one study available or inconsistent findings in multiple (≥2) studies.

- **No evidence:** consistent findings for no relationship in at least one high-quality study and at least one low-quality study, or consistent findings for no relationship in multiple (≥2) low-quality studies.

Results between studies were considered to be consistent when at least 75% of the studies showed results in the same direction, which was defined by significance (P<0.05). If two or more high-quality studies were available, findings from low methodological quality studies were disregarded in the evidence synthesis. For each study, results from the most fully adjusted model were included in the evidence synthesis.

Some studies examined multiple relationships for the same biomedical health indicator (e.g. analysing the data for multiple subgroups or analysing change scores as well as follow-up values for the health indicator). These studies were considered to add evidence when consistently demonstrating a relationship with the same health indicator, i.e. when at least 75% of examined relationships showed results in the same direction (P<0.05). Exceptions were made for gender, ethnicity and weight status.
group differences, for those differences we included the results in separate samples in the evidence synthesis.

Meta-analysis
As included studies were heterogeneous in measurement of sedentary behaviour and health indicators, statistical analyses and reported types of effect sizes varied to a large extent, making statistical pooling of all studies assessing the same health indicator impossible. Therefore, we systematically examined options for performing an appropriate meta-analysis including a subgroup of sufficiently homogenous studies. The only subgroup that was large enough (≥5 studies) consisted of studies examining the relationship of TV viewing and computer use with indicators of weight status. As large variations in statistical analyses (e.g. change scores vs absolutes values) within this subgroup remained, we contacted all authors of studies including the relationship of TV viewing or computer use at baseline with indicators of weight status (both continuously measured) at follow-up. We asked them to perform regression analyses that were adjusted for baseline weight status, and, if possible, regression analyses that were additionally adjusted for diet and/or physical activity. Thirty-six authors were contacted and ten provided the requested data. Nine independent cohort studies, reported in eight different publications (16-23), examined BMI and one study examined fat mass index (FMI) (24). Therefore, we decided to focus the meta-analysis on BMI only, making standardization of the health indicator unnecessary and thereby providing a better interpretable pooled effect estimate. Studies were only included in the meta-analysis when BMI was assessed before the age of 23, to reduce heterogeneity. Sensitivity analyses were conducted in order to examine whether effects were moderated by diet and/or physical activity and follow-up duration (>1 year). Random effects models were used to pool data and forest plots were created. Pooled effect estimates represent the increase in BMI for each additional hour of baseline TV viewing or computer use a day. Heterogeneity was determined by I^2 and T^2. I^2 represents the variability in effect estimates that is due to heterogeneity rather than sampling error. T^2 is an estimate for the amount of total heterogeneity. An I^2 value >75% was considered to indicate considerable heterogeneity (25). Meta-analyses were conducted in R using the Metafor package.

Results
The Literature search identified 6648 articles of potential interest: 1,810 in Pubmed, 2,547 in Embase, 1,263 in Cochrane and 1028 in PsycINFO. After removing duplicates, titles and abstracts of 4481 articles were screened and subsequently 227 full-text articles were assessed. Finally, 78 articles met the inclusion criteria and were added to the 31 articles included in the review of Chinapaw et al. (7).
information flow and reasons for excluding articles are summarized in figure 1. The characteristics of all newly identified studies (n = 78) are summarized in the online supporting information table S1. Table S2 (online supporting information) shows the quality assessment of all newly identified studies. The results described below and the evidence synthesis of the current review are based on all published studies (n = 109; i.e. the newly identified studies plus the studies included in the review of Chinapaw et al. (7)). Details of the studies included in the previous review can be found elsewhere (Chinapaw et al. (7)).

Sample characteristics
The 109 articles examined 83 different study samples (41 North American, 29 European, 8 Oceanian, 3 Asian and 2 South American). Sample sizes ranged from 26 to 16,225 participants. Ten studies included only girls, one study only boys and five studies included only overweight/obese children. The mean age at baseline varied from 41 weeks up to 16.3 years. The majority of the studies (n=69) was performed in children (6-12 years at baseline), 26 studies were performed in preschool children (2-5 years at baseline), 22 in adolescents (12-18 years at baseline) and one study was performed in infants (1-23 months at baseline). Follow-up duration varied from 5 months up to 27 years, and mean age at follow-up ranged from 25 months to 43 years.

Methodological quality
The quality scores ranged from 25% to 100%, with 19 studies (17%) rated as high quality. Criteria regarding measurement of sedentary behaviour (12%), selective non-response at follow-up (25%) and participation rate at baseline (30%) were least often scored positively.

Measures of sedentary behaviour
Eleven studies (10 different samples) used accelerometers to assess total sedentary time, of which nine studies used a sedentary cut-point of <100 counts per minute and two studies (26;27) used a cut-point of <1100 counts per minute. TV viewing was objectively measured in two studies, using a TV/computer monitoring device (28) and observations (29). Most studies included self- or parent reported TV viewing time as type of sedentary behaviour (i.e. 53% of all included studies). Other frequently used measures were self- or parent reported screen time (35%) and computer use/games (12%). Seven studies (6%) examined combinations of various screen-based and non-screen-based activities. Self-reported time playing board games (30), reading (30), arts (30), total sitting time (31), time spent in a car (32) and time in baby seats (33) were assessed once.
Figure 1: PRISMA flow chart of study selection. *one article included in the previous review was not found with our search. SB, sedentary behaviour.
Evidence for health effects

Relationship between sedentary behaviour and biomedical health indicators

Prospective relationships with 43 different biomedical health indicators were reported. Table 2 presents the results of the evidence synthesis, stratified by type of sedentary behaviour. Table S3 (online supporting information) presents the evidence synthesis in more detail, with references added. Table 3 lists the health indicators that were studied only once, indicating insufficient evidence for a prospective relationship.

Anthropometrics

Regarding TV viewing time, we found strong evidence for a positive relationship with overweight/obesity (based on three high-quality studies), and moderate evidence for a relationship with overweight/obesity incidence (based on three low-quality studies). For the remaining anthropometric indicators, we found insufficient evidence for a relationship, mainly due to inconsistent findings. For girls and overweight/obese children, we found no evidence for a relationship of TV viewing with BMI.

The meta-analysis examining the relationship between baseline TV viewing and BMI at follow-up pooled the data of 24,257 participants from nine independent cohorts (16-23). The pooled effect for each additional hour of baseline TV viewing per day was insignificant in the analysis adjusted for baseline BMI (β = 0.01, 95% CI = [-0.002; 0.02]) (Figure 2). Although a high percentage of total variability could be attributed to heterogeneity across studies (I² = 89%), the estimated absolute amount of heterogeneity was low (T² = 0.00, SE = 0.00). The pooled effect remained insignificant after additional adjustment for diet and/or physical activity (6 cohorts, 14,747 participants; β = 0.01, 95% CI = [-0.03; 0.05], I² = 55%, T² = 0.00, SE = 0.00) (online supporting information figure S1). When stratifying for follow-up duration, the pooled effect was significant for studies with a follow-up duration longer than one year (7 cohorts, 21,826 participants; β = 0.12, 95% CI = [0.01; 0.22], I² = 88%, T² = 0.02, SE = 0.01) (online supporting information figure S2). This effect became insignificant after adjustment for diet and/or physical activity (5 cohorts, 13,418 participants, β = 0.07, 95% CI = [-0.04; 0.18], I² = 64%, T² = 0.01, SE = 0.01) (online supporting information figure S3).

For computer use/games, we found no evidence for a relationship with BMI/BMI z-score (based on ten low-quality studies) or WC/WC z-score (based on two low-quality studies), and insufficient evidence for a relationship with the remaining indicators.

The meta-analysse examining the relationship between computer use and BMI pooled the data of 6,971 participants from five independent cohorts (16;18-20;23). The pooled effect for each additional hour of baseline computer use per day was not significantly related to follow-up BMI in the analysis adjusted
for baseline BMI ($\beta = 0.00$, 95% CI = [-0.004; 0.01], $I^2 = 72\%$, $T^2 = 0.00$, SE = 0.00) (online supporting information figure S4). This effect remained insignificant after additional adjustment for diet and/or physical activity (4 cohorts, 5,759 participants; $\beta = -0.03$, 95% CI [-0.11; 0.05], $P = 0.51$, $I^2 = 64\%$, $T^2 = 0.00$ SE 0.00) (online supporting information figure S5).

Regarding screen time, we found strong evidence for a relationship with BMI/BMI z-score (based on two high-quality studies), and moderate evidence for a relationship with overweight/obesity (based on three low-quality studies). We found insufficient evidence for a relationship with the remaining anthropometric indicators.

For objectively assessed total sedentary time, we found no evidence for a relationship with BMI/BMI z-score, WC/WC z-score and body fat (based on three, three and five high-quality studies, respectively). We found insufficient evidence for a relationship with overweight/obesity incidence.

When combining all different sedentary measures, we found strong evidence for a relationship with overweight/obesity (based on three high-quality studies all examining TV viewing), moderate evidence for a relationship with weight-for-height (based on two low-quality studies). We found no evidence for a relationship with body fat, skinfolds and hip circumference (based on seven high-quality, four high-quality and two low-quality studies, respectively). Studies on the relationship with the remaining anthropometric indicators were inconsistent.

*Cardiometabolic biomarkers*

Regarding TV viewing time, we found no evidence for a relationship with triglycerides and glucose (both based on one high-quality and two low quality studies). We found insufficient evidence for a relationship with the remaining cardiometabolic biomarkers.

For screen time, we found no evidence for a relationship with LDL-cholesterol (based on one high-quality and one low-quality study), ratio of total cholesterol to HDL-cholesterol (based on one high-quality and two low-quality studies) and triglycerides (based on two low-quality studies). We found insufficient evidence for a relationship with the remaining cardiometabolic biomarkers.

For computer use/games and objective assessed total sedentary time we found insufficient evidence for a relationship with cardiometabolic biomarkers.
When combining all different sedentary measures, we found strong evidence for an inverse relationship with HDL-cholesterol (based on two high-quality studies). We found no evidence for a relationship with the remaining cardiometabolic biomarkers.

**Blood pressure**

Regarding TV viewing time, we found no evidence for a relationship with systolic blood pressure (based on one high-quality study and four low-quality studies). Studies on the relationship with diastolic blood pressure and arterial blood pressure/raised blood pressure were inconsistent.

For computer use/games and screen time, we found no evidence for a relationship with systolic blood pressure (based on two high-quality studies) and diastolic blood pressure (based on two low-quality studies).

Regarding objectively assessed total sedentary time we found insufficient evidence for a relationship with mean arterial blood pressure/raised blood pressure.

When combining all different sedentary measures, we found no evidence for a relationship with systolic blood pressure (based on three high-quality studies) and diastolic blood pressure (based on two high-quality studies). We found insufficient evidence for a relationship with mean arterial blood pressure/raised blood pressure (based on two high-quality studies reporting inconsistent findings).

**Fitness**

Regarding TV viewing time, we found strong evidence for an inverse relationship with cardiorespiratory fitness/VO2max (based on two high-quality studies). We found insufficient evidence for a relationship with strength and being unfit.

For screen time, we found insufficient evidence for a relationship with cardiorespiratory fitness/VO2max, due to inconsistent findings. We found insufficient evidence for a relationship with strength and being unfit.

When combining all different sedentary measures, we found strong evidence for an inverse relationship with cardiorespiratory fitness/VO2max (based on two high-quality studies examining TV viewing) and moderate evidence for a relationship with being unfit (based on two low-quality studies). Studies on the relationship with strength were inconsistent.
Other biomedical health indicators

For TV viewing time, we found insufficient evidence for a relationship with metabolic risk z-score, due to inconsistent findings and insufficient evidence for a relationship with asthma and bone mass indicators.

Regarding computer use/games we found no evidence for a relationship with metabolic risk z-score based on two low-quality studies.

For screen time, we found insufficient evidence for a relationship with asthma, due to inconsistent findings and insufficient evidence for a relationship with metabolic risk z-score, due to a lack of studies.

For objectively assessed total sedentary time we found insufficient evidence for a relationship with metabolic risk z-score and bone mass indicators.

When combining all different sedentary measures, we found no evidence for a relationship with bone mass indicators (based on three low quality studies). Studies on the relationship with asthma and metabolic risk z-score were inconsistent.

<table>
<thead>
<tr>
<th>Health Indicator</th>
<th>Evidence Synthesis Stratified by Main Sedentary Behaviour Types</th>
<th>Evidence Synthesis for Overall Sedentary Time*</th>
</tr>
</thead>
<tbody>
<tr>
<td>BMI / BMI z-score</td>
<td>Insufficient</td>
<td>Insufficient</td>
</tr>
<tr>
<td>WC / WC z-score</td>
<td>Insufficient</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Body fat</td>
<td>Insufficient</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Sk infolds</td>
<td>Insufficient</td>
<td>Insufficient</td>
</tr>
<tr>
<td>HC</td>
<td>Insufficient</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Weight</td>
<td>Insufficient</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Weight - for height</td>
<td>Insufficient</td>
<td>Insufficient</td>
</tr>
</tbody>
</table>
Table 2: Evidence synthesis stratified by main type of sedentary behaviour and overall sedentary time

<table>
<thead>
<tr>
<th>Health indicator</th>
<th>Evidence synthesis stratified by main sedentary behaviour types</th>
<th>Evidence synthesis for overall sedentary time*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TV viewing time</td>
<td>Computer use/games</td>
</tr>
<tr>
<td>Anthropometrics</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BMI / BMI z-score</td>
<td>Insufficient</td>
<td>No evidence</td>
</tr>
<tr>
<td>WC / WC z-score</td>
<td>Insufficient</td>
<td>No evidence</td>
</tr>
<tr>
<td>Body fat</td>
<td>Insufficient</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Skinfolds</td>
<td>Insufficient</td>
<td>Insufficient</td>
</tr>
<tr>
<td>HC</td>
<td>Insufficient</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Weight</td>
<td>Insufficient</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Weight-for-height</td>
<td>Insufficient</td>
<td></td>
</tr>
</tbody>
</table>

*Evidence synthesis for overall sedentary time:

When combining all different sedentary measures, we found no evidence for a relationship with bone mass initiators (based on three low-quality studies). Studies on the relationship with asthma and metabolic risk were inconsistent.

For TV viewing time, we found insufficient evidence for a relationship with metabolic risk due to inconsistent findings and insufficient evidence for a relationship with asthma and bone mass initiators.

For computer use/games, we found no evidence for a relationship with metabolic risk based on two low-quality studies.

For screen time, we found insufficient evidence for a relationship with asthma, due to inconsistent findings and insufficient evidence for a relationship with metabolic risk.

For objectively assessed total sedentary time, we found insufficient evidence for a relationship with metabolic risk and bone mass initiators.
<table>
<thead>
<tr>
<th>Health indicators</th>
<th>Evidence synthesis stratified by main sedentary behaviour types</th>
<th>Evidence synthesis for overall sedentary time*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TV viewing time</td>
<td>Computer use/games</td>
</tr>
<tr>
<td>Overweight/obesity</td>
<td>Strong</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>+++++</td>
<td>0</td>
</tr>
<tr>
<td>Overweight/obesity resolution</td>
<td>Insufficient</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Overweight/obesity incidence</td>
<td>Moderate</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>+++++</td>
<td>0</td>
</tr>
</tbody>
</table>

Cardiometabolic biomarkers

<table>
<thead>
<tr>
<th></th>
<th>LDL-C</th>
<th>HDL-C</th>
<th>TC/HDL-C</th>
<th>TG</th>
<th>Glucose</th>
<th>SBP</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Insufficient</td>
<td>Insufficient</td>
<td>Insufficient</td>
<td>No evidence</td>
<td>No evidence</td>
<td>No evidence</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

Blood pressure

<table>
<thead>
<tr>
<th></th>
<th>No evidence</th>
<th>No evidence</th>
<th>No evidence</th>
<th>No evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>+</td>
<td>+</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

* Evidence synthesis for overall sedentary time: Insufficient, No evidence
Table 2 continued

<table>
<thead>
<tr>
<th>Health indicator</th>
<th>Evidence synthesis stratified by main sedentary behaviour types</th>
<th>Evidence synthesis for overall sedentary time*</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>TV viewing time</td>
<td>Computer use/games</td>
</tr>
<tr>
<td>DBP</td>
<td>Insufficient</td>
<td>No evidence</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>0 0</td>
</tr>
<tr>
<td>Mean arterial blood pressure and raised blood pressure</td>
<td>Insufficient</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>0</td>
</tr>
<tr>
<td>Fitness</td>
<td>Strong</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Strength (push up, curl up, SLJ)</td>
<td>Insufficient</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>Being unfit</td>
<td>Insufficient</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td>Other</td>
<td>Insufficient</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>+</td>
<td>+</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Metabolic risk z-score or MeTs</td>
<td>Insufficient</td>
<td>No evidence</td>
</tr>
<tr>
<td></td>
<td>+ ++</td>
<td>0 0</td>
</tr>
<tr>
<td>Bone mass indicators</td>
<td>Insufficient</td>
<td>Insufficient</td>
</tr>
<tr>
<td></td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
Bold indicates a high-quality study. * Note that the amount of studies under the stratified evidence synthesis do not count up in the combined evidence synthesis, due to two reasons. First, some studies examined types of sedentary behaviour that could not be classified in one of the four main types (e.g. subjective sitting time). As these additional types were only examined in its relationship with one health indicator, they were not considered as an additional main type of sedentary behaviour. Second, studies reporting relationships of more than one measurement type were counted once in the combined evidence synthesis, and were considered to add evidence when consistent findings were reported (i.e. ≥75% of the relationships showing results in the same direction).

+, study adding evidence for a positive relationship; - study adding evidence for an inverse relationship; 0 study indicating no evidence for a relationship; BMI, body mass index; CRF, cardiorespiratory fitness; DBP, diastolic blood pressure; FMI, fat mass index; HC, hip circumference; HDL-C, high-density lipoprotein cholesterol; LDL-C, low-density lipoprotein cholesterol; MetS, metabolic syndrome; SBP, systolic blood pressure; SSF, sum of skinfolds; SJ, standing long jump; TC/HDL-c, ratio of total cholesterol to high-density lipoprotein cholesterol; TG, triglycerides; TV, television; WC, waist circumference.

### Table 3: Biomedical health indicators that were reported in only one study sample, i.e. indicating insufficient evidence

<table>
<thead>
<tr>
<th>Anthropometrics</th>
<th>Cardiometabolic biomarkers</th>
<th>Other health indicators</th>
</tr>
</thead>
<tbody>
<tr>
<td>FFMI</td>
<td>Total serum cholesterol</td>
<td>Constipation</td>
</tr>
<tr>
<td>WHR</td>
<td>Non-HDL-C</td>
<td>Persistent fatigue</td>
</tr>
<tr>
<td>Adiposity (composite variable)</td>
<td>TG/HDL-C</td>
<td>Severe fatigue</td>
</tr>
<tr>
<td>Persistent overweight</td>
<td>LDL-C/HDL-C</td>
<td>Functional somatic symptoms</td>
</tr>
<tr>
<td>Weight-for-age</td>
<td>HbA1c</td>
<td>Musculoskeletal pain</td>
</tr>
<tr>
<td>WHR</td>
<td>insulin</td>
<td>Airway hyper responsiveness</td>
</tr>
<tr>
<td></td>
<td>HOMA-IR</td>
<td>Benign breast disease</td>
</tr>
</tbody>
</table>

FFMI, fat free-mass index; HbA1c, hemoglobin A1c; HOMA-IR, homeostatic model assessment of central insulin resistance; LDL-C/HDL-C, ratio of low-density lipoprotein cholesterol to high-density lipoprotein cholesterol; MetS, metabolic syndrome; TG/HDL-C, ratio of triglycerides to high-density lipoprotein cholesterol; WHR, waist-to-hip ratio; WHRR, waist-to-height ratio.
Evidence for health effects

**Figure 2:** Forest plot for the relationship between baseline TV viewing and BMI at follow-up, adjusted for baseline BMI.

<table>
<thead>
<tr>
<th></th>
<th>Coefficient [95% CI]</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dumith</td>
<td>-0.020 [-0.069, 0.019]</td>
</tr>
<tr>
<td>Danner</td>
<td>0.243 [0.174, 0.312]</td>
</tr>
<tr>
<td>Grundved</td>
<td>0.260 [-0.054, 0.574]</td>
</tr>
<tr>
<td>Fuller-Tyszkiewicz (cohort B)</td>
<td>0.317 [0.101, 0.533]</td>
</tr>
<tr>
<td>Fuller-Tyszkiewicz (cohort K)</td>
<td>0.103 [-0.132, 0.338]</td>
</tr>
<tr>
<td>Berentzen</td>
<td>0.000 [-0.002, 0.002]</td>
</tr>
<tr>
<td>Altenburg</td>
<td>0.018 [-0.100, 0.136]</td>
</tr>
<tr>
<td>Burke</td>
<td>0.058 [-0.009, 0.125]</td>
</tr>
<tr>
<td>Gentile</td>
<td>0.004 [0.003, 0.006]</td>
</tr>
<tr>
<td>RE Model</td>
<td>0.008 [-0.002, 0.017]</td>
</tr>
</tbody>
</table>

**Developments over the last years: findings from the current and the previous review**

Within this update, we added 78 publications to the 31 previously identified. Sixty-seven of the 78 added publications were published within the last five years, the remaining 12 publications were published within the time period covered by the previous review (January 1989 – April 2010) but were not identified at that time. The number of studies published per 5-year increments is visualised in the online supporting information figure S6. Nine of the 67 publications (13%) published from April 2010 to January 2015 were rated as high quality compared to 10 of the 43 publications (23%) published up to April 2010, indicating that, overall, the methodological quality of studies did not improve over the last five years.

This update included relationships with a broader range of biomedical health indicators, adding 30 new indicators, of which six reflected an indicator of body composition. For most of the added indicators (21 out of 30), we found insufficient evidence because they were only examined in one study. More studies that were added in this update used accelerometry to measure sedentary time objectively (13%
of the added studies compared to 3% of the previously identified studies), yet self- or parent-reported TV viewing remained the most commonly used measure of sedentary behaviour (i.e. 47% of the added studies and 55% of the studies in the previous review).

Table 4 summarizes the evidence of the previous review (Chinapaw et al. (7)) and compares it to the accumulated evidence of the current review. Since the previous review did not stratify the evidence synthesis for type of sedentary behaviour, this comparison concerns the evidence synthesis when combining all different measures of sedentary behaviour.

Regarding BMI/BMI z-score, we added 42 study samples (four high-quality) to the 22 (seven high-quality) previously identified. Based on the 11 high-quality samples together, evidence for a positive relationship with BMI/BMI z-score remained insufficient due to inconsistent findings. For indicators of fat mass (combining waist circumference, body fat and skinfolds), we added 18 study samples (eight high-quality) to the nine (five high-quality) study samples previously identified. Based on the 13 high-quality samples together, the evidence changed from insufficient to no evidence for a relationship (predominantly based on seven high-quality samples examining objectively assessed total sedentary time). For overweight/obesity we added 15 study samples (two high-quality) to the four study samples (one high-quality) previously identified. Based on the three high-quality samples together (all examining TV viewing) the evidence for a relationship with overweight/obesity changed from insufficient to strong. The evidence for a negative relationship between sedentary behaviour and cardiorespiratory fitness changed from moderate-to-strong, based on two high-quality studies examining TV viewing, of which one was added in this update. Relationships with indicators of bone mass and blood pressure were previously examined in one study, resulting in insufficient evidence. We added two studies investigating indicators of bone mass (both low-quality) and ten studies investigating systolic, diastolic or mean blood pressure/raised blood pressure (four high-quality) and concluded no evidence for a relationship with bone mass or blood pressure.
Table 4: Comparison between findings from the previous review and the current review

<table>
<thead>
<tr>
<th>BMI/BMI-z</th>
<th>Fat mass indicators (WC/WC-z, BF and SF)</th>
<th>Overweight/obesity</th>
<th>CRF/VO_{2max}</th>
<th>Bone mass</th>
<th>Blood pressure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Previous review</td>
<td>Insufficient</td>
<td>Insufficient</td>
<td>Insufficient</td>
<td>Moderate</td>
<td>Insufficient</td>
</tr>
<tr>
<td>Current review</td>
<td>Insufficient</td>
<td>No evidence</td>
<td>Strong(^1)</td>
<td>Strong(^1)</td>
<td>No evidence</td>
</tr>
</tbody>
</table>

BF, body fat; BMI, body mass index; BMI-z, body mass index z-score; CRF, cardiorespiratory fitness; SF, skinfolds; WC, waist circumference; WC-z, waist circumference z-score.

\(^1\) Evidence based on high-quality studies examining TV viewing as specific sedentary behaviour.
Chapter 2

Discussion

This systematic review aimed to summarize and update the evidence regarding the prospective relationship between childhood sedentary behaviour and biomedical health indicators. Overall, the added studies in this update contribute to a better understanding of the potential adverse health effects of young people’s sedentary behaviour by adding evidence from high-quality studies and studies examining sedentary time objectively, and the addition of a broader range of health indicators. The evidence for a relationship with the most frequently examined anthropometric indicators remained insufficient due to inconsistent findings, despite the large amount of additional studies in this update.

Overall, when combining all sedentary measures, we found strong evidence for a relationship with overweight/obesity (although based on high-quality studies assessing TV viewing) and strong evidence for an inverse relationship with fitness (based on high-quality studies assessing TV viewing) and HDL-cholesterol. Our findings for overall sedentary time are partly in accordance with findings from a recent review of reviews of de Rezende et al. (6). In line with De Rezende et al. (6), we concluded strong evidence for a relationship with overweight/obesity. De Rezende et al. (6) found moderate evidence for a relationship of overall sedentary time with fitness and total cholesterol, whereas we found strong evidence for a relationship with fitness and HDL-cholesterol, but insufficient evidence for total cholesterol as this was only reported in one study. We found no evidence for a relationship with blood pressure, whereas de Rezende et al. (6) concluded moderate evidence. A recent meta-analysis found a significant association between overall sedentary time and blood pressure as well (34). Each additional hour spent sedentary was associated with an increase of 0.06 mmHg (95% CI = [0.01; 0.11]) of SBP and 0.20 mmHg (95% CI = [0.10; 0.29]) of DBP.

An explanation for differences in findings with previous reviews is that previous reviews mainly relied on cross-sectional studies and our conclusions are based on prospective studies only. Prospective and cross-sectional studies on the association of sedentary time with health indicators may draw different conclusions as high levels of sedentary time as well as health indicators remain relatively stable over time in children (15;35). In this case no longitudinal relationship will be found between sedentary time and health indicators, however, cross-sectionally they can still be significantly associated. Another explanation for differences in findings between reviews is the use of different measures of sedentary behaviour. In order to examine whether types of sedentary behaviour are differentially related to biomedical health indicators we stratified our evidence synthesis for type of sedentary behaviour.
Evidence for health effects

For TV viewing time, we found strong evidence for a relationship with overweight/obesity, moderate evidence for a relationship with overweight/obesity incidence, and insufficient evidence for a relationship with other anthropometric indicators, including BMI. This suggests that a possible relationship for TV viewing and anthropometry is non-linear. This is confirmed by findings of Mitchell et al. (43) who examined relationship of TV viewing with sum of skinfolds stratified by sum of skinfold percentiles and found that the relationship became stronger in the higher percentiles. Studies including solely overweight or obese children found no evidence for a relationship between TV viewing and anthropometrics (44-46). A recent meta-analysis of cross-sectional studies found that the risk for childhood obesity increased by 13% for every additional hour of TV viewing per day (47). Our meta-analysis revealed no evidence for a prospective relationship of TV viewing with BMI (β = 0.01, 95% CI = [-0.002; 0.02]). Marshall et al. (12) found in their meta-analysis of TV viewing with BMI a very small, significant effect (r = 0.041, 95% CI = [0.030; 0.052]) and concluded that it was unlikely to be of substantial clinical relevance. These differences in statistical significance might be explained by different methods used to pool the data. Marshall et al. (12) transformed all different effect sizes to Pearson’s correlations, while we used regression coefficients to enhance the interpretability of the pooled effect estimate.

For computer use/games, we found in general no evidence for a relationship with biomedical health indicators. Computer use was not related to BMI in our meta-analysis either. This is in agreement with a previous review (11) and meta-analysis (12). An explanation for the different conclusions regarding TV viewing and computer use/games could be that children spend more time watching TV than using the computer (8). Another explanation could be that especially TV viewing is associated with unhealthy eating behaviour, less is known about the association of computer use with eating behaviour (10). A third explanation could be that computer use might induce higher muscle activity and energy expenditure than TV viewing (9). Moreover, in most studies computer use/gaming incorporated both active and passive gaming. Active gaming generally includes light to moderate physical activity (48) and inclusion of active gaming may therefore attenuate possible adverse health effects.

For screen time, we found strong evidence for a relationship with BMI and moderate evidence for a relationship with overweight/obesity. This is consistent with the conclusions of Costigan et al. (42). However, this systematic review of mainly cross-sectional studies was limited to adolescent girls and is therefore not comparable to ours. Our finding of strong evidence for a relationship between screen time and BMI (based on 2 high-quality studies) is remarkable compared to the insufficient evidence we found for a relationship between TV viewing and BMI (based on 6 high-quality studies). More high-quality studies are needed to examine whether the relationship between screen time and BMI remains
strong when more high-quality studies are available. In the meantime, results should be interpreted with caution.

For objectively assessed total sedentary time, we found no convincing evidence for a relationship with biomedical health indicators, which is consistent with a previous review (36) and a previous meta-analysis (37). Adult studies indicate that especially prolonged uninterrupted sitting is detrimental for cardiometabolic health (38-40). A previous study found that children accumulate their total sedentary time in bouts of relative short duration (less than 20 minutes) (41). It is possible that frequent interruptions in children’s sedentary time may prevent them from negative health consequences. To date, no prospective studies have been published on the health effects of different sedentary patterns in young people.

We found that different types of sedentary behaviour were differentially related with biomedical health indicators, which contributes to inconsistent findings when different types are combined into overall sedentary time. Results were inconsistent as well, within different types of sedentary behaviour. One explanation for this finding is that all different types of sedentary behaviour suffer from measurement limitations that may impact the strength of relations found and contribute to the inconsistent findings. Ninety percent of the included studies assessed self- or parent-reported sedentary activities, mostly TV viewing, or different combinations of screen behaviours, summarized as screen time. Lubans et al. (49) systematically reviewed the validity and reliability of sedentary behaviour measures in young people and concluded that the validity and reliability of self-reports was often unknown. Poor validity and reliability will increase the likelihood of misclassification, which will bias relationships to null. Moreover, studies that assessed reliability and validity had important limitations. Studies used a large variety of methods to assess validity and reliability and self- or proxy reports were often compared to another method of unknown validity. Reliability results were mixed and probably dependent on the type of sedentary behaviour, as reliability for TV viewing was in general higher compared to computer use/games (49). Moreover, for the exploration of prospective associations measures need to be sensitive enough to pick up behavioural change. This measurement property is hardly examined. Therefore, we recommend more high-quality studies on the validity, reliability and responsiveness of sedentary behaviour questionnaires. Another possible source of misclassification by self-reports is that children and adolescents may do many sedentary activities simultaneously (media multitasking), for example playing games on a tablet while watching television (50).
Ten percent of the included studies assessed total sedentary time, all using accelerometers. Although objective methods are a major step forward compared to self-reported methods, there are some limitations to consider as well. First, sedentary time assessed by accelerometers may incorporate time spent standing, since standing still will not result in substantial acceleration. The use of posture-based measurement methods such as the activPAL may overcome this problem. The activPAL inclinometer has shown to provide a valid estimate of sitting and standing time in children (51), yet none of the included studies used a posture-based method to measure sitting time. Secondly, accelerometry is not completely objective since subjective decisions (e.g. cut-points, non-wear criteria) should be made in the data reduction process. Used cut-points to classify sedentary time varied for example from 100 to 1100 counts per minute. Data reduction decisions may influence the association with health indicators (52) and no consensus exists yet about optimal data reduction criteria (53). Third, till today is it not possible to distinguish between different sedentary activities based on accelerometer data, future studies should investigate possibilities for this.

Another possible explanation for the inconsistent findings among the included studies is the large variety in used statistical analysis and included covariates and confounders. Forty-four percent of the studies adjusted relationships for physical activity, five percent for sleep and five percent for physical activity and sleep. Twenty-six percent of the studies examining relationships with indicators of body composition added diet as a confounder in the analysis. We have to keep in mind that even when adjustments are made, residual confounding may still exist due to measurement error in factors adjusted for. In order to control for the influence of current levels of sedentary behaviour on the health indicators some studies adjusted for current levels of sedentary behaviour, whereas others analysed change in sedentary behaviour. Additionally, some studies analysed change in sedentary behaviour as well as change in the health indicator. Only a few studies applied longitudinal analyses including multiple repeated measures of sedentary behaviour. As this variation in statistical analyses complicates comparison between studies, we recommend future studies to apply longitudinal techniques including repeated measures of sedentary behaviour over time, especially when follow-up periods are long.

Differences in follow-up duration are another explanation for the inconsistency in findings among our included studies. The follow-up periods of the included studies varied to a large extent (ranging from 5 months up to 27 years). In our meta-analysis on the relationship between TV viewing and BMI, the pooled effect estimated increased from $\beta = 0.01 \ (95\% \ CI = [-0.00; 0.02])$ to $\beta = 0.12 \ (95\% \ CI = [0.01; 0.22])$, when we included only studies with a follow-up duration longer than 1 year in the analysis. This suggests that follow-up duration explains some of the inconsistency in findings among our included studies.
studies. Since we applied a best evidence synthesis was to qualitatively summarize the evidence, high-quality studies were given more weight. However, study quality was not an important contributor to inconsistencies in findings. In general, results of low-quality and high-quality studies showed the same inconsistent pattern. Results within different age groups showed the same inconsistent pattern as well.

A strength of this review is that only prospective studies were included, allowing us to evaluate prospective relationships, which provide stronger evidence for causality than cross-sectional associations. Moreover, we were able to critically evaluate developments in sedentary behaviour research by updating the review of Chinapaw et al. (7). Other strengths of this review include the systematic approach, the quality assessment and the best evidence synthesis. Stratifying the best evidence synthesis by type of sedentary behaviour further strengthens this review. By contacting authors when information on quality criteria was lacking, we optimized the quality assessment. The systematic search for possibilities to conduct a meta-analysis on prospective studies that were sufficiently homogeneous strengthened our meta-analysis. We requested all authors of relevant studies to provide regression coefficients, ensuring that the pooled effect estimate was easily interpretable, which is an important strength of our meta-analysis.

One limitation of this review is the lack of high-quality studies, i.e. only 19 out of 109 studies. Second, a limited number of studies were available for health indicators other than anthropometrics. The large heterogeneity of included studies limited our meta-analysis, making statistical pooling of all included studies inappropriate. The selection of studies included in our meta-analysis may therefore not be representative of all performed studies. Moreover, the amount of studies included in our meta-analysis was too limited to perform a meta-regression on possible moderators (e.g. age, study quality) of the relationship. Finally, publication bias may have influenced the results, since studies with significant findings are more likely to be published.

**Conclusion**

The results of this review suggest that the evidence for a prospective relationship between young people’s sedentary behaviour and biomedical health indicators is not convincing, as it varies by type of sedentary behaviour and biomedical health indicator. For most health indicators, we found no convincing evidence due to inconsistent or non-significant findings. However, this does not mean that there is evidence for absence of a relationship. In order to truly conclude on a potential relationship it is needed to (i) improve the measurement of sedentary behaviour, both the duration as well as the type and context of sedentary behaviour; (ii) distinguish between different sedentary patterns i.e.
prolonged versus interrupted sedentary time; (iii) examine relationships of 24-hour patterns of sedentary behaviour, physical activity and sleep; and (iv) use high-quality cohort studies with multiple time points and examining a range of health indicators.

**Supporting information**

Additional supporting information may be found in the online version of this article at the publisher’s website: https://doi.org/10.1111/obr.12426

**Figure S1.** Forest plot for the relationship between baseline TV viewing and BMI at follow-up, adjusted for baseline BMI, diet and/or physical activity.

**Figure S2.** Forest plot for the relationship between baseline TV viewing and BMI at follow-up, adjusted for baseline BMI and only including studies with a follow-up duration ≥1 year.

**Figure S3.** Forest plot for the relationship between baseline TV viewing and BMI at follow-up, adjusted for baseline BMI, diet and/or physical activity and only including studies with a follow-up duration ≥1 year.

**Figure S4.** Forest plot for the relationship between baseline computer use and BMI at follow-up, adjusted for baseline BMI.

**Figure S5.** Forest plot for the relationship between baseline computer use and BMI at follow-up, adjusted for baseline BMI, diet and/or physical activity.

**Figure S6.** Number of studies published per 5-year increments.

**Table S1.** Characteristics of included studies, by category of biomedical health indicator and sorted by quality score and alphabetically by first author.

**Table S2.** Quality assessment of all newly identified prospective studies on sedentary behaviour and biomedical health indicators in youth sorted by quality score.

**Table S3.** Evidence synthesis stratified by main type of sedentary behaviour and overall sedentary time, in detail with references of included studies added.
References


