Reviewer's report

Title: Trends in child and adolescent obesity prevalence according to socioeconomic position: protocol for a systematic review

Version: 1  Date: 25 March 2014

Reviewer: Paul Shekelle

Reviewer's report:

This is an important topic. I'm not sure a systematic review is the way to answer this question, particularly as what is being looked for is a recent change in trends. Given that it takes several years for original data to get into the published literature, and then another few years for a sufficient number of such studies to be published such that a systematic review can be meaningfully performed, doing a review now in 2014 would be looking at best at data from 2009, and more likely much earlier. But, nevertheless, the availability and currency of the data is one of the goals of a systematic review. So if the authors wish to go forward, it's justifiable to do so.

A few things the authors may wish to consider:

In their inclusion criteria, they state that "only studies from developed countries will be considered". I think this would benefit from some specificity here. Do they plan on using some external criterion standard what is and what is not a developed country? If, so they should probably reference it here. If not, they probably want to state here what they consider the "developed" countries. England, Canada, the USA, etc are easy. But what about Greece? Poland? The areas that were formerly East Germany? Estonia? There are probably examples in southeast asia that are similarly gray, so they will want to think this through. Differentiating Singapore from Nepal is going to be easy. But differentiating some other countries may not be so easy.

Also in the eligibility criteria: the authors state they will include studies with two more time points. In the abstract they describe their search for "cross sectional studies ...with data recorded at two or more time points....". I think I know what they mean here, but the juxtaposition of "cross sectional" (which implies no time component) with "two or more time points" will confuse some readers. I think what the authors mean here are studies that examine annual or biannual cross sectional data across a number of years, such as some national population survey of weight that occurs every year or two. Since the same people aren't necessarily participants in each year of the survey, it's not really correct to call this a longitudinal cohort, but nevertheless calling it all cross sectional is likely to confuse readers unless there's a little more explanation. I don't know of a good term for this kind of study - "serial cross sectional studies...with two or more time points..."?
the authors assessment of quality is certain to cause some concern, both in terms of the actual criteria and in terms of the aggregation into a score. the representativeness of the sample to some larger population would seem to be paramount here, and I don't see it anywhere in the list. Readers will not consider the eligibility criteria as sufficiently homogenous such that all included studies can be generalized to a larger population, which is also made apparent by the data extraction elements (schools, for example). Serial cross sectional studies of weight and socioeconomic position from classrooms at a school or schools in one part of one country have little basis for generalization to the total population of that country. So there needs to be some measure or assessment of how representative the intended population is (national, regional, local, etc.) in addition to the measure about the representativeness of participation (the rate to exceed 85% - and why was 85% chosen? I have often seen 70% or even 80% chosen as a threshold, this is the first paper I have seen that suggests 85% be chosen - readers will be curious where this came from). Next, the "sample size of at least 5000 participants" will cause readers to wonder - why 5000? why not 4000, or 1,000, or 10,000? Can the authors provide some kind of rationale for this choice? i presume it is related to some measure of power, but if so they ought to be able to defend this choice. I am also surprised that in the quality assessment there is no criterion about how the key SEP variable is measured. they have a criterion for the difference between objective and self-report for the weight, but there must be some similar concern about how SEP is assessed, not just that it is assessed the same way over time. I suggest the authors looks at both the QUADAS criteria and the criteria proposed by Hayden and colleagues in Annals of Internal Medicine 2013, since these topics (diagnosis and prognosis) share many similarities in terms of the risk of bias with the authors proposed assessment of the association between weight and SEP. And then of course the aggregation into a scale needs justification. if they are going to use a scale, then they will need to be able to reference that the threshold on the scale has some documented relationship with differences in study outcome, for examples of such in the literature see Moher 1998 in the Lancet or van Tulder in SPINE in 2009.

The synthesis of these data are going to be tricky, as the authors will need to make lots of judgments about which studies have populations and measures of obesity and measures of SEP are sufficiently similar to be treated together, in a pooled result or even in a narrative result. there's no recipe to follow here, so careful judgments and explanations are what readers are going to want to see.