Reviewer's report

Title: National, regional, and global trends in overweight and obesity

Version: 1 Date: 18 April 2012

Reviewer: Tanja Srebotnjak

Reviewer's report:

Comments for the authors:

The present manuscript explains how a comprehensive set of estimates of the prevalence of overweight and obesity has been produced for men and women in 199 countries and territories between 1980 to 2008. The study does not aim to identify the causes for the increases or decreases in overweight and obesity prevalence but has as its goal to produce accurate estimates using the best available data.

The topic of the paper is of high interest to the audience of Population Health Metrics. It responds to a clear and immediate need for comprehensive national level statistics on overweight and obesity for setting policy targets. It thereby provides the complementary “bigger picture” analysis to sub-national and local analyses of overweight and obesity as well as for analyses for different population strata.

The paper is well-written. I have a few comments, questions for clarification, and suggestions to further improve it.

Major Compulsory Revisions

(1) My first main comment relates to the two modeling approached used to first obtain estimates of mean BMI for each country, sex, age group, and year and second to get from estimated mean BMI to estimated prevalence by sex, age, country and year.

The first model is a Bayesian hierarchical model and is explained in more detail in Finucane et al. (2011).

The second is a regression model with logit of prevalence as the outcome variable and estimated mean BMI from the first model the relevant covariate.

Overall, I am having some difficulties understanding exactly what the first and second models look like and would suggest including the equations into the manuscript (since the study produces estimates for each sex, separately, this could be done by showing the equations only for men to safe space). It would probably go a long way to help readers understand the technical approach taken.

My second main comment relates to the lack of a section on the sensitivity (or robustness) of the results to the assumptions and other decisions made in the models and analysis. The estimates are based on (i) data from different sources,
levels of representativeness, and accuracy, (ii) rely on at least two modeling steps (perhaps missingness was modeled as well before running the Bayesian model) and (iii) also involve decisions such as how to model temporal trends, the relationship between mean BMI and prevalence.

My question, therefore, is whether the authors have tested the sensitivity of the results to at least the following sources of uncertainty and sensitivity:

a. Assumptions of both the Bayesian and logit regression models.
b. Missing data treatment
c. Convergence diagnostics of the Markov Chain algorithm
d. Treatment of YEAR in the prevalence regressions as zero between 1980-1993 and then as smoothed linear function.
e. Choice of cubic splines as opposed to, for example, including the variable in linear, quadratic and cubic form into the regression model.

I recommend including the sensitivity analysis in a separate section Limitation in the manuscript.

Minor Essential Revisions

Abstract, Background Section: yet data on trends in overweight and obesity prevalence by country are not available.

I believe the authors mean that comparable estimates are not available for the majority of countries. The data sets identified and used for this study suggest that such data is available for some countries.

Abstract, Background Section: We used a regression equation to predict overweight and obesity prevalence by age, country, year, and sex, accounting for the uncertainty of prediction.

I think it’s relevant to state that estimated mean BMI is the key covariate in the model to estimate the prevalences of overweight and obesity by age, country, year and sex and that the model accounts, inter alia, for the uncertainty inherent in the mean BMI estimates.

Abstract, Results Section: the Results section only presents findings for obesity. Please, include a brief presentation of the findings for overweight.

Methods:

What software and function or module was used to produce the Bayesian mean BMI estimates and the prevalence estimates?

We estimated prevalences using the relationship between prevalence and mean BMI, applied to estimates of mean BMI from a previously published systematic analysis of health examination surveys and epidemiological studies.

Is this the study referenced in [6] by Finucane et al. published in 2011 in the Lancet? If so, it was stated that this study only produced estimates for 1980 and
2008, so when were the annual mean BMI estimates generated, in the present study or the 2011 study?

... converting from mean to prevalence.

How was this conversion done?

The uncertainty of the estimated mean BMI accounted for sampling uncertainty in the original data sources; uncertainties ... and uncertainty due to using a model to estimate mean BMI by age group, country, and year when data were missing.

The authors state that the Bayesian model controls for several sources of uncertainty but does not explain how. In reference to my previous comment, it would be very helpful to see what the model looks like and to briefly summarize how the model manages to account for all of these sources of uncertainty.

... obtained samples from the posterior distribution of model parameters,
Can the authors state if the sampling is done from the joint or marginal posterior distribution(s)? I assume it's the joint distribution.

We developed regressions to estimate overweight and obesity prevalences from mean BMI.

Please state how prevalence was calculated from the available data, e.g., taking into account the survey design, sampling weights, representativeness of the population included in the survey or epidemiological study. This can also be done in the paragraph that starts with: The regressions were fit using 1884 age-sex observations from 243 health examination surveys...

We used a smoothed linear function of year between the years of 1993 and 2008.

What form, or provide the complete model equation.

We also sampled from the error term of the fitted regression.

What approach was used to do that, e.g., a Gaussian model with the error term’s estimated mean and variance or the error term’s empirical distribution? Was the assumption on the error term’s distribution met approximately after fitting the model?

How was the posterior probability that the estimated increase or decrease in prevalence was truly reflecting an increase or decrease in BMI calculated?

Results Section:
The Results section is very informative and extensive. It is also a bit overwhelming to read in that a lot of figures are presented and from various angles. Could the authors attempt to summarize the findings in a table, such as the one below?

<table>
<thead>
<tr>
<th>Statistic</th>
<th>Overweight (95% CI)</th>
<th>Obesity (95% CI)</th>
</tr>
</thead>
</table>

```
Biggest increase in %, country, females, 1980-2008
Smallest increase in %, country, females, 1980-2008
Biggest increase in %, country, males, 1980-2008
Smallest increase in %, country, males, 1980-2008
Biggest increase in million people, country, females, 1980-2008
Smallest increase in million people, country, females, 1980-2008
Biggest increase in million people, country, males, 1980-2008
Smallest increase in million people, country, males, 1980-2008
SAME FOR REGIONS…

Discretionary Revisions

Abstract, Results Section: in the Democratic Republic of Congo

Page numbering is off following page 9.

I find the continuous color scale used in the maps less useful than identifying a specific color with an interval of values because the continuous scale makes it more difficult to map the color seen to a value on the scale. So I preferred the maps showing the uncertainty of the estimates in 5 intervals.

**Level of interest:** An article of outstanding merit and interest in its field

**Quality of written English:** Acceptable

**Statistical review:** Yes, but I do not feel adequately qualified to assess the statistics.

**Declaration of competing interests:**

I declare that I have no competing interests.