Reviewer’s report

Title: National and subnational hypertension prevalence estimates for the Republic of Ireland: better subnational data are needed to produce better subnational prevalence estimates

Version: Date: 13 November 2013

Reviewer: Michael Soljak

Reviewer’s report:

Is the question posed by the authors well defined?
No. The authors do not mention in the aims e.g. identification of risk factors and their odds.

2. Are the methods appropriate and well described?
The statistical methods are appropriate. However I do not understand why only doctor-diagnosed hypertension was modelled (noting comment below about incidence vs prevalence). The authors say the aim was to develop "a consistent set of national and subnational estimates of the prevalence of hypertension in a country", but they have only modelled already diagnosed disease. Our attached paper shows high levels of under-diagnosis. So why not model risk factors for overall prevalence? There is no evidence of internal validation e.g. prediction of caseness by the model in the dataset. Area under ROC curves was not carried out to determine sensitivity and specificity, and to determine the loss between an optimal multivariable model and local estimates. Were there any interactions e.g. between age and sex? The method by which local prevalence was calculated is not described. Multiple imputation of predictors could be used for respondents who did not have a physical exam.

3. Are the data sound?
The dataset is a representative national survey, but the authors do not state if households or individuals are the sample frame. However one of the outcomes is patient-reported doctor diagnosed disease. According to the paper, patients were "asked if, in the last 12 months, they had been told by a doctor that they had high blood pressure". This question is appropriate for diagnosed incidence not prevalence, for the latter it should be ever told. The authors should say more about how BP was measured: once or several times? With what device? Also the definition should also include physically measured hypertension (#140mmHg SBP or #90mmHg DBP) AND were using anti-hypertensive medication as well as OR.

Does the manuscript adhere to the relevant standards for reporting and data deposition?
I was very surprised that the only local data available was an age breakdown. The local English model includes deprivation and ethnicity. Table 1 mentions
deprivation scores and social class so why were these not included in the models? This is likely to be an important predictor. If only age is available locally, why bother to fit a regression model?

5. Are the discussion and conclusions well balanced and adequately supported by the data?
Yes, a range of other surveys are discussed. Some of this material should be moved to the Background to give readers more context and understanding of other prevalence studies.

6. Are limitations of the work clearly stated?
The problem of small sample size for BP measurement is acknowledged. Multiple imputation could be used for missing i.e. unrecorded data.

7. Do the authors clearly acknowledge any work upon which they are building, both published and unpublished?
Generally, but they do not mention the English model (http://www.apho.org.uk/DISEASEPREVALENCEMODELS), or the validity of patient reports.

8. Do the title and abstract accurately convey what has been found?
No, it does not include the data source i.e. cross-sectional survey or method e.g. risk factors.

9. Is the writing acceptable?
Generally good, but the explanation of local variable selection is poor.

**Level of interest:** An article whose findings are important to those with closely related research interests

**Quality of written English:** Not suitable for publication unless extensively edited

**Statistical review:** No, the manuscript does not need to be seen by a statistician.

**Declaration of competing interests:**
'I declare that I have no competing interests'