Reviewer's report

Title: Exploring Genetic Variants Predisposing to Diabetes Mellitus and their Association with Indicators of Socioeconomic Status

Version: 1  Date: 28 April 2014

Reviewer: Jon Ivar Elstad

Reviewer's report:

1. The relevance of the topic

The topic of this paper is well motivated: Most major diseases have an inverse association with SES (higher incidence/prevalence in lower social strata), and since family studies suggest that genetics is involved in the aetiology of practically all such diseases, the hypothesis that disease-associated alleles have a higher frequency in low-SES than in high-SES persons should be considered. That genetic variations are part of the explanation for the ubiquitous social inequalities in health is, arguably, a not uncommon belief both within and outside science, but empirical assessment of this assumption is very rare.

In my opinion, a great value of the present paper is therefore that it directly addresses this issue which is strangely understudied. During the last decade, analyses of associations between health outcomes and genetic differences at the molecular level have multiplied. Differences in allele frequencies (usually focusing on SNPs and more and more using the GEWAS technique) between subpopulations defined by ethnicity and geography have been described in order to investigate sources of disease variation. Published studies on socioeconomic variations in frequency of disease-associated genetic variants are strikingly lacking, however. The authors are probably right when they say (p.11) that their study may be the first one addressing this topic with respect to diabetes. Thus, the paper is a welcome contribution to the clarification of an issue interesting both for public health (possible role of genetics in social health inequalities) and for the more specific issue of how the epidemiology of/inequalities in diabetes should be understood.

2. Overall comments to the paper

The study reported in this paper appears as solidly done, and the presentation is generally transparent, well structured, and clearly written. The main results are interesting: In the analyzed sample (1) diabetes has a clear SES gradient, (2) the analyzed genetic variants have some (although often weak) associations with diabetes, but (3) no distinct association between the three measurements of SES and the presence/absence of diabetes risk alleles emerged.

The sample (around 4700 middle-aged adults) seems adequate. A larger sample would of course be nice, but a sample of this size should be sufficient (larger samples might produce statistically significant results with little “real” value).
The focus of the study is on 11 selected SNPs with previously established associations with diabetes, analyzed both separately and as a combined “genetic risk score”. Over time, as studies have accumulated, the number of genetic variants with well-established associations with diabetes (i.e., findings robustly reproduced in several independent samples and studies) has increased. Thus, a 2011 study (Herder & Roden, Eur Journ Clinical Investigation) lists “36 diabetes-associated genes”, and the authors themselves mention that an even more recent 2012 study notes “66 genomic loci related to diabetes”. However, the 11 SNPs studied in the present paper are among the most established “candidate genes” for diabetes and the authors are probably right (p.12) that it is hardly plausible that analyses of other diabetes-associated SNPs would challenge the main results of the present study.

3. Suggestions for Discretionary Revisions

3a. The classification into “cases” (i.e., afflicted by diabetes) seems reasonable (p.5) – I think, however, that the characterization of diabetes as a “late onset complex disease” (p.4) should be qualified as diabetes type 1 often develops fairly early in life (type 2 is usually a late-onset disease, maybe the type 1/type 2 distinction of diabetes mellitus should be mentioned).

3b. The socioeconomic classification with respect to education, occupation and income is acceptable (possible misclassification will hardly be of a serious magnitude). The utilization of three different ways of measuring SES constitutes a kind of robustness check. P.6: “… dividing the total household net income by a weighting factor for each household member…” - I suggest that the equivalence scale used here should be specified.

3c. The statistical techniques employed in the paper appear to me as satisfactory and in line with standard practice. I think, however, that the use of the word “dichotomization” could be confusing? See p.7, last paragraph, p.8 upper paragraph, and a few other places: I believe that “dichotomize” usually means a partition into two parts, but here it is used to indicate that a categorical predictor (occupation, education) with K levels is represented by K-1 dummies.

3d. The description of the findings conforms to the statistical results, and the summary and discussion section (pp.10-12) has an adequate summary and a reasonable assessment of possible limitations of the study. I think however that it could also be underlined in the summary that although a number of genetic loci are found to be statistically related to diabetes, even all the established “risk alleles” together, in combination, seem to account for no more than a rather modest proportion (about 5 per cent?) of the total variation in diabetes occurrence (at least when standard statistical techniques, as in this study, is employed). This point is touched upon in the introduction p.3-4, but it could be developed in the summary.

3e. Overall the paper is clearly written, but a few sentences seem not “good English” and perhaps in need of clarification – I suggest, for instance, that the
latter part of p.3 and the upper part of p.4 should be checked, and the sentence in the middle of p.6 starting with “Compared to previous…”. Are “were” and “was” sometimes used when “have been/has been” would be better? P.6, last line, “windows” should be given capital letter “Windows”.

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

**Quality of written English:** Acceptable

**Statistical review:** Yes, and I have assessed the statistics in my report.

**Declaration of competing interests:**

I declare that I have no competing interests.