Reviewer’s report

Title: Diagnostic criteria for diabetes revisited: Taking advantage of sensitivity and specificity of current cut-offs

Authors:

Ali Parappil (aliparappil@yahoo.com)
Suhail A.R. Doi (sardoi@onebox.com)
Kamal A.S. Al-Shoumer (kshoumer@hsc.kuniv.edu.kw)

Version: 3 Date: 19 Dec 2001

Reviewer: Dr L Czupryniak

Level of interest: A paper of considerable general medical or scientific interest

Advice on publication: Accept after discretionary revisions

The paper by A. Parappil et al. deals with one of the most important issues in diabetology at the dawn of the new millennium, namely - how to make sure that a given person is actually suffering from diabetes and that initiation of the treatment is impartially indicated. Since 1997, when ADA changed the criteria for diabetes diagnosis, a - sometimes heated - dispute is going on whether the changes are really changes for better. In particular, as new criteria employ fasting rather than nonfasting plasma glucose values, they are criticized from the point that by using them one may miss a considerable proportion of subjects in whom postprandial glucose is elevated to the level fulfilling criteria of diabetes in oral glucose tolerance test.

There are numerous reports indicating that new criteria are burdened with severe imperfection and that are not able to identify all human subjects eligible for diabetes treatment. The paper I have had honour to review tries to find a solution to the debate and proposes the use of combined - fasting and 2-hour post-challenge glucose values - criteria for diagnosing of diabetes and impaired glucose tolerance. The findings of the authors, presented in good and clear English, are not immensely original and have been reported already by other groups, but the case for the combined criteria use is presented in a clear way, and is sufficiently justified. It seems more and more likely that in future we will all end up using plasma glucose measured in both fasting and postprandial states in order to make a valid diagnosis.

However, being perhaps in the first line of the diagnostic criteria reform, the paper is not flawless and cannot escape some criticism:

Major points:
1. The main drawback of the study is the study population:
   a) The authors studied a rather small, 44-person sample from general clinic referred to a diabetes centre with suspicion of glucose intolerance. It is interesting what diabetes risk factors were present in this small group. We have shown that high risk individuals are at particular risk of discordance between fasting and post-challenge criteria (Drzewoski, Czupryniak, Diabetic Med, 2001, 18: 29-31).
   b) The group was of highly unequal gender distribution (33 men and - only! - 11 women). Moreover, its ethnicity was mixed (Kuwaiti residents and expatriate), and the group was heterogenous in terms of age (30 to 70 years) and body weight (BMI ranged from 18.3 to 44.9 kg/m2). It has been shown that ethnic background and age may influence the agreement between old and new criteria (cf. Wahl et al, Lancet,

c) This large spectrum of the studied subjects may seem valuable, but only when the studied cohort is significantly more numerous. To be able to reach valid conclusions in the subject of diagnostic criteria and to have sufficient power whilst proposing new criteria, one has to study at least several hundreds, if not thousands, subjects (cf. DECODE study). The ADA Expert Committee, who published their recommendations for their new criteria in 1997, had analyzed data from the largest studied ever done in this field. Nevertheless, in my opinion this small sample number does not necessarily preclude the paper from publication, but the authors should definitely acknowledge this serious weakness of the study in the discussion to a considerably greater extent than they have done already in the last sentence of the paper.

2. The description of the statistical analysis needs some improvement: when comparing two (or more) sets of diagnostic criteria, McNemar's test and kappa statistics should be used. The authors do not state either which criteria they used as gold standard in the specificity and sensitivity analysis or how (with what test) they assessed the significance of differences between means.

Minor points:
1. It is not clear why mean BMI value is given only for 39 subjects out of 44 from the whole group. That should be corrected.
2. HbA1c is apparently more closely related with fasting rather than postprandial glucose (cf. Bonora et al, Diabetes Care, 2001, 24: 2023-2029) and this should be acknowledged in the discussion.
3. The paper lacks clear conclusion which should be given under its own subheading.
4. Abstract is too short (only around 100 words, while it could be twice as long) and less than informative as it lacks hard data, which should be included in its content. Also, clear conclusion should be added.

It might be useful to include the phrase 'combined criteria' in the title as they are the real value of the study.

**Competing interests:**

None declared.