Available online at
ScienceDirect
www.sciencedirect.com
Elsevier Masson France
EM|consulte
www.em-consulte.com/en
Diabetes & Metabolism
Diabetes & Metabolism xxx (2016) xxx–xxx

Research letter

Reproducibility and least significant differences of oral glucose tolerance test-derived parameters in a postmenopausal population without diabetes

1. Introduction

The prevalence of type 2 diabetes is increasing worldwide and is associated with important morbidity and mortality, necessitating reproducible diagnostic tools to establish early and reliable diagnosis. Similarly, reproducibility is indispensable to allow evaluation of the effect of therapeutic interventions. Widely used diagnostic tools for impaired glucose tolerance (IGT) and type 2 diabetes include the oral glucose tolerance test (OGTT) and fasting plasma glucose (FPG) levels [1].

Although the OGTT has the benefit of being practical and standardizable in clinical settings, previous studies have reported rather poor reproducibility. Most of those studies, however, had methodological shortcomings, such as study population heterogeneity, and the use of variable time intervals (which were not always realistic or clinically relevant) between OGTTs and archaic analytical methods [2–4]. This therefore prompted an investigation into the reproducibility of OGTT-derived parameters in a homogeneous postmenopausal population without diabetes within a clinically relevant time interval, using state-of-the-art analytical methods. In addition, least significant differences (LSDs) were calculated to allow better appreciation of any changes in glucose parameters.

2. Methods

Forty postmenopausal women without diabetes were enrolled in the control arm of a prospective randomized controlled trial. Of these women, six were excluded from our analysis: three had received a new diagnosis of diabetes; and three others had increased estradiol and sex hormone-binding globulin (SHBG) levels suggesting they were not in a true menopausal state.

The study was approved by the ethics review board of Ghent University Hospital and conducted according to the principles of the Declaration of Helsinki. Informed consent was obtained from all participants before entering the study (EudraCT number: 2008-003661-19).

Participants were seen on two visits with an interval of 6–8 weeks after an overnight fast. At baseline, anthropometric measurements were taken. At both visits, a catheter was placed and baseline blood samples were drawn. Each participant was then given a 75-g (200-mL) glucose solution (Novolab NV, Geraardsbergen, Belgium) that had to be drunk within 5 min. Thereafter, the participants were asked to rest in a chair while blood samples were taken after 30, 60, 90 and 120 min. Patients were deemed to have IFG if their FPG was between 5.6 mmol/L and 6.9 mmol/L, or IGT if their $glucose_{120min}$ was between 7.8 mmol/L and 11.0 mmol/L.

Also, serum C-peptide and insulin levels were determined by electrochemiluminescence using a cobas e411 analyzer (Roche Diagnostics, Basel, Switzerland), which had an analytical intra-assay coefficient of variation (intra-CV) between 2.93% ($\mu = 14.16\ \mu U/mL$) and 2.38% ($\mu = 62.6\ \mu U/mL$) for insulin, and between 3.11% ($\mu = 1.38\ \mu g/L$) and 2.55% ($\mu = 7.07\ \mu g/L$) for C-peptide. Glucose was analyzed by the hexokinase method (cobas, Roche Diagnostics), and its intra-CV ranged from 1.58% ($\mu = 64.7$ mg/dL) to 1.38% ($\mu = 369$ mg/dL). Homoeostasis model assessment for insulin resistance (HOMA-IR) was calculated as $HOMA - IR = \frac{fasting\ \ glucose\left(\frac{mmol}{l}\right) \times fasting\ \ insulin\left(\frac{mU}{l}\right)}{22.5}$. To investigate within-subject reproducibility, the coefficient of variation (CV) and percentage mean difference (%MDiff) were calculated as $CV = \frac{SD}{mean}$ and $\%MDiff = \frac{abs(V_1 - V_2)}{\frac{V_1 + V_2}{2}} \times 100$, respectively.

The CVs reported here comprise the sum of the biological, pre-analytical and analytical CVs. Good reproducibility was defined as a CV < 5%, moderate reproducibility as a CV of 5–10% and poor reproducibility as a CV > 10%. To assess differences in reproducibility, these CVs were compared using the related-samples Wilcoxon signed-rank test. Differences were considered statistically significant at P-values < 0.05. The LSD, which indicates how much two test results need to differ to be significantly different, was calculated as $LSD = t_{0.05/2, DFw}\sqrt{MSw\left(\frac{1}{n_A} + \frac{1}{n_B}\right)}$, with the mean square within (MSw) determined by analysis of variance (ANOVA) test. IBM

Abbreviations: $C\text{-peptide}_{120min}$, 2-h C-peptide; CV, coefficient of variation; FPG, fasting plasma glucose; $glucose_{60min}$, 1-h glucose; $glucose_{120min}$, 2-h glucose; HOMA-IR, homoeostasis model assessment of insulin resistance; IFG, impaired fasting glucose; IGT, impaired glucose tolerance; $insulin_{120min}$, 2-h insulin; LSD, least significant difference; MSw, mean square within; OGTT, oral glucose tolerance test; %MDiff, percentage mean difference.

http://dx.doi.org/10.1016/j.diabet.2016.12.005

SPSS software (version 23, Armonk, NY, USA) was used for all statistical analyses.

3. Results

All 34 subjects were free from diabetes and were postmenopausal white women [age 54 ± 3 years; body mass index (BMI): 24.12 ± 3.44 kg/m^2]; 35.3% were overweight (BMI > 25 kg/m^2) and 5.9% (two women) were obese (BMI > 30 kg/m^2).

Table 1 shows the CV, %MDiff and LSD of the OGTT-derived parameters. The best reproducibility (lowest CV) was observed for fasting glucose, whereas fasting C-peptide and fasting insulin were less reproducible ($P < 0.001$ vs. fasting glucose for both), with fasting insulin being the least reproducible ($P < 0.001$ vs. fasting C-peptide).

The reproducibility of either glucose$_{60min}$ or glucose$_{120min}$ did not differ from each other ($P = 0.791$), although both were less so than fasting glucose levels ($P < 0.001$ and $P = 0.001$, respectively). Similar results were found for both insulin [fasting CV lower than the CV at 60 min ($P = 0.009$) and at 120 min ($P = 0.019$), although the CV at 60 min did not differ from the CV at 120 min ($P = 0.469$)] and C-peptide levels [fasting CV lower than the CV at 60 min ($P = 0.002$) and 120 min ($P = 0.012$), although the CV at 60 min did not differ from the CV at 120 min ($P = 0.898$)].

The HOMA-IR score was less reproducible than any of the fasting parameters ($P < 0.001$ vs. fasting glucose and fasting C-peptide; $P = 0.005$ vs. fasting insulin). Overall, the LSD was greater for the 120-min parameters compared with the respective fasting parameters.

Greater reproducibility for glucose$_{60min}$ was observed in participants with a BMI > 25 kg/m^2 compared with those with a BMI < 25 kg/m^2 ($P = 0.033$). Likewise, when dividing participants according to their median HOMA-IR at baseline, greater reproducibility was found for glucose$_{60min}$ for participants with a HOMA-IR score > 1.61 vs. < 1.61 ($P = 0.041$).

The reproducibility of establishing a diagnosis of either IFG or IGT based on these consecutive OGTTs was also calculated. Four participants had IFG, and 23 participants had normal fasting glucose levels at both time points. Seven of the 34 participants, however, had one fasting glucose level above and one below the 5.6 mmol/L cut-off point, such that 21% of the diagnoses for IFG were inconsistent. Similarly, 29 participants had normal glucose tolerance and two participants had IGT on both OGTTs. Three other participants, however, had discrepant glucose$_{120min}$ values with respect to the 7.8 mmol/L cut-off. Thus, 9% of the diagnoses for IGT were inconsistent.

4. Discussion

In this homogeneous postmenopausal population without diabetes, our present analysis has demonstrated good to moderate reproducibility of fasting glucose, fasting C-peptide and glucose$_{120min}$ levels, but considerable variation in HOMA-IR scores and insulin levels. Overall, fasting levels were more reproducible than levels measured at later time points during an OGTT. These results, especially the poorer reproducibility of the 2-h parameters, add to the current literature, as a clinically relevant time interval between OGTTs was applied along with the use of state-of-the-art analytical methods. In addition, the calculated LSD may be helpful in making clinical diagnoses and evaluating therapeutic interventions. Indeed, as the LSD of glucose$_{120min}$ is 0.677 mmol/L, patients with values around the cut-off levels for a diagnosis of IGT (7.12–8.48 mmol/L) or diabetes (10.4–11.8 mmol/L) might benefit from repeat testing before establishing a diagnosis. This underlines the importance of the recommendation of the American Diabetes Association (ADA) to confirm diagnosis through repeat testing. Similarly, longitudinal changes in fasting glucose levels (such as during an intervention) need to exceed 0.24 mmol/L before they can be considered significantly different.

Our present findings on the reproducibility of glucose and insulin levels are largely in line with those of previous studies [2–5]. Moreover, the reproducibility of diagnostic outcomes (IFG, IGT) also agree with previous studies reporting reproducibility scores for a diagnosis of IGT or type 2 diabetes ranging from 65.6% to 78% [6,7].

The observed variation in OGTT-derived parameters can be explained by pre-analytical, analytical and biological variations [8]. Pre-analytical variation encompasses every variable before sample analysis, including handling of the sample (time between sampling and analysis), variations in test procedures (temperature of the soluble glucose load, ingestion time), and environmental and patient-related factors (time of day, duration of fasting, changes in body weight). Biological variation is due to day-to-day changes, which can vary in each individual (within variation), but also between people (between variation). Analytical variation represents the variation that arises when a sample is measured several times using the same analytical device. The analytical intra-CVs for glucose, insulin and C-peptide (see the above Methods section) were small compared with the overall CVs, thereby rendering the contribution of analytical variation to the total variation apparently minor.

The lesser reproducibility of the 2-h parameters did not come as a surprise, as the factors influencing biological variations during fasting can also have postprandial effects, together with additional factors such as gastrointestinal absorption rate, time allowed for drinking the glucose solution and biological variations in glucose metabolism. However, as analytical variation cannot be influenced and biological variation can only be minimally controlled, minimizing the pre-analytical variation of an OGTT is of some importance.

An explanation for the poor reproducibility of the HOMA-IR is its dependence on variations in both fasting glucose and fasting insulin levels. The wide variation in insulin and insulin-dependent OGTT-derived parameters can partly be explained by the pulsatile secretion of insulin [9]. Nevertheless, as the variation between consecutive measurements is random, repeat testing should lead to more reliable results. This idea has also been stated by Gordon et al. [10].

Thus, our present study indicates that both BMI and insulin resistance (HOMA-IR) affect the reproducibility of glucose$_{60min}$. While such results have never been reported

Please cite this article in press as: Van de Velde FP, et al. Reproducibility and least significant differences of oral glucose tolerance test-derived parameters in a postmenopausal population without diabetes. Diabetes Metab (2016), http://dx.doi.org/10.1016/j.diabet.2016.12.005