Accuracy of the one-hour plasma glucose during the oral glucose tolerance test to diagnose type 2 diabetes in adults: a meta-analysis

Accuracy of one-hour OGTT to detect diabetes

Vasudha Ahuja PhD[1], Pasi Aronen MSc[2], TA Pramod Kumar MPhil[3], Helen Looker MBBS[4], Angela Chetrit MA[5], Aini H Bloigu BSc[6], Auni Juutilainen MD[7], Cristina Bianchi MD[8], Lucia La Sala PhD[9], R M Anjana MD[3], R Pradeepa PhD[3], U Venkatesan MSc[3], S Jebarani EMBA[3], V Baskar MSc[3], Teresa Vanessa Fiorentino MD[10], Patrick Timpel MSc[11], Ralph A DeFronzo MD[12], Antonio Ceriello MD[9], Stefano Del Prato MD[8], Muhammad Abdul-Ghani MD[12], Sirkka Keinänen-Kiukaanniemi MD[6,13], Rachel Dankner MD[5,14], Peter H Bennett MB ChB[4], William C Knowler MD[4], Peter Schwarz MD[11,15,16], Giorgio Sesti MD[17], Rie Oka, MD[18], V Mohan, MD[3], Leif Groop, MD[1,19], Jaakko Tuomilehto, MD[20,21,22], Samuli Ripatti, PhD[1,23,24], Michael Bergman, MD[25]*, Tiinamaija Tuomi, MD[1,19,26]*

*** Equal contribution**

[1]Institute for Molecular Medicine Finland, University of Helsinki, Helsinki, Finland

[2]Biostatistics unit, Faculty of Medicine at the University of Helsinki and Helsinki University Hospital, University of Helsinki, Helsinki

[3]Madras Diabetes Research Foundation & Dr. Mohan's Diabetes Specialties Centre, ICMR Centre for Advanced Research on Diabetes and IDF Centre of Excellence in Diabetes Chennai, India

[4]Phoenix Epidemiology and Clinical Research Branch, National Institute for Diabetes and Digestive and Kidney Diseases, Phoenix, AZ, USA

[5]Unit for Cardiovascular Epidemiology, the Gertner Institute for Epidemiology and Health Policy Research, Ramat Gan, Israel

[6]Center for Life Course Health Research, University of Oulu, Oulu, Finland

[7]University of Eastern Finland, Kuopio University Hospital, Kuopio, Finland

[8]Section of Diabetes and Metabolic Diseases, Department of Clinical and Experimental Medicine, University Hospital of Pisa, Pisa, Italy

[9]Department of Cardiovascular and Dysmetabolic Diseases, IRCCS MultiMedica, Milan, Italy

[10]Department of Medical and Surgical Sciences, University Magna Graecia of Catanzaro, Catanzaro, Italy

[11]Department of Medicine III, Technical University of Dresden, Dresden, Germany

[12]Division of Diabetes, University of Texas Health Science Center at San Antonio, San Antonio, TX, USA

[13]Healthcare and Social Services of Selänne, Pyhäjärvi, Finland

[14]Department for Epidemiology and Preventive Medicine, Sackler Faculty of Medicine, Tel Aviv University, Tel Aviv, Israel

[15]Paul Langerhans Institute Dresden of the Helmholtz Center Munich at University Hospital and Faculty of Medicine, TU Dresden, Dresden, Germany

[16]German Center for Diabetes Research (DZD e.V.), Neuherberg, Germany

[17]Department of Clinical and Molecular Medicine, Sapienza University of Rome, Rome, Italy

[18]Department of Internal Medicine, Hokuriku Central Hospital, Toyama, Japan

[19]Lund University Diabetes Centre, Lund University, Malmö, Sweden

[20] Public Health Promotion Unit, Finnish Institute for Health and Welfare, Helsinki, Finland

[21]Department of Public Health, University of Helsinki, Helsinki, Finland

[22]Saudi Diabetes Research Group, King Abdulaziz University, Jeddah, Saudi Arabia

[23]Department of Public Health, Clinicum, University of Helsinki, Helsinki, Finland

[24]Broad Institute of the Massachusetts Institute of Technology and Harvard University, Cambridge, MA, USA

[25] Department of Medicine and of Population Health, Division of Endocrinology and Metabolism, NYU Langone Diabetes Prevention Program, NYU Grossman School of Medicine, New York, NY, USA

[26]Abdominal Centre, Endocrinology, Helsinki University Hospital; Folkhalsan Research Centre, Biomedicum, and Research Program Unit, Clinical and Molecular Medicine, University of Helsinki, Helsinki, Finland

Correspondence to: Dr. Vasudha Ahuja, Institute for Molecular Medicine Finland, Helsinki, University of Helsinki, PO Box 63, FI-00014, Finland. Email: vasudha.ahuja@helsinki.fi

Word count: 4144

Number of tables: two

Number of figures: two

Abstract

Objective One-hour plasma glucose (1-h PG) during the oral glucose tolerance test (OGTT) is an accurate predictor of type 2 diabetes. We performed a meta-analysis to determine the optimum cut-off of 1-h PG to detect type 2 diabetes using 2-h PG as the gold standard.

Research Design and Methods We included 15 studies with 35,551 participants from multiple ethnic groups (53.8% Caucasian) and 2705 newly detected cases of diabetes based on 2-h PG during OGTT. We excluded cases identified only by elevated fasting plasma glucose and/or HbA1c. We determined the optimal 1-h PG threshold and its accuracy at this cut-off to detect diabetes (2-h PG ≥ 11.1 mmol/L) using a mixed linear effects regression model with different weights to sensitivity/specificity (2/3, 1/2, and 1/3).

Results Three cut-offs of 1-h PG at 10.6 mmol/L, 11.6 mmol/L, and 12.5 mmol/L had sensitivities of 0.95, 0.92, and 0.87 and specificities of 0.86, 0.91, and 0.94 at weights 2/3, 1/2, and 1/3, respectively. The cut-off of 11.6 mmol/L (95% CI 10.6, 12.6) had a sensitivity 0.92 (0.87, 0.95), specificity of 0.91 (0.88, 0.93), AUC 0.939 (95% confidence region for sensitivity at a given specificity: 0.904, 0.946), and a positive predictive value of 45%.

Conclusions The 1-h PG of ≥ 11.6 mmol/L during OGTT has a good sensitivity and specificity for detecting type 2 diabetes. Prescreening with a diabetes-specific risk calculator to identify high-risk individuals is suggested to decrease the proportion of false-positive cases. Studies including other ethnic groups and assessing complication risk are warranted.

In 1979, the National Diabetes Data Group (NDDG) and the World Health Organization (WHO) established the current practice of diagnosing type 2 diabetes based on fasting and/or 2-hours threshold levels after a 75-g oral glucose tolerance test (OGTT) (1,2). The diagnostic criteria have since undergone two major changes by the WHO and American Diabetes Association (ADA), first, by lowering the diagnostic threshold of the fasting plasma glucose (FPG) from 7.8 mmol/L to 7.0 mmol/L in the late 1990s and second, by introducing HbA_{1c} as an additional diagnostic criterion in the late 2000s (3-6).

A similar consensus does not exist for diagnosing prediabetes, also referred to as intermediate hyperglycemia (IH). The WHO, ADA, and an ad hoc "International Expert Committee (IEC)" advocate different criteria to define IH based on the FPG (impaired fasting glucose or IFG), 2 hour plasma glucose (2-h PG) during the OGTT (impaired glucose tolerance or IGT), and/or HbA_{1c} (7). Nevertheless, multiple studies in various ethnicities have indicated that the one-hour plasma glucose (1-h PG) $\geq$ 8.6 mmol/L is a more accurate predictor of incident type 2 diabetes than IFG, IGT, HbA_{1C}, or their combination (7). Hence, an expert panel proposed a 1-h PG $\geq$ 8.6 mmol/L level to define IH (7). Since several studies have shown the association of the 1-h PG with cardiovascular disease and mortality and a better and independent association of post-challenge glucose concentration than FPG and HbA_{1c} with these outcomes, it is logical to evaluate the potential of the 1-h PG for the detection of type 2 diabetes (8-12).

Zhou et al and Paddock et al reported 1-h PG threshold values to detect type 2 diabetes in Chinese and American Indian populations, respectively (13,14). As the threshold could be affected by study design, differences in recruitment of participants, ethnicity as well other factors, we performed a meta-analysis of 15 studies comprising 35,551 participants derived from varied ethnicities to determine the optimum 1-h PG level equivalent to the 2-h PG $\geq$ 11.1 mmol/L diagnostic of type 2 diabetes.

Research Design and Methods

Search strategy and selection criteria

In this meta-analysis, principal investigators of 15 studies involved in cross-sectional or longitudinal studies from 1965 to the present with access to fasting, 1-h (index test) and 2-h (reference standard) PG data during an OGTT participated (10, 15-24). Cases included adults with newly detected type 2 diabetes defined as a 2-h PG ≥ 11.1 mmol/L during the OGTT. We excluded participants identified to have diabetes only based on FPG ≥ 7.0 mmol/L and/or HbA_{1C} ≥ 48.0 mmol/mol (6.5%, [≥ 43.0 (6.1%) mmol/mol for Japanese participants]) and/or who were on glucose-lowering medications (25). This is because we considered 2-h PG as the reference standard and including participants based on FPG and/or HbA1c criteria would have reduced the specificity of the 1-h PG to detect diabetes with a 2-h PG ≥ 11.1 mmol/L. Individuals without diabetes from the same cohorts (2-h PG < 11.1 mmol/L) constituted the control group.

Data analysis

In eight of the 15 studies included in the meta-analysis, analysts provided information on the study design, sample size, setting (e.g., primary health centers, diabetes clinics, population-based), recruitment procedures, percentage of women, mean age, mean 1-h PG, and percentage of diabetes cases based on the 2-h PG. Furthermore, they provided the numbers of true positives (TP), false positives (FP), true negatives (TN), false negatives (FN), and the 1-h PG cut-off to detect diabetes based with a 2-h PG ≥ 11.1 mmol/L **(Supplemental Material)**. We included 1-h PG thresholds at the maximum Youden's index and the minimum distance for each study, if they differed. The primary analyst (VA) performed the analyses using raw data in seven studies. Two authors (VA and MB) independently performed a quality assessment of the studies using the quality assessment of diagnostic accuracy studies-2 (QUADAS-2) tool through consensus **(Supplemental Material)** (26).

We constructed a forest plot displaying TP, FP, FN, TN, 1-h PG cut-offs, sensitivity, specificity, and their 95% CI of 1-h PG for each study using Review Manager 5.3. We further created a receiver operating characteristic (ROC) ellipse plot that depicts the estimate of each study with its 95% confidence region (CR) in the ROC area. Furthermore, we constructed a forest plot of 1-h PG showing the log diagnostic odds ratio (lnDOR) of each study with its summary estimate. In addition, we plotted a Fagan's nomogram that integrates prevalence, likelihood ratios (positive and negative likelihood ratio), and post-test probabilities (positive and negative predictive values [PPV and NPV]).

Having two cut-offs from each study and two outcomes as specificities and sensitivities make this a multi-level random effects model. To account for this structure, while meta-analyzing, we constructed a summary ROC curve (SROC) for the 1-h PG utilizing a class of weighted mixed linear effects regression model that modeled sensitivities and specificities separately for participants with and without diabetes across all studies considering fixed effects for studies, cut-offs, and their interactions and various random effects (27). Furthermore, we used three different lambdas while constructing the SROC curve in order to assign different weights to specificities and sensitivities: 1/2 that weighs specificity and sensitivity equally and thus resembles the maximum Youden's index, 2/3 enhancing sensitivity, and 1/3 emphasizing specificity. In addition, among models with the same fixed and different random effects, we chose the model that best described our data using the smallest restricted maximum likelihood (REML) criteria. Finally, to find the optimum cut-off of 1-h PG to detect 2-h PG ≥11.1 mmol/L on the chosen random effect model, we chose the lambda that provided the best combination of specificity and sensitivity.

We explored heterogeneity across the studies using meta-regression with the design of studies (cross-sectional vs. longitudinal), the setting of studies (diabetes clinic vs. population-based), the dose of glucose used for the OGTT (75 g vs. 100 g), ethnicity (South Asian, American Indian, Japanese, and Mexican American vs. Caucasian), and bias (studies with risk vs. low risk of bias) as

covariates. We also examined sample size-related effects by constructing a funnel plot and assessed its asymmetry using Deeks' test (28).

Finally, for the seven studies with available raw data, we performed certain sub-analyses. First, we restricted the analyses to the cases of diabetes with 2-h PG just above the diagnostic cut-off (≥ 11.1 to ≤ 13.0 mmol/L) both to increase the specificity of the analysis for that cut-off but also because they are likely to be of more recent onset than those with the 2-h PG ≥ 11.1 mmol/L. Second, we compared the cut-off obtained by meta-analyzing unadjusted cut-offs to the cut-off obtained by meta-analyzing age, sex, and BMI adjusted cut-offs in order to assess how these factors affect the cut-off of the 1-h PG. We utilized R version 3.6.3 for analyses unless mentioned otherwise.

Results

We included 15 studies with 35,551 participants representing Caucasian, American Indian, Japanese, Mexican American, and South Asian ethnicities (46.2% non-Caucasian). Four studies were longitudinal and 11 were cross-sectional; two were primary health care center-based, four diabetes clinic-based, and nine population-based. All but one study used a glucose dose of 75 g for the OGTT **(Table 1)**. The mean value of 1-h PG across studies ranged from 10.1- 18.5 mmol/L in individuals with and 7.4-9.2 mmol/L in those without diabetes. Of the newly detected cases of diabetes (N=3382), we excluded 677 (20.0%) who had diabetes based on FPG and/or HbA1c only and analyzed data for 2705 (80.4%) with diabetes based on a 2-h PG ≥ 11.1 mmol/ L: 1746 (51.6%) of them based on 2-h PG only and 959 (28.4%) based on both 2-h PG and FPG/HbA1c (**Supplemental Table S1**).

QUADAS-2 assessment showed a strong quality of evidence (**Supplemental Figure S1).** Eleven studies had low risk of bias or applicability concerns while two studies were at risk of bias in the domain of patient selection and two in the applicability concerns.

The forest plot shows that 1-h PG of 10.2-11.9 mmol/L had a sensitivity of 0.82-1.0 and specificity of 0.79-0.97 to detect a 2-h PG ≥11.1 mmol/L (**Figure 1**). In the ROC ellipse plot, the estimates from all studies positioned in the upper left portion of the ROC area demonstrating high diagnostic accuracy of the 1-h PG to detect a 2-h PG ≥11.1 mmol/L (**Figure 2**). The forest plot of lnDOR shows 4.6 times higher odds of obtaining a positive result using 1-h PG in individuals with than without diabetes (**Supplemental Figure S2**). The Fagan's nomogram indicates that the probability of having diabetes increases from 7% to 45% with a positive result. Furthermore, the probability of having diabetes drops from 7% to 1% with a negative result (**Supplemental Figure S3A**).

The selected different random slope (DS) model suggested a study-specific effect of 1-h PG on the accuracy to detect 2-h PG ≥ 11.1 mmol/L. **Supplemental Figure S4** shows three alternative cut-offs

at different lambda levels: 10.6 (95% CI: 10.0, 11.3) mmol/L at lambda 2/3 (higher sensitivity), 11.6 (10.6, 12.6) mmol/L at lambda 1/2 with equal weights for sensitivity and specificity (Youden's index), and 12.5 (11.3, 14.0) mmol/L at lambda 1/3 (higher specificity). At these cut-offs (10.6 vs. 11.6 vs. 12.5 mmol/L) the 1-h PG had sensitivity of 0.95 (0.91, 0.97) vs. 0.92 (0.87, 0.95) vs. 0.87 (0.79, 92) and specificity of 0.86 (0.82, 0.89) vs. 0.91 (0.88, 0.93) vs. 0.94 (0.92, 0.96), respectively. At all these cut-offs, the AUC of 1-h PG to detect 2-h PG ≥ 11.1 mmol/L was 0.939 (95% confidence region for sensitivity at given specificity: 0.904, 0.946). **Table 2** shows the numbers of TP, FN, FP TN, and PPV at these cut-offs. As expected, the number of FN increased, FP decreased, and PPV increased as cut-off levels of 1-h PG increased.

Cross-sectional studies are more likely to recruit long-standing undiagnosed cases of diabetes as "incident" compared to longitudinal studies, and clinic-based studies are likely to recruit more participants with IH compared to population-based studies. However, the meta-regression analysis did not show differences in sensitivity or specificity for the diagnostic accuracy of the 1-h PG to diagnose the 2-h PG ≥ 11.1 mmol/L when comparing cross-sectional and longitudinal studies (P=0.43 and 0.88, respectively) or diabetes clinic-based and population-based studies (P=0.58 and 0.46, respectively). Further, studies administering 75 g of glucose demonstrated no difference in diagnostic accuracy compared with the study using a 100 g dose (sensitivity P=0.88, specificity P=0.24). In addition, meta-regression analysis by ethnicity showed that American Indian had the highest sensitivity of 1-h PG followed by Japanese, Caucasians, South Asians, and Mexican Americans, respectively (P <0.0001). Again, American Indians had the highest specificity of 1-h PG, followed by Caucasians, Mexican Americans, Japanese, and South Asians, respectively (P <0.0001). Although studies with risk of bias demonstrated similar sensitivity to those with low risk of bias (P=0.19), they had lower specificity (P=0.001) (**Supplemental Table S2 & Supplemental Table S3**). Furthermore, the examination of the funnel plot using Deeks' test showed non-significant (P=0.21) results indicating the absence of sample size-related effects **(Supplemental Figure S5)**.

Using the raw data of seven studies in the sub-analysis restricted to detect diabetes with the 2h-PG ≥ 11.1 mmol/L to ≤13.0 mmol/L, presumed to have fairly recent duration, we found that the cut-off levels of the 1-h PG were higher than when individuals with 2-h PG ≥ 11.1 mmol/L were included: 12.6 mmol/L at lambda 2/3 (higher sensitivity), 13.5 mmol/L at lambda1/2 (Youden's index), and 14.5 mmol/L at lambda 1/3 (higher specificity). Furthermore, the sensitivity of the 1-h PG to detect diabetes with 2-h PG within ≥ 11.1 to ≤13.0 mmol/L was lower and specificity was higher than detecting diabetes with the 2-h PG ≥ 11.1 mmol/L **(Supplemental Table S4)**. Finally, we found that unadjusted cut-offs were either similar or lower than age, sex, and BMI adjusted cut-offs depending on the lambda used **(Supplemental Table S5)**.

Conclusion

In this meta-analysis of over 35,000 individuals across multiple ethnic groups, we demonstrate that the 1-h PG of 10.6 -12.5 mmol/L detects individuals with a 2-h PG level diagnostic of diabetes (≥11.1 mmol/L) with 87-95% sensitivity and 86-94% specificity. The choice of the 1-h PG cut-off depends on whether more weight is given to sensitivity or specificity. Thus, using the cut-off at the Youden's index (11.6 mmol/L) with sensitivity of 92% and specificity of 91%, the 1-h PG detected 2489 of 2705 (92%) cases of type 2 diabetes while missing 216 (8%). Whereas the 1-h PG correctly classified 31,164 of 32,246 (91%) individuals as not having diabetes, it classified as many as 3082 (9%) individuals who are non-diabetic by current criteria as having diabetes.

The OGTT is considered the "gold standard" for the diagnosis of diabetes despite having a large coefficient of variation and being inconvenient (29). It is noteworthy that it reflects the progressive failure of β-cell function, the primary phenomenon that drives the development of overt diabetes (30). While the clinical use of OGTT usually includes only fasting and 2-h PG levels, the deterioration of the insulin secretory response can be estimated from glucose and insulin concentrations either at 30 minutes post-challenge, as a proxy for first-phase insulin response, or at two hours reflecting both first-and second-phase insulin responses (31). Expectedly, 1-h PG, not currently measured during the OGTT, has a stronger correlation with the Matsuda index, the disposition index at 120 min, and glucose area under the curve than the 2-h PG (24). Considering that these proxy measurements of insulin secretion and insulin sensitivity are consistently lower in those that progress to type 2 diabetes, it is not surprising that the 1-h PG is also a more accurate predictor of progression to type 2 diabetes than IFG, IGT, and elevated HbA_{1C} (24, 32) .

When evaluating the 1-h PG for diagnosing diabetes, two approaches can be considered. We evaluated the 1-h PG level coincident with the 2-h PG diagnostic of diabetes (11.1 mmol/L). The

alternative and perhaps more biologically relevant approach would be to compare the 1-h PG with the 2-h PG value that best predicts diabetes-related complications. Regarding the approach described herein, the 1-h PG is convenient and strongly correlates with the 2-h PG. However, several factors affect the relationship between the 1-h and 2-h PG. Different pathogenic mechanisms in glucose-responsiveness and insulin secretion manifest in significantly different ratio of 1-h PG and 2-h PG in carriers of *GCK* and *HNF1A* mutations (33). Furthermore, the profile of the glucose response during the OGTT changes with progression from normoglycaemia to IGT to overt diabetes (30). Finally, glucose control may have an effect since chronic hyperglycaemia causes an insulin secretory defect that can result in different cut-off values in cohorts with recent-onset or long-standing diabetes (34). For this reason, in this study, we only included newly detected cases based on screening with OGTT and a 2-h PG value diagnostic of diabetes. This may partly explain the lower cut-off for the 1-h PG of 11.6 mmol/L in the present meta-analysis compared with 13.0 mmol/L in a Chinese hospital-based study (13).

The current diagnostic threshold values for diabetes originated based on the association of glycemic levels with increased prevalence of diabetic retinopathy, especially non-proliferative diabetic retinopathy (35). In this regard, the 1-h PG was significantly associated with prevalent and incident diabetic retinopathy in American Indians and with incident diabetic retinopathy in a Swedish cohort (8, 14). Furthermore, the 1-h PG was similarly associated with diabetic retinopathy as 2-h PG in the former population. Multiple studies have demonstrated an association of 1-h PG with cardiovascular outcomes and mortality (8-10). Moreover, among men without diabetes in the Malmö Preventive Project, the 1-h PG predicted cardiovascular death and all-cause mortality better than the 2-h PG (8).

As information relating to the presence of retinopathy or cardiovascular disease was not available for the cohorts analyzed in this meta-analysis, we could not evaluate the cut-off 1-h PG that would best

detect diabetic complications. However, Paddock et al, in a cross-sectional analysis, identified a 1-h PG threshold of 12.0 mmol/L for diagnosing type 2 diabetes in American Indians with retinopathy, comparable to 11.6 mmol/L in the present meta-analysis (14). In the same study, the cut-off based on a longitudinal analysis was 12.8 mmol/L, again similar to the 12.6 mmol/L at lambda 2/3 in our study when we restricted the analysis to presumably more recently diagnosed type 2 diabetes (14). Ideally, the comparisons should include information regarding the distribution of values as the results will differ substantially if the majority of individuals have 2-h PG near the cut-off of 11.1 mmol/L or much higher.

A meta-analysis only enables utilization of aggregate measures, e.g. proportion of females, therefore, assessing differences in diagnostic accuracy according to participant level variables may introduce bias. Nevertheless, using raw data from the available seven studies, we explored how these demographic factors affect the diagnostic accuracy of 1-h PG. First, we found that the unadjusted and age, sex, and BMI adjusted cut-offs of 1-h PG were significantly different in five out of seven studies **(Supplemental Table S6)**. Second, at meta-analytical level, we found the meta-analyzed unadjusted estimates to be either similar to or lower than meta-analyzed adjusted cut-offs **(Supplemental Table S5)**. Additionally, the cut-offs differed minimally according to ethnicities as did their sensitivities and specificities except for American Indian where the cut-off was lower and sensitivity and specificity were higher than in other groups, which may be due to exceptionally high-risk of type 2 diabetes in this population (36). Moreover, universal diagnostic cut-off values to diagnose diabetes apply for all glycaemic indices irrespective of age, sex, BMI, and ethnicity. This is true despite reported distinctive values of these indices in individuals of different demographic characteristics without diabetes (37). Thus, in line with the current universal diagnostic threshold values, we suggest utilizing the same cut-off value of 1-hPG to detect type 2 diabetes among different groups.

The very first criteria for usefulness of a diagnostic test is its ability to discriminate between individuals with and without disease, i.e. the sensitivity and specificity. These are adequately high for the 1-h PG of 11.6 mmol/L using the 2-h PG for defining disease status. However, 55% of individuals classified as having diabetes by this 1-h PG were non-diabetic according to the 2-h PG (**Table 2)**. Although the sensitivity and specificity are not mathematically dependent on prevalence, the number of FN increases and that of FP decreases as prevalence increases. Consequently, using the Genetic PHYsiopathology, and Evolution of Type 2 Diabetes study with a high prevalence of diabetes (27%) instead of all cohorts of the meta-analysis combined having a lower prevalence (7%) increased the PPV from 45% to 79% (and decreased the number of FP from 55% to 21%; **Supplemental Figure 3B**) (19).

Some aspects favor using a higher cut-off of 12.5 mmol/L. First, especially in populations with a low prevalence, a higher cut-off would be needed to increase the PPV; using 12.5 mmol/L instead of 11.6 mmol/L in the meta-analysis increased the PPV from 45% to 53% (**Table 2).** Second, in the sub-analysis of individuals presumably having more recent onset diabetes (2h-PG < 13 mmol/L), the cut-off of 1-h PG was higher > 12.5 mmol/L **(Supplemental Table S4)**. On the other hand, a large proportion of the FP individuals had 2-h PG values just below the current diagnostic cut-off for diabetes, and the majority (59%) had IGT **(Supplemental Figure S6)**. Previous studies have reported that ~8% of people with IGT in the US DPP study had evidence of diabetic retinopathy, suggesting a significant false negative rate for the FPG and 2-h PG criteria thereby underestimating detection of dysglycemia (38). Furthermore, the 1-h PG has a stronger association with cardiovascular outcome and all-cause mortality than the 2-h PG (8). Thus, we hypothesize that the so-called FP cases, i.e. having diabetes based on the 1-h PG but not the 2-hPG, may actually turn out to be true positive cases regarding high risk of complications and serve as a target group for prevention.

It needs to be stressed, that we are not proposing that the OGTT be performed as the initial screening test for type 2 diabetes (or prediabetes) as this would be highly infeasible and costly. In accordance with others, we advocate implementation of validated diabetes risk screening calculators (e.g. Finnish Diabetes Risk Score, ADA) to identify high-risk individuals (39). Further laboratory measurements would only be instituted in those identified as high-risk based on the outcome of the screening calculator. The diagnosis of diabetes would be confirmed with a second test as recommended by WHO and the ADA (40). With this procedure, the proportion of FP cases and high-risk individuals would likely be reduced that otherwise might incorrectly have been suggested to have diabetes. Furthermore, individuals that have been positively screened may have on-going abnormalities in glucose regulation and therefore still remain at high-risk for developing diabetes in the future, may benefit from lifestyle modification.

The strength of this meta-analysis resides in its size, comprising approximately 35,000 participants and diversity including populations from different countries. As we obtained raw estimates from studies in contrast to extracting published data, we achieved uniformity in defining type 2 diabetes and obtained complete information to assess the quality of studies. Two of the studies reported herein may have had volunteer-bias due to convenient sampling and two studies had significant loss-to-follow-up, which would have resulted in increased proportions of type cases and decreased specificity in these studies compared to others. Overall, the quality of evidence was strong. The meta-analysis also has certain weaknesses. The number of studies included is small. Although we included studies having participants with different ethnic backgrounds, major ethnic groups such as of African or South American origin were missing. Moreover, it is ideal to choose diagnostic thresholds using incident cases of diabetes from population-based longitudinal studies as differences in the characteristics of participants in non-population-based and cross-sectional studies might affect the

accuracy of a test. While we did not find significant differences in accuracy of the 1-h PG to detect type between longitudinal and cross-sectional studies or between population and non-population-based studies, a higher 1-h PG cut-off was obtained in the subgroup with presumably more recent onset (2-h PG < 13 mmol/L). Although examination of the funnel plot showed non-significant results, it displays an asymmetry that may point to a significant sample size-related effect. Here, a non-significant Deeks' test might reflect its low power in case of heterogeneous DOR. Of note, the presence of sample size-related effects may not only reflect the possibility of publication bias (in this meta-analysis would rather reflect sampling bias) but also the relation of the sample size of the studies to the type of study population or study quality. However, after excluding the studies that stood apart in Deeks' funnel plot, the cut-off of 1-h PG was similar with little change in sensitivity and specificity (11.3 mmol/L [0.91, 0.89]; data not shown).

In summary, a 1-h PG of 11.6 mmol/L detected the 2-h PG ≥ 11.1 mmol/L diagnostic of type 2 diabetes with high sensitivity and specificity among adults previously undiagnosed with diabetes but detected a high proportion of FP cases . At least three aspects warrant further research including other ethnic backgrounds. First, we suggest reproducibility studies of 1-h PG compared to 2-h PG in populations other than in American Indians in whom it is poorer and the distribution bimodal in contrast to most populations (41). Second, we recommend population-based longitudinal studies comparing the strength of association of the 1-h PG and 2-h PG with diabetic retinopathy and other microvascular complications, cardiovascular complications, and all-cause mortality.

1. Classification and diagnosis of diabetes mellitus and other categories of glucose intolerance. National Diabetes Data Group. Diabetes. 1979;28(12):1039-57.
2. World Health Organization. WHO Expert Committee on Diabetes Mellitus [meeting held in Geneva from 25 September to 1 October 1979] : second report. Geneva: WHO Press, 1980.
3. Report of the Expert Committee on the Diagnosis and Classification of Diabetes Mellitus. Diabetes Care. 1997;20(7):1183-97.
4. World Health Organization. Definition, diagnosis and classification of diabetes mellitus and its complications : report of a WHO consultation. Part 1, Diagnosis and classification of diabetes mellitus. Geneva: WHO Press, 1999.
5. Gillett MJ. International Expert Committee report on the role of the A1c assay in the diagnosis of diabetes: Diabetes Care 2009; 32(7): 1327-1334. Clin Biochem Rev. 2009;30(4):197-200.
6. World Helath Orgnaization, ed. Use of glycated haemoglobin (HbA1c) in the diagnosis of diabetes mellitus: abbreviated report of WHO consulatation. Geneva: WHO Press, 2011.
7. Bergman M, Manco M, Sesti G, Dankner R, Pareek M, Jagannathan R, et al. Petition to replace current OGTT criteria for diagnosing prediabetes with the 1-hour post-load plasma glucose>/=155mg/dl (8.6mmol/L). Diabetes Res Clin Pract. 2018;146:18-33.
8. Pareek M, Bhatt DL, Nielsen ML, Jagannathan R, Eriksson KF, Nilsson PM, et al. Enhanced Predictive Capability of a 1-Hour Oral Glucose Tolerance Test: A Prospective Population-Based Cohort Study. Diabetes Care. 2018;41(1):171-7.
9. Lind M, Tuomilehto J, Uusitupa M, Nerman O, Eriksson J, Ilanne-Parikka P, et al. The association between HbA1c, fasting glucose, 1-hour glucose and 2-hour glucose during an oral glucose tolerance test and cardiovascular disease in individuals with elevated risk for diabetes. PLoS One. 2014;9(10):e109506.
10. Bergman M, Chetrit A, Roth J, Dankner R. One-hour post-load plasma glucose level during the OGTT predicts mortality: observations from the Israel Study of Glucose Intolerance, Obesity and Hypertension. Diabet Med. 2016;33(8):1060-6.
11. Glucose tolerance and mortality: comparison of WHO and American Diabetes Association diagnostic criteria. The DECODE study group. European Diabetes Epidemiology Group. Diabetes Epidemiology: Collaborative analysis Of Diagnostic criteria in Europe. Lancet. 1999;354(9179):617-21.
12. Meigs JB, Nathan DM, D'Agostino RB, Sr., Wilson PW. Fasting and postchallenge glycemia and cardiovascular disease risk: the Framingham Offspring Study. Diabetes Care. 2002;25(10):1845-50.
13. Zhou W, Gu Y, Li H, Luo M. Assessing 1-h plasma glucose and shape of the glucose curve during oral glucose tolerance test. Eur J Endocrinol. 2006;155(1):191-7.
14. Paddock E, Looker HC, Piaggi P, Knowler WC, Krakoff J, Chang DC. One-Hour Plasma Glucose Compared With Two-Hour Plasma Glucose in Relation to Diabetic Retinopathy in American Indians. Diabetes Care. 2018;41(6):1212-7.
15. Tripathy D, Carlsson M, Almgren P, Isomaa B, Taskinen MR, Tuomi T, et al. Insulin secretion and insulin sensitivity in relation to glucose tolerance: lessons from the Botnia Study. Diabetes. 2000;49(6):975-80.
16. Succurro E, Arturi F, Caruso V, Rudi S, Sciacqua A, Andreozzi F, et al. Low insulin-like growth factor-1 levels are associated with anaemia in adult non-diabetic subjects. Thromb Haemost. 2011;105(2):365-70.
17. Abdul-Ghani MA, Abdul-Ghani T, Muller G, Bergmann A, Fischer S, Bornstein S, et al. Role of glycated hemoglobin in the prediction of future risk of T2DM. J Clin Endocrinol Metab. 2011;96(8):2596-600.
18. La Sala L, Mrakic-Sposta S, Tagliabue E, Prattichizzo F, Micheloni S, Sangalli E, et al. Circulating microRNA-21 is an early predictor of ROS-mediated damage in subjects with high risk of developing diabetes and in drug-naive T2D. Cardiovasc Diabetol. 2019;18(1):18.
19. Bianchi C, Miccoli R, Bonadonna RC, Giorgino F, Frontoni S, Faloia E, et al. Pathogenetic mechanisms and cardiovascular risk: differences between HbA(1c) and oral glucose tolerance test for the diagnosis of glucose tolerance. Diabetes Care. 2012;35(12):2607-12.
20. Pyorala M, Miettinen H, Laakso M, Pyorala K. Hyperinsulinemia predicts coronary heart disease risk in healthy middle-aged men: the 22-year follow-up results of the Helsinki Policemen Study. Circulation. 1998;98(5):398-404.

21. Sai Prasanna N, Amutha A, Pramodkumar TA, Anjana RM, Venkatesan U, Priya M, et al. The 1h post glucose value best predicts future dysglycemia among normal glucose tolerance subjects. J Diabetes Complications. 2017;31(11):1592-6.
22. Mutt SJ, Jokelainen J, Sebert S, Auvinen J, Jarvelin MR, Keinanen-Kiukaanniemi S, et al. Vitamin D Status and Components of Metabolic Syndrome in Older Subjects from Northern Finland (Latitude 65 degrees North). Nutrients. 2019;11(6).
23. Oka R, Yagi K, Sakurai M, Nakamura K, Moriuchi T, Miyamoto S, et al. Insulin secretion and insulin sensitivity on the oral glucose tolerance test (OGTT) in middle-aged Japanese. Endocr J. 2012;59(1):55-64.
24. Abdul-Ghani MA, Williams K, DeFronzo RA, Stern M. What is the best predictor of future type 2 diabetes? Diabetes Care. 2007;30(6):1544-8.
25. Kashiwagi A, Kadowaki T, Haneda M, Nawata S, Itoh H, Tominaga M, et al. Consensus and Statement on International Standardization of HbA1C in Japan : Committee Report on Diabetes Mellitus Laboratory Testing Standardization Journal of the Japan Diabetes Society. 2009;52(9):811-8.
26. Whiting PF, Rutjes AW, Westwood ME, Mallett S, Deeks JJ, Reitsma JB, et al. QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med. 2011;155(8):529-36.
27. Steinhauser S, Schumacher M, Rucker G. Modelling multiple thresholds in meta-analysis of diagnostic test accuracy studies. BMC Med Res Methodol. 2016;16(1):97.
28. Deeks JJ, Macaskill P, Irwig L. The performance of tests of publication bias and other sample size effects in systematic reviews of diagnostic test accuracy was assessed. J Clin Epidemiol. 2005;58(9):882-93.
29. World Health Orgnaization, International Diabetes Federation, eds. Definition and diagnosis of diabetes mellitus and intermediate hyperglycaemia: report of a WHO/IDF consultation. Geneva: WHO Press, 2006.
30. Bonadonna RC, Boselli L, Dei Cas A, Trombetta M. Methods to Assess In Vivo Insulin Sensitivity and Insulin Secretion. In: Bonora E, DeFronzo RA, editors. Diabetes Epidemiology, Genetics, Pathogenesis, Diagnosis, Prevention, and Treatment (Endocrinology). Switzerland: Springer; 2018.
31. Kettunen JLT, Tuomi T. Human Physiology of Genetic Defects Causing Beta-cell Dysfunction. J Mol Biol. 2020;432(5):1579-98.
32. Peddinti G, Bergman M, Tuomi T, Groop L. 1-Hour Post-OGTT Glucose Improves the Early Prediction of Type 2 Diabetes by Clinical and Metabolic Markers. J Clin Endocrinol Metab. 2019;104(4):1131-40.
33. Stride A, Vaxillaire M, Tuomi T, Barbetti F, Njolstad PR, Hansen T, et al. The genetic abnormality in the beta cell determines the response to an oral glucose load. Diabetologia. 2002;45(3):427-35.
34. White MG, Shaw JA, Taylor R. Type 2 Diabetes: The Pathologic Basis of Reversible beta-Cell Dysfunction. Diabetes Care. 2016;39(11):2080-8.
35. International Expert Committee Report on the Role of the A1C Assay in the Diagnosis of Diabetes. Diabetes Care. 2009;32(7):1327-34.
36. Knowler WC, Pettitt DJ, Saad MF, Bennett PH. Diabetes mellitus in the pima indians: Incidence, risk factors and pathogenesis. Diabetes/Metabolism Reviews. 1990;6(1):1-27.
37. Cavagnolli G, Pimentel AL, Freitas PA, Gross JL, Camargo JL. Effect of ethnicity on HbA1c levels in individuals without diabetes: Systematic review and meta-analysis. PLoS One. 2017;12(2):e0171315.
38. Diabetes Prevention Program Research G. The prevalence of retinopathy in impaired glucose tolerance and recent-onset diabetes in the Diabetes Prevention Program. Diabetic medicine : a journal of the British Diabetic Association. 2007;24(2):137-44.
39. Nowak C, Ingelsson E, Fall T. Use of type 2 diabetes risk scores in clinical practice: a call for action. The Lancet Diabetes & Endocrinology. 2015;3(3):166-7.
40. Zhang Y, Hu G, Zhang L, Mayo R, Chen L. A Novel Testing Model for Opportunistic Screening of Pre-Diabetes and Diabetes among U.S. Adults. PLOS ONE. 2015;10(3):e0120382.
41. Rushforth NB, Bennett PH, Steinberg AG, Miller M. Comparison of the value of the two- and one-hour glucose levels of the oral GTT in the diagnosis of diabetes in Pima Indians. Diabetes. 1975;24(6):538-46.

Acknowledgements

Contributions

VA, TT, and MB designed the meta-analysis. LG, TT, GS, ACeriello, PS, SDP, AChetrit, RDankner, VM, RO, and SKM MAG, RDeFronzo designed the individual studies. GS, TVF, PS, ACeriello, LLS, SDP, AChetrit, RDankner, TAPK, SJ, WK, PB, RO, MAG, and RDeFronzo contributed to data collection. VA and MB performed the risk-of-bias analysis. VA, PA, LLS, CB, AJ, AChetrit, RDankner, HL, RP, UV, VB, and AB performed the analyses of the individual studies. VA and PA performed the meta-analysis supervised by SR and TT. VA, PA, TT, and MB wrote the report. All authors contributed to data interpretation and revised the report.

Guarantors

Vasudha Ahuja and Tiinamaija Tuomi are the guarantors of the work

Funding

VA was funded through a grant from the Academy of Finland (grant no 312072). The Botnia studies (LG, TT) have been financially supported by grants from Folkhälsan Research Foundation, the Sigrid Juselius Foundation, The Academy of Finland (grants no. 263401, 267882, 312063 to LG, 312072 to TT), University of Helsinki, Nordic Center of Excellence in Disease Genetics, EU (EXGENESIS, MOSAIC FP7-600914), Ollqvist Foundation, Swedish Cultural Foundation in Finland, Finnish Diabetes Research Foundation, Foundation for Life and Health in Finland, Signe and Ane Gyllenberg Foundation, Finnish Medical Society, Paavo Nurmi Foundation, State Research Funding via the Helsinki University Hospital, Perklén Foundation, Närpes Health Care Foundation and Ahokas Foundation. The study has also been supported by the Ministry of Education in Finland, Municipal Heath Care Center and Hospital in Jakobstad and Health Care Centers in Vasa, Närpes and Korsholm. The Botnia Study has received research funding from Pfizer Inc. and Regeneron Pharmaceuticals. The research leading to these results has received funding from the European Research Council under the European Union's Seventh Framework Programme (FP7/2007-2013) / ERC grant agreement no 269045. TVF was supported by the European Foundation for the Study of Diabetes. DIAPASON study was supported by the Fondazione Romeo ad Enrica Invernizzi – Milano, EFSD/Sanofi programme (to LLS), and Italian Ministry of Health-Ricerca Corrente. The GENFIEV study has been supported by FoRiSID, Rome, Italy and with an unconditional grant from Eli Lilly, Italy. The GOH study was funded by The DCURE Foundation. Oulu45 and Oulu45P studies were funded by the Finnish Cultural Foundation - Central fund. The PIBS study was funded by the Division of Intramural Research of NIDDK, NIH. The PSW and PSWP study were funded by the Japan Health Promotion Foundation and Grant-in-Aid from Toyama Medical Association.

Conflict of interest

ACeriello receives research support from Mitsubishi; is a member of the advisory boards of Abbott, BD, Eli Lilly, Janssen, and Mundipharma; and is a member of the speakers' bureaus of Astra Zeneca, Berlin Chemie, Boehringer Ingelheim, Novo Nordisk, and Roche Diagnostics. CB receives honoraria for consulting from Novonordisk and research support from Roche Diagnostic. JT received research grants from Bayer Pharma, Boehringer Ingelheim, Merck, and Sanofi, received consulting and travel fees from Eli Lilly, Merck, MSD, Novo Nordisk, and owns stocks of Orion Pharma.VA, AP, TAPK, LH, AC, AB, AJ, LLS, RMA, RP, UV, SJ, VB, VTF, PT, RDF, SDP, MAG, SKK, RD, PB, WK, PS, GS, OR, VM, LG, SR, MB, and TT have no potential conflict of interest relevant to this article.

Table 1: Characteristics of included studies

Study*	Design	Setting	Ethnicity	N (% females)	Age at baseline (mean ± SD)	Glucose dose in OGTT (g)	1-h PG mmol/L (mean ± SD)	Type 2 diabetes cases with 2-h PG ≥11.1 mmol/L
BFS 1990[15†‡]	cross-sectional	primary health care	Caucasian	2995 (55%)	46.2 ± 13.7	75	7.9 ±2.7	126 (4.2%)
BPS 1990[16†‡]	longitudinal	primary health care	Caucasian	3168 (55%)	54.0 ± 14.7	75	8.0 ± 2.7	85 (2.7%)
CATAMERI 2005[17]	cross-sectional	diabetes-clinic	Caucasian	3324 (54%)	48.4 ± 13.9	75	8.8 ±2.7	249 (7.5%)
DIAGEN 1996[18‡]	cross-sectional	population	Caucasian	2679 (56%)	52.6 ± 16.5	75	9.2 ± 2.9	204 (7.6%)
DIAPASON 2014[19‡]	cross-sectional	diabetes-clinic	Caucasian	531 (57%)	59.4 ±9.9	75	8.4 ± 2.6	34 (6.4%)
GENFIEV 2003[20]	cross-sectional	diabetes-clinic	Caucasian	916 (57%)	49.3 ± 11.3	75	9.8 ± 2.8	116 (26.6)
GOH 1979[10]	cross-sectional	population	Caucasian	2092 (48%)	51.3 ± 8.0	100	8.6 ± 3.2	149 (7.1%)
HPS 1966[21†]	cross-sectional	population	Caucasian	1026 (0%)	44.0 ± 7.7	75	7.1 ± 2.0	11 (0.9%)
MDRF 1991[22]	cross-sectional	diabetes-clinic	South Asian	9651 (45%)	45.0 ± 12.0	75	9.4 ± 2.5	802 (8.3%)
Oulu45 2001[23†]	cross-sectional	population	Caucasian	933 (56%)	56.8 ± 0.6	75	8.6 ± 2.3	33 (3.6%)
Oulu45 2001[23†]	longitudinal	population	Caucasian	825 (58%)	56.8 ± 0.6	75	8.0 ± 1.9	44 (5.3%)
PIBS 1966[14]	longitudinal	population	American Indian	2664 (50%)	32.2 ± 15.1	75	8.2 ± 4.1	399 (15.1%)
PSW 2006[24‡]	cross-sectional	population	Japanese	2085 (32%)	52.6 ± 7.2	75	8.5 ± 2.6	70 (3.4%)
PSWP 2006[25‡]	longitudinal	population	Japanese	1997 (28%)	52.4 ± 6.9	75	8.6 ± 2.7	65 (3.23%)
SAHS 1992[26]	cross-sectional	population	Mexican American	689 (66%)	49.8 ± 12.1	75	12.1 ± 4.2	318 (46.2%)

*Studies with their initiation year; †Blood glucose converted to plasma glucose using a conversion factor of 1.13; ‡Studies with data for $HbA1_c$

N, numbers; SD, standard deviation; OGTT, oral glucose tolerance test; 1-hPG, one-hour plasma glucose; 2-hPG, two-hour plasma glucose; BFS, Botnia Family Study; BPS, Botnia Prospective Study; CATAMERI, CATAnzaro MEtabolic RIsk factors; DIAGEN, DIAbetes GENetic study; GENFIEV, Genetic PHYsiopathology, and Evolution of Type 2 Diabetes; DIAPASON, Diabetes Prediction and Screening Observational Study; GOH, Israel Study of Glucose Intolerance, Obesity and Hypertension; HPS, Helsinki Policemen Study; MDRF, Madras Diabetes Research Foundation study; Oulu45P, Oulu45 Prospective; PIBS, Pima Indian Biennial Study; PSW, Public School Worker; PSWP, Public School Worker Prospective; SAHS, San Antonio Heart Study

Table 2. The number of true and false positive (TP, FP) or negative (TN, FN) cases with three different one-hour plasma glucose cut-offs to diagnose type 2 diabetes of 2-h PG ≥ 11.1 mmol/L and the associated PPV

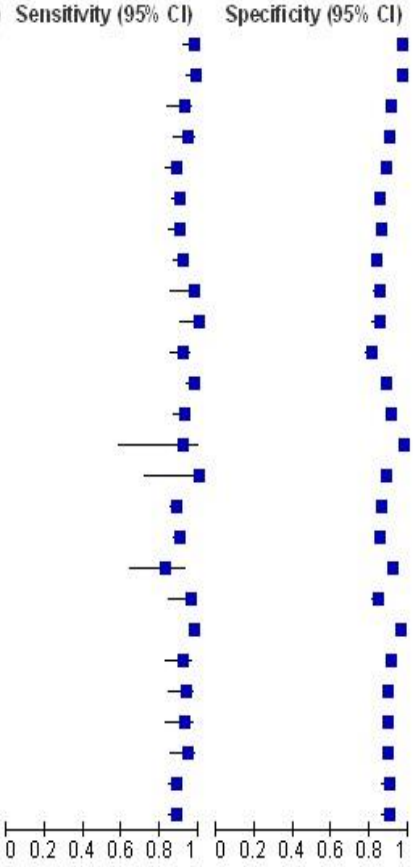
			Cases of type 2 diabetes by 2-h PG (N=2705)		Controls by 2-h PG (N =34,246)		PPV %
Cut-off in mmol/L (Se, Sp)	**Weight ratio for Se vs. Sp**	**lambda (λ)**	**TP**	**FN**	**FP**	**TN**	**FP/TP + FP**
10.6 (0.95, 0.86)	more	2/3	2570	135	4794	29,452	34.9
11.6 (0.92, 0.91)	equal	1/2	2489	216	3082	31,164	44.6
12.5 (0.87, 0.94)	less	1/3	2353	352	2055	32,191	53.4

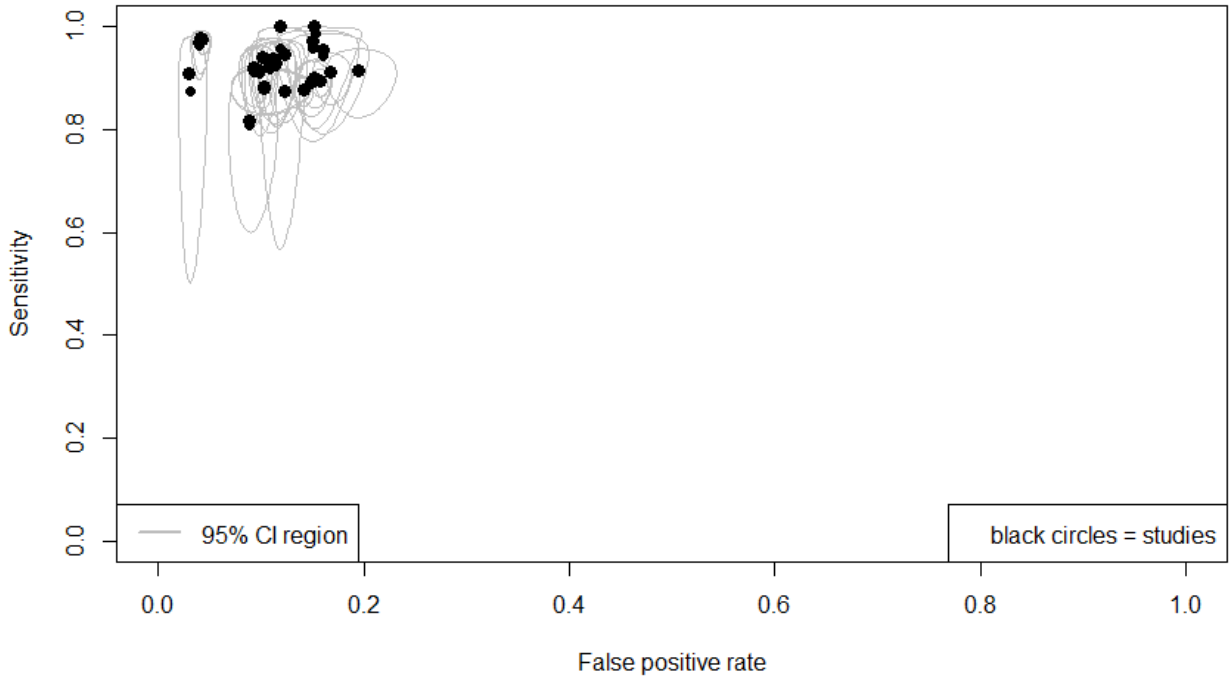
2-hPG, 2-h plasma glucose; N, numbers; Se, sensitivity; Sp, specificity; PPV, positive predictive value

Figure 1. A forest plot showing the sensitivity and specificity of the obtained 1-hour plasma glucose cut-offs to detect diabetes (defined as 2-h glucose ≥11.1 mmol/L) in the individual studies together with the number of true positive (TP); false positive (FP); false negative (FN) and true negative (TN) participants.
_M after the study name indicates the cut-off at the minimum distance, and_Y at the Youden index (in case of no postfix the cut-off is the same at the minimum distance and Youden index). BFS; Botnia Family Study; BPS, Botnia Prospective Study; CATAMERI, CATAnzaro MEtabolic RIsk factors; DIAGEN, DIAbetes GENetic study; GENFIEV, Genetic PHYsiopathology, and Evolution of Type 2 Diabetes; DIAPASON, Diabetes Prediction and Screening Observational Study; GOH, Israel Study of Glucose Intolerance, Obesity and Hypertension; HPS, Helsinki Policemen Study; MDRF, Madras Diabetes Research Foundation study; Oulu45P, Oulu45 Prospective; PIBS, Pima Indian Biennial Study; PSW, Public School Worker; PSWP, Public School Worker Prospective; SAHS, San Antonio Heart Study

Figure 2. ROC ellipse plot showing the cut-offs of studies with their 95% confidence regions.

Study	TP	FP	FN	TN	Cut-off	Sensitivity (95% CI)	Specificity (95% CI)
BFS_M 1990	122	112	4	2717	11.9	0.97 [0.92, 0.99]	0.96 [0.95, 0.97]
BFS_Y 1990	123	116	3	2713	11.9	0.98 [0.93, 1.00]	0.96 [0.95, 0.97]
BPS_M 1990	78	287	7	2796	11.4	0.92 [0.84, 0.97]	0.91 [0.90, 0.92]
BPS_Y 1990	80	313	5	2770	11.2	0.94 [0.87, 0.98]	0.90 [0.89, 0.91]
CATAMERI_M 2005	218	378	31	2696	11.3	0.88 [0.83, 0.91]	0.88 [0.86, 0.89]
CATAMERI_Y 2005	224	469	25	2605	10.8	0.90 [0.86, 0.93]	0.85 [0.83, 0.86]
DIAGEN_M 1996	182	369	22	2106	11.4	0.89 [0.84, 0.93]	0.85 [0.84, 0.86]
DIAGEN_Y 1996	186	416	18	2059	11.2	0.91 [0.86, 0.95]	0.83 [0.82, 0.85]
DIAPASON_M 2014	33	71	1	403	10.7	0.97 [0.85, 1.00]	0.85 [0.81, 0.88]
DIAPASON_Y 2014	34	72	0	402	10.6	1.00 [0.90, 1.00]	0.85 [0.81, 0.88]
GENFIEV 2003	106	156	10	644	11.1	0.91 [0.85, 0.96]	0.81 [0.78, 0.83]
GOH_M 1979	141	239	4	1704	11.4	0.97 [0.93, 0.99]	0.88 [0.86, 0.89]
GOH_Y 1979	137	181	12	1762	11.4	0.92 [0.86, 0.96]	0.91 [0.89, 0.92]
HPS_M 1966	10	31	1	984	12.3	0.91 [0.59, 1.00]	0.97 [0.96, 0.98]
HPS_Y 1966	11	120	0	895	10.2	1.00 [0.72, 1.00]	0.88 [0.86, 0.90]
MDFR_M 1991	703	1254	99	7595	11.7	0.88 [0.85, 0.90]	0.86 [0.85, 0.87]
MDRF_Y 1991	717	1395	85	7454	11.5	0.89 [0.87, 0.91]	0.84 [0.83, 0.85]
Oulu45 2001	27	78	6	799	11.3	0.82 [0.65, 0.93]	0.91 [0.89, 0.93]
Oulu45P 2001	42	83	2	434	11.3	0.95 [0.85, 0.99]	0.84 [0.80, 0.87]
PIBS 1966	389	210	10	4640	10.8	0.97 [0.95, 0.99]	0.96 [0.95, 0.96]
PSW_M 2006	64	199	6	1816	11.8	0.91 [0.82, 0.97]	0.90 [0.89, 0.91]
PSW_Y 2006	65	231	5	1784	11.2	0.93 [0.84, 0.98]	0.89 [0.87, 0.90]
PSWP_M 2006	60	117	5	956	11.4	0.92 [0.83, 0.97]	0.89 [0.87, 0.91]
PSWP_Y 2006	61	121	4	952	11.3	0.94 [0.85, 0.98]	0.89 [0.87, 0.91]
SAHS_M 1992	281	38	38	333	11.8	0.88 [0.84, 0.91]	0.90 [0.86, 0.93]
SAHS_Y 1992	280	38	38	333	11.8	0.88 [0.84, 0.91]	0.90 [0.86, 0.93]




Sensitivity
0.0
0.2
0.4
0.6
0.8
1.0
95% CI region
black circles = studies
0.0
0.2
0.4
0.6
0.8
1.0
False positive rate

Online-Only Supplemental Material

Approaches

Data analysis

Analyses of individual studies

Cross-sectional studies

We assessed the ability of the one-hour plasma (1-hPG) alone or in combination with covariates such as age, sex, ethnicity, and BMI to detect a two-hour plasma glucose (2-h PG) ≥ 11.1 mmol/L using logistic regression analyses. For Botnia Family Study (BFS) and Botnia Prospective Study (BPS), we used the Huber-White method to adjust the variance-covariance matrix for correlated errors. We considered the thresholds of 1-h PG to detect (2-h PG) ≥ 11.1 mmol/L at the maximum Youden's index and the minimum distance for each study, if they differed. We used Receiver Operator Characteristic (ROC) curve analyses to assess the ability of the 1-h PG to discriminate between cases of type 2 diabetes and non-cases. On a ROC curve, maximum Youden's index is the maximum vertical height above the chance line and the minimum distance is the point from the left-upper corner of the unit square. Further, at these indices, the sensitivity and specificity of a test are equal. We performed bootstrapping resampling to validate our model in order to prevent over-interpretation of the study data. We utilized R version 3.6.3 for the analyses of following studies; BFS, BPS, CATAnzaro MEtabolic RIsk factors, DIAbetes GENetic, Public School Worker, Public School Worker Prospective, and San Antonio Heart Study. Additionally, we used SAS version 9.4 for following studies; Diabetes Prediction and Screening Observational Study, Israel Study of Glucose Intolerance, Obesity and Hypertension, Madras Diabetes Research Foundation, and Pima Indian Biennial Study. Further, we employed SPSS version 23 for Genetic Physiopathology and Evolution of Type 2 Diabetes, SPSS version 25 for Helsinki Policemen Study, and SPSS version 24 for Oulu45.

Longitudinal studies

For participants with diabetes, we considered the value of 1-h PG at the first visit when they had a 2-h PG ≥11.1 mmol/L. In addition, for persons without diabetes, we considered the 1-h PG value at the last visit.

Meta-analysis

Quality assessment of diagnostic accuracy studies-2 (QUADAS-2) is a standardized evidence-based tool to assess quality of primary diagnostic accuracy studies. We applied it to this meta-analysis to understand how differences in design and conduct of studies might affect the accuracy of 1-h PG. It assesses two major areas, risk of bias and concerns regarding applicability. The "risk of bias" assesses the degree to which the estimates of diagnostic accuracy avoided risk of bias that might occur due to systemic flaws in the design and conduct of studies. The "concerns regarding applicability" assess the extent to which studies are applicable to the research question, e.g. regarding clinical and demographic features, the definition of target condition etc. It has four key domains: patient selection, index test, reference standard, and flow and timing. Patient selection aims to assess how studies recruited participants and their demographic and clinical characteristics, index test the conduct and interpretation of index test, reference standard the conduct and interpretation of reference test, and flow and timing the difference in the number of participants recruited to the number used in analyses. Each study underwent assessment of every domain in terms of risk of bias and the first three domains in terms of concerns regarding applicability.

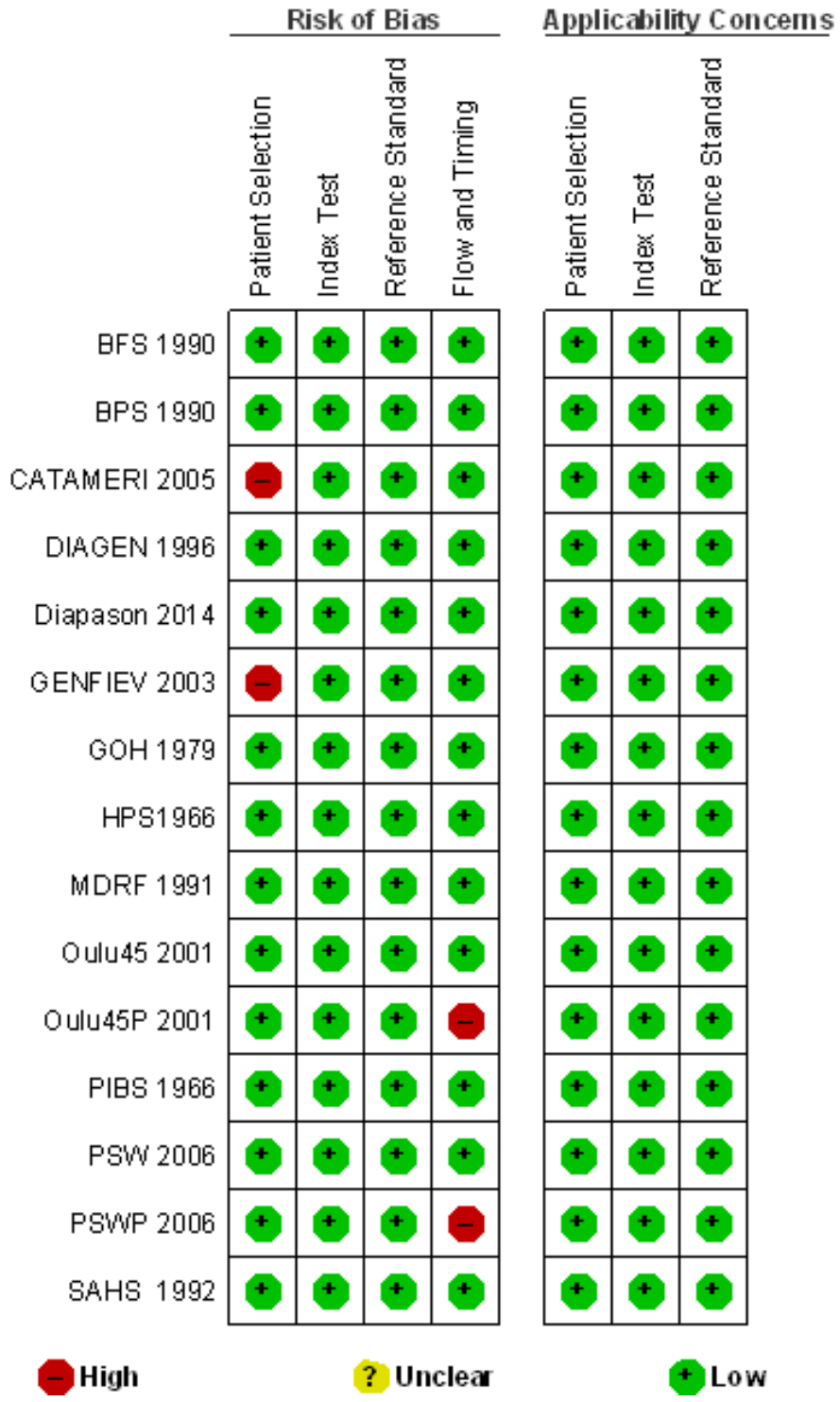


Figure S1. QUADAS-2 methodological assessment summary
BFS, Botnia Family Study; BPS, Botnia Prospective Study; CATAMERI, CATAnzaro MEtabolic RIsk factors; DIAGEN, DIAbetes GENetic study; GENFIEV, Genetic PHYsiopathology, and Evolution of Type 2 Diabetes; DIAPASON, Diabetes Prediction and Screening Observational Study; GOH, Israel Study of Glucose Intolerance, Obesity and Hypertension study; HPS, Helsinki Policemen Study; MDRF, Madras Diabetes Research Foundation study; Oulu45P, Oulu45 Prospective study; PIBS, Pima Indian Biennial Study; PSW, Public School Worker study; PSWP, Public School Worker Prospective study; SAHS, San Antonio Heart Study

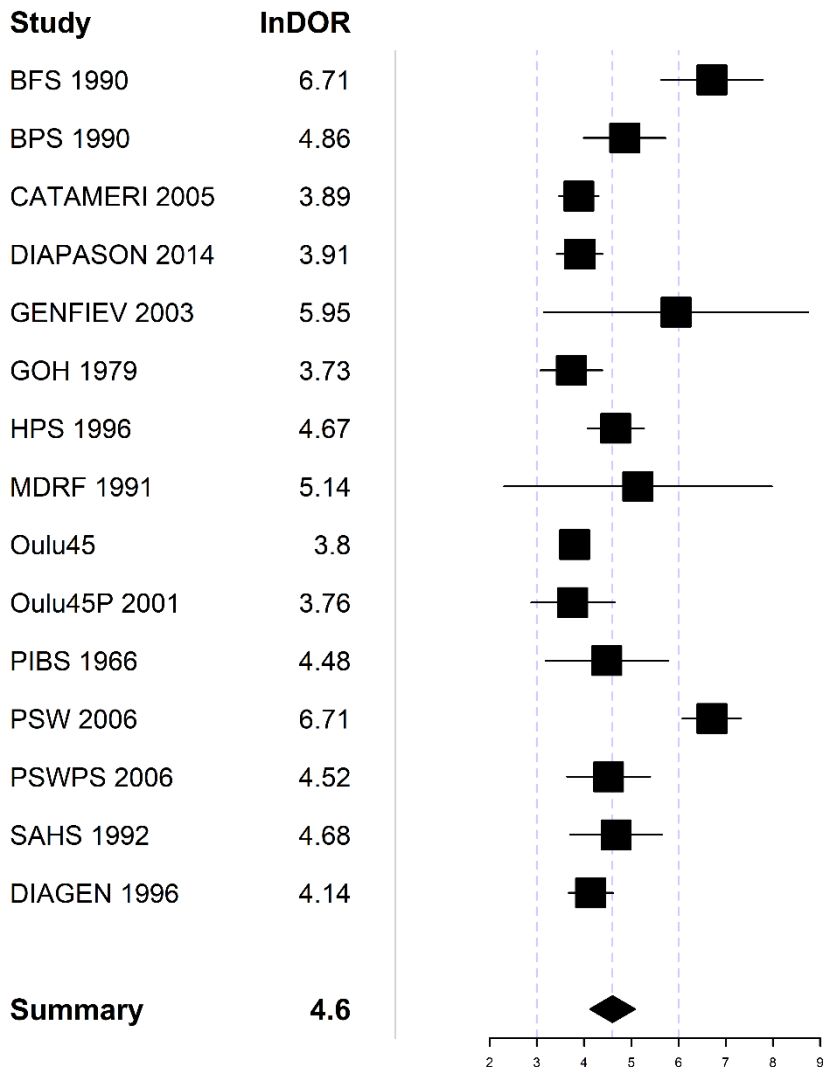


Figure S2. Forest plot of the log diagnostic odds ratio (DOR) of the individual studies with the summary DOR
BFS, Botnia Family Study; BPS, Botnia Prospective Study; CATAMERI, CATAnzaro MEtabolic RIsk factors; DIAGEN, DIAbetes GENetic study; GENFIEV, Genetic PHYsiopathology, and Evolution of Type 2 Diabetes; DIAPASON, Diabetes Prediction and Screening Observational Study; GOH, Israel Study of Glucose Intolerance, Obesity and Hypertension study; HPS, Helsinki Policemen Study; MDRF, Madras Diabetes Research Foundation study; Oulu45P, Oulu45 Prospective study; PIBS, Pima Indian Biennial Study; PSW, Public School Worker study; PSWP, Public School Worker Prospective study; SAHS, San Antonio Heart Study

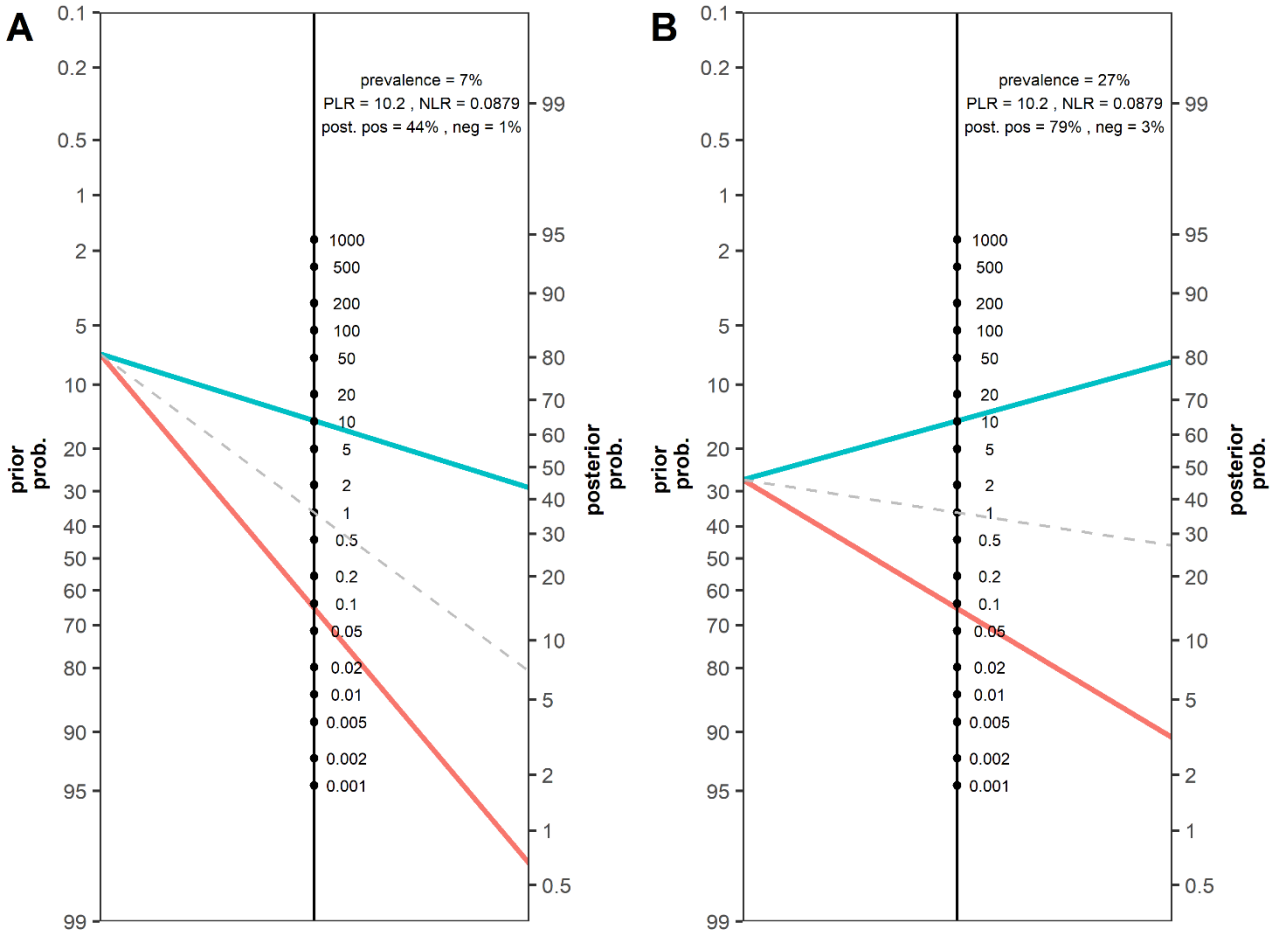


Figure S3. Fagan's nomogram displaying pre and post-test probabilities of 1-h PG at the 7% prevalence of diabetes in the meta-analysis (A) and at 27% prevalence in the Genetic PHYsiopathology, and Evolution of Type 2 Diabetes Study (B)

Prior prob., prior probability or prevalence; PLR, positive likelihood ratio; NLR, negative likelihood ratio; post. pos; posterior positive predictive value; post. neg, posterior negative predictive value

Upper line = positive predictive value, middle dashed line = null line; lower line = negative predictive value

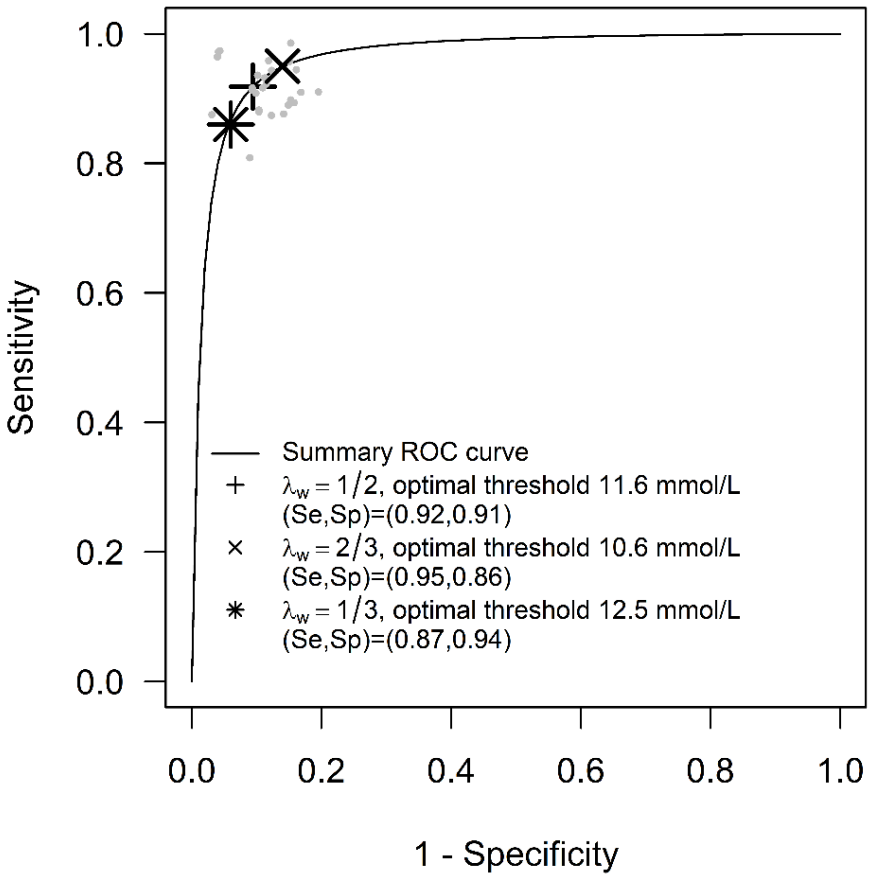


Figure S4. The Summary Receiver Operator Characteristic (SROC) curve displays three different 1-h plasma glucose cut-offs to detect 2-h plasma glucose of ≥11.1 mmol/L and associated sensitivities and false positive rates from the different random slope models (DS) with different weight ratios (λ) for sensitivity (Se) and specificity (Sp)
Grey circles = estimate of individual studies

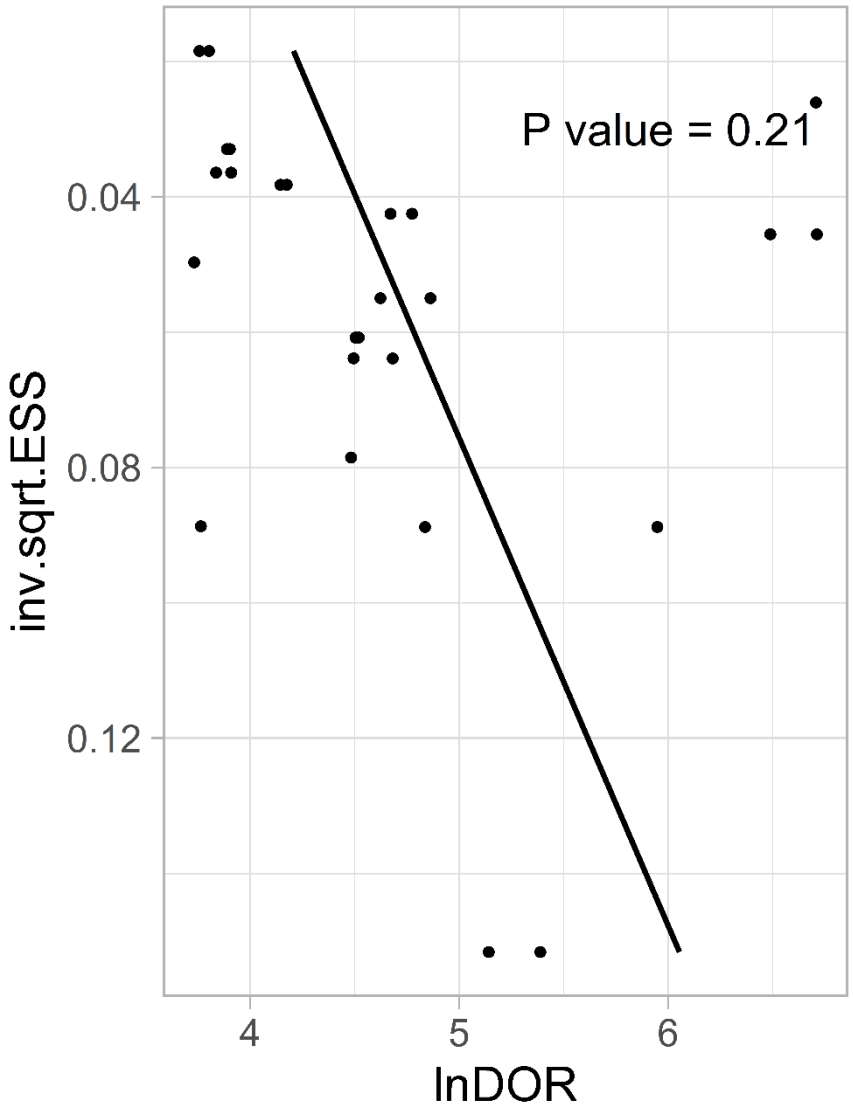


Figure S5. Funnel plot to examine sample size-related effects
lnDOR, log diagnostic odds ratio; inv.sqr.EES, inverse squared effective sample size
Black circles = individual study estimates

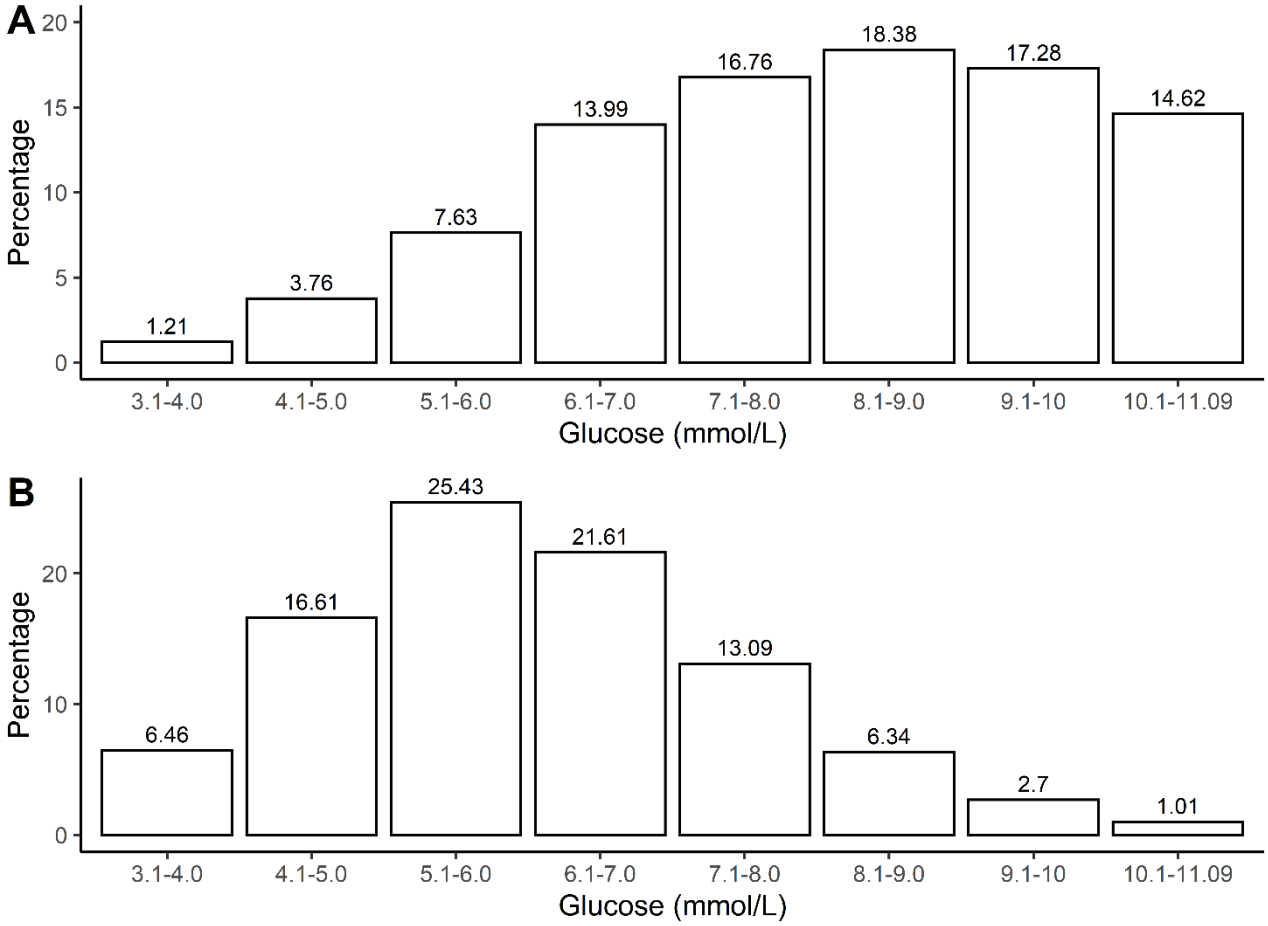


Figure S6. The bar chart shows distribution of 2h-PG values among participants classified as not having diabetes by the 2h-PG (<11.1 mmol/L), who were considered false positive (A) or true negatives (B) by the one-hour plasma glucose cut-off of 11.6 mmol/L in the studies with raw data*

* Botnia Family Study, Botnia Prospective Study, CATanzaro METabolic RIsk factors, DIAbetes GENetic study, Public School Worker Study, Public School Worker Prospective Study, San Antonio Heart Study

Table S1. Characteristics of the included studies

Study*	N without cases on glucose-lowering medications	N, new type 2 diabetes (2-h PG or FPG or HbA_{1c})	Cases by 2-h PG only, N (%)	Cases by 2-h PG and FPG/HbA_{1c}, N (%)	Excluded cases by FPG and/or HbA1c, N (%)	N without cases identified by only FPG/ HbA_{1c}	1-h PG in cases (2-h PG ≥11.1 mmol/L) (mean ± SD)	1-h PG in controls (mmol/L) (mean ± SD)
BFS 1990[15†‡]	3022	193	47 (24.4)	79 (40.9)	67 (34.7)	2995	15.3 ± 3.2	7.6 ± 2.2
BPS 1990[16†‡]	3253	170	49 (28.8)	36 (21.2)	85 (50.0)	3168	14.1 ± 2.3	7.8 ± 2.5
CATAMERI 2005[17]	3340	265	188 (70.9)	61 (23.0)	16 (6.0)	3324	13.5 ± 2.3	8.4 ± 2.4
DIAGEN 1996[18‡]	2770	294	113 (38.4)	91 (31.0)	90 (30.6)	2679	14.2 ± 2.9	8.7 ± 2.4
DIAPASON 2014[19‡]	531	34	34 (100)	0 (0)	0 (0)	531	13.1 ±1.5	8.1 ±2.3
GENFIEV 2003[20]	931	131	86 (65.7)	30 (22.9)	15 (11.5)	916	13.1 ± 1.8	9.2 ± 2.5
GOH 1979[10]	2126	183	77 (42.1)	72 (39.3)	34 (18.6)	2092	15.3 ± 3.8	8.1 ± 2.4
HPS 1966[21†]	1033	18	6 (33.3)	5 (27.8)	7 (38.9)	1026	15.9 ±3.7	7.4 ± 2.2
MDRF 1991[22]	9872	1023	583 (57.0)	219 (21.4)	221 (21.6)	9651	13.4 ± 1.6	9.1 ± 2.2
Oulu45 2001[23†]	959	59	20 (33.9)	13 (22.0)	26 (44.1)	933	13.9 ± 2.9	8.4 ± 2.1
Oulu45 2001[23†]	846	65	35 (4.1)	9 (1.1)	21 (2.5)	825	10.1 ± 1.9	7.9 ± 1.8
PIBS 1966[14]	2644	417	297 (71.2)	102 (24.5)	18 (4.3)	2640	18.5 ± 6.5	7.4 ± 2.3
PSW 2006[24‡]	2157	118	48 (40.7)	22 (18.6)	48 (40.7)	2085	13.5 ± 1.8	8.3 ± 2.4
PSWP 2006[25‡]	2015	83	49 (59.0)	16 (19.3)	18 (21.7)	1997	13.9 ±2.4	8.3 ± 2.4
SAHS 1992[26]	700	329	114 (34.6)	204 (62.0)	11 (3.3)	689	15.5 ±3.5	9.2 ± 2.0
All	36,199	3382	1746	959	677	35,551		

*Studies with their initiation years. †Blood glucose converted to plasma glucose using a conversion factor of 1.13. ‡Studies that determined HbA_{1c}; N, numbers; 2-hPG, two-hour plasma glucose; FPG, fasting plasma glucose; 1-hPG, one-hour plasma glucose; FPG/HbA1c, FPG and/or HbA1c in diabetic range; SD, standard deviation; BFS, Botnia Family Study; BPS, Botnia Prospective Study; CATAMERI, CATAnzaro MEtabolic RIsk factors; DIAGEN, DIAbetes GENetic study; GENFIEV, Genetic PHYsiopathology, and Evolution of Type 2 Diabetes; DIAPASON, Diabetes Prediction and Screening Observational Study; GOH, Israel Study of Glucose Intolerance, Obesity and Hypertension study; HPS, Helsinki Policemen Study; MDRF, Madras Diabetes Research Foundation study; Oulu45P, Oulu45 Prospective study; PIBS, Pima Indian Biennial Study; PSW, Public School Worker study; PSWP, Public School Worker Prospective study; SAHS, San Antonio Heart Study

Table S2. Meta-regression of sources of heterogeneity in the meta-analysis of 1-h PG to detect 2-h PG ≥ 11.1 mmol/L [*]

	Sensitivity		Specificity	
	Q coefficient	***P* value**	**Q coefficient**	***P* value**
Design of study[†]	0.63	0.43	0.02	0.88
Setting[‡]	0.31	0.58	0.54	0.46
Dose of glucose used for OGTT[§]	0.02	0.88	1.39	0.24
Ethnicity[‖]	25.90	<0.0001	355.53	<0.0001
Bias[¶]	1.69	0.19	6.73	0.001

[*]We used the cut-offs at the Youden's index for the meta-regression analyses; [†]Longitudinal vs. cross-sectional studies; [‡]Population-based vs diabetes clinic-based studies; [§]75 g vs 100 g; [‖]Ethnicity (Caucasians vs. South Asians vs. American Indians vs. Japanese vs. Mexican Americans); [¶]Studies with low risk of bias vs. with risk of bias
1-hPG, one-hour plasma glucose; 2-hPG, two-hour plasma glucose; OGTT, oral glucose tolerance test

Ethnicity (study)	1-h PG cut-off[†]	Sensitivity	Specificity
Caucasians[†§]	11.7	0.91 (0.84, 0.95)	0.92 (0.89, 0.94)
South Asians[‡] **(MDRF)**	11.5	0.90 (0.87, 0.91)	0.84 (0.83, 0.85)
American Indians[‡] **(PIBS)**	10.8	0.97 (0.95, 0.99)	0.96 (0.95, 0.96)
Japanese[†] **(PSW, PSWP)**	11.3	0.92 (0.87, 0.96)	0.89 (0.88, 0.90)
Mexican Americans[‡] **(SAHS)**	11.8	0.88 (0.83, 0.91)	0.90 (0.86, 0.93)

Table S3. Comparison of 1-h PG cut offs to detect 2-h PG ≥ 11.1 mmol/L among ethnicities*

*Statistical test not available to compare cut-offs among different groups; [†]The cut-off for Caucasians and Japanese obtained after meta-analysing studies with Caucasian and Japanese participants because of availability of sufficient sample size; [‡]the cut-off at the Youden's index is displayed for South Asians; American Indians, and Mexican Americans (the cut-offs at the minimum distance were 11.7, 10.8, and 11.8, respectively). [§] Studies with Caucasian participants (Botnia Family Study; Botnia Prospective Study; CATAMERI, CATAnzaro METabolic RIsk factors; DIAbetes GENetic study; Genetic Physiopathology and Evolution of Type 2 Diabetes study; Diabetes Prediction and Screening Observational Study; Israel Study of Glucose Intolerance, Obesity and Hypertension study; Helsinki Policemen Study; Oulu45P study, and Oulu45 Prospective study); 1-hPG, one-hour plasma glucose
2-hPG, two-hour plasma glucose; MDRF, Madras Diabetes Research Foundation study; PIBS, Pima Indian Biennial Study; PSW, Public School Worker; PSWP, Public School Worker Prospective; SAHS, San Antonio Heart Study

Table S4. The comparison of cut-offs of the 1-h PG to detect 2-h PG ≥ 11.1 mmol/L to ≤ 13.0 mmol/L with cut-offs to detect a 2-h PG ≥ 11.1 mmol/L in the sensitivity analysis*

		Cut-off in mmol/L (Se, Sp)		
		Studies with raw data		meta-analysis
		2-h PG ≥ 11.1 to ≤ 13.0	2-h PG ≥ 11.1	2-h PG ≥ 11.1
Weight ratio for Se vs. Sp	**lambda (λ)**			
more	2/3	12.6 (0.88, 0.96)	10.7 (0.94, 0.86)	10.6 (0.95, 0.86)
equal	1/2	13.5 (0.87, 0.98)	12.1 (0.90, 0.92)	11.6 (0.92, 0.91)
less	1/3	14.5 (0.86, 0.99)	13.5 (0.85, 0.95)	12.5 (0.87, 0.94)

1-hPG, one-hour plasma glucose; 2-hPG, two-hour plasma glucose; Se, sensitivity; Sp, specificity

Table S5. The comparison of unadjusted and adjusted cut-offs of 1-h PG to detect 2-hPG ≥ 11.1 mmol/L for the studies with raw data[*]

Weight ratio for Se vs. Sp	lambda (λ)	Unadjusted[†] cut-off (se, sp)		AUC[‡] (CR for se at given sp)		Adjusted cut-off[†] cut-off (se, sp)		AUC[‡] (CR for se at given sp)	
more	2/3	11.0	(0.93, 0.86)	0.962	(0.885, 0.985)	10.4	(0.93, 0.89)	0.973	(0.277, 0.996)
equal	1/2	12.3	(0.89, 0.90)	0.962	(0.885, 0.985)	14.8	(0.90, 0.94)	0.973	(0.277, 0.996)
less	1/3	13.7	(0.84, 0.95)	0.962	(0.885, 0.985)	20.7	(0.86, 0.96)	0.973	(0.277, 0.996)

[*] Botnia Family Study, Botnia Prospective Study, CATanzaro METabolic RIsk factors, DIAbetes GENetic study, Public School Worker Study, Public School Worker Prospective Study, San Antonio Heart Study; [†]Unadjusted cut-offs obtained after meta-analyzing unadjusted cut-offs from studies and adjusted cut-offs obtained after meta-analyzing age, sex, body-mass index (BMI; available for five out of seven studies) adjusted cut-offs from studies. [‡]Statistical test not available to compare AUC between unadjusted and adjusted cut-offs

1-h PG, 1-h plasma glucose; 2-hPG, 2-h plasma glucose; Se, sensitivity; Sp, specificity; AUC, area under the curve for the summary receiver operator characteristic curve; CR, confidence region

Table S6. The comparison of unadjusted and age, sex, body-mass index adjusted cut-offs of 1-h PG to detect 2-h PG ≥ 11.1 mmol/L in the studies with available raw data*

Study	Unadjusted		Adjusted†		*P* value
	cut-off (se, sp)	AUC	cut-off (se, sp)	AUC	
BFS	11.9 (0.98, 0.96)	0.998	11. 3 (0.98, 0.97)	0.994	0.04
BPS	11.2 (0.94, 0.90)	0.959	10.8 (0.94, 0.90)	0.964	0.03
CATAMERI	10.8 (0.90, 0.85)	0.940	11.2 (0.91, 0.87)	0.944	0.02
DIAGEN	11.2 (0.91, 0.83)	0.941	11.2 (0.96, 0.85)	0.957	0.01
PSW	11.2 (0.93, 0.89)	0.956	11.8 (0.94, 0.89)	0.960	0.70
PSPW	11.3 (0.94, 0.89)	0.963	11.3 (0.92, 0.90)	0.964	0.45
SAHS	11.8 (0.88, 0.90)	0.956	11.5 (0.91, 0.90)	0.964	0.01

*Botnia Family Study, Botnia Prospective Study, CATAMERI, CATAnzaro METabolic RIsk factors, DIAbetes GENetic studyPublic School Worker Study, Public School Worker Prospective Study, San Antonio Heart Study;
†Body-mass index available for five out of seven studies
1-h PG, 1-h plasma glucose; 2-hPG, 2-h plasma glucose; Se, sensitivity; Sp, specificity; AUC, area under curve