Use of Alternative Thresholds Defining Insulin Resistance to Predict Incident Type 2 Diabetes Mellitus and Cardiovascular Disease

Martin K. Rutter, MD, **Peter W.F. Wilson, MD**, **Lisa M. Sullivan, PhD**, **Caroline S. Fox, MD, MPH**, **Ralph B. D'Agostino Sr, PhD**, and **James B. Meigs, MD, MPH**

The Cardiovascular Research Group, Division of Cardiovascular and Endocrine Sciences, University of Manchester, and The Manchester Diabetes Centre, Manchester Royal Infirmary, Manchester, United Kingdom (M.K.R.); Emory University School of Medicine (P.W.F.W.), Atlanta, Ga; Department of Biostatistics (L.M.S.), Department of Mathematics and Statistics/Consulting Unit (R.B.D.), Boston University, Boston, Mass; the National Heart, Lung, and Blood Institute's Framingham (Mass) Heart Study (C.S.F.), Harvard Medical School and the Department of Endocrinology and Metabolism, Brigham and Women's Hospital, Boston, Mass (C.S.F); and Harvard Medical School and the General Medicine Division, Department of Medicine, Massachusetts General Hospital, Boston, Mass (J.B.M.).

Abstract

Background—The performance characteristics of surrogate insulin resistance (IR) measures, commonly defined as the top 25% of the measure's distribution, used to predict incident type 2 diabetes mellitus (DM) and cardiovascular disease (CVD) have not been critically assessed in community samples.

Methods and Results—Baseline IR was assessed among 2720 Framingham Offspring Study subjects by use of fasting insulin, the homeostasis model assessment of IR (HOMA-IR), and the reciprocal of the Gutt insulin sensitivity index, with 7- to 11-year follow-up for incident DM (130 cases) or CVD (235). Area under the receiver operating characteristic curve, sensitivity, specificity, and positive likelihood ratio were estimated at 12 diagnostic thresholds (quantiles) of IR measures. Positive likelihood ratios for DM or CVD increased in relation to IR quantiles; risk gradients were greater for DM than for CVD, with no 9th to 10th quantile (76th centile) threshold effects. IR had better discrimination for incident DM than for CVD (HOMA-IR area under the receiver operating characteristic curve: DM 0.80 versus CVD 0.63). The HOMA-IR ≥76th centile threshold was associated with these test-performance values: sensitivity (DM 68%, CVD 40%), specificity (DM 77%, CVD 76%), and positive likelihood ratio (DM 3.0, CVD 1.7). The HOMA-IR threshold that yielded >90% sensitivity was the 6th quantile for DM prediction and the 3rd quantile for CVD. Compared with the ≥76th centile threshold, these alternative thresholds yielded lower specificity (DM 43%, CVD 17%) and positive likelihood ratios (DM 1.6, CVD 1.1).

Conclusions—Surrogate IR measures have modest performance at the 76th centile, with no threshold effects. Different centile thresholds might be selected to optimize sensitivity versus specificity for DM versus CVD prediction if surrogate IR measures are used for risk prediction.

Keywords

insulin resistance; cardiovascular diseases; diabetes mellitus; risk factors; prospective studies; prognosis

---

Insulin resistance (IR) is associated with incident type 2 diabetes mellitus (DM) and cardiovascular disease (CVD)[1,2] and may be part of the causal pathway that links obesity with these conditions. The "gold standard" method for quantifying IR is the hyperinsulinemic euglycemic clamp, but this is impractical in clinical settings or in epidemiological studies. Surrogate measures of IR provide approximations of clamp IR derived from plasma glucose and insulin levels. Surrogate measures commonly used in research include the simplest (the fasting insulin concentration); the most commonly used (homeostasis model assessment of IR [HOMA-IR], derived from the product of fasting glucose and insulin levels); and the more sophisticated insulin sensitivity index (ISI) described by Gutt and coworkers, which is derived from body weight and fasting and post–oral glucose tolerance test (OGTT) glucose and insulin levels.[2–4] IR is often defined in population studies as greater than or equal to the 75th centile of the surrogate IR-measure distribution in subjects without DM[5]; however, the test performance of this value or any other centile threshold for predicting incident DM or CVD has not been examined critically. Our aims were therefore (1) to assess test performance at different diagnostic thresholds of surrogate IR measures in the prediction of incident DM or CVD and (2) to compare the discriminatory capacity (measured as the area under the receiver operating characteristic curve [aROC]) for different surrogate measures of IR in the prediction of both DM and CVD in a population-based cohort.

CLINICAL PERSPECTIVE

In population-based studies, insulin resistance (IR) predicts type 2 diabetes mellitus (DM) and cardiovascular disease (CVD) and is commonly defined as the top 25% of the distribution of surrogate IR measure, such as the homeostasis model assessment of IR (HOMA-IR). The performance characteristics of surrogate IR measures for DM or CVD prediction have not been assessed. We assessed baseline IR using fasting insulin, HOMA-IR, and the reciprocal of the Gutt insulin sensitivity index (ISI) among 2720 Framingham Offspring Study subjects followed up for 7 to 11 years for incident DM (130 cases) or CVD (235 cases) and estimated test performance at 12 diagnostic thresholds (quantiles) of IR measures. Risk for DM or CVD increased in relation to IR quantiles, with no apparent 76th centile threshold effects; risk gradients were greater for DM than for CVD prediction. Surrogate IR measures demonstrated limited performance at this centile, especially for CVD, and alternative thresholds improved sensitivity or specificity at the cost of higher false-positive or false-negative rates. HOMA-IR and 1/ISI had similar test performance, and both measures outperformed fasting insulin for DM prediction but were essentially equivalent with regard to CVD prediction. The data suggest that in the community, HOMA-IR may have value for DM prediction, with a positive likelihood ratio of 3 and good discrimination for incident events, especially in multivariate models. However, surrogate IR measures had limited performance for CVD prediction, with a low positive likelihood ratio gradient across quantiles and values for the area under the receiver operating characteristic curve that were well below those for the Framingham CHD risk score.

Methods

The Framingham Offspring Study is a community-based prospective observational study of CVD and its risk factors.[6] Offspring subjects are white and of mixed European ancestry. During

the fifth examination cycle (baseline examination 1991 to 1995), 3799 participants fasted overnight and had a standardized medical examination, which included a 2-hour OGTT. From 3799 participants, we excluded those with prevalent DM (n=400), missing information on covariates (n=410), or prevalent CVD (n=269), which left 2720 subjects for analysis. Subjects were followed up from baseline over a mean of 6.8 years for new cases of DM and a mean of 11.6 years for first CVD events. The Institutional Review Board of Boston University approved the study protocol, and all subjects gave informed consent at each examination.

Clinical Definitions and Laboratory Methods

Plasma glucose was measured in fresh specimens with a hexokinase reagent kit (A-gent glucose test; Abbott, South Pasadena, Calif). Glucose assays were run in duplicate; the intra-assay coefficient of variation was <3%. Fasting insulin levels were measured in plasma as total immunoreactive insulin and were standardized to serum levels for reporting purposes. The lower limit of sensitivity was 8.0 pmol/L (1.1 μU/mL), and the intra-assay and interassay coefficients of variation ranged from 5.0% to 10.0%. Surrogate measures of IR, assessed by validated methods, included fasting insulin,[7] HOMA-IR, and Gutt's $ISI_{0,120}$. HOMA-IR was calculated as [fasting glucose (mmol/L) × fasting insulin (μU/mL)]/22.5.[8,9] HOMA-IR formula values[8] are highly correlated with computer-derived HOMA-IR model values[10] in the Framingham study (r=0.98, P<0.0001); only results using the former are presented. Gutt's $ISI_{0,120}$ was calculated as (m/MPG)/log MSI, where m= [75 000 mg + (fasting glucose−2-hour glucose)×0.19×body weight (kg)]/120 minutes, MPG is the mean of fasting and 2-hour glucose concentrations (mg/dL), and MSI is the mean of fasting and 2-hour insulin concentrations (mU/L).[3] This index measures insulin sensitivity, so for the present analysis, we used the inverse, 1/ISI, such that the index measured IR and was positively correlated with other surrogate IR measures. We divided each surrogate IR-measure distribution into 12 equally sized quantile groups, with the upper 3 quantiles representing "insulin-resistant" subjects. The boundary values for the 12 quantiles (Q1 to Q12) of the population distribution of HOMA-IR were as follows: Q1 2.21 to 4.19, Q2 4.20 to 4.69, Q3 4.70 to 5.09, Q4 5.10 to 5.44, Q5 5.45 to 5.79, Q6 5.80 to 6.19, Q7 6.20 to 6.60, Q8 6.61 to 7.17, Q9 7.18 to 7.87, Q10 7.88 to 8.80, Q11 8.81 to 10.71, and Q12 10.72 to 30.80 U. The lower boundary of Q10 is the 76% centile, which is essentially the same as the 75% centile commonly used to define IR.[5]

DM and CVD Assessment

We defined diabetes at the baseline examination as a fasting plasma glucose level ≥7.0 mmol/L, a 2-hour OGTT glucose level ≥11.1 mmol/L, or current use of hypoglycemic drug therapy. Impaired fasting glucose was defined as a fasting plasma glucose level of 5.6 to 6.9 mmol/L and impaired glucose tolerance as a 2-hour OGTT glucose level of 7.8 to 11.0 mmol/L. Subjects were followed up from baseline through the seventh (1998–2001) examination for DM and through December 2004 for CVD events. For DM incidence, we used the examination visit date on which a new case of DM was identified as the date of diagnosis. For CVD events, we used the actual date of the event as the date of diagnosis, and for subjects without events, the date of their last follow-up examination was used as the censoring date. We defined DM at follow-up as development of a fasting plasma glucose level ≥7.0 mmol/L or new use of hypoglycemic drug therapy during the study interval. Of the 400 patients with DM excluded at baseline, 54 (13.5%) were not undergoing treatment and had a fasting plasma glucose <7.0 mmol/L but a 2-hour OGTT glucose ≥11.1 mmol/L, which indicates the approximate proportion of incident DM cases that might be missed at follow-up by not performing an OGTT. More than 99% of diabetes cases among Framingham Offspring are type 2 diabetes mellitus.[11] We defined baseline and follow-up CVD by standard Framingham Heart Study criteria as any of the following: new-onset angina, fatal and nonfatal myocardial infarction or stroke, transient ischemic attack, heart failure, or intermittent claudication.[12]

Statistical Analysis

We divided the population distributions for each surrogate IR measure into 12 quantile groups, each of which included 227 individuals, and determined the number of DM or CVD events within each group. For each quantile group, we estimated risk for DM and CVD, sensitivity, specificity, and positive likelihood ratios (PLRs) using the lower boundary of each quantile as the "threshold" value. For example, the sensitivity for DM prediction associated with the 10th quantile of HOMA-IR was calculated from DM events in subjects with HOMA-IR values greater than the lower boundary of the 10th quantile (subjects in the 10th to 12th quantiles of the HOMA-IR distribution). The lower boundary of the 10th quantile is the 76th centile of the HOMA-IR distribution, and therefore, the measures of test performance associated with the 10th quantile are those associated with greater than or equal to the 76th centile threshold.[5]

We used logistic regression analysis to assess the relationship between IR thresholds, considered simultaneously, and incident DM or incident CVD. Cox proportional hazards regression models yielded nearly identical results; only results from logistic regression are presented here. Separate regression models were used for DM or CVD prediction. Primary analyses were performed without covariate adjustment to reflect standard use of blood test results in clinical practice. Subsidiary analyses of surrogate measures considered additional adjustment of all for age and sex. For fasting insulin, we also considered additional adjustment for fasting glucose (to assess adjusted discrimination compared with the discrimination with HOMA-IR) and for fasting and 2-hour OGTT insulin, 2-hour glucose, and weight (to compare with 1/ISI). For HOMA-IR, we also considered additional adjustment for 2-hour OGTT insulin and glucose levels and weight (to compare HOMA-IR with 1/ISI). For each surrogate measure, we compared the aROC of the fuller model with that of the sparser model.[13] Another subsidiary analysis considered the discrimination by surrogate measures in 2 strata, normal glucose tolerance versus impaired fasting glucose and/or impaired glucose tolerance.[14] To assess population risk prediction, we calculated aROCs and associated 95% confidence intervals (CIs). aROCs are interpreted as the probability that the modeled phenotype(s) can correctly discriminate subjects developing end points from those without end points, where 0.5 is chance discrimination and 1.0 is perfect discrimination. To address individual prediction, we calculated the likelihood ratio, which summarizes how likely patients with the disease are to have a specified test result compared with patients without the disease.[15] We used conventional definitions for PLR [sensitivity/(100%−specificity)]. We defined the false-positive rate as (100%−specificity) and the false-negative rate as (100%−sensitivity). We performed all analyses using SAS software (SAS Institute, Cary, NC).

The authors had full access to the data and take full responsibility for its integrity. All authors have read and agree to the manuscript as written.

Results

Baseline characteristics of the study sample are shown in Table 1. The mean age of the study sample was 54 years (range 26 to 82 years), and 55% were women. Correlations were significant (all $P<0.001$) among fasting insulin and HOMA-IR (Spearman $r=0.95$), HOMA-IR and 1/ISI ($r_s=0.50$), and fasting insulin and 1/ISI ($r_s=0.42$).

Impact of Different IR Diagnostic Thresholds on DM or CVD Prediction

The numbers of events occurring in the 12 quantile HOMA-IR groups are shown in Table 2 along with sensitivity, specificity, and PLR values associated with different IR thresholds for new DM or CVD. More than two thirds (68%) of incident DM events and two fifths (40%) of CVD events occurred in subjects classified as being "insulin resistant" (HOMA-IR levels in the upper 24% of the population distribution). Model performance was of similar magnitude

for HOMA-IR and 1/ISI, and both measures outperformed fasting insulin. For instance, in Figure 1, the likelihood of DM or CVD increased steadily with increasing centile of HOMA-IR, apparently without a threshold at the 76th centile (10th quantile and above) or elsewhere. Likelihood ratios for DM were higher than for CVD across the range but especially at higher centiles of HOMA-IR. The PLRs associated with the 76th centile threshold for HOMA-IR were 3.0 for 7-year DM prediction and 1.7 for 11-year CVD prediction.

The 76th centile diagnostic threshold for IR yielded false-positive rates of ≈ 1 in 4 for DM or CVD prediction for all surrogate measures (specificity: DM 77%, CVD 76%) and missed one to two thirds of cases (sensitivity 60% to 68% for DM and 36% to 40% for CVD; Table 2; Figure 2). The values of sensitivity and specificity that might be considered "acceptable" may differ depending on the clinical situation; Table 2 can be used to assess varying combinations. For instance, assume that >90% sensitivity and specificity represents acceptable test performance for both DM and CVD prediction. The HOMA-IR quantile threshold associated with >90% sensitivity was higher for DM than for CVD prediction (DM ≥6th quantile; CVD ≥3rd quantile), which corresponds to the 42nd centile or greater for DM and the 17th centile or greater for CVD, respectively. Compared with the ≥76th centile threshold, these alternative thresholds yield inferior specificity (DM 43%, CVD 17%) and PLR (DM 1.6, CVD 1.1). The HOMA-IR threshold that yields >90% specificity was the 12th quantile (≥92% centile) for both DM and CVD prediction. Compared with the ≥76th centile threshold, this alternative threshold yields inferior sensitivity (DM 38%, CVD 17%).

Surrogate IR Measures and Population Prediction of DM or CVD

Table 2 also provides data for aROC analyses that show that discrimination of DM is better with HOMA-IR or 1/ISI than with fasting insulin and that IR by any surrogate predicts incident DM better than CVD (Table 2; Figure 2). For example, in Table 2, the HOMA-IR aROC (95% CI) for incident DM was 0.80 (0.76–0.83) compared with 0.63 (0.59–0.66) for incident CVD; a similar pattern was observed for fasting insulin and 1/ISI. Results were generally similar in subsidiary analyses stratified by normal glucose tolerance versus impaired fasting glucose and/or impaired glucose tolerance, For instance, the HOMA-IR aROC (95% CI) for incident DM was 0.73 (0.62 to 0.83) in subjects with normal glucose tolerance and 0.71 (0.65 to 0.76) in those with impaired fasting glucose or impaired glucose tolerance; the HOMA-IR aROCs for incident CVD were 0.60 (0.54 to 0.66) and 0.60 (0.54 to 0.66), respectively. Although aROCs for DM were numerically lower in substrata than the corresponding aROC in the sample overall, these aROCs were not statistically different (probability value comparing aROCs ≥0.20).

One reason that fasting insulin underperformed HOMA-IR and 1/ISI is that the latter also included glucose information. Variation in age and sex may also contribute. We examined this in subsidiary analyses of surrogate measures that considered additional adjustment for age, sex, and the additional information in the incrementally more complex surrogate measures.

Adjustment for age and sex made no difference in discrimination of DM; crude versus age- and sex-adjusted aROCs were identical to the values shown in Table 2. Next, we found that aROC values for fasting glucose levels were 0.84 (0.80 to 0.88) for DM and 0.60 (0.56 to 0.64) for CVD prediction. For DM, adjustment of fasting insulin models for fasting glucose levels (to assess adjusted discrimination similar to HOMA-IR) produced a similar increase in aROC (age-sex-glucose-insulin–adjusted aROC [95% CI] 0.86 [0.82 to 0.89] versus crude insulin aROC, P<0.0001). Adjustment of fasting insulin for age, sex, fasting and 2-hour OGTT glucose and insulin, and weight (to compare with $ISI_{0,120}$) did not materially improve discrimination of DM beyond simply adjusting for age, sex, and fasting glucose (fully adjusted aROC 0.87). Adjustment of HOMA-IR for age, sex, 2-hour OGTT glucose and insulin, and weight (to

compare with $ISI_{0,120}$) slightly improved the aROC (adjusted aROC 0.84 [0.80 to 0.87], $P<0.001$ versus crude HOMA-IR aROC).

In CVD prediction models, adjustment for age and sex increased model aROCs from the range 0.60 to 0.63 to a value of 0.70 for all 3 surrogate IR measures. Additional adjustment as above for the DM models did not further alter CVD discrimination; fully adjusted aROCs were 0.70 to 0.71 for all surrogate IR measures.

Discussion

We have shown that risk for DM or CVD increases in relation to IR quantiles, with no apparent 76th centile threshold effects. Surrogate measures of IR demonstrated limited performance at this centile, especially for CVD, and alternative thresholds improved sensitivity or specificity at the cost of higher false-positive or false-negative rates. HOMA-IR and 1/ISI had similar test performance, and both measures outperformed fasting insulin for DM prediction but were essentially equivalent with regard to CVD prediction. We observed that with increasing IR quantiles, the risk gradients were greater for DM than for CVD prediction and that in the aROC analysis, IR predicted incident DM better than CVD.

Absence of Threshold Effects and Utility for Disease Prediction

That there are no 76th centile threshold effects has biological plausibility and is not surprising, because the decision to adopt the 76th centile as the diagnostic threshold for IR was arbitrary and not based on performance to predict disease. The observation that test performance at the 76th centile is limited is in keeping with the modest relative risk associated with surrogate IR measures used to predict incident DM (relative risk 1.6 to 6.5)[16,17] or CVD (relative risk 1.4 to 2.2)[1,17–19] in other studies. The limited performance of surrogate IR measures to predict incident DM or CVD could be explained by several factors, including the multifactorial causation of disease (in which IR is only weakly related or is unrelated to major risk factors, including low-density lipoprotein cholesterol and smoking) and that the surrogate measures are imprecise, which causes misclassification that biases the point estimates toward the null.

Choice of Surrogate IR Measure

Our main analysis in the present study focused on HOMA-IR because it is the most commonly used surrogate IR measure in epidemiological studies. We also focused on fasting insulin because of its simplicity and found that it performed surprisingly well compared with other measures in the population prediction of CVD. Our observation that it did not perform as well as HOMA-IR for DM prediction probably reflects the strong prognostic information associated with fasting glucose. We also studied Gutt's ISI both because Hanley and coworkers[4] showed that it was more strongly related to incident DM than several other surrogate IR measures and because our group has shown good performance of ISI for predicting CVD.[2] Compared with HOMA-IR, ISI is more difficult to estimate, because it requires measurement of body weight and tests for glucose and insulin before and after OGTT. In the present study, we used the reciprocal of ISI to estimate IR, and we showed similar test performance of 1/ISI and HOMA-IR. This was somewhat surprising, because compared with HOMA-IR, ISI contains additional prognostic information that might be expected to influence incident DM or CVD. The present findings suggest that for population DM or CVD risk prediction with surrogate IR measures, there may be little to be gained by performing an OGTT to estimate ISI.

Strengths and Limitations

Strengths of the study include analysis of a large representative population-based sample, analysis of several surrogate IR measures, and analysis of both incident DM and CVD. Because there is currently no standardization of insulin assays,[20] we have presented data for centiles

of surrogate IR measures that could be applied in settings in which normative data are available for fasting glucose and insulin (using any assay). The centiles approach circumvents to a large degree the problem that absolute insulin concentrations cannot be compared easily across samples regardless of the insulin assay kit. We used a nonspecific total immunoreactive insulin assay that cross-reacts as much as 40% with proinsulin, whereas many newer studies use insulin-specific assays, but regardless of assay type, our approach is likely to produce very similar rankings from low to high relative insulin concentration.

Furthermore, our purpose was to evaluate key surrogate IR measures themselves, to address the common assumption that the top 25% of the distribution of various surrogate IR measures adequately identifies a useful threshold that links IR proxies to key disease end points at the population level, and not to produce the "best" DM or CVD prediction models. In other work, we have published "best models" for DM and CVD prediction; interestingly, in the present analysis, we found that the discrimination afforded by consideration of age, sex, and fasting glucose and insulin was equivalent to that of the "best" DM prediction models but lower than that of the "best" CHD prediction models.[21,22] Also, we do not address the marginal risk for DM or CVD associated with surrogate IR measures after accounting for standard risk factors; we have addressed this issue in other work.[2,17]

Limitations of the present study beyond those discussed above include analysis of a white study sample, which limits generalizability to other race ethnic groups; lack of OGTT data to identify some cases of DM at follow-up; limited power to assess performance in age- or sex-stratified groups; and the likely large CIs around risk estimates at the extremes of the IR distribution. Also, we did not assess positive and negative predictive values, but these parameters are strongly cohort-specific and are influenced by follow-up duration.

Clinical Implications

A current uncertainty is the clinical value of HOMA-IR or any surrogate IR measure for use in management or clinical prediction of metabolic disorders. We document the expected value of IR surrogate measure thresholds for individual (PLR) and population (aROC) prediction of the 2 major chronic diseases associated with IR. The lack of risk thresholds at the 76th centile suggests that different centile thresholds might be selected to optimize sensitivity versus specificity depending on the diagnostic or screening situation. For example, a DM screening test requires high specificity (>95%) and moderate sensitivity (≈70%), whereas a diagnostic test requires a much higher specificity. Different diagnostic thresholds could be used for CVD than for DM prediction. Additional research would be required to determine whether this approach might be more useful in subjects at higher pretest risk of disease, such as those with impaired glucose tolerance, metabolic syndrome, or obesity.

Although measurement of IR is commonly discussed as a clinical or public health strategy to identify metabolic risk, the present data suggest that in the community, HOMA-IR in particular may have some apparent value for DM prediction, with a PLR of 3 and an aROC of 80%, as well as an aROC of 86% in a multivariate (age, sex, fasting glucose, and insulin) model. Remarkably, this very simple metabolic model has a similar discrimination capacity as a slightly more complex DM risk model based on family history of DM and metabolic syndrome variables.[22] However, surrogate IR measures appear to have limited performance for CVD prediction, with a low PLR gradient across centiles and no aROC >63%, well below aROCs for the Framingham CHD risk score.[21]

Conclusions

Population risk for DM or CVD increases in relation to centiles of surrogate measures of IR. There is modest performance at the 76th centile and no apparent threshold effects, and there

are no alternative thresholds at which population prediction for both DM and CVD is equally satisfactory. Different centile thresholds might be selected to optimize sensitivity versus specificity for DM and, perhaps, CVD prediction in clinical or epidemiological settings.

References

1. Hanley AJ, Williams K, Stern MP, Haffner SM. Homeostasis model assessment of insulin resistance in relation to the incidence of cardiovascular disease: the San Antonio Heart Study. Diabetes Care 2002;25:1177–1184. [PubMed: 12087016]

2. Rutter MK, Meigs JB, Sullivan LM, D'Agostino RB Sr, Wilson PW. Insulin resistance, the metabolic syndrome, and incident cardiovascular events in the Framingham Offspring Study. Diabetes 2005;54:3252–3257. [PubMed: 16249452]

3. Gutt M, Davis CL, Spitzer SB, Llabre MM, Kumar M, Czarnecki EM, Schneiderman N, Skyler JS, Marks JB. Validation of the insulin sensitivity index (ISI(0,120)): comparison with other measures. Diabetes Res Clin Pract 2000;47:177–184. [PubMed: 10741566]

4. Hanley AJ, Williams K, Gonzalez C, D'Agostino RB Jr, Wagenknecht LE, Stern MP, Haffner SM. Prediction of type 2 diabetes using simple measures of insulin resistance: combined results from the San Antonio Heart Study, the Mexico City Diabetes Study, and the Insulin Resistance Atherosclerosis Study. Diabetes 2003;52:463–469. [PubMed: 12540622]

5. Balkau B, Charles MA. European Group for the Study of Insulin Resistance (EGIR). Comment on the provisional report from the WHO consultation. Diabet Med 1999;16:442–443. [PubMed: 10342346]

6. Kannel WB, Feinleib M, McNamara PM, Garrison RJ, Castelli WP. An investigation of coronary heart disease in families: the Framingham offspring study. Am J Epidemiol 1979;110:281–290. [PubMed: 474565]

7. Laakso M. How good a marker is insulin level for insulin resistance? Am J Epidemiol 1993;137:959–965. [PubMed: 8317453]

8. Matthews DR, Hosker JP, Rudenski AS, Naylor BA, Treacher DF, Turner RC. Homeostasis model assessment: insulin resistance and beta-cell function from fasting plasma glucose and insulin concentrations in man. Diabetologia 1985;28:412–419. [PubMed: 3899825]

9. Bonora E, Targher G, Alberiche M, Bonadonna RC, Saggiani F, Zenere MB, Monauni T, Muggeo M. Homeostasis model assessment closely mirrors the glucose clamp technique in the assessment of insulin sensitivity: studies in subjects with various degrees of glucose tolerance and insulin sensitivity. Diabetes Care 2000;23:57–63. [PubMed: 10857969]

10. Levy JC, Matthews DR, Hermans MP. Correct homeostasis model assessment (HOMA) evaluation uses the computer program. Diabetes Care 1998;21:2191–2192. [PubMed: 9839117]

11. Meigs JB, Mittleman MA, Nathan DM, Tofler GH, Singer DE, Murphy-Sheehy PM, Lipinska I, D'Agostino RB, Wilson PW. Hyperinsulinemia, hyperglycemia, and impaired hemostasis: the Framingham Offspring Study. JAMA 2000;283:221–228. [PubMed: 10634338]

12. Cupples, LA.; D'Agostino, RB. Section 34: Some Risk Factors Related to the Annual Incidence of Cardiovascular Disease and Death Using Pooled Repeated Biennial Measurements: Framingham Heart Study, 30-Year Followup. Washington, DC: US Department of Commerce; 1988.

13. DeLong ER, DeLong DM, Clarke-Pearson DL. Comparing the areas under two or more correlated receiver operating characteristic curves: a nonparametric approach. Biometrics 1988;44:837–845. [PubMed: 3203132]

14. Harrell, FEJ. Regression Modelling Strategies With Applications to Linear Models, Logistic Regression, and Survival Analysis. New York, NY: Springer; 2001.

15. Deeks JJ, Altman DG. Diagnostic tests 4: likelihood ratios. BMJ 2004;329:168–169. [PubMed: 15258077]

16. Hanley AJ, Karter AJ, Williams K, Festa A, D'Agostino RB Jr, Wagen-knecht LE, Haffner SM. Prediction of type 2 diabetes mellitus with alternative definitions of the metabolic syndrome: the Insulin Resistance Atherosclerosis Study. Circulation 2005;112:3713–3721. [PubMed: 16344402]

17. Meigs JB, Rutter MK, Sullivan LM, Fox CS, D'Agostino RB Sr, Wilson PW. Impact of insulin resistance on risk of type 2 diabetes and cardiovascular disease in people with metabolic syndrome. Diabetes Care 2007;30:1219–1225. [PubMed: 17259468]

18. Hedblad B, Nilsson P, Engstrom G, Berglund G, Janzon L. Insulin resistance in non-diabetic subjects is associated with increased incidence of myocardial infarction and death. Diabet Med 2002;19:470–475. [PubMed: 12060058]

19. Robins SJ, Rubins HB, Faas FH, Schaefer EJ, Elam MB, Anderson JW, Collins D. Insulin resistance and cardiovascular events with low HDL cholesterol: the Veterans Affairs HDL Intervention Trial (VA-HIT). Diabetes Care 2003;26:1513–1517. [PubMed: 12716814]

20. Meigs JB, Haffner SM, Nathan DM, D'Agostino RB, Wilson PW. Sample exchange to compare insulin measurements between the San Antonio Heart Study and the Framingham Offspring Study. J Clin Epidemiol 2001;54:1031–1036. [PubMed: 11576815]

21. Wilson PW, D'Agostino RB, Levy D, Belanger AM, Silbershatz H, Kannel WB. Prediction of coronary heart disease using risk factor categories. Circulation 1998;97:1837–1847. [PubMed: 9603539]

22. Wilson PW, Meigs JB, Sullivan L, Fox CS, Nathan DM, D'Agostino RB Sr. Prediction of incident diabetes mellitus in middle-aged adults: the Framingham Offspring Study. Arch Intern Med 2007;167:1068–1074. [PubMed: 17533210]

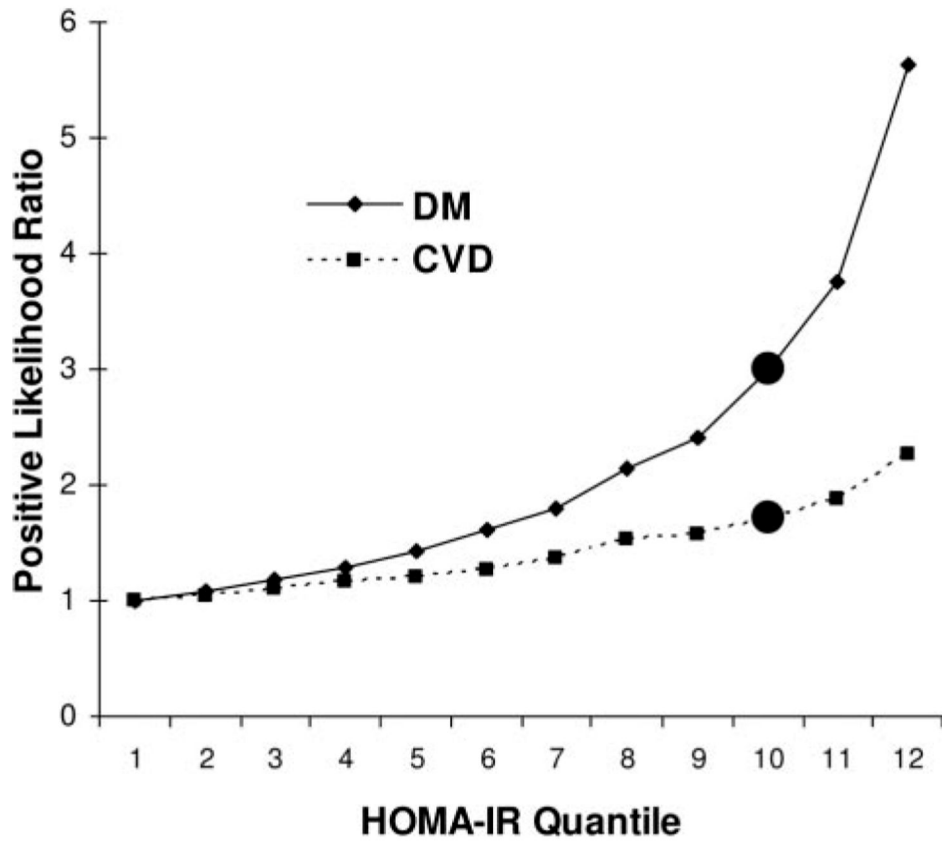**Figure 1.**
PLRs associated with 12 quantiles of HOMA-IR predicting incident DM or CVD. ● Indicates 76th centile point, at and above which subjects are "insulin resistant"; DM, 7-year incident DM; and CVD, 11-year incident CVD.

Figure 2.
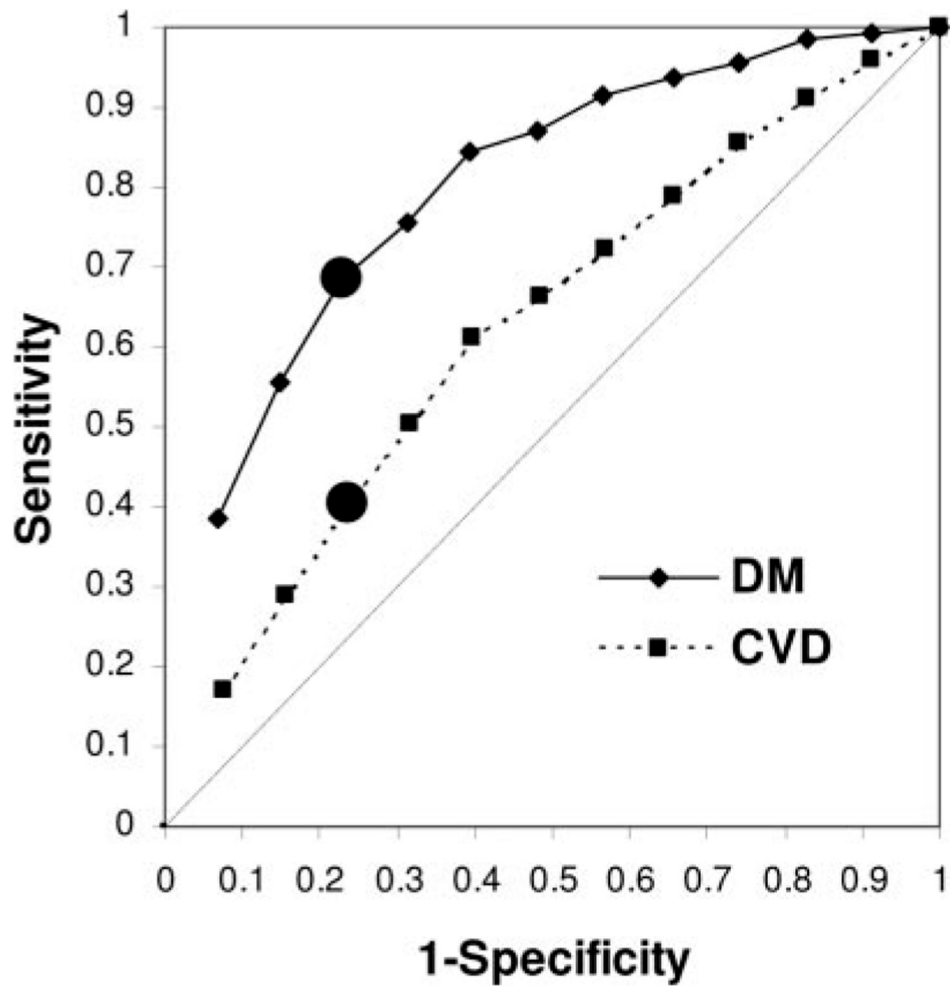aROC curves for HOMA-IR predicting incident DM or CVD. ● Indicates 76th centile point, at and above which subjects are "insulin resistant." At the 76th centile, sensitivity and specificity values for incident DM are 68% and 77%; for CVD, they were 40% and 76%, respectively. aROC (95% CI) was 0.80 (0.76 to 0.83) for DM and 0.63 (0.59 to 0.66) for CVD. DM indicates 7-year incident DM; CVD, 11-year incident CVD.

Table 1

Baseline Characteristics

Characteristic	Value
Age, y	53.5 (9.7)
Men, %	45
BMI, kg/m^2	27.1 (4.7)
Fasting glucose, mmol/L	5.2 (0.5)
2-h Glucose, mmol/L	5.8 (1.5)
No. with NGT	1864
No. with IFG and/or IGT	856
Fasting insulin, pmol/L	51.6 (44.5)
2-h Insulin, pmol/L	320.7 (275.6)
HOMA-IR	6.9 (2.9)
Gutt's ISI ($ISI_{0,120}$), mg · L^2/mmol · mU · min	26.6 (6.8)
7-y Cumulative incidence of DM, %	4.8
11-y Cumulative incidence of CVD, %	8.6

NGT indicates normal glucose tolerance; IFG, impaired fasting glucose; and IGT, impaired glucose tolerance.

Data are number, mean (SD), or %.

Table 2

DM and CVD Prediction: Test Performance by 12 Quantile Thresholds of Surrogate IR Measures

Outcome and Predictor	aROC (95% CI)	12 Quantiles of IR Measures											
		Insulin Sensitive									Insulin Resistant		
		1	**2**	**3**	**4**	**5**	**6**	**7**	**8**	**9**	**10**	**11**	**12**
DM, No. of events		1	1	4	2	3	6	3	12	9	17	22	50
Fasting insulin	0.73 (0.69–0.78)												
Sensitivity		100	97	94	90	88	84	80	75	66	60	47	31
Specificity		0	10	19	26	34	45	52	61	70	77	84	93
HOMA-IR	0.80 (0.76–0.83)												
Sensitivity		100	99	98	95	94	92	87	85	75	68	55	38
Specificity		0	9	17	26	35	43	52	61	69	77	85	93
1/ISI	0.79 (0.75–0.83)												
Sensitivity		100	100	98	95	92	90	85	82	75	65	57	38
Specificity		0	9	17	26	35	43	52	60	69	77	85	93
CVD, No. of events		10	11	13	16	15	14	12	26	23	27	28	40
Fasting insulin	0.62 (0.58–0.66)												
Sensitivity		100	94	87	84	80	71	68	58	47	40	29	15
Specificity		0	10	19	26	35	45	52	61	69	76	84	93
HOMA-IR	0.63 (0.59–0.66)												
Sensitivity		100	96	91	86	79	72	66	61	50	40	29	17
Specificity		0	9	17	26	34	43	51	60	68	76	85	92
1/ISI	0.60 (0.56–0.64)												
Sensitivity		100	95	91	86	80	69	61	52	43	36	27	19
Specificity		0	9	17	26	35	43	51	59	68	76	84	93

DM indicates 7-year incident DM; CVD, 11-year incident CVD.

Sensitivity and specificity data for each quantile, shown as percentages, were calculated with the lower boundary of each quantile as the threshold value. For example, the 10th quantile of HOMA-IR has a sensitivity of 68% for DM prediction. This is calculated from events that occur in subjects with HOMA-IR values between the lower boundary of the 10th quantile and the upper boundary of the 12th quantile. The lower boundary of the 10th quantile is the 76th centile of the HOMA-IR distribution, the usual threshold for defining IR.