BMJ Open Diabetes Research & Care

Enhanced prediction of abnormal glucose tolerance using an extended non-invasive risk score incorporating routine renal biochemistry

Jie He,[1] Baoqi Fan,[1,2] Eric S H Lau ,[1] Natural Chu,[1] Noel Yat Hey Ng,[1] Kathy Ho Ting Leung,[1] Emily W M Poon,[1] Alice Pik Shan Kong,[1,2,3] Ronald Ching Wan Ma ,[1,2,3] Andrea O Y Luk,[1,2,3,4] Juliana C N Chan,[1,2,3] Elaine Chow [1,2,3,4]

To cite: He J, Fan B, Lau ESH, *et al*. Enhanced prediction of abnormal glucose tolerance using an extended non-invasive risk score incorporating routine renal biochemistry. *BMJ Open Diab Res Care* 2024;**12**:e003768. doi:10.1136/bmjdrc-2023-003768

► Additional supplemental material is published online only. To view, please visit the journal online (https://doi.org/10.1136/bmjdrc-2023-003768).

Prior presentation. Parts of this work were presented in abstract from at the 2022 International Diabetes Federation Congress, Lisbon, Portugal, December 8, 2022.

Received 13 September 2023
Accepted 20 January 2024

For numbered affiliations see end of article.

Correspondence to
Dr Elaine Chow;
e.chow@cuhk.edu.hk

ABSTRACT

Introduction Type 2 diabetes is preventable in subjects with impaired glucose tolerance based on 2-hour plasma glucose (2hPG) during 75 g oral glucose tolerance test (OGTT). We incorporated routine biochemistry to improve the performance of a non-invasive diabetes risk score to identify individuals with abnormal glucose tolerance (AGT) defined by 2hPG≥7.8 mmol/L during OGTT.

Research design and methods We used baseline data of 1938 individuals from the community-based "Better Health for Better Hong Kong - Hong Kong Family Diabetes Study (BHBHK-HKFDS) Cohort" recruited in 1998–2003. We incorporated routine biochemistry in a validated non-invasive diabetes risk score, and evaluated its performance using area under receiver operating characteristics (AUROC) with internal and external validation.

Results The AUROC of the original non-invasive risk score to predict AGT was 0.698 (95% CI, 0.662 to 0.733). Following additional inclusion of fasting plasma glucose, serum potassium, creatinine, and urea, the AUROC increased to 0.778 (95% CI, 0.744 to 0.809, p<0.001). Net reclassification improved by 31.9% (p<0.001) overall, by 30.8% among people with AGT and 1.1% among people without AGT. The extended model showed good calibration (χ^2=11.315, p=0.1845) and performance on external validation using an independent data set (AUROC=0.722, 95% CI, 0.680 to 0.764).

Conclusions The extended risk score incorporating clinical and routine biochemistry can be integrated into an electronic health records system to select high-risk subjects for evaluation of AGT using OGTT for prevention of diabetes.

WHAT IS ALREADY KNOWN ON THIS TOPIC

⇒ Incorporation of laboratory markers such as fasting plasma glucose (FPG) and lipids improved the discrimination of non-invasive risk scores in predicting undiagnosed diabetes.

WHAT THIS STUDY ADDS

⇒ Type 2 diabetes can be prevented but a 75 g oral glucose tolerance test (OGTT) is needed to identify people with impaired glucose tolerance (IGT), the main beneficiaries of diabetes prevention programs. Inclusion of routine renal biochemistry (serum potassium, creatinine, and urea) and FPG to a validated non-invasive risk score can significantly discriminate subjects with abnormal glucose tolerance (2-hour plasma glucose ≥7.8 mmol/L) who will benefit from early intervention.

HOW THIS STUDY MIGHT AFFECT RESEARCH, PRACTICE OR POLICY

⇒ The incorporation of clinical and routine biochemistry in the extended risk score within an electronic health records system will improve the efficiency of using 75 g OGTT to identify subjects with IGT who will benefit from the diabetes prevention program.

INTRODUCTION

Type 2 diabetes (T2D) could be delayed and prevented by lifestyle modification and/or pharmacotherapy in individuals with impaired glucose tolerance (IGT).[1 2] In the China Daqing study, the rate of T2D conversion was reduced by 51% at 6 years via lifestyle modification which was translated to reductions in retinopathy and cardiovascular disease during 30-year follow-up.[1 3] In the latest China Diabetes Prevention Program, the combined use of lifestyle modification and metformin reduced progression to diabetes by 18% compared with lifestyle modification only, mainly in subjects with IGT.[4] The majority of Chinese individuals have isolated IGT.[5] A significant proportion of subjects with diabetes mellitus are only identified by abnormal 2-hour plasma glucose (2hPG) during 75 g oral glucose tolerance test (OGTT), with normal fasting plasma glucose (FPG) and HbA1c (glycated hemoglobin A1c).[6 7] Apart from progression to diabetes,

elevated post-OGTT plasma glucose strongly predicts cardiovascular events.[8 9] Given the resource implication in performing OGTT, developing a method to identify subjects with abnormal glucose tolerance (AGT) including IGT and undiagnosed diabetes will reduce the number of OGTT and improve the cost-effectiveness of diabetes prevention and early intervention programs.

There are only a few non-invasive risk scores used to identify Chinese subjects at high risk of having diabetes and pre-diabetes based on 75 g OGTT. With a modified cut-off point of 27 out of 51, the New Chinese Diabetes Risk Score (NCDRS, including age, sex, body mass index (BMI), systolic blood pressure, waist circumference, family history of diabetes) was used to predict undiagnosed diabetes and pre-diabetes based on the American Diabetes Association (ADA) criteria with a sensitivity of 0.694 and positive predictive value (PPV) of 36.2%.[10] The Risk-Understanding-By-Yourself (RUBY) risk score consists of self-assessment items (including age, sex, BMI, family history of diabetes, gestational diabetes, hypertension, and dyslipidemia) which had been validated in Chinese adults of working age for predicting T2D.[11]

Many individuals with diabetes or pre-diabetes have concomitant conditions such as hypertension, dyslipidemia, or cardiovascular disease managed in an electronic health records (EHR) system. These individuals have regular blood tests which may provide opportunities for early diagnosis of dysglycemia. In the Atherosclerosis Risk in Communities (ARIC) study, the area under receiver operating characteristics (AUROC) of a diabetes risk score increased from 0.71 to 0.80 (p<0.001) when FPG and lipids were added to non-invasive measurements.[12]

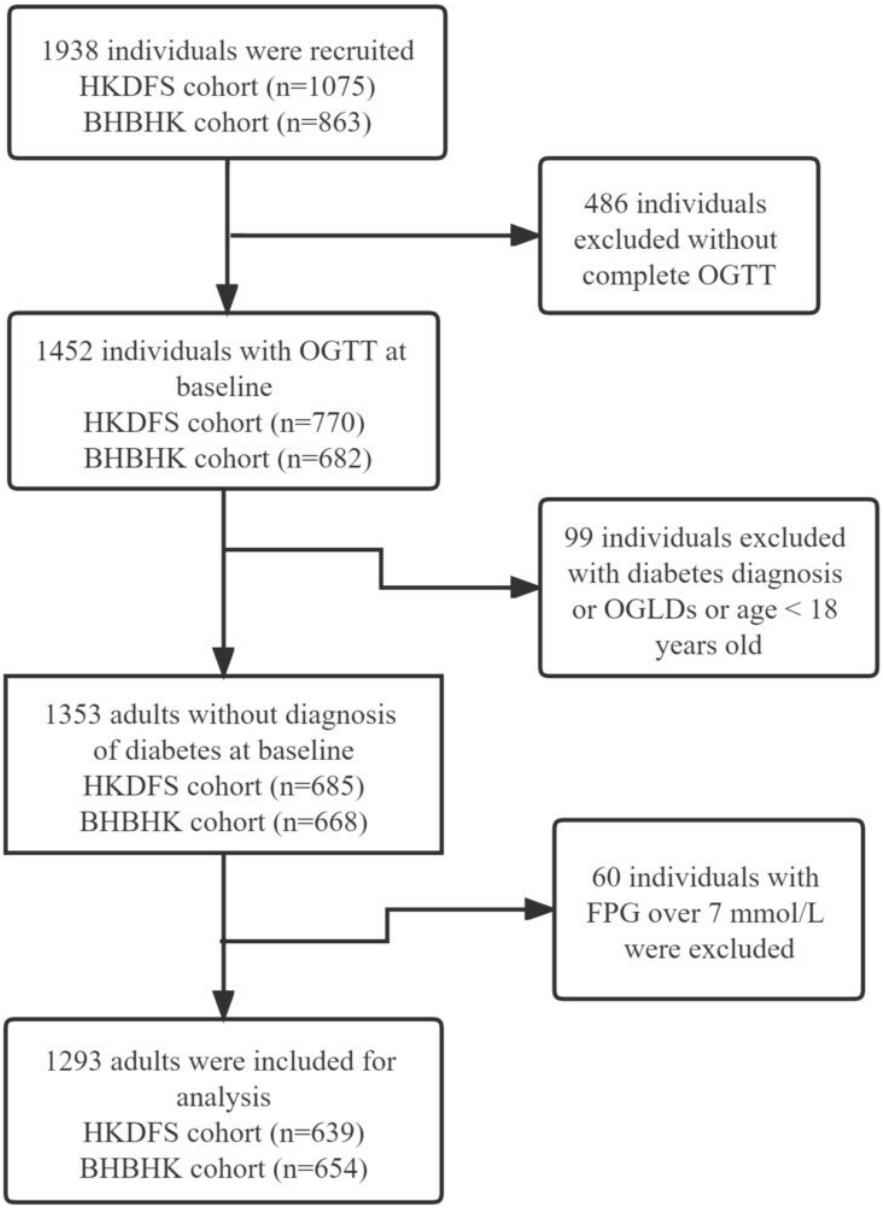


Figure 1 Flow chart of the study. BHBHK, Better Health for Better Hong Kong; FPG, fasting plasma glucose; HKDFS, Hong Kong Family Diabetes Study; OGTT, oral glucose tolerance test; OGLD, oral glucose lowering drugs.

In the Framingham Offspring study, the inclusion of FPG, high-density lipoprotein cholesterol (HDL-c), and triglycerides (TG) increased the AUROC of a diabetes risk score from 0.72 to 0.85.[13] In the European Prospective Investigation into Cancer and Nutrition (EPIC)-Potsdam study, the addition of FPG, lipids and liver enzymes to non-invasive measurements improved the discriminative power of a diabetes risk score.[14] In a prediction model for diabetes in Chinese, the inclusion of FPG, lipids and white blood cell (WBC) count also improved its performance.[15] Other risk scores that included laboratory measurements were used to predict AGT (diabetes and IGT). In the National Health and Nutritional Examination Survey, a risk score including clinical risk factors, FPG and TG had a total score of 8. The lowest cut-off point of 2 yielded high sensitivity (0.86) and the highest cut-off point of 6 yielded high specificity (0.96) with an AUROC of 0.74 for predicting AGT.[16]

Few studies to date have evaluated the addition of routine laboratory measurements, especially renal parameters, in improving discrimination for AGT in the Chinese population. In this study, we investigated whether the inclusion of routine biochemical measurements into a validated non-invasive diabetes risk score could enhance the performance and discrimination of AGT in community-dwelling Chinese adults.

METHODS

Study population

A total of 1,938 community-dwelling subjects were curated from the Better Health for Better Hong Kong (BHBHK) Survey and the Hong Kong Family Diabetes Study (HKFDS). Both cohorts were established in 1998–2003 to define the phenotypes of diabetes in Hong Kong Chinese adults. The BHBHK cohort was established during a city-wide health promotion and screening campaign targeting the workforce.[17] The HKFDS recruited first-degree (parents and children) and second-degree relatives of patients with T2D, the majority of whom were diagnosed before 40 years old.[18] In this cross-sectional study, we included 1,452 subjects who underwent OGTT in the BHBHK-HKFDS cohort. We excluded participants with known diagnoses of diabetes or using glucose-lowering drugs or less than 18 years old (n=99). We excluded 60 subjects with undiagnosed diabetes based on FPG≥7.0 mmol/L. A number of 1,293 individuals were included in this analysis (figure 1).

Data collection

All participants in the BHBHK-HKFDS cohort underwent structured assessments including anthropometric measurements, demographic status, medical history, lifestyle information, clinical measurements, and laboratory tests. Age, sex, body height, and weight (for calculation of BMI), waist circumference, blood pressure, smoking status, family history of diabetes, and medical conditions including hypertension, dyslipidemia, and gestational diabetes were documented. After an overnight 8-hour fast, all participants underwent standard 75 g OGTT with measurement of PG at 0, 60, and 120 min. Lipid profiles (total cholesterol (TC), TG, HDL-c and calculated low-density lipoprotein cholesterol (LDL-c)), renal function tests (serum sodium, potassium, urea, creatinine levels, and urinary albumin creatinine ratio (uACR)), total protein and albumin were also measured.

All laboratory assays with external accreditation were performed at the Departments of Chemical Pathology and Haematology Laboratory of the Prince of Wales Hospital. Plasma glucose was measured by a glucose oxidase method (Diagnostic Chemicals reagent Kit). Plasma TC and TG were measured enzymatically with commercial reagents (Dimension; DuPont Instruments, Delaware, USA) and HDL-c was measured by the same enzymatic assay after precipitation of HDL-c by the heparin/manganese method. LDL-c was calculated using Friedewald *et al* equation.[19]

External validation

We used data from 550 subjects without diabetes diagnosis or FPG≥7.0 mmol/L for external validation. This included data was from three diabetes prevention-related studies, a lifestyle modification study (NCT04588896, n=324), a polycystic ovary syndrome follow-up study,[20] (n=158) and a study evaluating traditional Chinese medicine in people with pre-diabetes (NCT04441216, n=68). We used anthropometric characteristics, medical history, and biochemical parameters collected before any intervention. All subjects similarly had a 75 g OGTT with FPG and 2hPG measurements after an 8-hour overnight fast. -

Statistical analysis

Data were summarized as mean±SD or median (IQR) as appropriate. We compared anthropometric and demographic characteristics, as well as the laboratory parameters between the three groups (stratified by 2hPG<7.8, 7.8–11.0, and ≥11.1 mmol/L) using analysis of variance or Kruskal-Wallis test for continuous data, and χ^2 test for categorical data.

Construction of the extended RUBY risk score with laboratory variables

The original RUBY included age, sex, BMI, family history of diabetes, hypertension, dyslipidemia, and gestational diabetes in the non-invasive risk score. Hypertension was defined as known hypertension with or without taking anti-hypertensive drugs, or blood pressure ≥140/90 mm Hg. Dyslipidemia was defined as the history of any lipid abnormalities (including TC≥5.2 mmol/L, TG≥2.3 mmol/L, and/or HDL-c<1.0 mmol/L) with or without lipid-modifying drugs. The original RUBY score was derived from community-dwelling adults in Hong Kong with AUROC 0.735 (0.705 to 0.765) for the prediction of diabetes. A cut-off point of 16 out of 30 had 40% PPV for detecting undiagnosed diabetes. It was externally validated in two cohorts of 3,743 subjects with known risk

factors for diabetes with an AUROC of 0.681 and another cohort (n=1,513) with an AUROC of 0.772.[11]

To identify laboratory variables for inclusion in the extended RUBY models, we used a forward stepwise logistic regression to select relevant laboratory predictors for AGT. These included renal function tests (serum levels of sodium, potassium, creatinine, urea, and uACR), total protein, albumin, and complete blood test (red blood cell, platelet, WBC). Multicollinearity diagnostics were tested among the final chosen prediction factors by comparing the variance inflation factor (VIF). A VIF≥10 indicates multicollinearity. The importance of each risk factor was estimated by the partial χ^2 statistic minus the predictor df (Wald-df). The best model was defined by the Bayesian information criterion (BIC), which had a stronger penalty for additional parameters than the Akaike information criteria.

10-fold cross-validation for internal validation

We performed 10-fold cross-validation of the extended RUBY model within the BHBHK-HKDFS cohort for internal validation. First, the whole data was randomly separated into 10-folds. Second, in the first loop, an arbitrary onefold data was selected as the first test set to examine the first trained model built by the other ninefold data to yield the first receiver operating characteristic (ROC). Third, we ran the second loop. The next fold data was set as the test set and the remaining ninefold data was used to build the trained model to generate the second ROC curve. During the 10 loops, each fold had the opportunity to be the tested data to yield their respective ROCs. This method was used to control overfitting and an AUROC of 0.7 to 0.8 was considered to be good to excellent. Kappa statistic is a measurement of the accuracy of a model while considering randomness. The closer to 1, the better the model with 0.4–0.6 regarded as having moderate performance.[21]

Comparison of extended models against other non-invasive risk scores for AGT

The discriminative value was assessed using AUROC. Integrated Discrimination Index (IDI) and net reclassification improvement (NRI)[22] were used to compare the extended models and the original RUBY score. In this study, the AUROC and the IDI were used to compare the performance of different models. A comparison of AUROCs was done using the Delong test. NRI was used to assess the improvement of the best extended RUBY model compared with the original RUBY score. We defined three risk categories a priori: low risk: <30%, intermediate: 30–60%, and high risk: >60%. Calibration measures the agreement between predicted probabilities and observed risks and was assessed by the Hosmer-Lemeshow test. Sensitivity, specificity, PPV, and negative predictive value (NPV) at different cut-off points were presented. We assessed screening efficiency by defining the number needed to screen (NNS) as the number of OGTTs needed to be performed to detect one case of AGT=1/positive case detection rate of AGT.

Decision curve analysis (DCA) was used to assess the utility of models, including the default strategies of treating all or no patients for decision-making. It combines the number of true positives and false positives into a single one as a net benefit to determine whether the model is superior to any other models across the full range of reasonable threshold probabilities.[23 24] A risk threshold of 0.25 was used based on the prevalence of IGT and T2D based on the report of the International Diabetes Federation 2021,[25] which is consistent with the percentage of AGT in our cohort.

We compared the performance of the extended RUBY models versus the NCDRS which included age, gender, BMI, waist circumference, hypertension, and family history of diabetes as risk predictors. The NCDRS was derived from 7,675 community residents of Eastern China. The NCDRS used a cut-off point of 27 out of 51 to predict undiagnosed diabetes and prediabetes.[10] Youden Index was used to determine the optimal cut-off for maximal sensitivity and specificity for the extended RUBY models to estimate the proportion AGT in the extended models (0–100%). We used the original RUBY to predict AGT with a modified cut-off point determined by the Youden Index. We compared the discrimination expressed as sensitivity, specificity, PPV, and NPV of the extended RUBY models using different thresholds for risk prediction. We also compared the extended RUBY models versus the NCDRS.

A p value<0.05 was considered significant. The analysis was conducted using IBM SPSS Statistics V.23.0 and R software V.4.1.1.

RESULTS

Participant characteristics

Among the 1,293 adults without known diabetes or FPG≥7.0 mmol/L, 317 had AGT, including 85 with 2hPG≥11.1 mmol/L and 232 with 2hPG between 7.8 mmol/L and 11.0 mmol/L (figure 1). Anthropometric measurements were compared among the three groups with 2hPG<7.8, 7.8–11.0, ≥11.1 mmol/L. The 2hPG≥11.1 mmol/L were the oldest, had the highest BMI, and highest rate of family history of diabetes. They were most likely to be active smokers and treated with anti-hypertensive drugs (all p<0.001). They also had higher plasma total protein, TC, TG, LDL-c, and uACR but lower HDL-c and serum potassium (p=0.016) (table 1).

Selection of variables in the extended model

Apart from variables in the original RUBY score, FPG, serum creatinine, WBC count, serum urea, potassium, and total protein were significant in the multivariable logistic regression model without collinearity (online supplemental table S1).

Table 1 Cross-sectional participant characteristics in Better Health for Better Hong Kong_Hong Kong Family Diabetes Study cohort

	Total	Non-IGT	AGT		
	(n=1293)	2hPG<7.8 (n=976)	2hPG 7.8–11.0 (n=232)	2hPG≥11.1 (n=85)	P value
Age (years)	41.9±12.0	40.2±11.5	46.2±12.1	49.9±12.1	**<0.001**
Female, n (%)	728 (56.3)	542 (55.5)	138 (59.5)	48 (56.5)	0.552
Weight (kg)	62.1±12.2	61.4±11.8	64.2±13.1	65.3±13.9	**<0.001**
Height (cm)	161.1±8.8	161.6±8.8	159.5±8.5	159.1±8.7	**<0.001**
BMI (kg/m^2)	23.9±3.9	23.4±3.7	25.1±4.1	25.6±4.1	**<0.001**
Waist circumference (cm)	78.9±10.4	77.6±10.0	82.5±10.4	84.3±10.8	**<0.001**
Female	74.7±9.4	73.2±9.0	79.3±9.1	79.1±9.4	**<0.001**
Male	84.3±9.1	83.1±8.4	87.1±10.5	91.4±8.4	**<0.001**
Hip (cm)	95.0±7.1	94.5±6.8	97.0±7.8	96.7±7.6	**<0.001**
SBP (mm Hg)	120.3±19.3	117.1±17.6	128.1±19.7	136.2±23.9	**<0.001**
DBP (mm Hg)	75.0±12.0	73.8±11.3	78.1±12.3	81.1±15.3	**<0.001**
DM FH, n (%)	721 (55.8)	516 (52.9)	149 (64.2)	56 (65.9)	**<0.001**
Smoking, n (%)					**0.028**
Current smoker	159 (12.3)	120 (12.3)	26 (11.2)	13 (15.3)	
Ex-smoker	49 (3.8)	28 (2.9)	16 (6.9)	5 (5.9)	
Use of antihypertensives, n (%)	58 (4.5)	19 (1.9)	19 (8.2)	20 (23.5)	**<0.001**
Glycemic measures					
FPG (mmol/L)	5.0±0.6	4.8±0.5	5.2±0.6	5.9±0.7	**<0.001**
1hPG (mmol/L)	8.7±2.7	7.7±2.2	10.9±1.8	13.6±1.7	**<0.001**
2hPG (mmol/L)	6.6±2.4	5.5±1.2	9.0±0.9	12.5±1.4	**<0.001**
Routine parameters					
Albumin (g/L)	41.1±3.1	41.2±3.2	41.0±3.0	40.8±2.6	0.405
Total protein (g/L)	76.2±4.3	75.9±4.3	77.0±4.6	77.4±3.9	**<0.001**
Total cholesterol (mmol/L)	5.2±1.0	5.1±1.0	5.4±1.0	5.5±0.9	**<0.001**
Triglyceride (mmol/L)	1.0 (0.8)	0.9 (0.7)	1.3 (0.9)	1.4 (1.0)	**<0.001**
HDL-c (mmol/L)	1.5±0.4	1.5±0.4	1.4±0.4	1.3±0.3	**<0.001**
LDL-c (mmol/L)	3.1±0.9	3.0±0.9	3.3±0.9	3.4±0.8	**<0.001**
Sodium (mmol/L)	140.9±1.9	140.8±1.8	141.0±1.8	141.2±2.0	0.065
Potassium (mmol/L)	4.09±0.35	4.10±0.34	4.09±0.37	3.97±0.36	**0.016**
Serum creatinine (µmol/L)	74.0±17.4	74.6±17.3	72.5±17.9	70.8±18.0	0.067
Urea (mmol/L)	5.0±1.26	5.0±1.2	4.9±1.2	4.8±1.17	0.320
uACR (mg/mmol)	0.8 (1.1)	0.7 (1.0)	0.9 (1.7)	1.3 (2.2)	**<0.001**

Data were shown as mean±SD or median (IQR) or number (%). The one-way analysis of variance test or Kruskal-Wallis test was used for continuous variables, χ^2 test was used for categorical variables.
Bold values signifies p <0.05
AGT, abnormal glucose tolerance; BMI, body mass index; DBP, diastolic blood pressure; DM FH, family history of diabetes; FPG, fasting plasma glucose; HDL-c, high-density lipoprotein cholesterol; 1hPG, 1-hour post-challenge glucose; 2hPG, 2-hour post-challenge glucose; IGT, impaired glucose tolerance; LDL-c, low-density lipoprotein cholesterol; PLT, platelet; RBC, red blood cells; SBP, systolic blood pressure; uACR, urinary albumin creatinine ratio; WBC, white blood cells.

The performance of the original RUBY score in predicting AGT (2hPG≥7.8 mmol/L) was 0.698 (95% CI, 0.662 to 0.733). We constructed extended RUBY model 1 and added FPG to the original RUBY with an increased AUROC to 0.756 (95% CI, 0.722 to 0.791, p<0.001) and improvement of IDI by 20.1% (p<0.001). In extended RUBY model 2, we included serum creatinine, potassium, and urea to extended RUBY model 1 with an AUROC of 0.778 (95% CI, 0.744 to 0.812, p=0.002) and improved IDI by 6.4% (p<0.001). However, when further including WBC or total protein to extended RUBY model 2, the AUROC

Table 2 Beta-coefficients and OR of extended RUBY models in predicting AGT

	Extended model 1		Extended model 2	
	β	OR (95% CI)	β	OR (95% CI)
Original RUBY	0.097	1.10 (1.07 to 1.14)	0.122	1.13 (1.09 to 1.17)
FPG (in mmol/L)	1.329	3.78 (2.83 to 5.10)	1.449	4.26 (3.15 to 5.82)
Serum potassium (in mmol/L)			−0.580	0.56 (0.35 to 0.89)
Serum urea (in mmol/L)			−0.210	0.81 (0.70 to 0.93)
Serum Creatinine (in µmol/L)			−0.019	0.98 (0.97 to 0.99)

FPG range, 3.6–6.9 mmol/L; potassium range, 3.0–5.5 mmol/L; urea range, 1.6–9.0 mmol/L; creatinine range, 30–141 µmol/L.
Extended RUBY model 1=0.097 × original RUBY + 1.329 × FPG – 8.767.
Extended RUBY model 2=0.122 × original RUBY + 1.449 × FPG – 0.580 × potassium – 0.210 × urea – 0.019 × creatinine – 4.755.
AGT, abnormal glucose tolerance; FPG, fasting plasma glucose; RUBY, Risk-Understanding-By-Yourself.

and IDI were not significantly improved (online supplemental table S2).

Although the extended RUBY model with FPG, creatinine, potassium, urea, and WBC had the lowest BIC (online supplemental table S3), however, since most individuals do not always have both WBC and renal biochemistry measured in routine clinical settings, the optimal model including RUBY score, and biochemistry parameters of FPG, creatinine, potassium and urea was chosen for practical purposes (table 2).

Discrimination of extended RUBY models compared with the original RUBY score

Figure 2 shows ROC curves of four models. Compared with the original RUBY score, the extended RUBY models including FPG showed better discrimination. The extended RUBY model 2 including additional potassium, urea, and creatinine showed best. There was no difference in discrimination of AGT between the model with FPG only and the original RUBY risk score. The AUROC for a model containing RUBY+creatinine+urea+potassium without FPG was 0.725 (95% CI, 0.690 to 0.761) which still showed improved discrimination compared with RUBY alone (IDI 0.047 (95% CI, 0.024 to 0.069, p<0.001) online supplemental table S2).

In the DCA, the extended RUBY model 2 achieved the highest net benefit across a range of risk thresholds for AGT. At the risk probability of 0.25, the net benefit was superior in extended RUBY model 2 than the original RUBY, and the extended model 1 was intermediate (online supplemental figure S5).

Calibration, re-classification and cross-validation compared with an original RUBY score

Online supplemental figure S1 compares observed and predicted probabilities of AGT by extended RUBY model 2 as assessed by the Hosmer-Lemeshow test with good calibration (χ^2=11.315, p=0.1845, F=8). After 10,000 bootstraps, compared with the original RUBY model, the extended RUBY model improved net reclassification by 31.9% (p<0.001), with 30.8% among persons with AGT and 1.1% among persons without AGT (online supplemental figure S2). Online supplemental figure S3 shows 10-fold cross-validation of the extended RUBY model 2. The average AUROC was 0.814 and the kappa was 0.415 indicating no overfitting in the extended RUBY model 2 construction.

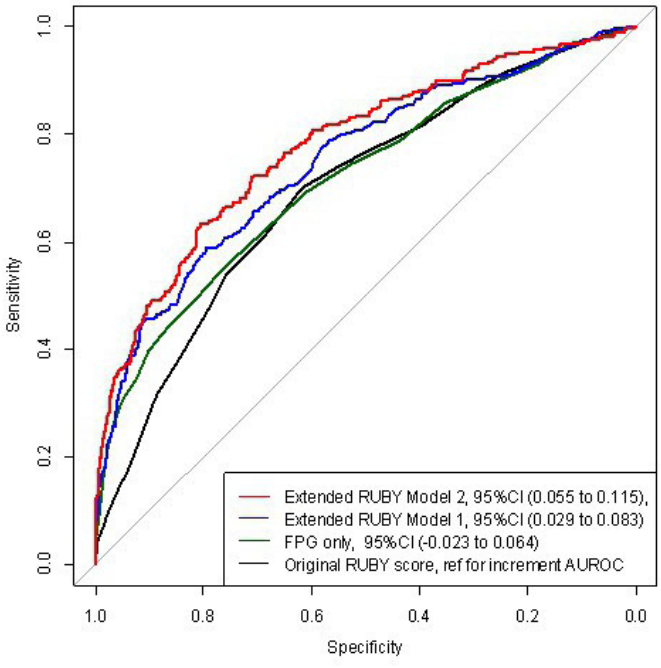


Figure 2 Receiver operating characteristic curves of the performance of original RUBY, FPG only, extended RUBY model 1 and 2 in predicting abnormal glucose tolerance in the Better Health for Better Hong Kong-Hong Kong Family Diabetes Study cohort. AUROC, area under receiver operating characteristics; FPG, fasting plasma glucose; RUBY, Risk-Understanding-By-Yourself.

Table 3 Performance of two extended RUBY models in predicting AGT and compared with NCDRS at optimal cutoffs in the Better Health for Better Hong Kong_Hong Kong Family Diabetes Study cohort (n=1054)

Model	AUROC (95% CI)	Sensitivity	Specificity	PPV	NPV	NNS
Original RUBY score	0.698 (0.662 to 0.733)	0.485	0.785	0.424	0.822	2.4
Extended RUBY model 1	0.756 (0.722 to 0.791)	0.588	0.794	0.484	0.855	2.1
Extended RUBY model 2	0.778 (0.744 to 0.812)	0.635	0.805	0.516	0.871	1.9
NCDRS	0.648 (0.609 to 0.687)	0.603	0.693	0.389	0.843	2.6

Optimal risk cut-offs: original RUBY score (12 out of 30), extended RUBY model 1 (risk threshold, 29.4%), and extended RUBY model 2 (risk threshold, 29.5%) for predicting AGT were determined by the Youden Index. NCDRS with cut-off points 27 for predicting pre-diabetes and diabetes based on ADA criteria. Extended RUBY models used corresponding absolute risk in percentage as the optimal risk cutoffs while NCDRS and original RUBY score provided risk score.
ADA, American Diabetes Association; AGT, abnormal glucose tolerance; AUROC, area under receiver operating characteristics; NCDRS, New Chinese Diabetes Risk Score; NNS, number needed to screen; NPV, negative predictive value; PPV, positive predictive value; RUBY, Risk-Understanding-By-Yourself.

Comparison of extended model 2 against other non-invasive risk scores

We compared the performance of extended RUBY model 2 versus other non-invasive risk scores. We selected the cut-off value using the Youden Index to yield high PPV and sensitivity (online supplemental table S4). With a modified cut-off point of 12 out of 30, the original RUBY had a sensitivity of 0.485 and 2.4 NNS. The NCDRS had a sensitivity of 0.603 and 2.6 NNS. The extended RUBY model 2 achieved the highest sensitivity of 0.635 with NNS 1.9, as well as the best specificity, PPV, and NPV (table 3).

External validation

For external validation, we included 550 adults at high risk for diabetes but without known diabetes or FPG≥7.0 mmol/L during 75 g OGTT. Their mean (±SD) age was 53±10 years and their BMI was 26.1±4.7 kg/m^2 (online supplemental table S5). Of these, 242 participants had AGT with 2hPG≥7.8 mmol/L. In the external validation set, the AUROC for the extended RUBY model 1 was 0.716 (95% CI, 0.674 to 0.759), while the extended RUBY model 2 (including FPG, serum potassium, urea, creatinine) achieved an AUROC of 0.722 (95% CI, 0.680 to 0.764). Both extended RUBY models had higher AUROCs than the original RUBY of 0.637 (95% CI, 0.591 to 0.683). Applying NCDRS to the external cohort yielded an AUROC of 0.632 (95% CI, 0.586 to 0.679) (online supplemental table S6). The Hosmer and Lemeshow test showed good calibration (χ^2=3.972, p=0.860, F=8) of the extended RUBY model 2 in the external cohort (online supplemental figure S4).

DISCUSSION AND CONCLUSIONS

The efficiency of the implementation of diabetes prevention and early treatment program relies on the effective identification of individuals with AGT (2hPG≥7.8 mmol/L) detected by OGTT. In this study, we have shown that the inclusion of renal parameters to validated non-invasive risk scores can improve the reclassification of AGT by 31.9%. To our knowledge, this is the first study to validate a model incorporating clinical risk factors, FPG, and renal parameters to predict AGT in Chinese people.

By adding FPG to the original RUBY model, we can substantially enhance its performance in predicting AGT. Other studies have also shown that clinical parameters combined with FPG and lipid profiles achieved the best performance in diabetes or IGT prediction.[12 14 16] We did not include lipid profiles in our extended RUBY models as the history of dyslipidemia was already contained in the original RUBY score.

Additional inclusion of serum potassium, creatinine, and urea on top of FPG further improved the performance of extended RUBY model 2 with low serum potassium being associated with an increased risk of AGT. In the 9-year prospective ARIC study, adults with serum potassium levels lower than 4.0 mmol/L had an adjusted HR of incident diabetes of 1.64 (95% CI, 1.29 to 2.08) compared with those with high-normal serum potassium level (5.0–5.5 mmol/L).[26] In the Toranomon Hospital Health Management Center Study 1 (TOPICS 1), during 5-year follow-up, the lowest tertile of serum potassium (2.8–3.9 mmol/L) was independently associated with incident diabetes with an HR of 1.57 (95% CI, 1.15 to 2.15) compared with the highest tertile (4.2–5.4 mmol/L) in Japanese.[27] Low serum potassium may be associated with poorer beta cell function and impaired insulin secretion.[28] In addition, potassium homeostasis is regulated by dietary intake and renal excretion, the activity of the renin-angiotensin-aldosterone system, and transmembrane cellular shift which can be activated by hyperinsulinemia and sympathetic nervous system.[29 30] Potassium supplements and herbal medicines, such as licorice, might affect serum potassium. Mineralocorticoid antagonist and potassium-sparing diuretics have been associated with a lower risk of diabetes[31] while thiazide diuretics have been associated with higher risk. We did not capture the details of medications or dietary intake in our cohort. However in this cohort of working age, 4.5% of subjects reported using anti-hypertensive drugs which was less likely to contribute to the risk association of low serum potassium and AGT in our study.

Mechanistically, serum creatinine and urea are inversely associated with a higher risk of AGT via different pathways. The lower serum creatinine might reflect hyperfiltration as an early marker of nephropathy which has been reported in people with pre-diabetes.[32 33] Serum creatinine is a product of muscle catabolism. A low muscle mass might be a proxy of low physical activity, poor peripheral insulin sensitivity, and glucose disposal. In a 7-year prospective study including 36,304 Koreans who did not have diabetes, the lowest sex-specific skeletal muscle index tertile measured by bioelectrical impedance was associated with increased risk of incident T2D (HR 1.31, 95% CI, 1.18 to 1.45).[34] On the other hand, urea is the primary metabolite derived from dietary protein and tissue protein turnover. A higher serum urea might reflect higher dietary protein intake, which has been associated with a lower risk of development of diabetes in some randomized controlled feeding studies.[35 36] Taken together, the association of renal parameters notably low serum potassium with AGT are biologically plausible and warrant further investigations. Future studies may examine the association between serum potassium, creatinine, and urea and body composition as well as measured insulin resistance.

These routine biochemical tests are commonly performed among people receiving medical care in clinical settings. By incorporating clinical and laboratory measurements to automatically generate risk scores, such as the extended RUBY model 2, we can increase the efficiency of diagnosing AGT based on OGTT. In a separate territory-wide EHR-based Hong Kong Diabetes Surveillance Database with 1,028,931 people with pre-diabetes, 96% of individuals had an FPG, 96% had routine biochemistry testing performed, and 93% had both FPG and serum creatinine before pre-diabetes diagnosis.[37] Since these are routinely captured data, this would not incur additional laboratory or blood-taking costs. Application of extended RUBY 2, as compared with original RUBY would reduce the number needed to screen from 2.4 to 1.9 (PPV 51.6% vs 42.4%) to diagnose one case of AGT. Although a formal cost-effectiveness analysis is beyond the scope of this paper, assuming a hypothetical population of 1 million, OGTT costs of $200 Hong Kong dollars (US$25) per test, and a prevalence of AGT of 24%, we estimate the extended RUBY model 2 would reduce the cost per case identified by 19%.

Our study has several limitations. First, our sample size was small although external validation of the extended RUBY model 2 showed good discrimination. Second, HbA1c was used for diagnosis of diabetes only since 2003 while the BHBHK-HKFDS cohort was established in 1998–2003. Thus, baseline HbA1c was not available in our analysis. However, the use of HbA1c to diagnose AGT is challenging in resource-restrained settings, apart from issues related to standardization of assays, there are other confounding factors such as ethnicity, anemia, hemoglobinopathy, or other coexisting conditions.[38] In our study, the BHBHK was a territory-wide community-based survey that covered 50% of the working population including manual and unskilled workers.[17] Nevertheless, selection criteria based on OGTT may have led to a potential selection bias, favoring participants with higher adherence to medical protocols. Selecting such a cohort might lead to an underestimation of AGT prevalence, as healthier habits could underestimate the incidence of glucose intolerance. Consequently, the actual prevalence in a more general population might be higher. Our study was conducted in middle-aged Chinese community-dwelling individuals which might limit its generalizability to other ethnicities or age groups.

In summary, the inclusion of routine renal parameters (serum potassium, urea, creatinine) and FPG, greatly improved the performance of a non-invasive risk score in identification of subjects with AGT. Pending further studies, the use of our extended RUBY model may improve the yield of using OGTT to diagnose subjects with IGT and diabetes with high 2hPG for early intervention.

Author affiliations
[1]Department of Medicine and Therapeutics, The Chinese University of Hong Kong, Hong Kong Special Administrative Region, China
[2]Li Ka Shing Institute of Health Sciences, The Chinese University of Hong Kong Faculty of Medicine, Hong Kong Special Administrative Region, China
[3]Hong Kong Institute of Diabetes and Obesity, The Chinese University of Hong Kong Faculty of Medicine, Hong Kong Special Administrative Region, China
[4]Phase 1 Clinical Trial Centre, The Chinese University of Hong Kong, Hong Kong Special Administrative Region, China

Acknowledgements We thank all study participants, physicians and nurses at the Diabetes and Endocrine Centre at the Prince of Wales Hospital. Special acknowledgment is extended to Ms Cherry Chiu, the nursing officer who coordinated the recruitment of the BHBHK-HKFDS Cohorts.

Contributors JH, BF, ESHL, NC, NN, KHTL and EWMP collected the data. JH, EC and JCNC drafted the manuscript. EC conceived and designed the manuscript. JH, BF and ESHL analyzed and interpreted the data. All authors critically reviewed the manuscript and provided final approval of the version to be published. EC is the guarantor of the work and, as such, had full access to all the data in the study and takes responsibility for the integrity of the data and the accuracy of the data analysis.

Funding The establishment of the cohorts in this analysis was supported by the Li Ka Shing Charitable Foundation, Asia Diabetes Foundation, CUHK Research Sustainability for RGC Research Schemes, Health and Medical Research Fund (Ref 17180431), Azziz-Baumgartner Young Travel Award from the Androgen Excess and PCOS Society.

Competing interests EC has received research grants and/or honoraria for lectures from Sanofi, Lee Powder, Medtronic Diabetes, and Novartis. JCNC has received research grants and/or honoraria for consultancy and/or giving lectures from AstraZeneca, Bayer, Boehringer Ingelheim, Celltrion, Eli-Lilly, Hua Medicine, Powder Pharmaceuticals, Merck Serono, Merck Sharp & Dohme, Pfizer, Servier, Sanofi and Viatris. AL has served as an advisory committee member for AstraZeneca, Boehringer Ingelheim, MSD, Novartis, Novo Nordisk, Sanofi and Amgen. RCWM has received research grants for clinical trials from AstraZeneca, Bayer, MSD, Novo Nordisk, Sanofi, Tricida and honoraria for consultancy or lectures from AstraZeneca, and Boehringer Ingelheim. APSK has received research grants and/or speaker honoraria from Abbott, AstraZeneca, Bayer, Boehringer Ingelheim, Kyowa Kirin, Eli-Lilly, Merck Serono, Nestle, Novo Nordisk, Pfizer and Sanofi. All other authors declare that they have no competing interests.

Patient consent for publication Not applicable.

Ethics approval The study has received ethical approval from the Joint New Territories East Cluster-Chinese University of Hong Kong ethics committee (CREC 2009.421). Participants gave informed consent to participate in the study before taking part. The studies have also received local ethical approvals (2019.310).

Provenance and peer review Not commissioned; externally peer reviewed.

Data availability statement Data are available upon reasonable request. Anonymised data are available upon reasonable request to the corresponding author.

ORCID iDs
Eric S H Lau http://orcid.org/0000-0003-1581-5643
Ronald Ching Wan Ma http://orcid.org/0000-0002-1227-803X
Elaine Chow http://orcid.org/0000-0002-4147-3387

REFERENCES

1 Pan XR, Li GW, Hu YH, *et al*. Effects of diet and exercise in preventing NIDDM in people with impaired glucose tolerance. *Diabetes Care* 1997;20:537–44.
2 Knowler WC, Barrett-Connor E, Fowler SE, *et al*. Reduction in the incidence of type 2 diabetes with lifestyle intervention or metformin. *N Engl J Med* 2002;346:393–403.
3 Gong Q, Zhang P, Wang J, *et al*. Morbidity and mortality after lifestyle intervention for people with impaired glucose tolerance: 30-year results of the DA Qing diabetes prevention outcome study. *Lancet Diabetes Endocrinol* 2019;7:452–61.
4 Zhang L, Zhang Y, Shen S, *et al*. Safety and effectiveness of metformin plus lifestyle intervention compared with lifestyle intervention alone in preventing progression to diabetes in a Chinese population with impaired glucose regulation: a multicentre, open-label, randomised controlled trial. *Lancet Diabetes Endocrinol* 2023;11:567–77.
5 Li Y, Teng D, Shi X, *et al*. Prevalence of diabetes recorded in Mainland China using 2018 diagnostic criteria from the American diabetes association: national cross sectional study. *BMJ* 2020;369:m997.
6 Cowie CC, Rust KF, Byrd-Holt DD, *et al*. Prevalence of diabetes and high risk for diabetes using A1C criteria in the U.S. population in 1988-2006. *Diabetes Care* 2010;33:562–8.
7 Bonora E, Tuomilehto J. The pros and cons of diagnosing diabetes with A1C. *Diabetes Care* 2011;34 Suppl 2:S184–90.
8 Meigs JB, Nathan DM, D'Agostino RB Sr, *et al*. Fasting and postchallenge glycemia and cardiovascular disease risk: the Framingham offspring study. *Diabetes Care* 2002;25:1845–50.
9 Chien K-L, Hsu H-C, Su T-C, *et al*. Fasting and postchallenge hyperglycemia and risk of cardiovascular disease in Chinese: the chin-shan community cardiovascular cohort study. *Am Heart J* 2008;156:996–1002.
10 Mao T, Chen J, Guo H, *et al*. The efficacy of new Chinese diabetes risk score in screening undiagnosed type 2 diabetes and prediabetes: a community-based cross-sectional study in Eastern China. *J Diabetes Res* 2020;2020:7463082.
11 Ko G, So W, Tong P, *et al*. A simple risk score to identify Southern Chinese at high risk for diabetes. *Diabet Med* 2010;27:644–9.
12 Schmidt MI, Duncan BB, Bang H, *et al*. Identifying individuals at high risk for diabetes: the atherosclerosis risk in communities study. *Diabetes Care* 2005;28:2013–8.
13 Wilson PWF, Meigs JB, Sullivan L, *et al*. Prediction of incident diabetes mellitus in middle-aged adults: the Framingham offspring study. *Arch Intern Med* 2007;167:1068–74.
14 Schulze MB, Weikert C, Pischon T, *et al*. Use of multiple metabolic and genetic markers to improve the prediction of type 2 diabetes: the EPIC-Potsdam study. *Diabetes Care* 2009;32:2116–9.
15 Chien K, Cai T, Hsu H, *et al*. A prediction model for type 2 diabetes risk among Chinese people. *Diabetologia* 2009;52:443–50.
16 Nelson KM, Boyko EJ, Third National Health and Nutrition Examination Survey. Predicting impaired glucose tolerance using common clinical information: data from the third national health and nutrition examination survey. *Diabetes Care* 2003;26:2058–62.
17 Ko GTC, Chan JCN, Chan AWY, *et al*. Association between sleeping hours, working hours and obesity in Hong Kong Chinese: the 'better health for better Hong Kong' health promotion campaign. *Int J Obes* 2007;31:254–60.
18 Li JKY, Ng MCY, So WY, *et al*. Phenotypic and genetic clustering of diabetes and metabolic syndrome in Chinese families with type 2 diabetes mellitus. *Diabetes Metab Res Rev* 2006;22:46–52.
19 Friedewald WT, Levy RI, Fredrickson DS. Estimation of the concentration of low-density lipoprotein cholesterol in plasma, without use of the preparative ultracentrifuge. *Clin Chem* 1972;18:499–502.
20 Ng NYH, Jiang G, Cheung LP, *et al*. Progression of glucose intolerance and cardiometabolic risk factors over a decade in Chinese women with polycystic ovary syndrome: a case-control study. *PLoS Med* 2019;16:e1002953.
21 Zhang T, Yu B. Boosting with early stopping: convergence and consistency. 2005. Available: https://ui.adsabs.harvard.edu/abs/2005math.....8276Z [Accessed 01 Aug 2005].
22 Pencina MJ, D'Agostino RB, Pencina KM, *et al*. Interpreting incremental value of markers added to risk prediction models. *Am J Epidemiol* 2012;176:473–81.
23 Vickers AJ, Elkin EB. Decision curve analysis: a novel method for evaluating prediction models. *Med Decis Making* 2006;26:565–74.
24 Vickers AJ, van Calster B, Steyerberg EW. A simple, step-by-step guide to interpreting decision curve analysis. *Diagn Progn Res* 2019;3:18.
25 International Diabetes Federation. *International Diabetes Federation (IDF) Diabetes Atlas, 10th ed*. Brussels, Belgium: International Diabetes Federation, 2021.
26 Chatterjee R, Zelnick L, Mukamal KJ, *et al*. Potassium measures and their associations with glucose and diabetes risk: the multi-ethnic study of atherosclerosis (MESA). *PLoS One* 2016;11:e0157252.
27 Heianza Y, Hara S, Arase Y, *et al*. Low serum potassium levels and risk of type 2 diabetes: the Toranomon hospital health management center study 1 (TOPICS 1). *Diabetologia* 2011;54:762–6.
28 Rowe JW, Tobin JD, Rosa RM, *et al*. Effect of experimental potassium deficiency on glucose and insulin metabolism. *Metabolism* 1980;29:498–502.
29 Palmer BF, Clegg DJ. Physiology and pathophysiology of potassium homeostasis. *Adv Physiol Educ* 2016;40:480–90.
30 Tal B, Sack J, Yaron M, *et al*. Increment in dietary potassium predicts weight loss in the treatment of the metabolic syndrome. *Nutrients* 2019;11:1256.
31 Chatterjee R, Yeh H-C, Edelman D, *et al*. Potassium and risk of type 2 diabetes. *Expert Rev Endocrinol Metab* 2011;6:665–72.
32 Schlesinger S, Neuenschwander M, Barbaresko J, *et al*. Prediabetes and risk of mortality, diabetes-related complications and comorbidities: umbrella review of meta-analyses of prospective studies. *Diabetologia* 2022;65:275–85.
33 Jerums G, Panagiotopoulos S, Premaratne E, *et al*. Integrating albuminuria and GFR in the assessment of diabetic nephropathy. *Nat Rev Nephrol* 2009;5:397–406.
34 Jun JE, Lee S-E, Lee Y-B, *et al*. Low skeletal muscle mass accompanied by abdominal obesity additively increases the risk of incident type 2 diabetes. *J Clin Endocrinol Metab* 2023;108:1173–80.
35 Kitabchi AE, McDaniel KA, Wan JY, *et al*. Effects of high-protein versus high-carbohydrate diets on markers of B-cell function, oxidative stress, lipid peroxidation, proinflammatory cytokines, and adipokines in obese, premenopausal women without diabetes. *Diabetes Care* 2013;36:1919–25.
36 McAuley KA, Hopkins CM, Smith KJ, *et al*. Comparison of high-fat and high-protein diets with a high-carbohydrate diet in insulin-resistant obese women. *Diabetologia* 2005;48:8–16.
37 Zhang X, Wu H, Fan B, *et al*. The role of age on the risk relationship between prediabetes and major morbidities and mortality: analysis of the Hong Kong diabetes surveillance database of 2 million Chinese adults. *Lancet Reg Health West Pac* 2023;30:100599.
38 Bergman M, Abdul-Ghani M, Neves JS, *et al*. Pitfalls of HbA1C in the diagnosis of diabetes. *J Clin Endocrinol Metab* 2020;105:2803–11.

Supplementary materials

Table S1- Significant variables predicting AGT in BHBHK_HKFDS cohort using forward stepwise logistic regression (n=1054).

	β	OR (95% CI)	P value	**Wald-df**	VIF
Original RUBY score	0.116	1.123 (1.086 to 1.161)	<0.001	**45.292**	1.199
FPG (mmol/L)	1.391	4.02 (2.921 to 5.532)	<0.001	**71.910**	1.149
Creatinine (μmol/L)	-0.023	0.977 (0.967 to 0.988)	<0.001	**16.002**	1.192
WBC (10^9/L)	0.167	1.181 (1.066 to 1.309)	0.001	**9.123**	1.038
Urea (mmol/L)	-0.218	0.804 (0.693 to 0.934)	0.004	**7.167**	1.226
Potassium (mmol/L)	-0.667	0.513 (0.316 to 0.834)	0.007	**6.260**	1.028
Total protein (g/L)	0.051	1.053 (1.011 to 1.095)	0.012	**5.337**	1.020

Abbreviations: WBC, white blood cells; FPG, fasting plasma glucose; Wald-df, partial chi-square statistic minus the predictor degrees of freedom, assessed the importance of predictors; VIF, variance inflation factors, assessed the multicollinearity.

Table S2 - Relative contribution of RUBY and biochemical parameters in predicting AGT.

	AUROC		IDI	
	(95% CI)	P value	(95% CI)	P value
RUBY only	0.698 (0.662 to 0.733)	ref	ref	ref
RUBY + FPG (Extended RUBY model 1)	0.756 (0.722 to 0.791)	**<0.001**	0.201 (0.154 to 0.247)	**<0.001**
RUBY + Creatinine	0.726 (0.692 to 0.760)	**0.002**	0.024 (0.006 to 0.042)	**0.009**
RUBY + WBC	0.711 (0.676 to 0.746)	0.071	0.009 (-0.003 to 0.021)	0.132
RUBY + Creatinine + Urea + Potassium	0.725 (0.690 to 0.761)	**0.001**	0.047 (0.024 to 0.069)	**<0.001**
RUBY + FPG (Extended RUBY model 1)	0.756 (0.722 to 0.791)	ref	ref	ref
RUBY + FPG + Creatinine	0.776 (0.742 to 0.809)	**0.013**	0.029 (0.010 to 0.048)	**0.003**
RUBY + FPG + WBC	0.763 (0.729 to 0.797)	0.210	0.003 (-0.008 to 0.014)	0.565
RUBY + FPG + Total protein	0.759 (0.725 to 0.794)	0.413	0.003 (-0.005 to 0.010)	0.512
RUBY + FPG + Creatinine + Urea + Potassium (Extended RUBY model 2)	0.778 (0.744 to 0.812)	**0.002**	0.064 (0.039 to 0.089)	**<0.001**
RUBY + FPG + Creatinine + Urea + Potassium (Extended RUBY model 2)	**0.778 (0.744 to 0.812)**	ref	ref	ref
RUBY + FPG + Creatinine + Urea + Potassium + WBC	0.785 (0.752 to 0.819)	0.092	0.008 (-0.004 to 0.021)	0.161
RUBY + FPG + Creatinine + Urea + Potassium + Total protein	0.780 (0.746 to 0.815)	0.462	0.005 (-0.004 to 0.013)	0.315

Abbreviations: AUROC, area under receiver operating characteristics; IDI, integrated discrimination index; FPG, fasting plasma glucose; WBC, white blood cells.

Table S3 - AIC and BIC for different models in predicting AGT.

	AIC	BIC
RUBY	1087.0	1097.0
FPG only	1039.6	1049.5
RUBY + FPG	1001.0	1015
RUBY + FPG + Creatinine	977.9	997.7
RUBY + FPG + WBC	993.4	1013.3
RUBY + FPG + Total protein	997.8	1017.6
RUBY + FPG + Creatinine + Urea + Potassium	965.3	995.1
RUBY + FPG + Creatinine + Urea + Potassium + Total protein	960.8	995.5
RUBY + FPG + Creatinine + Urea + Potassium + WBC	954.4	989.1
RUBY + FPG + Creatinine + Urea + Potassium + Total protein + WBC	951.6	991.3

Abbreviations: AIC, akaike information criterion; BIC, bayesian information criterion

Table S4 - Sensitivity, specificity, PPV and NPV for various cutoff values by Extended RUBY model 2 for AGT at different corresponding absolute risk in percentage.

Risk threshold	**Sensitivity**	**Specificity**	**PPV**	**NPV**
9.5%	0.919	0.314	0.305	0.922
19.5%	0.765	0.642	0.412	0.893
29.5%	0.635	0.805	0.516	0.871
39.5%	0.492	0.897	0.610	0.844
49.5%	0.373	0.946	0.693	0.822

Abbreviation: PPV, positive predictive value; NPV, negative predictive value. 29.5% was the optimal risk threshold of the Extended model 2 using Youden index.

Table S5 - Baseline characteristics of the cohort for external validation.

	Total	Non-IGT	AGT		P value
	(n = 550)	2hPG < 7.8 (n = 308)	2hPG 7.8 - 11.0 (n = 179)	2hPG ≥ 11.1 (n = 63)	
Age (years)	53 ± 10	51 ± 11	55 ± 10	56 ± 9	<0.001
Female, n (%)	382 (69.5)	220 (71.4)	122 (68.2)	40 (63.5)	0.414
Weight (kg)	67.9 ± 14.5	66.8 ± 15.2	69.1 ± 14.2	69.4 ± 11.6	0.155
Height (cm)	160.9 ± 8.4	161.2 ± 9.1	160.6 ± 7.3	160.3 ± 7.7	0.668
BMI (kg/m^2)	26.1 ± 4.7	25.6 ± 4.8	26.8 ± 4.9	26.9 ± 3.5	0.012
Waist (cm)	89.0 ± 12.7	86.9 ± 13.4	91.6 ± 11.9	91.8 ± 9.6	<0.001
SBP (mmHg)	128.5 ± 17.3	125 ± 16.7	131.6 ± 17.2	136.8 ± 16.3	<0.001
DBP (mmHg)	80.9 ± 11.5	79.5 ± 11.5	81.7 ± 11.3	85.8 ± 11.1	<0.001
DM FH, n (%)	289 (52.5)	156 (50.6)	101 (56.4)	32 (50.8)	0.449
FPG (mmol/L)	5.2 ± 0.6	5.0 ± 0.5	5.3 ± 0.5	5.7 ± 0.5	<0.001
2hPG (mmol/L)	7.8 ± 2.5	6.1 ± 1.2	9.1 ± 0.9	12.6 ± 1.2	<0.001
Creatinine (μmol/L)	68.2 ± 17	68.4 ± 17.2	67.4 ± 16.1	69.1 ± 18.5	0.745
Urea (mmol/L)	5.1 ± 1.3	5.1 ± 1.4	5.2 ± 1.3	5.2 ± 1.1	0.783
Potassium (mmol/L)	4.2 ± 0.4	4.1 ± 0.4	4.2 ± 0.4	4.2 ± 0.4	0.225

Data were shown as mean ± SD or number (%). The one-way ANOVA test was used for continuous variables, chi-square test was used for categorical variables. FPG range, 3.8 - 6.9 mmol/L; Potassium range, 3.2 - 5.5 mmol/L; Urea range, 1.9 - 9.9 mmol/L; Creatinine range, < 129 μmol/L. Abbreviations: BMI, body mass index; FPG, fasting plasma glucose; 2hPG, two-hour post-prandial glucose.

Table S6 - Performance of the extended RUBY models, original RUBY and NCDRS in the external validation cohort (n = 550).

Model	**AUROC (95%CI)**
Original RUBY	0.637 (0.591, 0.683)
Extended RUBY model 1	0.716 (0.674, 0.759)
Extended RUBY model 2	0.722 (0.680, 0.764)
NCDRS	0.632 (0.586, 0.679)

Abbreviation: AUROC, area under receiver operating characteristics.

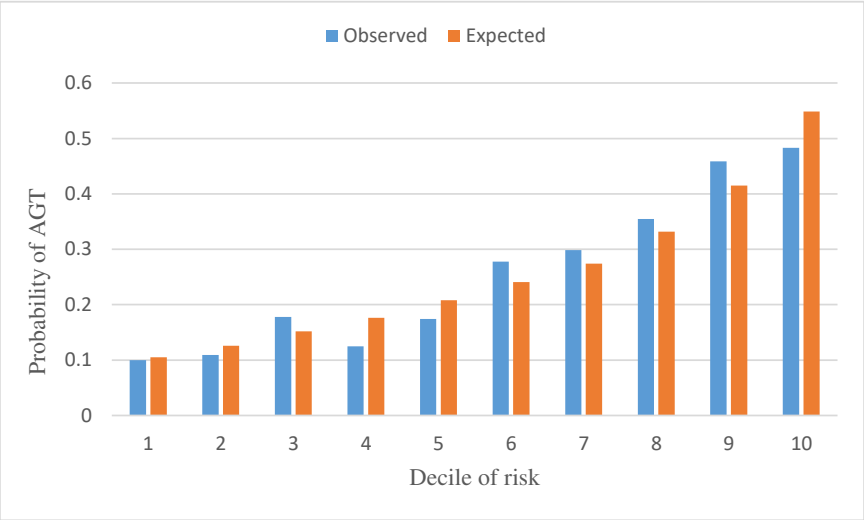
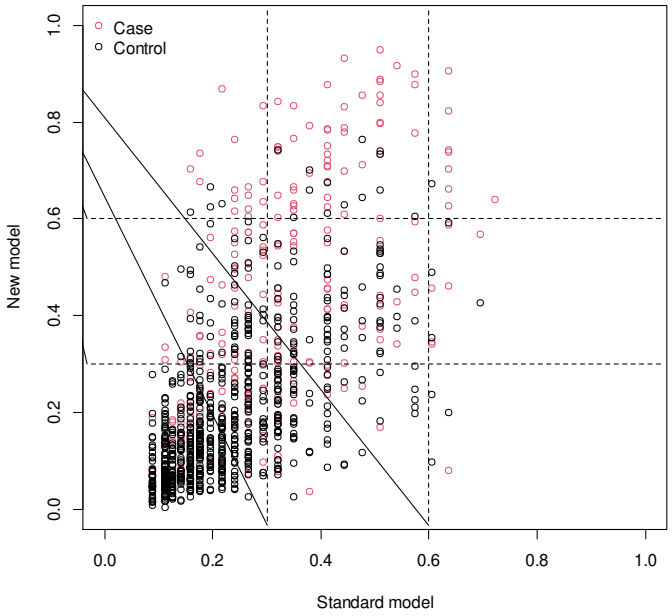


Fig. S1 - Observed vs expected probabilities of AGT as predicted by Extended RUBY model 2 for AGT by decile of risk in the BHBHK-HKFDS cohort.

A.



B.

	Risk threshold of original RUBY score	Risk threshold of Extended RUBY model 2				Cases superiorly detected by Extended RUBY model 2	Cases superiorly detected by RUBY score	NET	NRI
		< 0.3	0.3-0.6	> 0.6	Total				
AGT	< 0.3	79	40	15	134	103, 39.6%	23, 8.8%	30.8%	31.9%
	0.3-0.6	16	47	48	111				
	> 0.6	1	6	8	15				
	Total	96	93	71	260				
Non- IGT	< 0.3	556	63	4	623	89, 11.2%	80, 10.1%	1.1%	
	0.3-0.6	82	68	13	163				
	> 0.6	3	4	1	8				
	Total	641	135	18	794				

Fig. S2 - Net reclassification index of Extended RUBY model 2 and original RUBY score. A. Figure of NRI. new model, Extended RUBY model 2; standard model, original RUBY score; Case, AGT with 2hPG ≥ 7.8 mmol/L; control, non-IGT with 2hPG < 7.8 mmol/L; dash lines, prespecified risk threshold 0.3-0.6. B. Table of NRI indicating the numbers of AGT or non-IGT detected by Extended RUBY model 2 and original RUBY score.

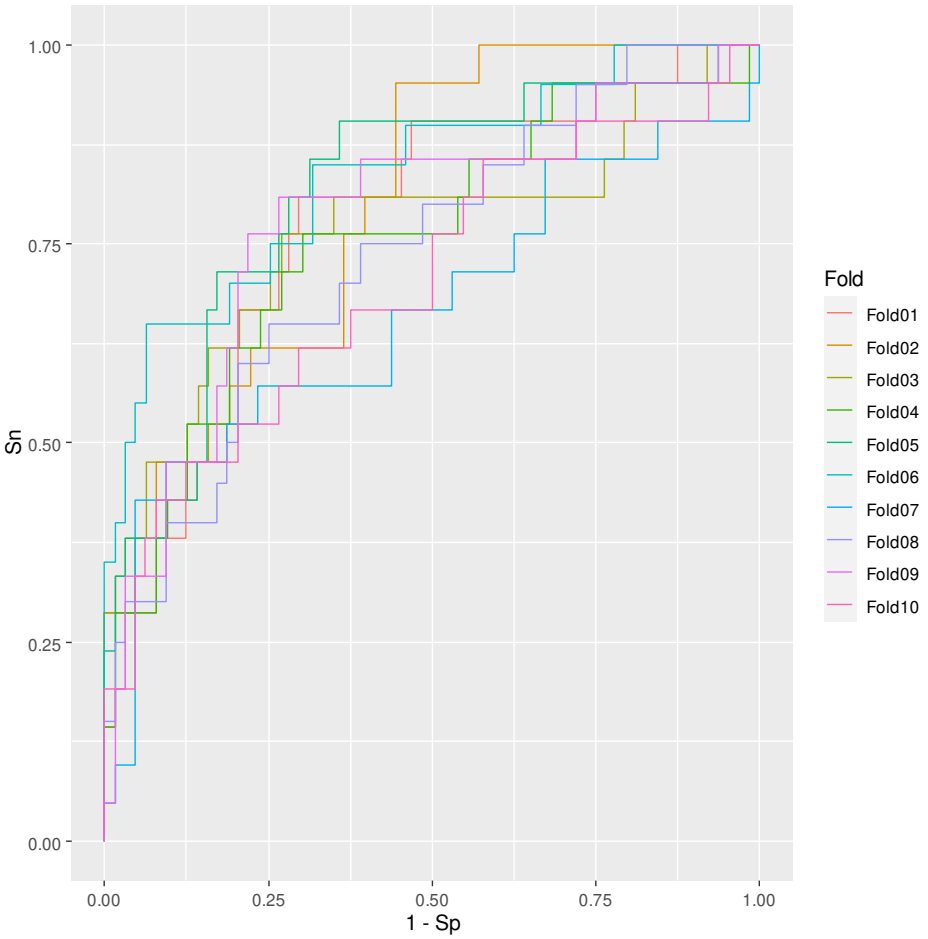


Fig. S3 - ROC of 10-fold resampling of the dataset divided into testing (1-fold) and training datasets (9-fold) in cross-validation for Extended RUBY model 2. Every fold sample had a chance to be the testing data to generate a ROC curve. AUROC of the 1st testing data (Fold 01): 0.788; AUROC of 2nd testing data (Fold 02): 0.780; AUROC of 3rd testing data (Fold 03): 0.810; AUROC of 4th testing data (Fold 04): 0.810; AUROC of 5th testing data (Fold 05): 0.800; AUROC of 6th testing data (Fold 06): 0.843; AUROC of 7th testing data (Fold 07): 0.800; AUROC of 8th testing data (Fold 08): 0.786; AUROC of 9th testing data (Fold 09): 0.812; AUROC of 10th testing data (Fold 10): 0.788. The average ROC was 0.814.

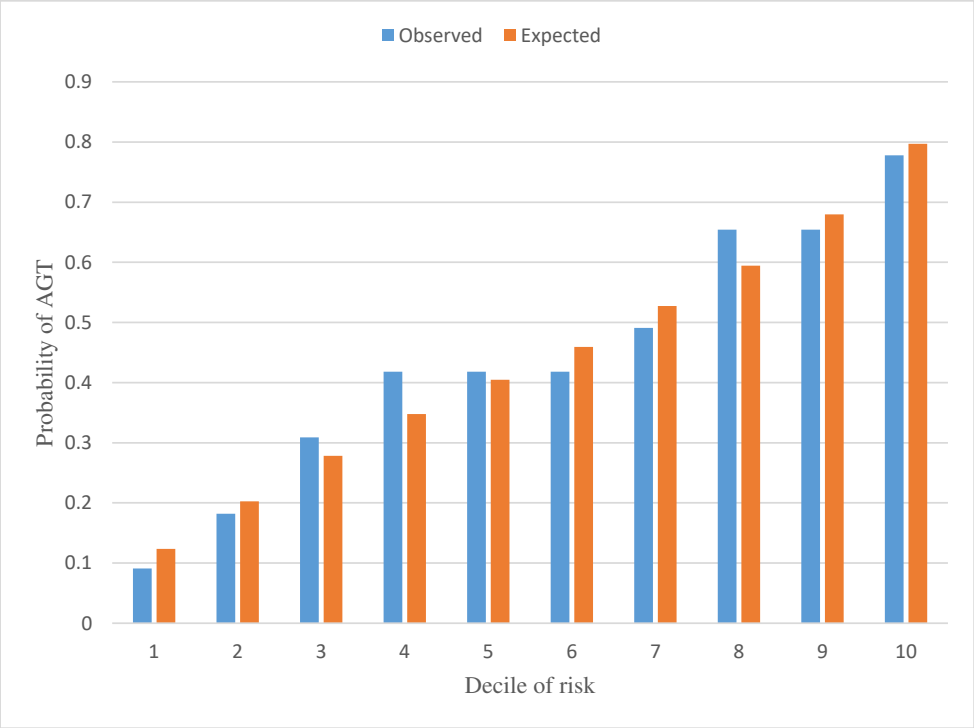


Fig. S4 - Observed vs expected as predicted by Extended RUBY model 2 for AGT by decile of risk in the external cohort.

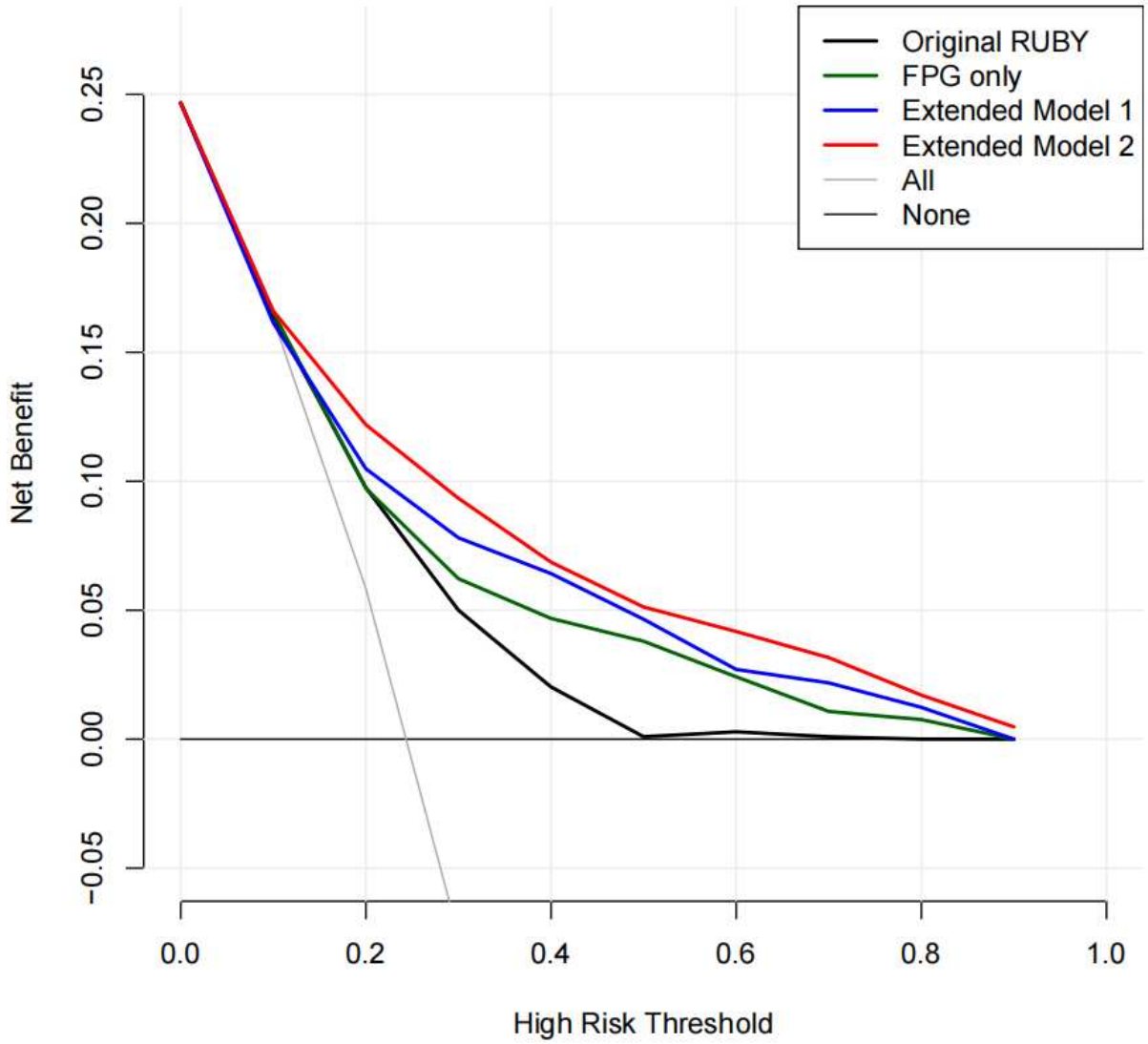


Fig. S5 – Decision curve analysis (DCA) of original RUBY, FPG only, Extended RUBY Model 1 and 2 in predicting AGT in the BHBHK-HKFDS cohort. Black line represents original RUBY score, the green one represents the model with FPG only, the blue and red for the Extended RUBY Model 1 and 2 respectively, and the light grey line represents treating all patients and dark grey line for treating no patient. Risk threshold is defined as 0.25 in this study.