Association of evaluated glomerular filtration rate and incident diabetes mellitus: a secondary retrospective analysis based on a Chinese Cohort Study

Running title: eGFR and incident diabetes

Haofei Hu¹,²,⁵#, Mijie Guan¹,²,⁵#, Zhuangsen Chen³,⁴,⁵, Yang Wu³,⁴,⁵, Qijun Wan¹,²,⁵*

¹Department of Nephrology, The First Affiliated Hospital of Shenzhen University, Shenzhen 518035, Guangdong Province, China
²Department of Nephrology, Shenzhen Second People’s Hospital, Shenzhen 518035, Guangdong Province, China
³Department of Endocrinology, The First Affiliated Hospital of Shenzhen University, Shenzhen 518035, Guangdong Province, China
⁴Department of Endocrinology, Shenzhen Second People’s Hospital, Shenzhen 518035, Guangdong Province, China
⁵Shenzhen University Health Science Center, Shenzhen 518000, Guangdong Province, China

Haofei Hu and Mijie Guan have contributed equally to this work.

*Corresponding author

Qijun Wan

Department of Nephrology,

The First Affiliated Hospital of Shenzhen University,
No.3002 Sungang Road, Futian District,
Shenzhen 518000,
Guangdong Province,
China
Tel:+86-755-83366388
E-mail: wanqijun123@126.com
Abstract

Background: Previous studies have revealed that chronic kidney disease (CKD) is one of major risk factors of insulin resistance and diabetes. However, there are few investigations of the correlations between the estimated glomerular filtration rate (eGFR) and incident diabetes, especially in Chinese population. This study was taken to explore the relationship between eGFR and incident diabetes in a large cohort in Chinese community population.

Methods: The present study was a retrospective cohort study. A total of 199,435 adults from Rich Healthcare Group in China, which includes all medical records for participants who received a health check from 2010 to 2016. The target independent variable and the dependent variable were eGFR measured at baseline and incident diabetes mellitus appeared during follow-up respectively. Covariates involved in this study included age, gender, body mass index, diastolic blood pressure, systolic blood pressure, fasting plasma glucose, total cholesterol, low density lipoprotein cholesterol, high density lipoprotein cholesterol, triglyceride, alanine aminotransferase, aspartate aminotransferase, smoking and drinking status and family history of diabetes. Cox proportional-hazards regression was used to investigate the association between eGFR and incident diabetes. Generalized additive model was used to identify non-linear relationships. Additionally, we also performed a subgroup analysis. It was stated that the data had been uploaded to the DATADRYAD website.

Result: After adjusting gender, body mass index, systolic blood pressure, diastolic blood pressure, fasting blood glucose, total cholesterol, low density lipoprotein cholesterol, high density lipoprotein cholesterol, triglyceride, alanine aminotransferase, aspartate aminotransferase, smoking and drinking status and family history of diabetes, result showed eGFR was negatively
associated with incident diabetes (HR=0.987, 95%CI (0.984, 0.989)). A J shape relationship was detected between eGFR and incident diabetes, which had an inflection point of eGFR was 97.967 mL/min\(^1\cdot(1.73 \text{ m}^2)^{-1}\). The effect sizes and the confidence intervals on the left and right sides of the inflection point were 0.999 (0.994,1.004) and 0.977 (0.974,0.981), respectively. Subgroup analysis showed, the stronger association can be found in the population with FPG<6.1mmol/L, BMI<24kg/m\(^2\), SBP<140mmHg, DBP<90mmHg, HDL in middle level and family history without diabetes. The same trend was also seen in men and in the population with never or ever smoking.

**Conclusion:** eGFR is independently associated with incident diabetes. The relationship between eGFR and incident diabetes is also non-linear. eGFR was strong negatively related to incident diabetes when eGFR is above 97.967 mL/min\(^1\cdot(1.73 \text{ m}^2)^{-1}\).

**Keywords:** Estimated glomerular filtration rate, Incident diabetes, Nonlinearity

**Background**

Diabetes has become one of the most common chronic diseases all over the world. In recent decades, the prevalence of diabetes among Chinese adults has increased significantly [1]. According to a large, nationally representative survey of Chinese adults, the estimated overall prevalence of diabetes had risen to be 10.9% in 2013[2]. Therefore, it is important to explore and intervene in the risk factors of diabetes. Diabetes is a debilitating disease that may cause various complications, thereby reducing the quality of life and causing serious socio-economic impacts. Therefore, the identification of risk factors is essential to prevent diabetes.

Patients with chronic kidney disease (CKD) and diabetes mellitus (DM) share common risk
factors[3] [4], which suggests that CKD may increase the risk of diabetes. Estimated glomerular filtration rate (eGFR) is a simpler and more applicable surrogate marker, used to describe the flow rate of filtrate through the kidney and has been widely used clinically to diagnose CKD and assess renal function[5]. A cardiovascular health study was conducted among 4680 U.S. participants without diabetes, the result[6] showed that a decreased eGFR was associated with increased insulin resistance. In their study, Mean eGFR was 72.2 (SD 17.1) ml/min per 1.73 m2. However, over a median follow-up of 12 years, participants with a decreased eGFR did not have an increased risk of incident diabetes. In another cohort study with 864 American participants, the result suggests that the relationship between GFR and the occurrence of diabetes was not linear, and within the upper and lower ranges of GFR, the risk of diabetes increases [7]. However, most of these studies did not perform subgroup analysis, and the relatively small sample size and regional population limited to generalizable other people. Moreover, findings from previous studies regarding the relationship between eGFR and incident diabetes was still limited in Chinese population. Therefore, this study set out to investigate whether eGFR was independently related to incident diabetes in a large cohort population across 32 sites and 11 cities in China.

Methods

Data source and participants

Date were obtained from ‘DATADRYAD’ database (www.Datadryad.org). This website permitted users to freely download the raw data. According to Dryad Terms of Service, we cited Dryad data package in the present study. (Dryad data package: Ying Chen, Xiao-Ping Zhang, Jie Yuan, Bo Cai, Xiao-Li Wang, Xiao-Li Wu, Yue-Hua Zhang, Xiao-Yi Zhang, Tong Yin, Xiao-Hui
Data from: Association of body mass index and age with incident diabetes in Chinese adults: a population-based cohort study. Dryad Digital Repository. http://dx.doi.org/10.1136/bmjopen-2018-021768). Variables included in the database file were as follows: age, gender, body mass index (BMI), diastolic blood pressure (DBP), systolic blood pressure (SBP), fasting plasma glucose (FPG), Triglyceride(TG), total cholesterol (TC), low density lipoprotein cholesterol(LDL-C), high density lipoprotein cholesterol (HDL-C), Serum urea nitrogen(BUN), Serum creatinine(Scr), Alanine aminotransferase(ALT), Aspartate aminotransferase(AST), smoking status, drinking status, family history of diabetes, year of follow up and censor of diabetes during follow up. In our research, we added evaluated glomerular filtration rate(eGFR), which was calculated based on age, gender and Scr according to CKD-EPI equation[8]. This new Asian modified CKD-EPI equation could lead to more accurate GFR estimation in Chinese patients with CKD in general practice, especially in the higher GFR group. Authors of the original study have waived all copyright and related ownership of these data. Therefore, we could use these data for secondary analysis without infringing on the authors’ rights. As research ethics approved was obtained in the previous research, no longer needed for this secondary study.

Data were obtained from a database provided by the Rich Healthcare Group in China, and the study recruited 685,277 participants who underwent a health check and were at least 20 years old with at least two visits between 2010 and 2016 across 32 sites and 11 cities in China(Shanghai, Beijing, Nanjing, Suzhou, Shenzhen, Changzhou, Chengdu, Guangzhou, Hefei, Wuhan, Nantong). The data we obtained has been preliminarily screened, as follows:(1) no available information about weight, height, gender, fasting plasma glucose value at baseline,(2) extreme BMI values.
(<15 kg/m² or >55 kg/m²), (3) excluded participants with visit intervals less than 2 years, (4) participants diagnosed with diabetes at baseline and participants with undefined diabetes status at follow-up. Finally, Ying Chen, et al [9] selected 211,833 participants in the analysis. The inclusion/exclusion criteria and outcome measures of the trial was specifically explained in the previous study [9]. In our research, we further excluded participants with missing values of baseline eGFR (n=11,175) from the analysis cohort. To reduce interference, we excluded outliers in eGFR which were less than means minus three standard deviation (SD) or greater than the means plus three SD (n=1223) [10]. The final analysis included 199,435 subjects (109,690 male and 89,745 female) for data analysis in our study.

Study design and measurement of variables

Researchers have obtained information (values) for our retrospective cohort study. The design of the study was documented in the original study [9]. To gave you a clearer understanding of the entire research process, we outlined the research steps here. At each visit to the health examination center, participants were asked to fill in a detailed questionnaire about demographic characteristics, lifestyle factors, personal medical history and family history of chronic diseases. Subjects were measured for height, weight and blood pressure by trained staff. Body weight was measured in light clothing with no shoes to the nearest 0.1 kg. Height was measured to the nearest 0.1 cm. BMI was derived from weight in kilograms divided by height in metres squared. Blood pressure was measured by standard mercury sphygmomanometers. Fasting venous blood samples were collected after at least a 10 hours fast at each visit. Scr, ALT, AST, TG, TC, LDL-C and HDL-C were measured on an autoanalyzer (Beckman 5800). Plasma glucose levels were measured by the glucose oxidase method on an autoanalyzer (Beckman 5800). The target independent variable is
eGFR obtained at baseline. The dependent variable is incident diabetes obtained in the follow up.

As this is a retrospective cohort study, it reduced the possibility of selection bias and observation bias.

Ascertainment of incident diabetes

Diagnosis of incident diabetes was defined as fasting plasma glucose of >7.00 mmol/L and/or self-reported diabetes during the follow-up period. Patients were censored at the date of diagnosis of diabetes or the final visit, whichever came first.

Statistical analysis

First, we dealt with the missing values of covariates. Missing continuous variables were mainly supplemented by means or medians. Since the missing values of HDL-C, LDL-C and AST were about 50%, we converted them as categorical variables based on the tertiles. Besides, missing categorical variables in each covariate are considered as a independent group[11]

Next, the participants were stratified by quartiles of eGFR. Continuous variables were expressed as the means ± standard deviations (normal distribution) or medians (quartiles) (skewed distribution), and categorical variables were expressed as a frequency or percentages. Differences between means and proportions of the groups were tested by the one-way ANOVA test for normally distributed quantitative variables, the Kruskal-Wallis test for skewed quantitative variables, and the chi-square test for categorical variables. The person-years of follow-up were calculated from the date of baseline interview to the date of incident diabetes or follow-up interview, whichever came first[12]. We used cumulative incidence and person-years incidence to describe the incidence rate[13]. Cox proportional hazard regression models were used to investigate the prognostic value of eGFR on incident diabetes, and adjusted hazard ratios (HRs)
with 95% confidence intervals (CIs) were estimated to evaluate the diabetes risk. According to the recommendation of the STROBE statement[14], we simultaneously showed the results from unadjusted, minimally adjusted analyses and those from fully adjusted analyses. Whether the covariances were adjusted determined by the following principle: when added to this model, changed the matched hazard ratio by at least 10% [15]. To ensure the robustness of data analysis, we did a sensitivity analysis. We converted the eGFR into a categorical variable, and calculated the P for trend. The purpose was to verify the results of eGFR as the continuous variable and to observe the possibility of nonlinearity. We also tried to use generalized additive models (GAM) to identify non-linear relationships because eGFR was a continuous variable. If a non-linear correlation was observed, a two-piecewise linear regression model was performed to calculate the threshold effect of the eGFR on incident of diabetes in terms of the smoothing plot. When the ratio between eGFR and incident diabetes appeared obvious in a smoothed curve, the recursive method automatically calculates the inflection point, where the maximum model likelihood will be used. Moreover, the cox proportional hazard models were applied to explore robustness of the results in various subgroups (age, gender, BMI, SBP, DBP, FPG, HDL-C, LDL-C, family history of diabetes, smoking and drinking status). For continuous variable, we first converted it to a categorical variable according to the clinical cut point or binary. Each stratification was adjusted for all the factors, except for the stratification factor itself. The modifications and interactions of subgroups were inspected by likelihood ration tests. Survival estimates and cumulative event rates were compared using the Kaplan–Meier method by using the time-to-first event for each endpoint. The log-rank test was used to compare the Kaplan–Meier hazard ratios (HR) for research events, and their corresponding 95% confidence intervals (CIs).
All of the analyses were performed with the statistical software package R (http://www.R-project.org, The R Foundation) and Empower-Stats (http://www.empowerstats.com, X&Y Solutions, Inc., Boston, MA). P values less than 0.05 (two-sided) were considered statistically significant.

## Results

A total of 199,435 participants (55.0% men and 45.0% women) were included in the analysis, the mean age of the population was 42.6 ± 12.5 years old. The mean year of follow up was 3.13 ± 0.94 years, and 3919 people developed diabetes during follow-up. The mean eGFR was 110.44 ± 15.13 mL/min⁻¹·(1.73 m²)⁻¹, and the mean FPG, BMI, SBP and DBP were 4.91 ± 0.61 mmol/L, 23.23 ± 3.34 kg/m², 118.92 ± 16.31 mmHg and 74.13 ± 10.78 mmHg respectively. The number of participants with missing value of TC, TG, HDL-C and LDL-C were 3114, 3119, 84402 and 83314, respectively. Besides, the missing value of SBP, DBP, ALT and AST were 19, 20, 1110 and 115231, respectively. In addition, the missing value of smoking and drinking status were 142,038 and 142,038.

### Baseline characteristics of the study participants

Table 1 depicted the baseline characteristics of the total population and by quartiles of the eGFR. We divided participants into subgroup using eGFR quartiles(≤100.44, 100.44-112.40, 112.40-122.06, >122.06). The results showed that in highest eGFR group, participants generally had lower age, BMI, blood pressure levels (including both systolic and diastolic blood pressures), fasting blood glycemic, TC, TG, LDL-C, ALT, AST and lower rates of current smoker and drinker. Besides, in the top eGFR group, there were fewer people with low HDL-C level and more people
with low LDL-C level. In addition, the group (eGFR>122.06 mL/min\(^1\cdot(1.73 \text{ m}^2)^{-1}\)) had higher incidence of family history of diabetes and higher proportion of women.

**Incidence rate of incident diabetes**

Table 2 revealed that a total of 3919 participants developed incident diabetes. The total incidence rate of all participants was 628.73 per 100,000 person-years. Specifically, the incidence rates of the four eGFR groups were 1140.90, 776.55, 386.46 and 211.19 per 100,000 person-years, respectively. Compared with the lowest eGFR group, participants with increased eGFR had a lower cumulative incidence (p<0.001 for trend). The cumulative incidence of total incident diabetes and each of the eGFR groups was 1.965 (1.904-2.026), 3.541 (3.379-3.703), 2.440 (2.305-2.576), 1.227 (1.131-1.324) and 0.652 (0.581-0.722), respectively.

**Univariate analysis**

The results of univariate analysis were shown in Table 3. The results of univariate analysis showed that age, BMI, SBP, DBP, FPG, TG, TC, LDL, ALT, AST, family history of diabetes, smoking and drinking status were positively correlated with incident of diabetes. In contrast, HDL-C and eGFR negatively correlated with incident diabetes. We also found that women have a lower risk of developing diabetes than men.

Figure 1 showed the Kaplan-Meier curves of the cumulative hazards of diabetes incident risk stratified by eGFR categories. The risk of developing diabetes between each of the four eGFR groups was significantly different (log-rank test, p<0.0001). With the increase of eGFR, the cumulative risk of incident diabetes gradually decreased, rendering the top quartile group with the minimum risk of incident diabetes.

**The results of relationship between eGFR and incident diabetes**
We used Cox proportional hazard regression model to evaluate the associations between eGFR and incident diabetes. We simultaneously showed the non-adjusted and two adjusted models in Table 4. In crude model, eGFR showed negative correlation with incident diabetes (HR=0.964, 95% confidence interval (CI): 0.962 to 0.966, P < 0.00001). In minimally adjusted model (adjusted gender, BMI, SBP, DBP, family history of diabetes, smoking and drinking status), the result did not change significantly (HR: 0.977, 95%CI: 0.975-0.979). After adjusting for the full model (adjusted gender, BMI, SBP, DBP, FPG, TC, TG, LDL, HDL, ALT, AST, smoking and drinking status, family history of diabetes), we found the relationship still exists (HR=0.987, 95%CI: 0.984 to 0.989, P < 0.00001). The results showed that for every 1 mL/min\(^{-1}\cdot(1.73\text{ m}^2)^{-1}\) increased in eGFR, the risk of diabetes decreased by 1.3%. For the purpose of sensitivity analysis, we also handled eGFR as categorical variable (Quartile), the top quartile had 48.7 percent decrement of diabetes risk when compared with the bottom quartile in the full model, and found that the trend across the quartiles was significant (P for trend=0.00001).

The analyses of non-linear relationship

In the present study, we also performed a generalized additive model (GAM) to explore the non-linear relationship eGFR and incident diabetes because eGFR was a continuous variable (Fig. 2). We found that the relationship between eGFR and incident of diabetes was also non-linear (after adjusting gender, BMI, SBP, DBP, FPG, TC, TG, LDL, HDL, ALT, AST, smoking and drinking status, family history of diabetes). By using a two-piecewise linear regression model, we calculated that the inflection point of eGFR was 97.967 mL/min\(^{-1}\cdot(1.73\text{ m}^2)^{-1}\) (Log-likelihood ratio test P<0.001). On the left of the inflection point, we observed a tiny negative relationship between eGFR and incident diabetes (HR: 0.999, 95%CI: 0.994-1.004, P=0.6928). On the right side
of the inflection point, however, we could found a obvious negative relationship between eGFR and incident diabetes (HR: 0.977, 95%CI: 0.974-0.981, P<0.0001). (Table 5).

The results of subgroup analyses

We used subgroup analysis to detect other potential risks in the associations between eGFR and incident diabetes to assess the factors that might influence the results. We treated age, gender, FPG, BMI, SBP, DBP, HDL, LDL, family history of diabetes, smoking and drinking status as the stratification variables to observe the trend of effect sizes in these variables (table 6). We noted a great number of interactions were observed based on our a priori specification including: gender, FPG, SBP, DBP, BMI, HDL, smoking status and family history of diabetes (all P values for interaction <0.05). In this study, the stronger association were detected in the population with FPG<6.1mmol/L, BMI<24kg/m^2, SBP<140mmHg, DBP<90mmHg, HDL in middle level and family history without diabetes. Moreover, we could also found a stronger association between eGFR and incident diabetes in women and in the population with never or ever smoking. In contrast, the weaker association were probed in men, current smoker and the population with FPG≥6.1mmol/L, BMI ≥28 kg/m^2, SBP≥140mmHg, DBP≥90mmHg, HDL in low level and family history with diabetes.

Discussion

The present retrospective cohort study showed that eGFR was negatively associated with incident diabetes after adjusting some covariates (Cox proportional hazard models). Furthermore, the trend of the effect sizes on the left and right sides of the inflection point was inconsistent [left (HR: 0.999, 95%CI: 0.994-1.004, P=0.6928); right (HR: 0.977, 95%CI: 0.974-0.981, P]
The results indicated that a J shape relationship on the independent association between eGFR and new onset diabetes. Subgroup analysis showed a stronger association in female, never or ever smoker, and the population with FPG<6.1mmol/L, BMI<24kg/m², SBP<140mmHg, DBP<90mmHg, HDL in middle level and family history without diabetes. In contrast, the weaker association were probed in male, current smoker and the population with PG≥6.1mmol/L, BMI ≥28 kg/m², SBP≥140mmHg, DBP≥90mmHg, HDL in low level and family history with diabetes. A few previous studies have probed the association between eGFR and incident diabetes. However, most of these studies are not conducted in the Chinese population[6, 7, 16, 17]. In one of such studies with 864 adults in USA, C. Lorenzo et al[7] found that the relationship between GFR and incident diabetes was not linear, which suggests that individuals in the upper and lower ranges of GFR are at increased risk of future diabetes, GFR and type 2 diabetes may share common pathogenic mechanisms. Some other studies have explored the association between CKD and incident diabetes. A prospective cohort study, focus on 1,713 American participants with reduced glomerular filtration rates and without diabetes at baseline, found that T2DM incidence rate among individuals with CKD is markedly higher than in the general population[18][18]. Another population-based cohort study in Taiwan found that CKD was a significant and independent predictor of incident diabetes(adjusted HR 1.204; 95% CI 1.11, 1.31)[19]. Consistently the same result that, we obtained cox proportional hazard regression model showed a negative association between eGFR and incident diabetes. Moreover, our research had a larger sample(199435) and from 32 sites and 11 cities in China, which was more representative of the Chinese population.

In contrast to the results of these studies, eGFR could not predict the risk of incident diabetes
in one study with 1,337,452 veterans conducted in United States[16], the researchers found that every 10 ml/min/1.73m² decrease in eGFR was not associated with risk of incident diabetes (1.00; 1.00-1.01). A similar study in a lean, normoglycemic healthy women population in Israel, showed that in a logistic regression model adjusted for age, body mass index, socioeconomic status, smoking, baseline glucose and serum uric acid, eGFR was associated with increased risk for incident diabetes (1.02; 1.01-1.03)[17]. We compared these studies mentioned above, the inconsistent results may come from the following: (1) the research population is different. These studies, which were inconsistent with our findings, were targeted at Israel and America,(2) many of these different conclusions do not clarify the nonlinear relationship and use different regression models, (3) compared with our work, the study did not take into account the effect of SBP, DBP, TC, TG, HDL,LDL, ALT, AST, Scr, drinking status and family history of diabetes, on the eGFR and incident diabetes relationships when adjusting covariates. However, previous studies have confirmed that these variables are related to eGFR or incident of diabetes,(4) this may be related to different renal function, some studies showed that the association of eGFR and insulin resistance or incident diabetes differ between different CKD stages[6][20] [21].

In the present study, we found using two-piecewise linear regression model to show a nonlinear relation is different to that obtained by Lorenzo et al[7].In their study, they used subgroup analysis stratified by GFR categories to assess a U shape relation between eGFR and risk of T2DM, they found individuals in the upper and lower ranges of GFR are at increased risk of future diabetes. In contrast, results of the present study indicated a J shape relationship on the independent association between eGFR and new onset diabetes. The difference may be caused by race, renal function level and different methods for evaluating eGFR. They chose the Modification of Diet in
Renal Disease (MDRD) equation based on six variables to estimate GFR[22, 23], however, we used CKD-EPI equation[8] to estimate GFR, which could lead to more accurate GFR estimation in Chinese patients with CKD in general practice, especially in the higher GFR group. Because of the participants in our study was all with eGFR above 60 ml/min/1.73m². Our study showed that when eGFR is above 97.967 ml/min/1.73m², the risk of diabetes decreases obviously with increasing eGFR levels, these people could pay more attention to prevent the risk of diabetes. The results of this study should be helpful for future studies on the establishment of diagnostic or predictive models of the risk of diabetes.

In recent years, researches have elucidated the correlations between eGFR and insulin resistance. In a community-based cohort study, the result suggested that insulin sensitivity measured with euglycemic clamp is independently associated with eGFR, and impaired insulin sensitivity may be related to the development of renal dysfunction at an early stage, before the onset of diabetes[21]. In another community-based cohort study in US older adults, researchers found that lower eGFR was associated with insulin resistance. However, with lower eGFR, risks of impaired glucose tolerance and incident diabetes were not increased[6]. The disturbances of glucose and insulin homeostasis in CKD are complex and represent two opposite effects. On the one hand, CKD reduces insulin sensitivity (and increases insulin resistance) and, results in beta-cell dysfunction and defective insulin secretion in advanced stages[24]. On the other hand, CKD leads to decreased insulin clearance, thus prolonging its half-life[25, 26]. The balance of these two opposing forces determines the state of glucose metabolism and ultimately the risk of diabetes mellitus in any individual.

Our study have some strengths. (1) Our sample size is relatively large compared with previous
similar studies; (2) we expounded the non-linear relationship and found the inflection point; (3) this study was an observational study, so it is likely to cause potential confusion. We used strict statistical adjustment to minimize residual confounders; (4) We handled target independent variable as both continuous variable and categorical variable. Such an sensitivity analysis can reduce the contingency in the data analysis and enhance the reliability of results; (5) the effect modifier factor analysis enable the use of data better and yield stable conclusion in different subgroups in this study; (6) in order to control bias, we did not exclude missing values for covariates, we converted them as categorical variables and included in the cox proportional hazard regression models.

There are still some potential limitations. Firstly, the raw data was from the Chinese population, which limits the generalizability of our findings. Besides, as this study is based on a secondary analysis of published data, there may be other related factors that are not included in the data, such as medication history, socioeconomic factors, etc. We cannot adjust those variables. Similarly, we could not distinguish between type 1, type 2, and other types of diabetes. Secondly, they did not perform 2-hr oral glucose tolerance test. According to the 1999 WHO recommendations for the diagnosis of diabetes, the definition of diabetes in our study might lead to miss some diabetic patients[27]. However, oral glucose tolerance tests were not feasible in such a large cohort. Thirdly, we only measured eGFR and other parameters at baseline, which changes over time are not concerned in this study. Finally, although we have adjusted for a number of confounding factors to the possible influences, residual confounding may exit and further investigations are needed.

**Conclusion**

eGFR is independently associated with incident diabetes. The relationship between eGFR and
incident diabetes is also non-linear. eGFR is obvious negatively related with incident diabetes when eGFR is above 97.967mL/min^1·(1.73 m2)^1. In addition, the stronger association of eGFR and incident diabetes were detected in female, never or ever smoker, and the population with FPG<6.1mmol/L, BMI<24kg/m², SBP<140mmHg, DBP<90mmHg, HDL in middle level and family history without diabetes.

**Abbreviations**

- BMI, Body mass index; SBP, Systolic blood pressure; DBP, Diastolic blood pressure; FPG, Fasting plasma glucose; Scr, Serum creatinine; eGFR, evaluated glomerular filtration rate; TC, Total cholesterol; TG, Triglyceride; HDL-C, High-density lipoprotein cholesterol; LDL-C, Low-density lipid cholesterol; ALT, Alanine aminotransferase; AST, Aspartate aminotransferase; GAM, Generalized additive models; CKD, Chronic kidney disease; T2DM, Type 2 diabetes mellitus; HR, hazard ratios; CI, Confidence intervals.

**Authors’ contributions**

Haofei HU and Mijie GUAN contributed to the study concept and design, researched and interpreted the data and drafted the manuscript. Zhuangsen CHEN and Yang WU oversaw the progress of the project, contributed to the discussion and reviewed the manuscript. Qijun WANG revised the manuscript. Haofei HU are the guarantors of this work and, as such, had full access to all the data in the study and take responsibility for the integrity of the data and the accuracy of the data analysis. All authors read and approved the final manuscript.

**Acknowledgements**

Not applicable
Competing interests

The authors declare that they have no competing interests

Availability of data and materials

Data can be downloaded from ‘DATADRYAD’ database (www.Datadryad.org).

Consent for publication

Not applicable.

Ethics approval and consent to participate

In the previously published article[9], Ying Chen, et al. has clearly stated that: the study was conducted in accordance with the Declaration of Helsinki. Informed consent was obtained from all Participants.

Funding

This study was supported by Discipline Construction Ability Enhancement Project of Shenzhen Municipal Health Commission (SZXJ2017031).

References:

1. Weng J, Ji L, Jia W, Lu J, Zhou Z, Zou D, Zhu D, Chen L, Chen L, Guo L et al: Standards of care for type 2 diabetes in China. Diabetes/Metabolism Research and Reviews 2016, 32(5):442-458.
2. Wang L, Gao P, Zhang M, Huang Z, Zhang D, Deng Q, Li Y, Zhao Z, Qin X, Jin D et al: Prevalence and Ethnic Pattern of Diabetes and Prediabetes in China in 2013. JAMA 2017, 317(24):2515.
3. Aschner P: Metabolic syndrome as a risk factor for diabetes. Expert Review of Cardiovascular Therapy 2010, 8(3):407-412.
4. Laguardia HA, Hamm LL, Chen J: The metabolic syndrome and risk of chronic kidney disease: pathophysiology and intervention strategies. J Nutr Metab 2012, 2012:652608.
5. Warnock DG: Estimated Glomerular Filtration Rate: Fit for What Purpose? NEPHRON 2016, 134(1):43-49.
6. Pham H, Robinson-Cohen C, Biggs ML, Ix JH, Mukamal KJ, Fried LF, Kestenbaum B, Siscovick DS, de Boer IH: Chronic kidney disease, insulin resistance, and incident diabetes in older adults.
1. Lorenzo C, Nath SD, Hanley AJ, Abboud HE, Gelfond JA, Haffner SM: Risk of type 2 diabetes among individuals with high and low glomerular filtration rates. *DIABETOLOGIA* 2009, 52(7):1290-1297.

2. Wang J, Xie P, Huang J, Qu Y, Zhang F, Wei L, Fu P, Huang X: The new Asian modified CKD-EPI equation leads to more accurate GFR estimation in Chinese patients with CKD. *INT UROL NEPHROL* 2016, 48(12):2077-2081.

3. Chen Y, Zhang X, Yuan J, Cai B, Wang X, Wu X, Zhang Y, Zhang X, Yin T, Zhu X et al: Association of body mass index and age with incident diabetes in Chinese adults: a population-based cohort study. *BMJ OPEN* 2018, 8(9):e21768.

4. Zhang N, Hu X, Zhang Q, Bai P, Cai M, Zeng TS, Zhang JY, Tian SH, Min J, Huang HT et al: Non-high-density lipoprotein cholesterol: High-density lipoprotein cholesterol ratio is an independent risk factor for diabetes mellitus: Results from a population-based cohort study. *J DIABETES* 2018, 10(9):708-714.

5. Erviti J, Alonso Á, Oliva B, Gorricho J, López A, Timoner J, Huerta C, Gil M, De Abajo F: Oral bisphosphonates are associated with increased risk of subtrochanteric and diaphyseal fractures in elderly women: a nested case–control study. *BMJ OPEN* 2013, 3(1):e2091.

6. Guo C, Qin P, Li Q, Zhang D, Tian G, Liu D, Liu L, Cheng C, Chen X, Qie R et al: Association between mean arterial pressure and risk of type 2 diabetes mellitus: The Rural Chinese Cohort Study. *PRIM CARE DIABETES* 2020.

7. Qin H, Chen Z, Zhang Y, Wang L, Zhang Y: Triglyceride to high density lipoprotein cholesterol ratio is associated with incident diabetes in male: A retrospective study in Chinese. *J DIABETES INVEST* 2019.

8. Yokoyama M, Watanabe T, Otaki Y, Takahashi H, Arimoto T, Shishido T, Miyamoto T, Konta T, Shibata Y, Daimon M et al: Association of the Aspartate Aminotransferase to Alanine Aminotransferase Ratio with BNP Level and Cardiovascular Mortality in the General Population: The Yamagata Study 10-Year Follow-Up. *DIS MARKERS* 2016, 2016:1-9.

9. Shani M, Vinker S, Dinour D, Leiba M, Twig G, Holtzman EJ, Leiba A: High Normal Uric Acid Levels Are Associated with an Increased Risk of Diabetes in Lean, Normoglycemic Healthy Women. *J Clin Endocrinol Metab* 2016, 101(10):3772-3778.

10. Jepson C, Hsu JY, Fischer MJ, Kusek JW, Lash JP, Ricardo AC, Schelling JR, Feldman HI, Appel LJ, Feldman HI et al: Incident Type 2 Diabetes Among Individuals With CKD: Findings From the Chronic Renal Insufficiency Cohort (CRIC) Study. *AM J KIDNEY DIS* 2019, 73(1):72-81.

11. Lin CY, Hsieh MC, Kor CT, Hsieh YP: Association and risk factors of chronic kidney disease and incident diabetes: a nationwide population-based cohort study. *DIABETOLOGIA* 2019, 62(3):438-447.

12. Wu PP, Kor CT, Hsieh MC, Hsieh YP: Association between End-Stage Renal Disease and Incident Diabetes Mellitus-A Nationwide Population-Based Cohort Study. *J CLIN MED* 2018,
21. Nerpin E, Riserus U, Ingelsson E, Sundstrom J, Jobs M, Larsson A, Basu S, Arnlov J: Insulin sensitivity measured with euglycemic clamp is independently associated with glomerular filtration rate in a community-based cohort. *Diabetes Care* 2008, 31(8):1550-1555.

22. Levey AS, Coresh J, Balk E, Kausz AT, Levin A, Steffes MW, Hogg RJ, Perrone RD, Lau J, Eknoyan G: National Kidney Foundation practice guidelines for chronic kidney disease: evaluation, classification, and stratification. *Ann Intern Med* 2003, 139(2):137-147.

23. Levey AS, Bosch JP, Lewis JB, Greene T, Rogers N, Roth D: A more accurate method to estimate glomerular filtration rate from serum creatinine: a new prediction equation. *Modification of Diet in Renal Disease Study Group. Ann Intern Med* 1999, 130(6):461-470.

24. Siew ED, Ikizler TA: Insulin resistance and protein energy metabolism in patients with advanced chronic kidney disease. *Semin Dial* 2010, 23(4):378-382.

25. Koppe L, Pelletier CC, Alix PM, Kalbacher E, Fouque D, Soulage CO, Guebre-Egziabher F: Insulin resistance in chronic kidney disease: new lessons from experimental models. *Nephrol Dial Transplant* 2014, 29(9):1666-1674.

26. Katz AI, Rubenstein AH: Metabolism of proinsulin, insulin, and C-peptide in the rat. *J Clin Invest* 1973, 52(5):1113-1121.

27. Alberti KG, Zimmet PZ: Definition, diagnosis and classification of diabetes mellitus and its complications. Part 1: diagnosis and classification of diabetes mellitus provisional report of a WHO consultation. *Diabet Med* 1998, 15(7):539-553.
Table 1 The Baseline Characteristics of participants

| eGFR (quartile) | Q1(≤100.44) | Q2(100.44 to 112.40) | Q3(112.40 to 122.06) | Q4(>112.06) | P-value |
|-----------------|-------------|----------------------|-----------------------|-------------|---------|
| Participants    | 49843       | 49874                | 49858                 | 49860       | <0.001  |
| AGE(years)      | 52.91 ± 14.33 | 45.24 ± 10.58 | 38.27 ± 6.57 | 31.83 ± 4.54 | <0.001  |
| GENDER          |             |                      |                       |             | <0.001  |
| Male            | 33460 (67.13%) | 29983 (60.12%) | 26243 (52.64%) | 20004 (40.12%) |         |
| Female          | 16383 (32.87%) | 19891 (39.88%) | 23615 (47.36%) | 29856 (59.88%) |         |
| BMI(kg/m²)      | 24.12 ± 3.14 | 23.63 ± 3.19 | 23.05 ± 3.29 | 22.14 ± 3.41 | <0.001  |
| FPG(mmol/L)     | 5.05 ± 0.64 | 4.96 ± 0.63 | 4.85 ± 0.59 | 4.79 ± 0.55 | <0.001  |
| SBP(mmHg)       | 124.66 ± 18.11 | 120.19 ± 16.17 | 116.39 ± 14.72 | 114.46 ± 14.04 | <0.001  |
| DBP(mmHg)       | 76.85 ± 11.21 | 75.41 ± 10.89 | 73.15 ± 10.38 | 71.11 ± 9.69 | <0.001  |
| eGFR(mL/min·(1.73 m²)⁻¹) | 89.71 ± 8.51 | 106.68 ± 3.44 | 117.37 ± 2.77 | 127.98 ± 4.65 | <0.001  |
| TC(mmol/L)      | 4.95 ± 0.92 | 4.83 ± 0.90 | 4.63 ± 0.84 | 4.43 ± 0.82 | <0.001  |
| TG(mmol/L)      | 1.28 (0.90-1.88) | 1.16 (0.80-1.73) | 1.02 (0.71-1.51) | 0.88 (0.63-1.24) | <0.001  |
| HDL-C(tertile)  |             |                      |                       |             | <0.001  |
| Low             | 11327 (22.73%) | 10803 (21.66%) | 9350 (18.75%) | 6609 (13.26%) |         |
| Medium          | 11190 (22.45%) | 10066 (20.18%) | 8959 (17.97%) | 7339 (14.72%) |         |
| High            | 10106 (20.28%) | 10186 (20.42%) | 9755 (19.57%) | 9343 (18.74%) |         |
| Not recorded    | 17220 (34.55%) | 18819 (37.73%) | 21794 (43.71%) | 26569 (53.29%) |         |
| LDL-C(tertile)  |             |                      |                       |             | <0.001  |
| Low             | 8381 (16.81%) | 9057 (18.16%) | 10354 (20.77%) | 10560 (21.18%) |         |
| Medium          | 10951 (21.97%) | 10510 (21.07%) | 9512 (19.08%) | 7695 (15.43%) |         |
| High            | 14380 (28.85%) | 11556 (23.17%) | 8165 (16.38%) | 5000 (10.03%) |         |
| Not recorded    | 16131 (32.36%) | 18751 (37.60%) | 21827 (43.78%) | 26605 (53.36%) |         |
| ALT(U/L)        | 19.60 (14.50-28.00) | 19.20 (14.00-29.00) | 18.00 (12.50-28.00) | 15.20 (11.10-24.00) | <0.001  |
| AST(tertile)    |             |                      |                       |             | <0.001  |
| Low             | 5371 (10.78%) | 6413 (12.86%) | 7421 (14.88%) | 8710 (17.47%) |         |
| Medium          | 8037 (16.12%) | 7400 (14.84%) | 6612 (13.26%) | 6176 (12.39%) |         |
| High            | 9382 (18.82%) | 7399 (14.84%) | 6296 (12.63%) | 4987 (10.00%) |         |
| Not recorded    | 27053 (54.28%) | 28662 (57.47%) | 29529 (59.23%) | 29987 (60.14%) |         |
| Smoking status  |             |                      |                       |             | <0.001  |
| Never smoker    | 9741 (19.54%) | 9821 (19.69%) | 11386 (22.84%) | 12657 (25.39%) |         |
| Ever smoker     | 661 (1.33%) | 591 (1.18%) | 682 (1.37%) | 530 (1.06%) |         |
| Current smoker  | 3555 (7.13%) | 3628 (7.27%) | 2571 (5.16%) | 1574 (3.16%) |         |
| Not recorded    | 35886 (72.00%) | 35834 (71.85%) | 35219 (70.64%) | 35099 (70.40%) |         |
| Drinking status |             |                      |                       |             | <0.001  |
| Never drinker   | 11201 (22.47%) | 11198 (22.45%) | 12067 (24.20%) | 13026 (26.13%) |         |
| Ever drinker    | 2301 (4.62%) | 2419 (4.85%) | 2291 (4.60%) | 1612 (3.23%) |         |
| Current drinker | 455 (0.91%) | 423 (0.85%) | 281 (0.56%) | 123 (0.25%) |         |
| Not recorded    | 35886 (72.00%) | 35834 (71.85%) | 35219 (70.64%) | 35099 (70.40%) |         |
| Family history of diabetes |             |                      |                       |             | <0.001  |
| NO              | 49025 (98.36%) | 48807 (98.36%) | 48587 (97.45%) | 48820 (97.91%) |         |
| YES             | 818 (1.64%) | 1067 (2.14%) | 1271 (2.55%) | 1040 (2.09%) |         |
Values are n(%), mean±SD or medians (quartiles)

BMI, body mass index; SBP, Systolic blood pressure; DBP, Diastolic blood pressure; FPG, fasting plasma glucose; TC, Total cholesterol;

TG, Triglyceride; LDL-C, Low-density lipoprotein cholesterol; HDL-C, high-density lipoprotein cholesterol; ALT, Alanine aminotransferase;

AST, Aspartate aminotransferase; eGFR evaluated glomerular filtration rate

Scr, Serum creatinine.
Table 2 Incidence rate of incident diabetes

| eGFR(mL/min⁻¹·(1.73 m²)⁻¹) | Participants(n) | DM events(n) | Cumulative incidence(95% CI) | Per 100,000 person-year |
|-----------------------------|-----------------|-------------|-------------------------------|-------------------------|
| Total                       | 199435          | 3919        | 1.965 (1.904-2.026)           | 628.73                  |
| Q1                          | 49843           | 1765        | 3.541 (3.379-3.703)           | 1140.90                 |
| Q2                          | 49874           | 1217        | 2.440 (2.305-2.576)           | 776.55                  |
| Q3                          | 49858           | 612         | 1.227 (1.131-1.324)           | 386.46                  |
| Q4                          | 49860           | 325         | 0.652 (0.581-0.722)           | 211.19                  |

P for trend <0.001

MAP, Mean arterial pressure; CI, confidence interval; DM, diabetes mellitus.
Table 3 The results of univariate analysis

|                  | Statistics | HR (95% CI)               | P value |
|------------------|------------|---------------------------|---------|
| **Age (years)**  | 42.064 ± 12.530 | 1.066 (1.064, 1.069)     | <0.00001|
| **Gender**       |            |                           |         |
| Male             | 109690 (55.000%) | Ref.                     |         |
| Female           | 89745 (45.000%) | 0.476 (0.444, 0.511)     | <0.00001|
| **BMI(Kg/m^2)**  | 23.235 ± 3.339 | 1.237 (1.228, 1.246)     | <0.00001|
| **SBP(mmHg)**    | 118.922 ± 16.308 | 1.039 (1.037, 1.041)    | <0.00001|
| **DBP(mmHg)**    | 74.126 ± 10.782 | 1.047 (1.044, 1.049)     | <0.00001|
| **eGFR(mL/min\(1.73 \text{m}^2\))** | 110.438 ± 15.109 | 0.964 (0.962, 0.966)     | <0.00001|
| **FPG(mmol/L)**  | 4.913 ± 0.612 | 10.494 (10.035, 10.974)  | <0.00001|
| **TC(mmol/L)**   | 4.713 ± 0.892 | 1.428 (1.386, 1.471)     | <0.00001|
| **TG(mmol/L)**   | 1.338 ± 1.025 | 1.264 (1.252, 1.276)     | <0.00001|
| **HDL-C(mmol/L)**|            |                           |         |
| Low              | 38089 (19.098%) | Ref.                     |         |
| Medium           | 37554 (18.830%) | 0.836 (0.763, 0.917)     | 0.00013 |
| High             | 39390 (19.751%) | 0.735 (0.669, 0.809)     | <0.00001|
| Not recorded     | 84402 (42.321%) | 0.585 (0.541, 0.633)     | <0.00001|
| **LDL-C(mmol/L)**|            |                           |         |
| Low              | 38352 (19.230%) | Ref.                     |         |
| Medium           | 38668 (19.389%) | 1.157 (1.046, 1.281)     | 0.00462 |
| High             | 39101 (19.606%) | 1.701 (1.548, 1.868)     | <0.00001|
| Not recorded     | 85314 (41.775%) | 0.824 (0.751, 0.903)     | 0.0004  |
| **ALT(U/L)**     | 23.973 ± 22.031 | 1.005 (1.004, 1.005)     | <0.00001|
| **AST(U/L)**     | 27915 (13.997%) | Ref.                     |         |
| Low              | 28225 (14.152%) | 1.419 (1.235, 1.630)     | <0.00001|
| Medium           | 28064 (14.072%) | 2.775 (2.447, 3.147)     | <0.00001|
| High             | 115231 (57.779%) | 1.380 (1.228, 1.551)     | <0.00001|
| **Smoking status**|            |                           |         |
| Never smoker     | 43605 (21.864%) | Ref.                     |         |
| Ever smoker      | 2464 (1.235%) | 1.884 (1.483, 2.394)     | <0.00001|
| Current smoker   | 11328 (5.680%) | 2.308 (2.034, 2.618)     | <0.00001|
| Not recorded     | 142038 (71.220%) | 1.356 (1.244, 1.477)     | <0.00001|
| **Drinking status**|            |                           |         |
| Never drinker    | 47492 (23.813%) | Ref.                     |         |
| Ever drinker     | 8623 (4.324%) | 1.037 (0.883, 1.218)     | 0.65585 |
| Current drinker  | 1282 (0.643%) | 2.179 (1.620, 2.931)     | <0.00001|
| Not recorded     | 142038 (71.220%) | 1.072 (0.993, 1.156)     | 0.07433 |
| **Family history of diabetes**|| | |
| No               | 195239 (97.896%) | Ref.                     |         |
| Yes              | 4196 (2.104%) | 1.695 (1.448, 1.984)     | <0.00001|
Fig 1. Kaplan–Meier event-free survival curve. Kaplan–Meier analysis of incident of diabetes based on eGFR quartiles (log-rank, \( P < 0.0001 \)).
## Table 4 Relationship eGFR and the incident of diabetes in different models

| Exposure | Crude model (HR, 95% CI, P) | Adjust I (HR, 95% CI, P) | Adjust II (HR, 95% CI, P) |
|----------|-----------------------------|--------------------------|---------------------------|
| eGFR     | 0.964 (0.962, 0.966) < 0.00001 | 0.977 (0.975, 0.979) < 0.00001 | 0.987 (0.984, 0.989) < 0.00001 |
| eGFR (quartile) |                         |                          |                           |
| Q1       | 1.0                         | 1.0                      | 1.0                       |
| Q2       | 0.633 (0.589, 0.681) < 0.00001 | 0.765 (0.711, 0.824) < 0.00001 | 0.824 (0.765, 0.888) < 0.00001 |
| Q3       | 0.306 (0.279, 0.336) < 0.00001 | 0.444 (0.405, 0.488) < 0.00001 | 0.611 (0.555, 0.672) < 0.00001 |
| Q4       | 0.181 (0.160, 0.203) < 0.00001 | 0.314 (0.278, 0.354) < 0.00001 | 0.513 (0.453, 0.581) < 0.00001 |
| P for trend | < 0.00001                   | < 0.00001                | < 0.00001                 |

Crude model: we did not adjust other covariates.

Model I: we adjust gender, BMI, SBP, DBP, family history of diabetes, smoking and drinking status.

Model II: we adjust gender, BMI, SBP, DBP, FPG, TC, TG, HDL-C, LDL-C, ALT, AST, family history of diabetes, smoking and drinking status.

CI: confidence interval, Ref: reference.
Fig 2. The non-linear relationship between eGFR and incident of diabetes.

A non-linear relationship between them was detected after adjusting for gender, BMI, SBP, DBP, FPG, TC, TG, HDL-C, LDL-C, ALT, AST, family history of diabetes, smoking and drinking status.
| Fitting model by standard linear regression | 0.987 (0.984, 0.989) | <0.0001 |
|--------------------------------------------|----------------------|---------|
| Fitting model by two-piecewise linear regression |                       |         |
| Inflection point of eGFR ≤ 97.769 | 97.769 | 0.999 (0.994, 1.004) | 0.6928 |
| > 97.769 | 0.977 (0.974, 0.981) | <0.0001 |
| P for log likelihood ratio test | <0.001 |         |

Cl: Confidence interval

We adjusted gender, BMI, SBP, DBP, FPG, TC, TG, HDL-C, LDL-C, ALT, AST, family history of diabetes, smoking and drinking status.
### Table 6 Effect size of eGFR on incident diabetes in prespecified and exploratory subgroups

| Characteristic            | No of participants | HR (95% CI)       | P value      | P for interaction |
|---------------------------|--------------------|-------------------|--------------|------------------|
| **Age (years)**           |                    |                   |              |                  |
| 20 to <30                 | 26726              | 0.999 (0.978, 1.019) | 0.8890       |                  |
| 30 to <40                 | 78227              | 1.011 (1.003, 1.019) | 0.0098       |                  |
| 40 to <50                 | 42909              | 1.005 (0.998, 1.012)  | 0.1342       |                  |
| 50 to <60                 | 28388              | 1.002 (0.997, 1.008)  | 0.3874       |                  |
| 60 to <70                 | 16795              | 1.004 (0.997, 1.010)  | 0.2783       |                  |
| ≥70                       | 6390               | 1.010 (1.001, 1.020)  | 0.0389       |                  |
| **Gender**                |                    |                   |              |                  |
| Male                      | 109690             | 0.989 (0.987, 0.992)  | <0.0001      |                  |
| Female                    | 89745              | 0.981 (0.977, 0.985)  | <0.0001      |                  |
| **BMI (Kg/m²)**           |                    |                   |              | <0.0001          |
| <18.5                     | 11367              | 0.968 (0.946, 0.990)  | 0.0045       |                  |
| ≥18.5, <24                | 109984             | 0.980 (0.976, 0.984)  | <0.0001      |                  |
| ≥24, <28                  | 61047              | 0.986 (0.983, 0.989)  | <0.0001      |                  |
| ≥28                       | 17037              | 0.998 (0.994, 1.002)  | 0.4033       |                  |
| **FPG (mmol/L)**          |                    |                   |              | <0.0001          |
| <6.1                      | 192704             | 0.978 (0.975, 0.981)  | <0.0001      |                  |
| ≥6.1                      | 6731               | 0.994 (0.991, 0.997)  | 0.0004       |                  |
| **HDL-C (mmol/L)**        |                    |                   |              | 0.0204           |
| Low                       | 38089              | 0.991 (0.988, 0.995)  | <0.0001      |                  |
| Medium                    | 37554              | 0.983 (0.978, 0.988)  | <0.0001      |                  |
| High                      | 39390              | 0.987 (0.982, 0.992)  | <0.0001      |                  |
| Not recorded               | 84402              | 0.984 (0.981, 0.988)  | <0.0001      |                  |
| **LDL-C (mmol/L)**        |                    |                   |              | 0.0560           |
| Low                       | 38352              | 0.984 (0.979, 0.988)  | <0.0001      |                  |
| Medium                    | 38668              | 0.990 (0.985, 0.994)  | <0.0001      |                  |
| High                      | 39101              | 0.989 (0.985, 0.993)  | <0.0001      |                  |
| Not recorded               | 83314              | 0.984 (0.980, 0.988)  | <0.0001      |                  |
| **Smoking status**        |                    |                   |              | 0.0038           |
| Never smoker              | 43605              | 0.983 (0.977, 0.988)  | <0.0001      |                  |
| Ever smoker               | 2464               | 0.979 (0.963, 0.996)  | 0.0175       |                  |
| Current smoker            | 11328              | 0.999 (0.991, 1.007)  | 0.7830       |                  |
| Not recorded               | 142038             | 0.986 (0.984, 0.989)  | <0.0001      |                  |
| **Drinking status**       |                    |                   |              | 0.7418           |
| Never drinker             | 47492              | 0.986 (0.981, 0.990)  | <0.0001      |                  |
| Ever drinker              | 8623               | 0.987 (0.976, 0.998)  | 0.0216       |                  |
| Current drinker           | 1282               | 1.001 (0.974, 1.028)  | 0.9536       |                  |
| Not recorded               | 142038             | 0.986 (0.984, 0.989)  | <0.0001      |                  |
| **Family history of diabetes** |            |                   |              | 0.0416           |
| No                        | 195239             | 0.986 (0.984, 0.988)  | <0.0001      |                  |
| Yes                       | 4196               | 0.998 (0.987, 1.009)  | 0.6964       |                  |
|        |       |       |       |       |
|--------|-------|-------|-------|-------|
| **SBP** |       |       |       |       |
| <140   | 179686| 0.985 (0.982, 0.987) | <0.0001|       |
| ≥140   | 19749 | 0.992 (0.988, 0.997) | 0.0003 |       |
| **DBP** |       |       |       | <0.0001|
| <90    | 183594| 0.984 (0.982, 0.986) | <0.0001|       |
| ≥90    | 15841 | 0.999 (0.994, 1.005) | 0.8486 |       |

**Note 1:** Above model adjusted for gender, BMI, SBP, DBP, FPG, TC, TG, HDL-C, LDL-C, ALT, AST, family history of diabetes, smoking and drinking status.

**Note 2:** In each case, the model is not adjusted for the stratification variable.