Synergistic effect of non-alcoholic fatty liver disease and history of gestational diabetes to increase risk of type 2 diabetes

Yoosun Cho, MD, PhD[1]; Yoosoo Chang, MD, PhD[2,3,4*]; Seungho Ryu, MD, PhD[2,3,4*]; Sarah H. Wild, MB, BChir, PhD [5]; Christopher D. Byrne, MB, BCh, PhD[6,7]

[1]Total Healthcare Center, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

[2]Center for Cohort Studies, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

[3]Department of Occupational and Environmental Medicine, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Seoul, Republic of Korea

[4]Department of Clinical Research Design & Evaluation, Samsung Advanced Institute for Health Sciences & Technology, Sungkyunkwan University, Seoul, Republic of Korea

[5]Usher Institute, University of Edinburgh, Edinburgh, U.K.

[6]Nutrition and Metabolism, Faculty of Medicine, University of Southampton, Southampton, U.K.

[7]National Institute for Health and Care Research etc Southampton Biomedical Research Centre, University Hospital Southampton, Southampton, U.K.

[*] **Correspondence:**

Seungho Ryu, MD, PhD, Department of Occupational and Environmental Medicine, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Samsung Main Building B2, 250, Taepyung-ro 2ga, Jung-gu, Seoul 04514, Republic of Korea

Email: sh703.yoo@gmail.com

Yoosoo Chang, MD, PhD, Department of Occupational and Environmental Medicine, Kangbuk Samsung Hospital, Sungkyunkwan University School of Medicine, Samsung Main Building B2, 250, Taepyung-ro 2ga, Jung-gu, Seoul 04514, Republic of Korea

Tel: +82-2-2001-5139; Fax: +82-2-757-0436; Email: yoosoo.chang@gmail.com

Methods

Before constructing our prediction model, we split the entire cohort data into two parts: a training set and a validation set in an 8:2 ratio. Using the multivariable-adjusted Cox-proportional models, which included elements of ADA score plus GDM and NAFLD, we applied the models to a random selection of 80% of the entire cohort's data. To validate the prediction ability of our model, we used the remaining 20% of the cohort data. The derived estimated coefficient of the predicted model using Cox regression was then implemented to calculate the prognostic index (PI). The PI is a numerical score that reflects the likelihood of an individual developing T2DM. It was calculated using the following equation, where PI stands for the prognostic index [9].

$$PI = \sum_{i=1}^{k}\sum_{j=1}^{k}\hat{\beta}_{ij} * (Risk\ Factor)_{ij}\ \ , if\ j\ is\ a\ base\ category, \hat{\beta}_{ij} = 0$$

To evaluate the prediction ability of our model, we used the Harrell's C statistic, which is an estimate of the concordance probability adapted for survival analysis. This allowed us to assess the model's ability to discriminate between individuals who will develop T2DM and those who will not. We also used calibration to evaluate the prediction ability of the model. Decile categories are typically used, and the observed risk (proportion) is compared to the average predicted risk within each category using a chi-square statistic. Specifically, we compared the observed and predicted 5-year diabetes incidence risks in each decile using the Hosmer-Lemeshow test separately in the training set and internal validation set.

Results

We used Harrell's C-index to assess discrimination of our model, which included elements of the ADA score plus pGDM and NAFLD. The C statistic was 0.842 (0.830-0.853) for training set and 0.817 (0.788-0.845) for internal validation set, indicating good discrimination (**Supplementary Figure 1**). The C-statistics were 0.821 (0.810-0.833) for the ADA risk and 0.825 (0.813-0.853) for Leicester Diabetes Risk (UK risk) (**Supplementary Table 3**) after adding NAFLD and pGDM to those base models. The ADA and Leicester Diabetes Risk (UK risk) prediction models overestimated the actual risk in our population; the chi-square statistics were 217.0 (p value <0.001) and 214.8 (p value <0.001), respectively, indicating poor calibration. However, the calibration plot of our model for the training set and validation set showed that the predicted and observed probabilities were similar (**Supplementary Figure 2**). Although predicted probability was slightly higher than the observed risk, and this overestimation was slightly greater in the validation set than in the

training set, chi-square statistics for the training set and validation set were 13.16 (p value= 0.16) and 15.09 (p value = 0.09), respectively, indicating no significant difference between predicted and observed probability.

Supplementary Table and Figure of Contents

Supplementary Table 1. Baseline characteristics of the study participants who were included in the analysis, as well as those who were excluded due to missing information

Characteristic	Subjects excluded due to missing information	Subjects included in the analysis	*P* value	Standardized difference
Number	12,311 (11.2%)	97,347 (88.8%)		
Age (years)	43.7 (11.0)	39.0 (7.8)	<0.001	0.484
Alcohol intake (%) †	11.2	10.4	0.063	0.026
Current smoker (%)	1.8	1.3	0.096	0.041
Education level (%) ‡	71.6	80.6	<0.001	0.213
HEPA (%) §	14.2	12.7	<0.001	0.045
History of hypertension (%)	6.8	2.8	<0.001	0.190
History of CVD (%)	1.5	0.7	<0.001	0.082
Anti-lipid medication use (%)	4.5	1.4	<0.001	0.180
Family history of diabetes (%)	15.2	15.0	0.559	0.006
Obesity (%)‖	16.4	13.2	<0.001	0.089
BMI (kg/m^2)	22.1 (3.2)	21.8 (3.1)	<0.001	0.103
Waist circumference (cm)	75.6 (8.6)	74.6 (8.0)	<0.001	0.123
Systolic BP (mmHg)*	105.1 (12.4)	103.0 (10.7)	<0.001	0.184
Diastolic BP (mmHg)*	66.7 (8.7)	65.5 (8.2)	<0.001	0.141
Glucose (mg/dl)*	92.2 (8.2)	91.1 (7.6)	<0.001	0.131
Total cholesterol (mg/dl)*	191.8 (34.0)	186.4 (31.6)	<0.001	0.165
LDL-C (mg/dl)*	120.6 (32.4)	114.8 (29.7)	<0.001	0.188
HDL-C (mg/dl)*	66.5 (15.9)	68.0 (15.6)	<0.001	0.097

Data are expressed as *means (standard deviations) or percentages.
† ≥10 g/day; ‡ ≥college graduate; § health-enhancing physical activity; ‖ BMI ≥25 kg/m^2
Abbreviations: BMI, body mass index; BP, blood pressure; HDL-C, high-density lipoprotein cholesterol

Supplementary Table 2. Development of diabetes for population strata defined by prior gestational diabetes mellitus, and NAFLD after multiple imputation for missing data [a]

Group	Multivariable-adjusted HR[b](95% CI)
Neither pGDM nor NAFLD	1.00 (reference)
pGDM without NAFLD	2.50 (1.99-3.14)
NAFLD without pGDM	2.13 (1.86-2.43)
Both pGDM and NAFLD	5.80 (4.69-7.16)

This analysis included study participants who were included in the final analysis (n=97 347), as well as those who were excluded due to missing information (n=12 311)

[a] Missing data were replaced using multiple imputation by chained equations, resulting in 20 imputed data sets. Combined inferences were based on Rubin rules [3].

[b] Estimated from Cox proportional hazards models; multivariable Model 1 was adjusted for age, centre, year of screening examination, alcohol consumption, smoking status, physical activity, education level, hyperlipidaemia medication, family history of diabetes, history of hypertension, BMI, SBP; total cholesterol, HDL-C, triglyceride levels, HOMA-IR, and hs-CRP level.

Abbreviations: NAFLD, non-alcoholic fatty liver disease; pGDM, prior gestational diabetes mellitus

Supplementary Table 3. Comparison of the discriminatory power of prior history of gestational diabetes and NAFLD for incident type 2 diabetes using the ADA risk score and Leicester Diabetes Risk Score (UK risk score) as the base models

	Harrell's C−index		IDI	
	Value (95% CI)	***P*-value**	**value**	***P*-value**
ADA risk score[*]				
Base model	0.774 (0.762–0.787)	reference		reference
+ pGDM only	0.790 (0.778–0.803)	<0.001	0.028	<0.001
+ NAFLD only	0.812 (0.800–0.824)	<0.001	0.254	<0.001
+ Both pGDM and NAFLD	0.821 (0.810–0.833)	<0.001	0.279	<0.001
Leicester Diabetes Risk Score (UK risk score) [†]				
Base model	0.780 (0.767–0.792)	reference		reference
+ pGDM only	0.793 (0.781–0.805)	<0.001	0.063	<0.001
+ NAFLD only	0.816 (0.804–0.828)	<0.001	0.318	<0.001
+ Both pGDM and NAFLD	0.825 (0.813–0.836)	<0.001	0.340	<0.001

[*] Age, family history of diabetes, hypertension, BMI, and waist circumference were used as components.
[†] Age, family history of diabetes, hypertension, BMI, and physical activity were used as components.
Abbreviations: ADA, American Diabetes Association; BMI, body mass index; CI, confidence interval; IDI, integrated discrimination improvement; NAFLD, non-alcoholic fatty liver disease; pGDM, prior gestational diabetes mellitus

Supplementary Table 4. Number needed to screen in primary prevention of diabetes

Group	Case (n)	Incident cases	5-year cumulative Incidence (per 10^3 persons)	ARR (%)	NNS
Neither pGDM nor NAFLD	82,681	609	10.9	0.0	reference
pGDM without NAFLD	3,887	78	28.9	1.8	56
NAFLD without pGDM	10,088	712	102.3	9.1	11
Both pGDM and NAFLD	691	116	229.2	21.8	5

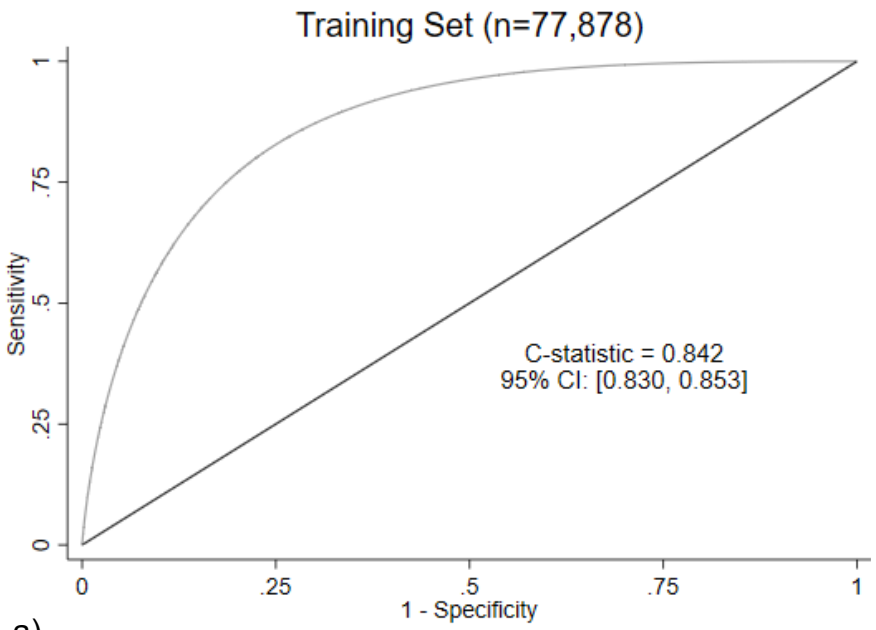
Number needed to screen (NNS) = 1/ARR
Abbreviation: ARR, absolute risk reduction; NNS, number needed to screen

Supplementary Table 5. Development of diabetes for population strata defined by prior gestational diabetes mellitus, and NAFLD according to age or BMI

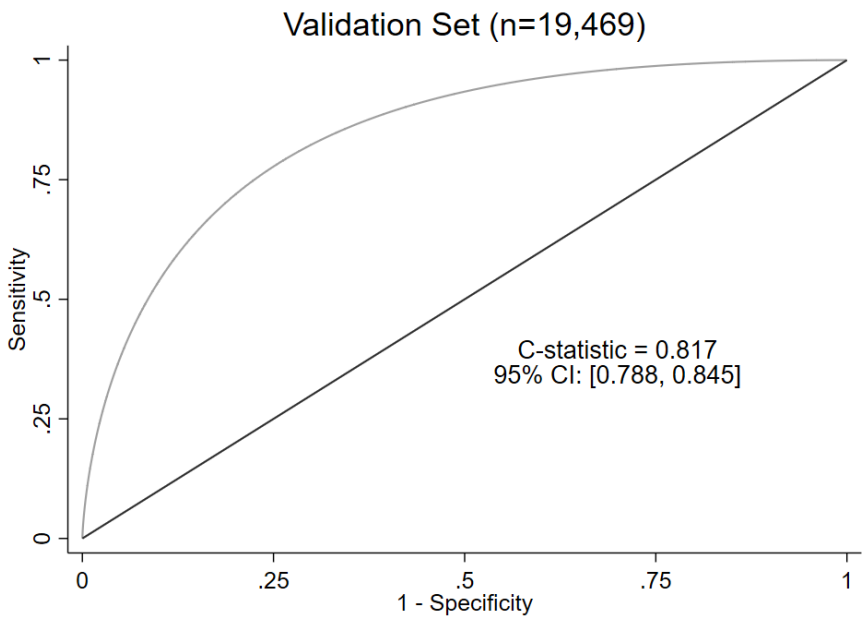
Group	Multivariable-adjusted HR* (95% CI)				
	Age <35 years (n = 30,967)	Age 35-39 years (n = 25,705)	Age ≥40 years (n = 40,675)	BMI <23kg/m^2 (n = 70,355)	BMI ≥23kg/m^2 (n = 26,992)
Neither pGDM nor NAFLD	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)	1.00 (reference)
pGDM without NAFLD	3.11 (1.65-5.87)	2.94 (1.96-4.40)	2.29 (1.64-3.20)	2.16 (1.53-3.05)	3.13 (2.26-4.33)
NAFLD without pGDM	2.95 (1.99-4.38)	2.16 (1.55-3.01)	2.14 (1.81-2.53)	1.97 (1.43-2.72)	2.29 (1.95-2.69)
Both pGDM and NAFLD	7.61 (4.29-13.50)	6.83 (4.39-10.62)	6.02 (4.54-7.99)	10.52 (6.33-17.47)	6.07 (4.76-7.74)
Measures of interaction on an additive scale					
P-value for the multiplicative interaction	0.649	0.802	0.344	0.006	0.403
P-value for the additive interaction	0.233	0.058	0.003	0.006	0.039
Relative excess risk due to interaction (PERI)	2.55 (-1.64 to 6.73)	2.73 (- 0.10 to 5.55)	2.59 (0.88-4.29)	7.39 (2.12-12.66)	1.65 (0.09-3.22)
Attributable proportion due to interaction (AP)	0.33 (-0.07 to 0.74)	0.40 (0.12-0.64)	0.43 (0.24-0.62)	0.70 (0.54-0.87)	0.27 (0.06-0.49)
Synergistic index	1.63 (0.45-2.81)	1.88 (0.82-2.94)	2.06 (1.17-2.95)	4.47 (1.44-7.51)	1.48 (0.93-2.04)

* Estimated from Cox proportional hazards models; multivariable Model 1 was adjusted for age, centre, year of screening examination, alcohol consumption, smoking status, physical activity, education level, hyperlipidaemia medication, family history of diabetes, history of hypertension, BMI, SBP; total cholesterol, HDL-C, triglyceride levels, HOMA-IR, and hs-CRP level.
Abbreviations; BMI, body mass index; NAFLD, non-alcoholic fatty liver disease

Supplementary Figure 1. a) Harrell's C-statistic for training set and b) validation set. Predictive performance of our model, which included elements of the ADA score plus pGDM and NAFLD, we used the Harrell's C-index, which measures discrimination ability in survival analysis.

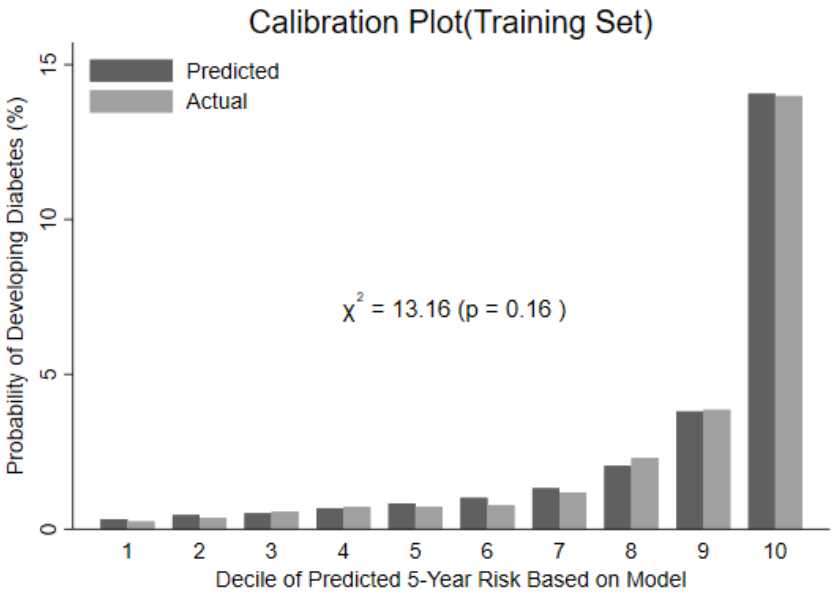


a)

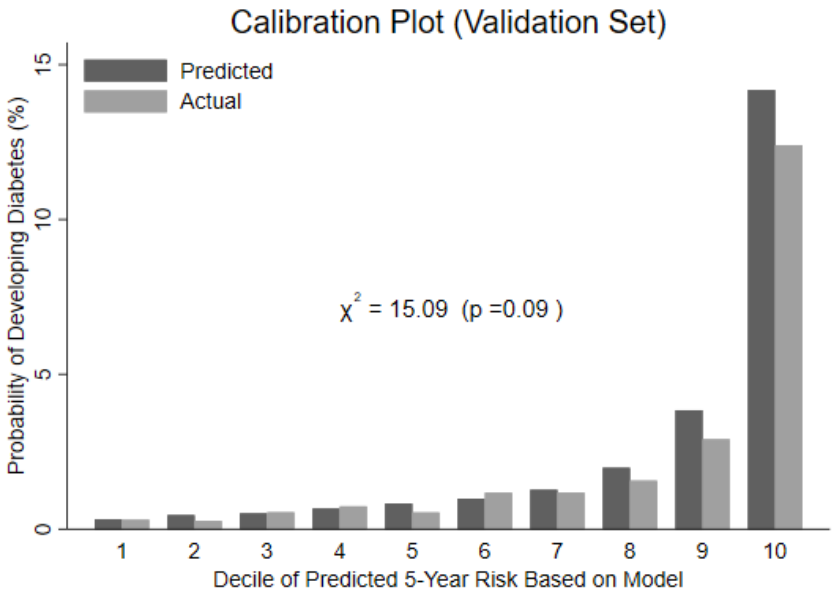


b)

Supplementary Figure 2. Calibration plots for a) training set and b) validation set. Calibration to evaluate the prediction ability of present model. Decile categories are typically used, and the observed risk (proportion) is compared to the average predicted risk within each category using a chi-square statistic.



a)



b)