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

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

Acknowledgements

We thank our staff members at the Kangbuk Samsung Health Study for their hard work, dedication, and support. This study was supported by the SKKU Excellence in Research Award Research Fund, Sungkyunkwan University (2021), and the National Research Foundation of Korea, funded by the Ministry of Science, ICT, and Future Planning (NRF-2021R1A2C1012626). CDB was supported in part by the Southampton National Institute for Health Research Biomedical Research Centre (NIHR203319), U.K.

Word count: abstract 242 words, **text** 5,261 words

ABSTRACT

Background: Whether non-alcoholic fatty liver disease (NAFLD) improves risk prediction for subsequent type 2 diabetes mellitus (T2DM) in women with prior gestational diabetes mellitus (pGDM) is uncertain. We examined the combined effects of NAFLD and pGDM on risk prediction for incident T2DM.

Methods: This retrospective cohort study included 97,347 Korean parous women without diabetes mellitus at baseline whose mean (SD) age was 39.0 (7.8) years. Cox proportional hazards models were used to estimate hazard ratios (HRs) for incident T2DM according to self-reported pGDM and ultrasound-diagnosed NAFLD at baseline. When combined with conventional diabetes risk factors, the incremental predictive ability of NAFLD and pGDM to identify incident T2DM was assessed.

Results: During a median follow-up of 3.9 years, 1,515 cases of incident T2DM occurred. Multivariable-adjusted HRs (95% confidence intervals [CIs]) for incident T2DM comparing pGDM alone, NAFLD alone, and both NAFLD and pGDM to the reference (neither NAFLD nor pGDM) were 2.61 (2.06-3.31), 2.26 (1.96-2.59), and 6.45 (5.19-8.00), respectively (relative excess risk of interaction=2.58 [95% CI, 1.21–3.95]). These associations were maintained after adjusting for insulin resistance, body mass index, and other potential confounders as time-dependent covariates. The combination of NAFLD and pGDM improved risk prediction for incident T2DM (based on Harrell's C–index, also known as the concordance index, and net reclassification improvement) compared to conventional diabetes risk factors.

Conclusion: NAFLD synergistically increases the risk of subsequent T2DM in women with pGDM. The combination of NAFLD and pGDM improves risk prediction for T2DM.

Keywords: Non-alcoholic fatty liver disease · Gestational diabetes mellitus · type 2 diabetes mellitus · cohort study

Introduction

Gestational diabetes mellitus (GDM), diabetes diagnosed in the second or third trimester of pregnancy that was not clearly overt diabetes prior to gestation [1], is a common endocrinopathy, and GDM prevalence has been rapidly increasing worldwide, with estimates of 14%, reaching up to 21% in South-East Asia [2]. GDM is recognized as a risk factor for type 2 diabetes mellitus (T2DM) [3] and cardiovascular disease (CVD) [4] and GDM is an established predictor of subsequent diabetes. One in five affected women reportedly develops subsequent T2DM [3]. Given the considerable rise in GDM prevalence, there is a need to identify modifiable risk factors of subsequent T2DM and apply preventive measures following GDM to minimize its adverse consequences.

Non-alcoholic fatty liver disease (NAFLD), the hepatic manifestation of metabolic syndrome, is a common chronic liver disease [5], leading to increased risks of hepatic complications, type 2 diabetes, CVD, CVD-related mortality, and all-cause mortality [6]. Interestingly, NAFLD in women with prior gestational diabetes mellitus (pGDM) predicts further deterioration of insulin resistance over time [7], along with altered lipid metabolism or obesity [8], which may persist or worsen in the postpartum period and lead to the development of T2DM. Growing epidemiological evidence supports the notion that pGDM is associated with an increased risk of NAFLD development [9, 10]. NAFLD increases the risk of incident T2DM [11], and combining NAFLD with conventional diabetes risk factors improves the risk prediction for incident T2DM [12].

Given that insulin resistance, T2DM, and NAFLD worsen the severity of each other as part of a vicious cycle, we hypothesized that the combination of NAFLD and pGDM would synergistically increase the risk of incident T2DM, and this synergistic increase in risk would be markedly greater than the risk attributed to either pGDM or NAFLD alone. Thus, we aimed to evaluate 1) the combined effect of NAFLD and pGDM on the risk of incident T2DM and 2) whether the combination of NAFLD and pGDM improved risk prediction for subsequent T2DM, over and above conventional diabetes risk factors, in a large sample of healthy parous Korean women.

Methods

Study population

As part of the Kangbuk Samsung Health Study, the current cohort study comprised Korean parous women who had annual or biennial health screenings at Kangbuk Samsung Hospital Total Healthcare Centers in Seoul and Suwon, South Korea [13]. Our study was limited to women who had one or more births, underwent a comprehensive health examination between 2015 and 2019, and had at least one follow-up visit before December 2020 (n=143,255). We excluded women who met the following criteria: were pregnant at baseline; or had known liver disease; prevalent diabetes; history of malignancy; significant alcohol consumption [14]; positive serologic markers for hepatitis B or C, use of steatogenic mediations. or had missing information on GDM, alcohol intake, menopausal status, body mass index (BMI), liver ultrasound and laboratory data, including blood glucose and HbA1c; Ultimately, 97,347 eligible women were included (**Figure 1**). The sample size for this study was based on the fixed available sample; specifically, the number of parous women enrolled during the study period determined our sample size. This study was approved by the Institutional Review Board of Kangbuk Samsung Hospital (IRB No. KBSMC_2022-06-007) and exempt from the requirement for informed consent due to the use of anonymised retrospective data that were routinely collected during health examinations.

Data collection

The dataset consists of questionnaire-based information on sociodemographic and health-related behaviors, medical history, parity (number of deliveries including stillbirths and live births), pregnancy history, and other reproductive characteristics, as well as anthropometric and laboratory measurements [13]. Smoking status was assessed by questionnaire, in which participants were asked to report whether they were never smokers, former smokers, or. The questionnaire also included questions related to smoking habits, such as lifetime and current smoking status, smoking duration, and the average number of cigarettes smoked per day. Participants who had smoked fewer than 100 cigarettes during their lifetime were considered never-smokers. Participants who had smoked more than 100 cigarettes were classified as either current smokers, who were actively smoking at the time of the questionnaire, or former smokers, who no longer smoked at the time of their screening examination. The frequency of alcohol consumption and amount of alcohol consumed per drinking day were recorded in standard units and used to calculate the average alcohol consumption per day. Physical activity levels were recorded using the validated Korean version of the International Physical Activity Questionnaire short form, converted to metabolic equivalents (min/week), and categorized as inactive,

minimally active, or health-enhancing physical activity (HEPA) [15]. Menopause was defined as ≥1 year of amenorrhoea. Obesity was defined as a BMI of ≥25 kg/m^2 according to Asian-specific criteria [16]. Hypertension was defined as blood pressure (BP) ≥140/90 mmHg or the use of BP-lowering medication. Blood samples were obtained after at least 10 h of fasting to measure laboratory glycaemic parameters, fasting serum lipid profiles, liver enzymes, and high-sensitivity C-reactive protein levels (hs-CRP). Homeostatic model assessment of insulin resistance (HOMA-IR) was calculated using the following equation: fasting blood insulin (mU/mL) × fasting blood glucose (mmol/L) / 22.5.

Definition of GDM history and type 2 diabetes

Screening for GDM is performed in almost all pregnant women, either through a two-step approach using the 50-g oral glucose tolerance test (OGTT) followed by the 100-g OGTT (diagnosis of GDM requires at least two abnormal values, such as fasting glucose ≥95 mg/dl, 1-h glucose ≥180 mg/dl, 2-h glucose ≥155 mg/dl, and 3-h ≥140 mg/dl) [17] or a one-step approach using the 75-g OGTT (diagnosis of GDM requires only one abnormal value, such as fasting glucose ≥92 mg/dl, 1-h glucose ≥180 mg/dl, and 2-h glucose ≥153 mg/dl), according to the International Association of Diabetes and Pregnancy Study Groups criteria in 2011 [18]. While there is no validation study available to assess the accuracy of self-reported GDM in our population, self-reported GDM diagnosis showed a sensitivity of 93% and a specificity of 100% when medical records were used as the reference standard, and these results did not differ by ethnicity or socioeconomic status [19, 20].

During the health screening examination, pGDM was assessed using a self-report questionnaire. The pGDM-specific question was, "Have you ever been diagnosed with gestational diabetes?" with two response options (yes or no). Women who responded ‘yes’ were considered to have pGDM. Importantly, all pregnant women in South Korea are recommended to undergo GDM screening at 24–28 weeks, regardless of the underlying GDM risk [21]. Due to about 97% health care coverage by the Korean National Health Insurance Service (KNHIS) [22], almost all pregnant women are likely to undergo GDM screening during pregnancy. Although no available validation study for the accuracy of self-reported pGDM in our population is available, the sensitivity and specificity of self-reported GDM diagnosis using medical records as the reference standard

were 93%–95% and 100%, respectively, and did not differ by ethnicity or socioeconomic status [19, 20, 23].

T2DM was defined as fasting serum glucose level ≥126 mg/dL, HbA1c ≥6.5% (48 mmol/mol), a history of physician-diagnosed diabetes, or the current use of glucose-lowering medications for diabetes treatment.

NAFLD diagnosis

Abdominal ultrasonography was conducted by experienced radiologists who were unaware of the study's objectives; they diagnosed fatty liver according to the following standard criteria: a diffuse increase in fine echoes in the liver parenchyma compared with those in the kidney or spleen parenchyma, deep beam attenuation, and bright vessel walls [24]. The inter-observer and intra-observer reliability values for fatty liver diagnosis were substantial (kappa statistic = 0.74) and excellent (kappa statistic = 0.94), respectively [13]. Since secondary causes of steatosis, including excessive alcohol use (≥20 g/d for women), were already excluded (see the flow chart in **Figure 1**), we considered an US-diagnosed fatty liver as a diagnosis of NAFLD.

Statistical analysis

We assessed the combined effects of pGDM and NAFLD on the development of T2DM. During follow-up, the primary endpoint was incident T2DM. The incidence was described as the number of cases per 1,000 person-years. Follow-up started from the baseline visit and was terminated at the endpoint or the last health screening exam (31 December 2020), whichever occurred first. Cox proportional hazard models were used to estimate hazard ratios (HRs) with 95% confidence intervals (CIs) for incident T2DM, comparing women with and without (the reference group) pGDM.

The models were adjusted as follows. Model 1 was adjusted for age, center (Seoul or Suwon), examination year, alcohol consumption (<10 or ≥10 g/day), smoking status (never, former, current, or unknown), physical activity level (inactive, minimally active, HEPA, or unknown), education level (below college graduate, college graduate or higher, or unknown), family history of diabetes, history of hypertension, use of lipid-lowering medication, and BMI. Model 2 was further adjusted for systolic (SBP), lipid profile (total cholesterol, high-density lipoprotein cholesterol [HDL-C], and triglyceride levels), HOMA-IR, and hs-CRP levels. To incorporate the effects of changes in NAFLD status, pGDM, BMI, and other covariates during the follow-up

period, we conducted time-dependent analyses in which the updated GDM history status, NAFLD, and changes in BMI and other covariates were treated as time-varying covariates. We adjusted for measured potential confounders that might affect the relationship between the prior history of GDM (pGDM), NAFLD and incident type 2 diabetes mellitus (T2DM). The confounding variables were defined using the following criteria: 1) association with the outcome (T2DM), 2) association with exposures (prevalent NAFLD or pGDM), and 3) not known to be a mediator in the causal pathway between exposure (and the outcome). To investigate the potential impact of unmeasured confounding, we estimated the E-value, a measure of the minimum strength of association that an unmeasured confounder would need to have with the exposures (pGDM and NAFLD) and incident diabetes, assuming equal strengths of association [25].

We assessed additive and multiplicative interactions to examine the interaction effect of pGDM and NAFLD on incident T2DM. We determined the relative excess risk due to interaction (RERI), attributable proportion due to interaction (AP), synergy index (S), and their CIs as follows [26]:

a. The relative excess risk due to interaction (RERI):

$$RERI = HR_{++} - HR_{+-} - HR_{-+} + 1$$

b. The attributable proportion due to interaction (AP):

$$AP = RERI/HR_{++}$$

c. The synergy index (S):

$$S = [HR_{++} - 1] / [(HR_{+-} - 1) + (HR_{-+} - 1)]$$

To investigate potential multiplicative interaction, the multivariable model included pGDM, NAFLD, and the product terms of pGDM and NAFLD, along with other confounders. Multiplicative interaction was evaluated using likelihood ratio tests that compared models with and without the product terms of pGDM and NAFLD.

Harrell's C−index, a measure of the concordance probability adapted for survival analysis, was use to assess whether the addition of pGDM and NAFLD alone or in combination to the base model, including conventional risk factors, improved prediction of T2DM. The concordance probability is the most commonly applied global measure of discrimination when the outcome is survival time [27]. We further calculated the and integrated discrimination improvement (IDI) to quantify the incremental predictive ability by adding pGDM, NAFLD, or both to conventional risk factors. Compared with two validated prediction tools, the Leicester Diabetes Risk Score (age, sex, family history of diabetes, hypertension, BMI, and waist circumference) [28] and

the American Diabetes Association (ADA) Risk Score (age, family history of diabetes, hypertension, BMI, and physical activity) [29], as the base model, we evaluated the incremental predictive ability of pGDM, NAFLD, or both added to the base model to predict incident type 2 diabetes. We also evaluated our model's discrimination, which included elements of the ADA score plus pGDM and NAFLD, and calibration through internal validation (refer to the **Supplementary Information** for details).

Moreover, to estimate clinical significance of women with pGDM and NAFLD, we calculated the number needed to screen (NNS) [30] according to the strata of pGDM and NAFLD. We explored potential effect modification by performing subgroup analyses stratifying by age group (<35 years, 35-39 years, and ≥40 years) and BMI (<23 kg/m2 and ≥23 kg/m2) at baseline with further adjustment for potential confounders. We used STATA version 17·0 (Stata Corp LP, College Station, TX, USA) for statistical analyses. Statistical significance was defined as a two-sided P-value <0·05.

Results

The baseline characteristics of the participants are depicted for the four strata according to pGDM and NAFLD status (**Table 1**). The prevalence of NAFLD in women with and without pGDM was 15.1% and 10.9%, respectively. Women with pGDM and NAFLD were more likely to be older; current smokers; have a family history of diabetes, history of CVD, hypertension; and have higher use of lipid-lowering medication, higher levels of BMI, waist circumference, glycaemic parameters, lipid profiles (total cholesterol, low-density lipoprotein cholesterol, and triglycerides), liver enzymes, hs-CRP, and HOMA-IR levels than those with neither NAFLD nor pGDM. Conversely, women with both NAFLD and pGDM were more likely to be physically inactive and have lower HDL-C levels than those without NAFLD or pGDM.

The proportion of people who were excluded due to missing information at baseline was 11.2%. **Supplementary Table 1** shows baseline characteristics of the study participants who were included in the analysis (n=97,347), as well as of those who were excluded due to missing information (n=12,311). Participants who were excluded were found to be older, to have lower levels of education and more unfavorable metabolic profiles in comparison to those who were included in the study. We have performed a sensitivity analysis using a multiple imputation method to account for missing data in our dataset [31]. The results of this analysis were consistent with our original findings (**Supplementary Table 2**).

With almost 350,000 person-years of follow-up (median 3·9 years, interquartile range 2.3–4.7 years, maximum 6.0 years), 1,515 cases of incident T2DM were identified (incidence rates per 10^3 person-years were 4.4 in total; 52.3 for NAFLD with pGDM) (**Table 2**). The 5-year cumulative incidence (per 10^3 persons) of T2DM for different strata defined by pGDM and NAFLD status (neither pGDM nor NAFLD, pGDM without NAFLD, NAFLD without pGDM, and both pGDM and NAFLD) were 10.9, 28.9, 102.3, and 229.2, respectively. After adjusting for potential confounders such as age, center, screening year, alcohol intake, smoking status, physical activity level, education level, lipid-lowering medication use, family history of diabetes, history of hypertension, and BMI (Model 1), the multivariable-adjusted HRs (95% CIs) for incident T2DM were 2.71 (2.14–3.43) for pGDM only, 3.67 (3.22–4.18) for NAFLD only, and 11.02 (8.94–13.57) for both pGDM and NAFLD compared to neither condition. These associations were only slightly attenuated after further adjustment for SBP and total cholesterol, HDL-C, triglyceride, hs-CRP, and HOMA-IR levels (Model 2). The multivariable-adjusted HRs (95% CIs) for incident T2DM were 2.61 (2.26–3.31) for pGDM only, 2.26 (1.96–2.59) for NAFLD only, and 6.45 (5.19–8.00) for both pGDM and NAFLD compared to neither condition. In our study, only 29 individuals (less than 0.03% of the study population) met the criteria for NAFLD with high fibrosis-4 scores of ≥2.67, which limited our ability to further evaluate the impact of fibrotic NAFLD on the risk of developing T2DM. The sensitivity analysis estimating the E value suggests that an unmeasured confounder with a risk ratio of 12.38 (lower CI: 9.85) for its association with pGDM, NAFLD and incident diabetes, would be required to explain our finding of an HR adjusted for known confounders of 6.45 (5.19-8.00) of pGDM and NAFLD for incident diabetes which seems unlikely.

pGDM and NAFLD status and risk of incident diabetes had an additive interaction (*P*-value <0·001). (**Table 2**). Specifically, the RERI was 2.58 (95% CI, 1.21-3.95). Moreover, the AP was 0.40 (95% CI, 0.25-0.55), that is, 40% of the risk of subsequent T2DM was due to the synergistic interaction between pGDM and NAFLD, and the synergy index was 1.90 (95% CI, 1.32-2.48). All measures of interaction on the additive scales were statistically significant, but the multiplicative interaction of pGDM and NAFLD with T2DM incidence was not (*P*-value = 0.561). In a time-dependent model that included changes in pGDM, NAFLD, BMI, HOMA-IR, hs-CRP, and other covariates as time-dependent covariates, aHRs (95% CIs) remained statistically significant.

The incremental predictive ability for incident T2DM was determined after adding pGDM, NAFLD, or both to the base model consisting of conventional T2DM risk factors [32, 33] (**Table 3**). Adding pGDM or

NAFLD separately, or both together, to the base model improved the Harrell's C–index for predicting incident T2DM. Adding both pGDM and NAFLD together markedly improved the prediction of incident T2DM (Harrell's C–index, 95% CI 0.825, 0.813-0.836). On comparing the predictive ability of pGDM or NAFLD alone, and both pGDM and NAFLD, a statistically significant increase in the category-based IDI was observed when added to the base model, showing the highest improvement with the addition of both pGDM and NAFLD to the base model. Similarly, a statistically significant improvement in predictive ability based on Harrell's C-index, and IDI was also observed when both pGDM and NAFLD were added to the pre-existing risk models (**Supplmental Table 3)**. When generating T2DM prediction models that incorporated elements of the ADA T2DM risk score as well as pGDM and NAFLD, our study found that these models, had good discrimination and calibration through internal validation **(Supplementary Figure 1 and 2**). The addition of NAFLD and pGDM to the elements of the ADA prediction model improved the accuracy of T2DM risk prediction in our population. However, further external validation studies in diverse cohorts are necessary to test the generalizability of our model to other ethnic groups (refer to the **Supplementary Information** for details).

The NNS values for individuals with pGDM without NAFLD, NAFLD without pGDM, and both pGDM and NAFLD in comparison to those with neither pGDM nor NAFLD were 56, 11, 5, respectively (**Supplementary Table 4**).

The stratified analyses are reported in **Supplementary Table 5** show similar findings across the age subgroups. The association between pGDM and NAFLD status and risk of incident diabetes tended to be stronger in women with BMI <23 kg/m2 than those with BMI ≥23 kg/m2 with evidence of both additive interaction and multiplicative interaction (P-value = 0.006). Moreover, the attributable proportion was 0.70 (95% CI, 0.54-0.87), that is, 70% of the risk of subsequent T2DM was due to the synergistic interaction between pGDM and NAFLD, and the synergy index was 4.47 (95% CI, 1.44-7.51). Since obesity is a potent risk factor for both NAFLD and T2DM, and attenuates the pGDM and NAFLD-associated T2DM risk, the association appeared to be more robust in non-overweight women.

Discussion

Whilst it is well established that NAFLD increases the risk of incident T2DM approximately two fold, we show for the first time that the presence of NAFLD in women with pGDM highly elevated its risk more than

297 11-fold, compared to neither condition. This association remained statistically significant after adjustment for
298 BMI, HOMA-IR levels, and other covariates at baseline and, importantly, after further adjustment for an
299 updated status of these variables during follow-up as time-dependent covariates in regression modeling. The
300 excess risk of T2DM in women with both NAFLD and pGDM was greater than that in women with either
301 condition alone, with a statistically significant additive interaction suggesting a synergistic effect of NAFLD and
302 pGDM on T2DM risk.

303 Moreover, the attributable proportion of risk due to the interaction between NAFLD and pGDM was
304 0.40, indicating that 40% of the risk of subsequent T2DM was due to the synergistic interaction between pGDM
305 and NAFLD. Furthermore, an incremental predictive ability analysis for T2DM showed that the highest
306 improvement in Harrell's C–index, and IDI occurred by adding both NAFLD and pGDM to the base model.
307 Thus, women with both NAFLD and pGDM represent a very high-risk group for subsequent T2DM, who would
308 potentially benefit most from intensive risk-minimizing measures.

309 The prevalence of NAFLD in women with a prior history of pGDM in our study was lower than that
310 reported in two previous studies [34, 35]. This discrepancy may be due to differences in ethnicity or in study
311 design. The two previous studies were undertaken in Indian and Canadian women, whose BMI was higher,
312 leading to a higher prevalence of insulin resistance and NAFLD. Previous studies were also conducted in a
313 hospital setting, whereas our study used retrospective data from the general population who underwent health
314 examinations, which likely included a larger proportion of participants without disease. Additionally, the
315 definition of NAFLD differed in the study for Canadians, with Mehmood et al. using liver fat scores instead of
316 ultrasound as used in our study [35]. Moreover, the precise timing between pGDM onset and NAFLD onset
317 remains unknown due to the design of the Kangbuk Samsung Health Study, whereas each previous study
318 utilized a median of 16 months of postpartum for Indians [34] and mean of 4.8 years of postpartum for
319 Canadians [35] as the timeframe. These factors may contribute to the lower prevalence of NAFLD in women
320 with pGDM in our study compared to previous studies.

321 It is important to establish how NAFLD interacts with other risk factors to further magnify the risk of
322 developing diabetes. The combined effect of pGDM and NAFLD on incident T2DM is poorly studied. In a
323 small cohort study comprising only 68 pGDM participants versus 29 healthy controls, postpartum NAFLD

(identified by the fatty liver index biomarker test) with pGDM was strongly associated with increased levels of insulin resistance and inflammatory cytokines, which conferred a high risk of diabetes during follow-up [7]. However, the fatty liver index is a relatively poor diagnostic test for NAFLD, and the very small sample size of the population limited the generalizability of the results [7]. In our large cohort of 97,347 middle-aged parous women, the coexistence of ultrasound-defined NAFLD and self-reported pGDM considerably increased the risk of incident T2DM compared with individual factors. Thus, our findings suggest that postpartum risk assessment for subsequent T2DM in women with pGDM identifies a subgroup with NAFLD who are at very high risk of subsequent T2DM, in whom early surveillance of metabolic risk factors (such as NAFLD) may be needed.

We found an additive interaction between NAFLD and pGDM on the subsequent T2DM; the coexistence of prevalent NAFLD and pGDM was associated with a higher risk of subsequent diabetes than the sum of the individual effects. Although the multiplicative interaction was not statistically significant, the positive interaction on the additive scales confirmed the synergistic effect of NAFLD and pGDM on incident T2DM, and the additive interaction is less prone to error in the estimates than multiplicative models [36]. Furthermore, there is increasing evidence of a bidirectional association between NAFLD and GDM, which could be explained by common shared pathways such as insulin resistance, obesity, and altered lipid metabolism [8]. Previous cohort studies have reported that NAFLD in early pregnancy was an independent risk factor for GDM [37, 38] and improved the predictability of incident GDM in 1,443 Korean pregnant women [39]. Women with pGDM are at increased risk of NAFLD development and vice versa, although inconsistent results have been observed upon further adjustment for incident T2DM during a follow-up period [9, 10].

Notably, adding conventional risk factors for diabetes to NAFLD and pGDM status, are more likely to improve risk prediction for incident T2DM (*C* statistics ~0.825), than pGDM or NAFLD alone (*C* statistics ~0.793 or ~0.816, respectively). On evaluating the incremental predictive ability of the model after adding either NAFLD or pGDM, or both, to the base model; the addition of both NAFLD and pGDM together showed the highest improvement in Harrell's C–index, and IDI. No previous studies have assessed whether the prevalence of NAFLD in women with pGDM improves the predictability of T2DM, although a few have estimated the discriminatory power of either pGDM or NAFLD alone, in predicting the risk of incident T2DM [12, 40, 41]. Another previous study of 256 women with pGDM demonstrated that moderate fatty liver was associated with 3.6 fold higher pre-diabetes/diabetes prevalence defined by fasting glucose or 2-hour glucose levels [35]. A

previous validation study testing 12 non-invasive prediction models for T2DM development showed that the *C* statistics ranged from 0.78 to 0.81 in women [41]. A result of 0.81 is similar to the results in our models showing that NAFLD alone improved the discrimination of traditional risk for diabetes (*C* statistics ~0.816). The only other cohort study of women with pGDM that estimated the discriminatory performance of pGDM in predicting T2DM development (the base model included six risk factors: age, BMI, fasting glucose in pregnancy and postpartum, triglyceride, and total cholesterol levels) showed that the C statistic was 0.756 (which was considerably lower than the C statistic in our study: 0.793) [40]. Thus, our novel findings suggest that, beyond conventional risk factors of T2DM, a diagnosis of NAFLD in women with pGDM provides important valuable additional information on who should be identified as being at risk for T2DM.

Our findings, when compared to those previously reported in the general population, suggest that T2DM screening may be particularly beneficial for women with NAFLD regardless of pGDM status. A previous study evaluating screening strategies to identify T2DM in Korea [42] estimated the NNS to detect one T2DM case to be 63-71 for ages 30-34 years and 34–42 for ages 35-39 years if the whole population was screened at the each age group. If screening was limited to patients with both pGDM and NAFLD, then only 5 individuals would need to be screened to identify one T2DM case. Early diagnosis and intervention have been shown to be effective in preventing or delaying the onset of T2DM in individuals at high risk [43], and moreover, stepwise primary care–based screening and treatment primarily accrue from early diagnosis, that is, by reducing the delay time between diabetes onset and clinical diagnosis and by hastening the treatment of cardiovascular risk factors [44]. Hence, it is helpful to identify women at increased risk of development of T2DM after a pregnancy complicated by GDM, and women who develop NAFLD who can be expected to receive the most benefit from surveillance and effective interventions, ultimately contributing to reducing the social cost of diabetes and its complications.

Our study found that the effect size for NAFLD on risk of subsequent T2DM appears to be greater than that of pGDM. It is important to note that the relative risk of T2DM associated with pGDM declines with time after pregnancy with GDM, although higher risk relative to women without pGDM remained increased even after >35 years [45]. Since we excluded participants with diabetes at baseline in our study design, the effect of pGDM on incident T2DM is likely be underestimated, as some participants may have developed diabetes after pregnancy but before the study baseline visit. Additionally, we used a self-report questionnaire,

specifically asking about a history of gestational diabetes, which, although accurate with a sensitivity of 93%–95% and specificity of 100% [19, 20, 23], could still result in misclassification bias. Any non-differential misclassification of the history of GDM would attenuate the strength of associations with T2DM.

Although the pathogenesis of NAFLD and pGDM in association with incident T2DM is not fully understood, common pathophysiological factors such as abnormal lipid metabolism, increased insulin resistance, and systemic inflammation, when combined, may interact synergistically to increase T2DM development [46]. Women with GDM are predisposed to inadequate pancreatic β-cell compensation and deterioration of β-cell function, amplified by the insufficient insulin secretion placed by chronic insulin resistance [47, 48]. Fatty liver is promoted by the accumulation of fat in hepatocytes in the form of hepatic diacyl glycerols (DAGs) and triglycerides, which specifically exacerbate the deterioration of β-cell function, resulting in impaired insulin sensitivity to meet the increased glucose demands determined by insulin resistance [35] and inflammation with the secretion of pro-inflammatory cytokines and hepatokines, that is, fetuin A, fetuin B, antiopoietin-like protein, or fibroblast growth factor 21 (FGF21), resulting in progression to T2DM [49]. Furthermore, a vicious cycle can develop when increased insulin resistance, a hallmark of GDM, and NAFLD coexist, resulting in pronounced deterioration of insulin resistance and inflammation, leading to a greater risk of subsequent T2DM than with either condition alone [8]. Interestingly, in our study, synergic effect of pGDM–NAFLD on T2DM remained statistically significant after adjusting for baseline or changing status of the pathophysiological factors previously described: BMI, HOMA-IR, fasting triglycerides, and hs-CRP. Thus, the effect of NAFLD with pGDM on T2DM development is not fully explained by these confounders, suggesting an independent and synergistic role of pGDM–NAFLD in T2DM development.

The current study has a few limitations that are worth mentioning. First, we used self-reported data to define pGDM using a self-administered structured questionnaire. However, self-reported diagnosis of GDM is accurate, with a sensitivity of 93%–95% and specificity of 100% [19, 20, 23]. Women who reported pGDM are highly likely to have been tested for diabetes during pregnancy, and any non-diagnosis of GDM due to missed testing, would result in misclassification bias. Any misclassification of the diagnosis of GDM (or NAFLD) would attenuate the strength of our strong associations toward the null. Second, we performed ultrasonography instead of liver biopsy, liver magnetic resonance imaging, or computed tomography to define NAFLD. However, liver biopsy is neither ethical nor feasible in health participants, and imaging techniques such as computed

tomography or magnetic resonance imaging are not practical and too expensive for routine health care check-ups in large populations. Third, we could not rule out the possibility of unmeasured or residual confounding by pre-pregnancy or pregnancy risk factors. It was therefore not possible to examine the role of these factors in the association between prevalent NAFLD in women with pGDM and subsequent risk of diabetes. The HOMA-IR, used in our study, is not sophisticated method to assess insulin resistance. However, HOMA methods for estimating insulin resistance, less invasive and relatively easy to perform and thus more applicable to the population studies, have been validated against insulin clamp studies [50]. Fourth, the definition of T2DM was based on fasting glucose and HbA1c measurements without data from a 2-hour glucose test. Since the participants' laboratory data were obtained only in a fasting state, 2-hour glucose levels were unavailable. However, in clinical practice around the globe, HbA1c is now widely accepted as a diagnostic test for type 2 diabetes diagnosis and monitoring. It is also advantageous to use HbA1c in large cohort studies such as ours because it has preanalytical stability and is not affected by acute perturbations (for example, diet, exercise, and smoking) [51]. Fifth, we could not obtain the GDM screening rate in all our participants' delivery cases. However, all Korean pregnant women are recommended to take GDM screening regardless of predisposing GDM risk, and almost all pregnant women undergo GDM screening during pregnancy for KNHIS coverage. Finally, since our study included relatively healthy, middle-aged Korean women with easy access to healthcare facilities, our findings need to be tested in other populations in the future.

Our cohort study is the first to demonstrate that NAFLD in women with pGDM is a much more potent risk factor and predictor of T2DM development, than NAFLD or pGDM alone. Importantly, this effect was independent of all measured potential confounders at baseline and also after taking account of change in these potential confounders between baseline and follow-up. The combination of NAFLD and pGDM increased the risk of incident T2DM 11 fold, compared to the absence of both risk factors. Also presence of NAFLD and pGDM improved the prediction of T2DM risk, over and above the presence of conventional diabetes risk factors. Thus, we suggest that the presence of NAFLD in women with pGDM might help identify a subgroup of women at very high risk of developing T2DM, who are particularly likely to benefit from prevention measures that are known to attenuate risk of both NAFLD and T2DM.

Funding

This research did not receive any specific grant from funding agencies in the public, commercial, or not-for-profit sectors.

Competing Interests

The authors have no relevant financial or non-financial interests to disclose.

Ethical approval

No ethical permission was required due to the use of anonymised data.

Informed consent

Informed consent was not obtained due to the use of anonymised data.

Consent to publish

Not applicable.

Authors' contributions

All authors planned, designed, and implemented the study, including quality assurance and control. Seongho Ryu analysed the data and developed the analytical strategy. Yoosoo Chang and Seongho Ryu supervised the field activities. Yoosun Cho and Yoosoo Chang drafted the manuscript, with additional writing input from Christopher D. Byrne and Sarah H. Wild. All authors interpreted the results and contributed to the critical revisions of the manuscript. All authors approved the final version of the manuscript.

Data Availability Statement

The data are not publicly available outside the hospital because of institutional review board restrictions (the data were not collected in a manner that could be widely distributed). However, the analytical methods are available from the corresponding author upon request.

REFERENCES

1. American Diabetes Association Professional Practice C. 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes-2022. Diabetes Care. 2022;45(Suppl 1):S17-S38.
2. Wang H, Li N, Chivese T, et al. IDF Diabetes Atlas: Estimation of Global and Regional Gestational Diabetes Mellitus Prevalence for 2021 by International Association of Diabetes in Pregnancy Study Group's Criteria. Diabetes Res Clin Pract. 2022;183:109050.
3. Vounzoulaki E, Khunti K, Abner SC, Tan BK, Davies MJ, Gillies CL. Progression to type 2 diabetes in women with a known history of gestational diabetes: systematic review and meta-analysis. BMJ. 2020;369:m1361.
4. Kramer CK, Campbell S, Retnakaran R. Gestational diabetes and the risk of cardiovascular disease in women: a systematic review and meta-analysis. Diabetologia. 2019;62(6):905-14.
5. Younossi ZM, Koenig AB, Abdelatif D, Fazel Y, Henry L, Wymer M. Global epidemiology of nonalcoholic fatty liver disease-Meta-analytic assessment of prevalence, incidence, and outcomes. Hepatology. 2016;64(1):73-84.
6. Targher G, Tilg H, Byrne CD. Non-alcoholic fatty liver disease: a multisystem disease requiring a multidisciplinary and holistic approach. Lancet Gastroenterol Hepatol. 2021;6(7):578-88.
7. Bozkurt L, Gobl CS, Tura A, et al. Fatty liver index predicts further metabolic deteriorations in women with previous gestational diabetes. PLoS One. 2012;7(2):e32710.
8. Forbes S, Godsland IF, Taylor-Robinson SD, et al. A history of previous gestational diabetes mellitus is associated with adverse changes in insulin secretion and VLDL metabolism independently of increased intrahepatocellular lipid. Diabetologia. 2013;56(9):2021-33.
9. Ajmera VH, Gunderson EP, VanWagner LB, Lewis CE, Carr JJ, Terrault NA. Gestational Diabetes Mellitus Is Strongly Associated With Non-Alcoholic Fatty Liver Disease. Am J Gastroenterol. 2016;111(5):658-64.
10. Lavrentaki A, Thomas T, Subramanian A, et al. Increased risk of non-alcoholic fatty liver disease in women with gestational diabetes mellitus: A population-based cohort study, systematic review and meta-analysis. J Diabetes Complications. 2019;33(10):107401.
11. Mantovani A, Petracca G, Beatrice G, Tilg H, Byrne CD, Targher G. Non-alcoholic fatty liver disease

and risk of incident diabetes mellitus: an updated meta-analysis of 501 022 adult individuals. Gut. 2021;70(5):962-9.

12. Kim Y, Chang Y, Ryu S, Wild SH, Byrne CD. NAFLD improves risk prediction of type 2 diabetes: with effect modification by sex and menopausal status. Hepatology. 2022;76(6):1755-65.

13. Chang Y, Ryu S, Sung KC, et al. Alcoholic and non-alcoholic fatty liver disease and associations with coronary artery calcification: evidence from the Kangbuk Samsung Health Study. Gut. 2019;68(9):1667-75.

14. Chalasani N, Younossi Z, Lavine JE, et al. The diagnosis and management of non-alcoholic fatty liver disease: practice guideline by the American Gastroenterological Association, American Association for the Study of Liver Diseases, and American College of Gastroenterology. Gastroenterology. 2012;142(7):1592-609.

15. Craig CL, Marshall AL, Sjostrom M, et al. International physical activity questionnaire: 12-country reliability and validity. Medicine and science in sports and exercise. 2003;35(8):1381-95.

16. World Health Organization, Regional Office for the Western Pacific. The Asia-Pacific perspective: redefining obesity and its treatment. Sydney: Health Communications Australia; 2000.

17. Committee ADAPP. 2. Classification and Diagnosis of Diabetes: Standards of Medical Care in Diabetes—2022. Diabetes Care. 2021;45(Supplement_1):S17-S38.

18. Weinert LS. International Association of Diabetes and Pregnancy Study Groups recommendations on the diagnosis and classification of hyperglycemia in pregnancy: comment to the International Association of Diabetes and Pregnancy Study Groups Consensus Panel. Diabetes Care. 2010;33(7):e97; author reply e8.

19. Beekers P, Jamaladin H, van Drongelen J, Roeleveld N, van Gelder M. Data From Web-Based Questionnaires Were Valid for Gestational Diabetes and Preeclampsia, but Not Gestational Hypertension. J Clin Epidemiol. 2020;125:84-90.

20. Donnelly SR, Hinkle SN, Rawal S, et al. Prospective study of gestational diabetes and fatty liver scores 9 to 16 years after pregnancy. J Diabetes. 2019;11(11):895-905.

21. Association KD. Treatment guideline for diabetes. Korean Diabetes Association. 2019;6 ed.

22. Cheol Seong S, Kim YY, Khang YH, et al. Data Resource Profile: The National Health Information

Database of the National Health Insurance Service in South Korea. Int J Epidemiol. 2017;46(3):799-800.

23. Hinkle SN, Rawal S, Zhu Y, Grewal J, Albert PS, Zhang C. Validation of Self-reported Diagnosis of Gestational Diabetes at 6-weeks Postpartum. Epidemiology. 2017;28(5):747-52.

24. Mathiesen UL, Franzen LE, Aselius H, et al. Increased liver echogenicity at ultrasound examination reflects degree of steatosis but not of fibrosis in asymptomatic patients with mild/moderate abnormalities of liver transaminases. Dig Liver Dis. 2002;34(7):516-22.

25. VanderWeele TJ, Ding P. Sensitivity Analysis in Observational Research: Introducing the E-Value. Ann Intern Med. 2017;167(4):268-74.

26. Erfanpoor S, Etemad K, Kazempour S, et al. Diabetes, Hypertension, and Incidence of Chronic Kidney Disease: Is There any Multiplicative or Additive Interaction? Int J Endocrinol Metab. 2021;19(1):e101061.

27. Uno H, Cai T, Pencina MJ, D'Agostino RB, Wei LJ. On the C-statistics for evaluating overall adequacy of risk prediction procedures with censored survival data. Stat Med. 2011;30(10):1105-17.

28. Gray LJ, Khunti K, Edwardson C, et al. Implementation of the automated Leicester Practice Risk Score in two diabetes prevention trials provides a high yield of people with abnormal glucose tolerance. Diabetologia. 2012;55(12):3238-44.

29. Woo YC, Lee CH, Fong CHY, Tso AWK, Cheung BMY, Lam KSL. Validation of the diabetes screening tools proposed by the American Diabetes Association in an aging Chinese population. PLoS One. 2017;12(9):e0184840.

30. Rembold CM. Number needed to screen: development of a statistic for disease screening. BMJ. 1998;317(7154):307-12.

31. Fraser G, Yan R. Guided multiple imputation of missing data: using a subsample to strengthen the missing-at-random assumption. Epidemiology. 2007;18(2):246-52.

32. Gray LJ, Taub NA, Khunti K, et al. The Leicester Risk Assessment score for detecting undiagnosed Type 2 diabetes and impaired glucose regulation for use in a multiethnic UK setting. Diabet Med. 2010;27(8):887-95.

33. Bang H, Edwards AM, Bomback AS, et al. Development and validation of a patient self-assessment

score for diabetes risk. Ann Intern Med. 2009;151(11):775-83.

34. Kubihal S, Gupta Y, Shalimar, et al. Prevalence of non-alcoholic fatty liver disease and factors associated with it in Indian women with a history of gestational diabetes mellitus. J Diabetes Investig. 2021;12(5):877-85.

35. Mehmood S, Margolis M, Ye C, et al. Hepatic fat and glucose tolerance in women with recent gestational diabetes. BMJ Open Diabetes Res Care. 2018;6(1):e000549.

36. Kendler KS, Gardner CO. Interpretation of interactions: guide for the perplexed. Br J Psychiatry. 2010;197(3):170-1.

37. Hagstrom H, Hoijer J, Ludvigsson JF, et al. Adverse outcomes of pregnancy in women with non-alcoholic fatty liver disease. Liver Int. 2016;36(2):268-74.

38. Lee SM, Kwak SH, Koo JN, et al. Non-alcoholic fatty liver disease in the first trimester and subsequent development of gestational diabetes mellitus. Diabetologia. 2019;62(2):238-48.

39. Lee SM, Hwangbo S, Norwitz ER, et al. Nonalcoholic fatty liver disease and early prediction of gestational diabetes mellitus using machine learning methods. Clin Mol Hepatol. 2022;28(1):105-16.

40. Lappas M, Mundra PA, Wong G, et al. The prediction of type 2 diabetes in women with previous gestational diabetes mellitus using lipidomics. Diabetologia. 2015;58(7):1436-42.

41. Kengne AP, Beulens JW, Peelen LM, et al. Non-invasive risk scores for prediction of type 2 diabetes (EPIC-InterAct): a validation of existing models. Lancet Diabetes Endocrinol. 2014;2(1):19-29.

42. Ha KH, Lee KA, Han KD, Moon MK, Kim DJ. Diabetes screening in South Korea: a new estimate of the number needed to screen to detect diabetes. Korean J Intern Med. 2023;38(1):93-100.

43. Tuomilehto J, Lindstrom J, Eriksson JG, et al. Prevention of type 2 diabetes mellitus by changes in lifestyle among subjects with impaired glucose tolerance. N Engl J Med. 2001;344(18):1343-50.

44. Herman WH, Ye W, Griffin SJ, et al. Early Detection and Treatment of Type 2 Diabetes Reduce Cardiovascular Morbidity and Mortality: A Simulation of the Results of the Anglo-Danish-Dutch Study of Intensive Treatment in People With Screen-Detected Diabetes in Primary Care (ADDITION-Europe). Diabetes Care. 2015;38(8):1449-55.

45. Diaz-Santana MV, O'Brien KM, Park YM, Sandler DP, Weinberg CR. Persistence of Risk for Type 2 Diabetes After Gestational Diabetes Mellitus. Diabetes Care. 2022;45(4):864-70.

46. Targher G, Corey KE, Byrne CD, Roden M. The complex link between NAFLD and type 2 diabetes mellitus - mechanisms and treatments. Nat Rev Gastroenterol Hepatol. 2021;18(9):599-612.

47. Xiang AH, Kjos SL, Takayanagi M, Trigo E, Buchanan TA. Detailed physiological characterization of the development of type 2 diabetes in Hispanic women with prior gestational diabetes mellitus. Diabetes. 2010;59(10):2625-30.

48. Xiang AH, Kawakubo M, Trigo E, Kjos SL, Buchanan TA. Declining beta-cell compensation for insulin resistance in Hispanic women with recent gestational diabetes mellitus: association with changes in weight, adiponectin, and C-reactive protein. Diabetes Care. 2010;33(2):396-401.

49. Tilg H, Moschen AR, Roden M. NAFLD and diabetes mellitus. Nat Rev Gastroenterol Hepatol. 2017;14(1):32-42.

50. McAuley KA, Mann JI, Chase JG, Lotz TF, Shaw GM. Point: HOMA--satisfactory for the time being: HOMA: the best bet for the simple determination of insulin sensitivity, until something better comes along. Diabetes Care. 2007;30(9):2411-3.

51. Bonora E, Tuomilehto J. The pros and cons of diagnosing diabetes with A1C. Diabetes Care. 2011;34 Suppl 2(Suppl 2):S184-90.

Figure legend

Figure 1. Flow chart of the study population