**Online Supplemental Material**

**Supplemental Methods**

Insulin value transformation

We transformed insulin values from Laboratory 1 into Laboratory 2-equivalent values using this equation: 0.8886 × [Laboratory 1 insulin result] − 2.2909.

Longitudinal modeling

The likelihood ratio test was used to compare the fit of a linear mixed effects model with random intercepts to account for participant-specific differences to a model that additionally included random slopes for participant-specific visit effects. If the null hypothesis that the nested model provided an adequate fit was rejected, random slopes were added. Throughout, random intercepts models were used for Stumvoll models and random intercepts and slopes models were used for Matsuda and PIP models.

Missing data and multiple imputation

Some participants missed visits because they left the study, while others were excluded from completing the mid-late pregnancy visit per protocol. Some participants who missed the mid-late pregnancy visit returned for the postpartum visit. At attended visits, some participants had incomplete glucose or insulin measures due to technical issues.

Missingness in insulin and glucose values at all attended visits is summarized in **Supplemental Figure 1**. Breastfeeding status at the postpartum visit was missing for 1 participant (no GDM).

We generated M=25 imputed datasets and included the following predictors in the imputation model: BMI at each visit, age at the early pregnancy visit, family history of diabetes, Hispanic ethnicity, marital status, smoking status, and completion of college. Missing values were imputed in the “wide” format, with one row per participant and separate columns for each of the 12 glucose and 12 insulin values (4 values for 3 visits). Glucose and insulin values were also log-transformed prior to imputation to address potential skewness and ensure imputed values were positive. Log-transformed imputed values were exponentiated after imputation prior to calculating Stumvoll and Matsuda indices. Stumvoll and Matsuda index calculations from imputed data that resulted in negative values were replaced with the smallest observed values for each in longitudinal and subsequent analyses. Only multiply imputed data for attended visits with at least 1 observed insulin measurement were used.

Rubin’s rules were used to combine results across imputed datasets (1). We assumed throughout that data was missing at random. Multiple imputation and linear mixed effects models that use all available data are both valid under the missing at random assumption when variables associated with missingness are included in the imputation model or are adjusted for in the model as fixed effects (2).

Rubin’s rules could not be readily applied to estimate uncertainty for the AUC in the presence of missing data. We used the “Boot MI” method for bootstrapping multiply imputed data (3). We drew 500 bootstrapped samples from the study population, generated 25 imputed datasets per bootstrapped dataset, calculated the AUC for each imputed dataset, and averaged over the estimates obtained from each of the 25 imputations to generate a distribution of 500 AUC estimates. The overall point estimate was derived by taking the median of that distribution, and the 2.5th and 97.5th percentiles were used to construct the 95% confidence intervals.

Susceptibility of the disposition index to model misspecification

We demonstrate that a linear regression model assessing the effect of an exposure on a disposition index-like measure (DI, defined as the product of insulin secretion and resistance) may be misspecified if the relationship between insulin secretion and resistance is not strictly rectangular hyperbolic. Instead, a linear regression model treating log-transformed insulin secretion as the outcome, adjusting for log-transformed insulin resistance, is correctly specified under a wider range of hyperbolic relationships.

We define R as insulin resistance (Matsuda index) and S as insulin secretory response (Stumvoll index). If there is a rectangular hyperbolic relationship between R and S, then RS is constant for individuals with the same beta-cell function. However, a rectangular hyperbolic relationship between two variables is not the only way that two variables can be hyperbolically related. For example, consider the wide range of relationships where SRδ is constant for individuals with the same beta-cell function and δ > 0 (Supplementary Figure 2).

Suppose it is the case that some binary exposure variable of interest X has the following relationship with R and S:

SRδ = α0 + αxX + ε

where δ, α0, and αx are constants and ε is an error term serving as the source of random noise. We will refer to this as the generative model (1).

After collecting some data, consider fitting the following linear regression model (2):

SR = β0 + βxX + ε

The implicit assumption of this model is that the mean difference in SR between individuals with X = 1 and X = 0 is equal to a constant, βx. If this assumption is true, then the model is correctly specified. To evaluate this assumption, we rearrange the terms in the generative model (1):

SR = R1-δ(α0 + αxX + ε)

Using this equation, we denote SR1 and SR0 as the product of S and R for individuals with X = 1 and X = 0, respectively:

SR1 = R1-δ(α0 + αx + ε)

SR0 = R1-δ(α0 + ε)

The mean difference in SR comparing two subjects with X = 1 and X = 0, E[SR1 - SR0], where the E[·] operator denotes the expected value, is

E[R1-δ(α0 + αx + ε) - R1-δ(α0 + ε)]

=E[R1-δαx]

= R1-δαx

If δ ≠ 1, this quantity depends on the value of R. In other words, unless δ = 1, the mean difference in SR between subjects with X = 1 and X = 0 is not constant. Therefore, unless the relationship between S and R is rectangular hyperbolic, model (2) is misspecified.

Rather than modeling SM as the outcome as in model (2), consider fitting an alternative linear regression model (3):

log S = **γ**0 + **γ**r log R + **γ**x X + ε

The implicit assumption of this model is that the mean difference in log S between individuals with X = 1 and X = 0 but have the same value of log R is a constant, **γ**x. Consider two individuals, both with R fixed at the same value R\*, but one individual has X = 1 and the other has X = 0. Rearranging terms in the generative model (1),

S=R\*-δ(α0 + αx + ε)

log S=log [R\*-δ(α0 + αx + ε)]

Then log S1 (corresponding to X = 1) and log S0 (corresponding to X = 0) are as follows:

log S1 = log[R\*-δ(α0 + αx + ε)]

log S0 = log[R\*-δ(α0 + ε)]

The mean difference in log Stumvoll between these two individuals, E[log S1 – log S0], is

E{log[R\*-δ(α0 + αx + ε)] - log[R\*-δ(α0 + ε)]} = E{log[{R\*- δ(α0 + αx + ε)}/{R\*- δ(α0 + ε)}]}

= E{log[(α0 + αx + ε)/(α0 + ε)]}

= E{log[1 + αx/(α0 + ε)]}

≈ log[1 + αx/α0]

where we have assumed that the contribution of the random error ε to the expectation is negligible. Because the mean difference is a function of constants, the regression model (3) is correctly specified: **γ**x = log[1 + αx/α0].

Not only is this true for any value of δ,the specific value of δ in the generative model (1) does not even factor into the expression for **γ**x . This means a statistically significant finding in model (3) for the association of X and log Stumvoll implies a statistically significant finding for the association of X and SRδ.

In summary, if we assume model (1) holds, then model (2) is correctly specified only when δ ≠ 1. However, model (3) is correctly specified under all values of δ, and does not require us to specify a value for δ. Therefore, in this paper, we use model (3): we treat log S as the outcome and adjust for log R, which is more robust to model misspecification than the model with SR as the outcome.

We also propose a quantity that utilizes our best estimate of δ, the Pregnancy Insulin Physiology (PIP) index. We derive our best estimate of δ by considering the following relationship between S and R for individuals with the same beta-cell function:

SRδ = c

where c is a constant. Rearranging terms,

log(SRδ) = log c

log S + δlog R = log c

log S = -δlog R + log c

Using this relationship, δ can be estimated by fitting a linear regression with log S as the outcome and log R as a predictor. The PIP index is defined as SRΔ, where Δ is the negative of the estimated slope of that regression line. We then model how the PIP index changes with each visit, and evaluate the ability of PIP to predict GDM.

**Supplemental References**

1. Little RJ, Rubin DB. Statistical analysis with missing data: John Wiley & Sons; 2019.

2. Fitzmaurice GM, Laird NM, Ware JH. Applied longitudinal analysis: John Wiley & Sons; 2012.

3. Schomaker M, Heumann C. Bootstrap inference when using multiple imputation. Stat Med. 2018;37(14):2252-66.

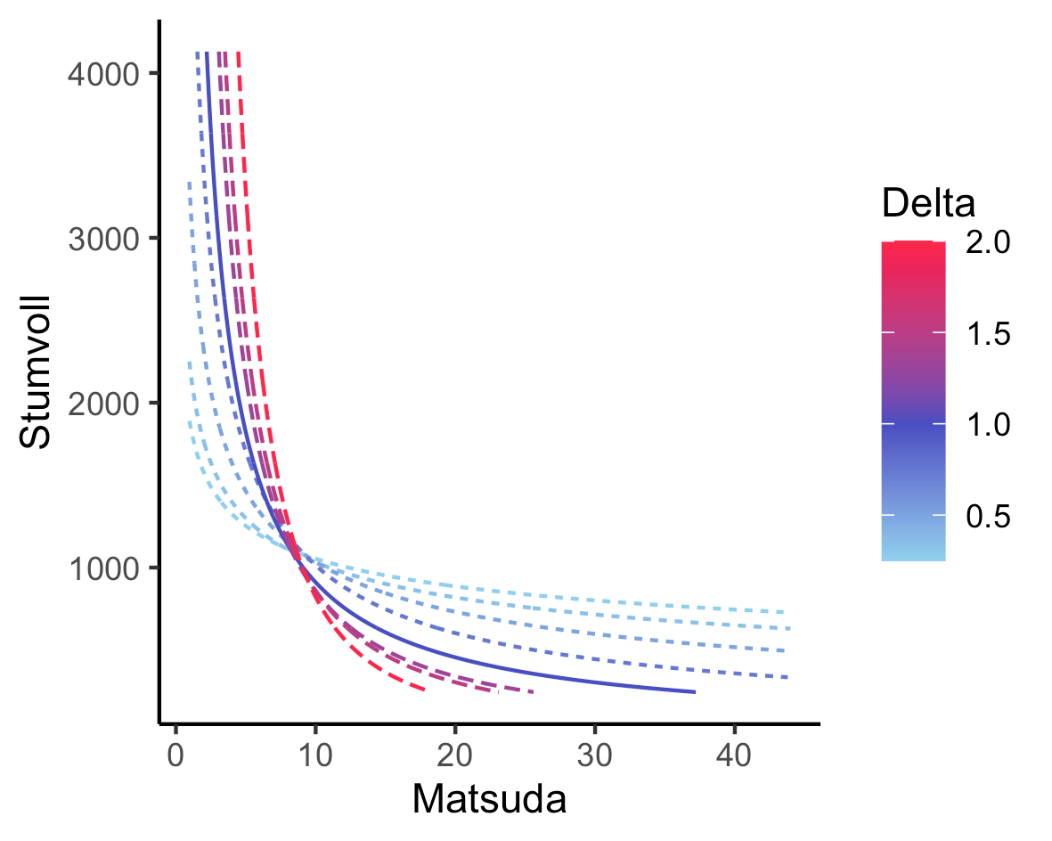
**Supplemental Figures**

**Supplemental Figure 1.** Visit attendance and data completeness

**Diagram

Description automatically generated**

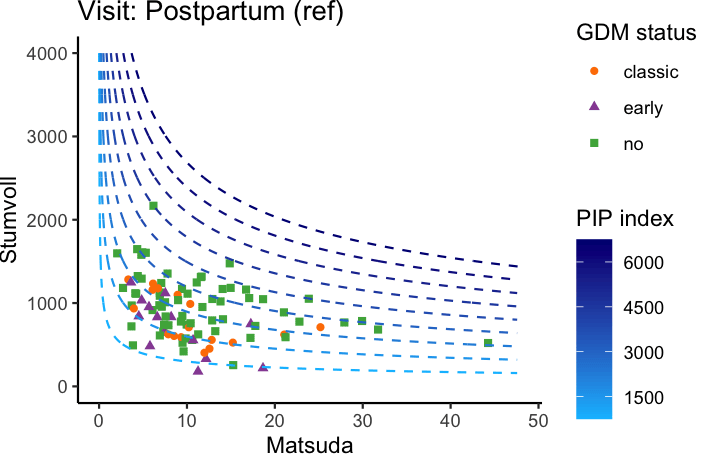
**Supplemental Figure 2.** Potential hyperbolic relationships between insulin sensitivity/resistance and insulin secretory response



These curves depict various hyperbolic relationships between insulin sensitivity/resistance (R: Matsuda) and insulin secretory response (S: Stumvoll). Each curve corresponds to combinations of Matsuda and Stumvoll where SRδ is constant. The solid dark blue curve represents the relationship where δ = 1 (rectangular hyperbolic). The longer dashed lines (dark blue to red) correspond to hyperbolic relationships where δ > 1. For these curves, changes in Matsuda are associated with larger changes in Stumvoll than when δ = 1. The shorter dashed lines (dark blue to light blue) correspond to 0 < δ < 1. When 0 < δ < 1, changes in Matsuda are associated with smaller changes in Stumvoll than when δ = 1.

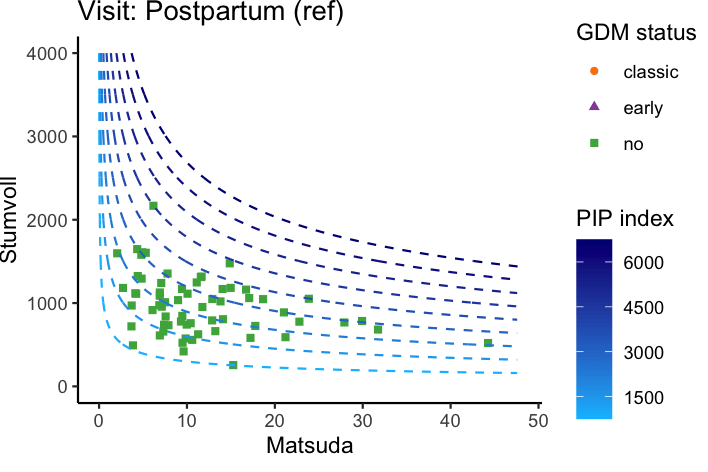
**Supplemental Video 1**

**Supplemental Video 1a. Animation demonstrating changes in Matsuda, Stumvoll, and PIP index values between visits, stratified by GDM status**

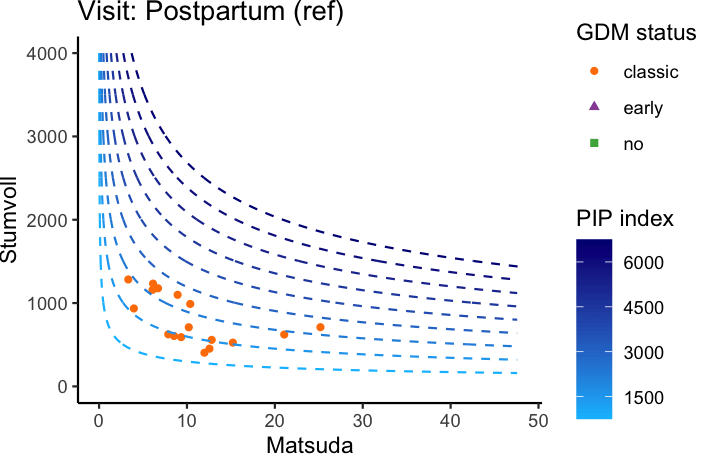


Only participants who attended both the early pregnancy and postpartum visits are included in this figure. The “No” group refers to participants who attended the mid-late pregnancy visit and were not diagnosed with GDM at either pregnancy visit. Participants who were diagnosed with early GDM but did not attend the mid-late pregnancy visit do not appear on the Mid-Late Pregnancy visit section of the animation, but do appear on the Early pregnancy and postpartum visit sections.

**Supplemental Video 1b. Animation demonstrating changes in Matsuda, Stumvoll, and PIP index values between visits, restricted to participants not diagnosed with GDM at the early or mid-late pregnancy visits**

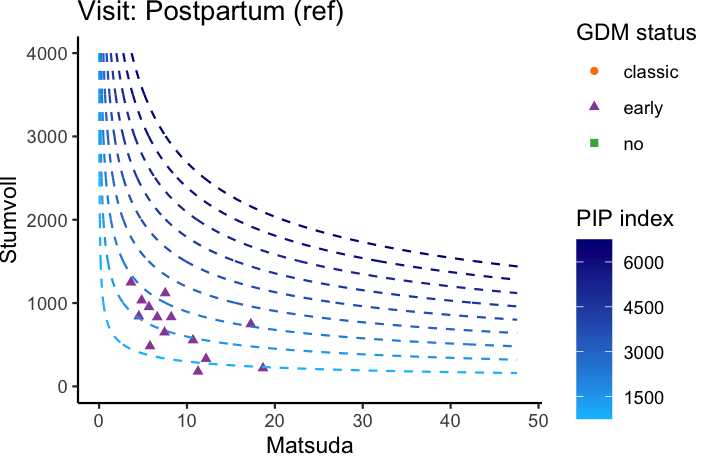


Only participants who attended all 3 visits are included in this figure.

**Supplemental Video 1c. Animation demonstrating changes in Matsuda, Stumvoll, and PIP index values between visits, restricted to participants diagnosed with GDM at the mid-late pregnancy visit (classic GDM)** 

Only participants who attended all 3 visits are included in this figure.

**Supplemental Video 1d. Animation demonstrating changes in Matsuda, Stumvoll, and PIP index values between the early pregnancy and postpartum visits, restricted to participants diagnosed with early GDM**



Only participants who attended both the early pregnancy and postpartum visits are included in this figure.

**Supplemental Tables**

**Supplemental Table 1**. Covariate-adjusted differences obtained from longitudinal models, comparing visits and GDM subgroups

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Outcome** | **Visit** | **All** | | | **No GDM** | | | **Early GDM** | | | | | **Classic GDM** | | | | |
| **Est. mean** | **Compared to postpartum** | | **Est. mean** | **Compared to postpartum** | | **Est. mean** | **Compared to postpartum** | | **Compared to No GDM** | | **Est. mean** | **Compared to postpartum** | | **Compared to No GDM** | |
| **Diff** | **p** | **Diff** | **p** | **Diff** | **p** | **Diff** | **p** | **Diff** | **p** | **Diff** | **p** |
| Log Stumvoll, not adjusted for Matsuda | Early | 6.86 (6.77, 6.95) | 0.21 | <0.001 | 6.92 (6.83, 7.01) | 0.20 | <0.001 | 6.30 (5.63, 6.97) | 0.43 | 0.002 | -0.25 | 0.018 | 6.69 (6.42, 6.97) | 0.14 | 0.183 | -0.23 | 0.022 |
| Mid-late | 6.91 (6.82, 7.01) | 0.27 | <0.001 | 6.96 (6.87, 7.05) | 0.24 | <0.001 | - | - | - | - | - | 6.80 (6.53, 7.06) | 0.24 | 0.030 | -0.14 | 0.050\* |
| Postpartum | 6.65 (6.55, 6.74) | - | - | 6.73 (6.63, 6.82) | - | - | 5.87 (5.19, 6.55) | - | - | -0.36 | 0.019 | 6.55 (6.26, 6.85) | - | - | -0.12 | 0.376 |
| Log Matsuda | Early | 2.16 (2.03, 2.29) | -0.20 | <0.001 | 2.29 (2.14, 2.44) | -0.14 | 0.030 | 1.85 (1.85, 2.41) | -0.59 | <0.001 | -0.71 | <0.001 | 1.88 (1.49, 2.28) | -0.23 | 0.058 | -0.31 | 0.035 |
| Mid-late | 1.88 (1.76, 2.00) | -0.47 | <0.001 | 2.01 (1.88, 2.14) | -0.42 | <0.001 | - | - | - | - | - | 1.49 (1.11, 1.87) | -0.63 | <0.001 | -0.42 | <0.001 |
| Postpartum | 2.36 (2.24, 2.48) | - | - | 2.43 (2.29, 2.57) | - | - | 2.45 (2.45, 2.98) | - | - | -0.23 | 0.121 | 2.12 (1.74, 2.50) | - | - | -0.26 | 0.053 |
| Log Stumvoll, adjusted for log Matsuda | Early | 6.84 (6.76, 6.93) | 0.16 | <0.001 | 6.94 (6.86, 7.02) | 0.16 | 0.002 | 6.20 (5.59, 6.81) | 0.25 | 0.140 | -0.52 | <0.001 | 6.54 (6.32, 6.76) | 0.04 | 0.698 | -0.35 | <0.001 |
| Mid-late | 6.84 (6.74, 6.93) | 0.16 | 0.001 | 6.91 (6.83, 6.99) | 0.13 | 0.012 | - | - | - | - | - | 6.48 (6.23, 6.72) | -0.02 | 0.852 | -0.29 | <0.001 |
| Postpartum | 6.68 (6.59, 6.77) | - | - | 6.78 (6.68, 6.87) | - | - | 5.95 (5.32, 6.57) | - | - | -0.42 | 0.006 | 6.50 (6.27, 6.72) | - | - | -0.19 | 0.158 |
| Pregnancy Insulin Physiology (PIP) Index | Early | 2439 (2255, 2624) | 215 | 0.040 | 2652 (2445, 2859) | 284 | 0.035 | 1378 (704, 2051) | 131 | 0.551 | -1029 | <0.001 | 1810 (1488, 2132) | -11 | 0.965 | -791 | <0.001 |
| Mid-late | 2280 (2110, 2449) | 55 | 0.616 | 2428 (2243, 2613) | 60 | 0.650 | - | - | - | - | - | 1663 (1369, 1958) | -158 | 0.550 | -681 | <0.001 |
| Postpartum | 2224 (1993, 2455) | - | - | 2368 (2112, 2625) | - | - | 1247 (483, 2011) | - | - | -716 | 0.009 | 1821 (1288, 2355) | - | - | -464 | 0.127 |

Estimated mean values for each outcome are obtained from covariate-adjusted linear mixed effects models, fit among all participants together and then in separate models stratified by GDM subgroup. The estimated mean value at each visit corresponds to a participant who is the mean age of the participant population, has the mean BMI of all participants at the early pregnancy visit, has no family history of diabetes, is not Hispanic, is married, and has completed college. For log Stumvoll models adjusted for insulin sensitivity, the reference individual also has the mean observed value of log Matsuda at the visit being modeled. Under “Compared to postpartum”, the “diff” column refers to the estimated value of the outcome at a given visit minus the estimated value at postpartum, and the p-value corresponds to the statistical test of the null hypothesis that that difference is 0. The results under “Compared to No GDM” corresponds to a different set of linear regression models, where the observations included in each model are from participants in all 3 subgroups at a single visit. The “diff” column refers to the difference between the value for a given subgroup minus the value for the no GDM group at the given visit, and the p-value corresponds to the null hypothesis that the difference is 0. These differences do not align exactly with the estimated mean values from the longitudinal models as they are obtained from a separate set of regression models that leverage data from all 3 subgroups simultaneously, but only at one visit at a time. The asterisk corresponds to a p-value of 0.04996.

**Supplemental Table 2.** Percent differences in pregnancy insulin physiology (PIP) indices between visits, overall and by GDM subgroup

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **PIP index comparisons** | **All** | **No GDM** | **Early GDM** | **Classic GDM** |
| N | 98 | 65 | 14 | 17 |
| Early pregnancy value >10% lower than postpartum value | 29 (30%) | 16 (25%) | 3 (21%) | 9 (53%) |
| Early pregnancy value <10% different from postpartum value | 24 (24%) | 19 (29%) | 4 (29%) | 1 (6%) |
| Early pregnancy value >10% higher than postpartum value | 45 (46%) | 30 (46%) | 7 (50%) | 7 (41%) |
| p-value comparing distribution to “No GDM” group | - | - | 0.958 | 0.036 |
|  |  |  |  |  |
| N | 87 | 65 |  | 18 |
| Mid-late pregnancy value >10% lower than postpartum value | 33 (38%) | 24 (37%) |  | 7 (39%) |
| Mid-late pregnancy value <10% different from postpartum value | 15 (17%) | 13 (20%) |  | 2 (11%) |
| Mid-late pregnancy value >10% higher than postpartum value | 39 (45%) | 28 (43%) |  | 9 (50%) |
| p-value comparing distribution to “No GDM” group | - | - | - | 0.676 |

Percent differences are calculated by taking the PIP index at a given visit, subtracting the value at the postpartum visit, and dividing by the value at the postpartum visit and multiplying by 100%. The distribution of participants across the 3 categories (>10% less, <10% different, >10% greater) are compared between subgroups using a chi-square test, with the no GDM group as the reference.

**Supplemental Table 3.** Evaluation of predictive models for early or classic GDM using the pregnancy insulin physiology (PIP) index

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| **Predictors** | **Adjustment for clinical characteristics** | **GDM types included in outcome** | **AUC (95% CI)** | **Early pregnancy PIP** | | **Postpartum PIP** | |
| **OR** | **p-value** | **OR** | **p-value** |
| None | Yes | Early + Classic | 0.704 (0.606, 0.794) | - | - | - | - |
| Classic | 0.757 (0.633, 0.867) | - | - | - | - |
| Early pregnancy PIP | No | Early + Classic | 0.827 (0.762, 0.892) | 0.816 | <0.001 | - | - |
| Classic | 0.812 (0.726, 0.893) | 0.832 | <0.001 | - | - |
| Early pregnancy PIP | Yes | Early + Classic | 0.867 (0.801, 0.928) | 0.805 | <0.001 | - | - |
| Classic | 0.874 (0.776, 0.950) | 0.816 | 0.002 | - | - |
| Postpartum PIP | No | Early + Classic | 0.717 (0.604, 0.808) | - | - | 0.913 | 0.037 |
| Classic | 0.685 (0.512, 0.798) | - | - | 0.923 | 0.081 |
| Postpartum PIP | Yes | Early + Classic | 0.789 (0.686, 0.877) | - | - | 0.906 | 0.041 |
| Classic | 0.832 (0.692, 0.938) | - | - | 0.919 | 0.099 |
| Early pregnancy and postpartum PIP | No | Early + Classic | 0.798 (0.710, 0.875) | 0.849 | 0.001 | 0.961 | 0.369 |
| Classic | 0.794 (0.690, 0.878) | 0.857 | 0.005 | 0.962 | 0.402 |
| Early pregnancy and postpartum PIP | Yes | Early + Classic | 0.855 (0.775, 0.923) | 0.840 | 0.001 | 0.953 | 0.323 |
| Classic | 0.886 (0.778, 0.960) | 0.837 | 0.009 | 0.957 | 0.400 |
| Early pregnancy fasting glucose | No | Early + Classic | 0.788 (0.702, 0.859) | - | - | - | - |
| Classic | 0.641 (0.433, 0.745) | - | - | - | - |
| Early pregnancy fasting glucose | Yes | Early + Classic | 0.843 (0.763, 0.908) | - | - | - | - |
| Classic | 0.782 (0.675, 0.897) | - | - | - | - |

Participants included in the prediction model are those who were diagnosed with early or classic GDM or who attended both the early and mid-late pregnancy visits and were not diagnosed with GDM (n=144 overall). Clinical characteristics include age at the early pregnancy visit, family history of diabetes, Hispanic ethnicity, Marital status (married or not), College educated, and BMI at the early pregnancy visit. AUC was also calculated after excluding the early GDM subgroup (n=21). The odds ratio (OR) corresponds to the association of a 100-unit increase in each PIP index with the odds of being diagnosed with GDM at either the early or the mid-late pregnancy visit.

**Supplemental Table 4.** Participant characteristics, overall and stratified by GDM subgroup, restricted to participants who attended all 3 visits

|  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- |
|  | Participants who attended all 3 visits | No GDM | Early GDM | Classic GDM | p-value |
| Participants | 98 | 73 | 4 | 21 |  |
| Age at early pregnancy visit (years) | 33.10 (4.31) | 32.92 (4.24) | 34.73 (5.47) | 33.40 (4.46) | 0.774 |
| Gestational age at early pregnancy visit (weeks) | 12.7 (1.53) | 12.7 (1.53) | 11.6 (0.56) | 12.6 (1.63) | 0.018 |
| Gestational age at mid-late pregnancy visit (weeks) | 26.3 (1.62) | 26.1 (1.44) | 25.6 (1.61) | 27.2 (1.97) | 0.123 |
| Weeks postpartum at postpartum visit | 10.9 (5.17) | 10.6 (4.67) | 17.2 (8.65) | 10.7 (5.60) | 0.402 |
| BMI, early pregnancy visit (kg/m2) | 29.0 (5.94) | 29.5 (5.16) | 35.3 (12.80) | 25.9 (5.66) | 0.085 |
| BMI, mid-late pregnancy visit (kg/m2) | 31.1 (5.67) | 31.6 (4.92) | 36.6 (12.28) | 28.4 (5.59) | 0.115 |
| BMI, postpartum visit (kg/m2) | 30.1 (5.87) | 30.8 (5.18) | 34.1 (12.23) | 26.7 (5.55) | 0.051 |
| Nulliparous | 43 (43.9%) | 34 (46.6%) | 2 (50.0%) | 7 (33.3%) | 0.542 |
| Family history of diabetes | 36 (36.7%) | 26 (35.6%) | 1 (25.0%) | 9 (42.9%) | 0.735 |
| Personal history of gestational diabetes among parous participants | 9 (16.4%) | 3 (7.7%) | 0 (0.0%) | 6 (42.9%) | 0.008 |
| Family history of gestational diabetes | 18 (18.4%) | 15 (20.5%) | 0 (0.0%) | 3 (14.3%) | 0.505 |
| Race/Ethnicity |  |  |  |  | 0.261 |
| *Hispanic/Latina* | 17 (17.3%) | 12 (16.4%) | 1 (25.0%) | 4 (19.0%) |
| *Non-Hispanic/Latina* |  |  |  |  |
| *White* | 60 (61.2%) | 46 (63.0%) | 2 (50.0%) | 12 (57.1%) |
| *Black* | 10 (10.2%) | 10 (13.7%) | 0 (0.0%) | 0 (0.0%) |
| *Asian* | 8 (8.2%) | 3 (4.1%) | 1 (25.0%) | 4 (19.0%) |
| *None of the Above* | 3 (3.1%) | 2 (2.7%) | 0 (0.0%) | 1 (4.8%) |
| Employed full-time | 70 (71.4%) | 51 (69.9%) | 3 (75.0%) | 16 (76.2%) | 0.841 |
| Married | 77 (78.6%) | 58 (79.5%) | 3 (75.0%) | 16 (76.2%) | 0.935 |
| Completed college | 89 (90.8%) | 66 (90.4%) | 3 (75.0%) | 20 (95.2%) | 0.426 |
| Breastfeeding, postpartum visit |  |  |  |  | 0.898 |
| *Exclusively breastfeeding* | 48 (49.5%) | 36 (50.0%) | 2 (50.0%) | 10 (47.6%) |
| *Some breastfeeding and some formula* | 33 (34.0%) | 24 (33.3%) | 2 (50.0%) | 7 (33.3%) |
| *Exclusively formula* | 16 (16.5%) | 12 (16.7%) | 0 (0.0%) | 4 (19.0%) |

**Supplemental Table 5**. Observed means and standard deviations of glucose, insulin, Stumvoll, and Matsuda for each visit and measurement time point (if applicable), overall and by GDM subgroup, restricted to participants who attended all 3 visits

|  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
|  | **Participants who**  **attended all 3 visits** | | **No GDM** | | **Early GDM** | | | **Classic GDM** | | |
|  | **Mean (SD)** | **p (PP)**a | **Mean (SD)** | **p (PP)** | **Mean (SD)** | **p (PP)** | **p (NG)**b | **Mean (SD)** | **p (PP)** | **p (NG)** |
| **Glucose (mg/dl)** |  |  |  |  |  |  |  |  |  |  |
| **Early pregnancy** |  |  |  |  |  |  |  |  |  |  |
| *Fasting* | 80.86 (6.23) | 0.001 | 79.82 (5.81) | 0.002 | 94.50 (1.73) | 0.492 | <0.001 | 81.86 (4.89) | 0.163 | 0.116 |
| *30-min* | 126.40 (22.42) | 0.067 | 123.46 (23.39) | 0.127 | 130.75 (27.28) | 0.670 | 0.634 | 135.10 (15.97) | 0.402 | 0.013 |
| *60-min* | 121.40 (30.31) | 0.700 | 115.23 (30.78) | 0.513 | 118.50 (22.52) | 0.798 | 0.797 | 143.38 (17.89) | 0.712 | <0.001 |
| *120-min* | 102.56 (25.56) | 0.603 | 97.88 (25.81) | 0.711 | 116.50 (18.77) | 0.745 | 0.139 | 116.19 (20.09) | 0.795 | 0.001 |
| **Mid-late pregnancy** |  |  |  |  |  |  |  |  |  |  |
| *Fasting* | 80.21 (6.73) | <0.001 | 78.70 (5.23) | <0.001 | - | - | - | 83.29 (8.23) | 0.656 | 0.023 |
| *30-min* | 133.93 (23.57) | 0.917 | 129.14 (22.11) | 0.738 | - | - | - | 149.71 (19.22) | 0.223 | <0.001 |
| *60-min* | 140.16 (32.87) | <0.001 | 128.96 (24.45) | 0.007 | - | - | - | 181.19 (25.30) | <0.001 | <0.001 |
| *120-min* | 116.02 (30.66) | <0.001 | 106.16 (21.64) | 0.003 | - | - | - | 148.76 (33.08) | <0.001 | <0.001 |
| **Postpartum** |  |  |  |  |  |  |  |  |  |  |
| *Fasting* | 83.33 (7.38) | - | 82.62 (6.96) | - | 91.00 (8.52) | - | 0.142 | 84.33 (8.01) | - | 0.381 |
| *30-min* | 131.28 (25.25) | - | 128.64 (26.16) | - | 139.75 (23.49) | - | 0.420 | 139.67 (20.40) | - | 0.063 |
| *60-min* | 122.84 (33.96) | - | 117.82 (31.79) | - | 122.00 (21.40) | - | 0.732 | 141.30 (38.32) | - | 0.018 |
| *120-min* | 101.23 (28.18) | - | 96.54 (23.75) | - | 111.50 (26.41) | - | 0.343 | 116.05 (37.60) | - | 0.038 |
| **Insulin (uIU/ml)** |  |  |  |  |  |  |  |  |  |  |
| **Early pregnancy** |  |  |  |  |  |  |  |  |  |  |
| *Fasting* | 6.53 (6.24) | 0.326 | 5.92 (3.80) | 0.905 | 11.68 (1.85) | 0.404 | 0.003 | 7.66 (11.33) | 0.299 | 0.495 |
| *30-min* | 55.93 (32.17) | 0.023 | 58.72 (33.55) | 0.038 | 60.15 (14.99) | 0.278 | 0.874 | 46.20 (28.84) | 0.631 | 0.104 |
| *60-min* | 57.49 (32.66) | <0.001 | 56.55 (32.32) | 0.001 | 47.41 (18.20) | 0.055 | 0.405 | 62.63 (36.23) | 0.023 | 0.494 |
| *120-min* | 44.16 (33.68) | <0.001 | 42.61 (32.22) | <0.001 | 39.55 (26.69) | 0.651 | 0.838 | 50.74 (40.31) | 0.087 | 0.413 |
| **Mid-late pregnancy** |  |  |  |  |  |  |  |  |  |  |
| *Fasting* | 8.59 (5.16) | <0.001 | 8.27 (5.01) | <0.001 | - | - | - | 8.78 (5.15) | 0.001 | 0.690 |
| *30-min* | 64.98 (39.39) | <0.001 | 66.26 (42.16) | 0.002 | - | - | - | 64.55 (32.65) | 0.004 | 0.847 |
| *60-min* | 76.02 (54.92) | <0.001 | 70.30 (45.22) | <0.001 | - | - | - | 100.91 (78.68) | <0.001 | 0.101 |
| *120-min* | 60.28 (37.74) | <0.001 | 55.29 (36.43) | <0.001 | - | - | - | 79.53 (38.44) | <0.001 | 0.015 |
| **Postpartum** |  |  |  |  |  |  |  |  |  |  |
| *Fasting* | 5.92 (3.53) | - | 5.97 (3.56) | - | 9.50 (5.24) | - | 0.271 | 5.04 (2.74) | - | 0.211 |
| *30-min* | 47.91 (29.78) | - | 49.59 (30.41) | - | 45.30 (28.67) | - | 0.789 | 41.98 (28.27) | - | 0.325 |
| *60-min* | 45.65 (33.91) | - | 46.57 (35.64) | - | 35.67 (14.65) | - | 0.250 | 44.29 (30.58) | - | 0.778 |
| *120-min* | 28.36 (21.56) | - | 26.56 (21.42) | - | 31.75 (22.51) | - | 0.681 | 34.13 (21.93) | - | 0.180 |
| **Stumvoll**c |  |  |  |  |  |  |  |  |  |  |
| *Early pregnancy* | 1072 (412) | 0.002 | 1111 (398) | 0.012 | 1201 (210) | 0.126 | 0.478 | 924 (460) | 0.233 | 0.103 |
| *Mid-late pregnancy* | 1162 (464) | <0.001 | 1206 (496) | <0.001 | - | - | - | 1038 (360) | 0.002 | 0.094 |
| *Postpartum* | 931 (351) | - | 972 (355) | - | 919 (273) | - | 0.731 | 774 (316) | - | 0.028 |
| **Matsuda**d |  |  |  |  |  |  |  |  |  |  |
| *Early pregnancy* | 10.54 (7.98) | 0.379 | 11.38 (8.63) | 0.778 | 4.95 (1.26) | 0.270 | <0.001 | 8.64 (5.19) | 0.154 | 0.082 |
| *Mid-late pregnancy* | 6.67 (4.46) | <0.001 | 7.17 (4.59) | <0.001 | - | - | - | 5.23 (3.96) | <0.001 | 0.065 |
| *Postpartum* | 11.29 (7.13) | - | 11.62 (7.56) | - | 8.30 (6.19) | - | 0.367 | 10.68 (5.63) | - | 0.546 |

Results are presented as mean (standard deviation). Only observed values (i.e., pre-imputation) are included in this table. Values obtained during visits in which participants were excluded per protocol are omitted. p (PP) refers to the p-value obtained from a paired t-test comparing the value of a given measure (e.g., fasting glucose) at a given visit compared to the postpartum visit within the same group of participants. p (NG) refers to the p-value obtained from a two-sample t-test comparing the value of a given measure at a particular visit for a given subgroup compared to the “no GDM” subgroup. There is no p-value for the rows and columns corresponding to the mid-late pregnancy visit for early GDM, as very few of these individuals attended those visits, per protocol.

**Supplemental Table 6.** Covariate-adjusted differences obtained from longitudinal models comparing visits, fit among participants who attended all 3 visits

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| **Outcome** | **Visit** | **Est. mean** | **Compared to postpartum** | |
| **Diff** | **p** |
| log Stumvoll, not adjusted for Matsuda | Early | 6.85 (6.75, 6.95) | 0.18 | <0.001 |
| Mid-late | 6.91 (6.81, 7.00) | 0.23 | <0.001 |
| Postpartum | 6.67 (6.57, 6.77) | - | - |
| log Matsuda | Early | 2.19 (2.04, 2.33) | -0.19 | 0.001 |
| Mid-late | 1.91 (1.78, 2.05) | -0.46 | <0.001 |
| Postpartum | 2.37 (2.24, 2.51) | - | - |
| log Stumvoll, adjusted for Matsuda | Early | 6.84 (6.75, 6.93) | 0.13 | 0.009 |
| Mid-late | 6.82 (6.73, 6.91) | 0.11 | 0.036 |
| Postpartum | 6.71 (6.62, 6.80) | - | - |
| PIP | Early | 2388 (2197, 2580) | 168 | 0.136 |
| Mid-late | 2226 (2056, 2396) | 5 | 0.962 |
| Postpartum | 2221 (1988, 2453) | - | - |

Estimated mean values for each outcome are obtained from covariate-adjusted linear mixed effects models, fit among participants who attended all 3 visits. The estimated mean value at each visit corresponds to a participant who is the mean age of the participant population, has the mean BMI of all participants at the early pregnancy visit, has no family history of diabetes, is not Hispanic, is married, and has completed college. For log Stumvoll models adjusted for insulin sensitivity, the reference individual also has the mean observed value of log Matsuda at the visit being modeled. Under “Compared to postpartum”, the “diff” column refers to the estimated value of the outcome at a given visit minus the estimated value at postpartum, and the p-value corresponds to the statistical test of the null hypothesis that that difference is 0.

**Supplemental Table 7.** Covariate-adjusted differences obtained from longitudinal models, comparing visits and GDM subgroups, without adjustment for BMI

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Visit** | **All** | | | **No GDM** | | | **Early GDM** | | | | | **Classic GDM** | | | | |
| **Est. mean** | **Compared to postpartum** | | **Est. mean** | **Compared to postpartum** | | **Est. mean** | **Compared to postpartum** | | **Compared to No GDM** | | **Est. mean** | **Compared to postpartum** | | **Compared to No GDM** | |
| **Diff** | **p** | **Diff** | **p** | **Diff** | **p** | **Diff** | **p** | **Diff** | **p** | **Diff** | **p** |
| log Stumvoll, not adjusted for Matsuda | Early | 6.84 (6.75, 6.94) | 0.19 | <0.001 | 6.90 (6.81, 7.00) | 0.16 | 0.003 | 6.30 (5.66, 6.94) | 0.43 | 0.002 | -0.22 | 0.045 | 6.60 (6.31, 6.89) | 0.13 | 0.233 | -0.29 | 0.005 |
| Mid-late | 6.95 (6.85, 7.05) | 0.30 | <0.001 | 7.00 (6.91, 7.10) | 0.26 | <0.001 | - | - | - | - | - | 6.78 (6.49, 7.08) | 0.32 | 0.005 | -0.20 | 0.011 |
| Postpartum | 6.65 (6.55, 6.76) | - | - | 6.75 (6.64, 6.85) | - | - | 5.87 (5.23, 6.51) | - | - | -0.39 | 0.013 | 6.47 (6.16, 6.78) | - | - | -0.23 | 0.091 |
| log Matsuda | Early | 2.20 (2.05, 2.34) | -0.16 | 0.005 | 2.33 (2.15, 2.52) | -0.05 | 0.485 | 1.91 (1.91, 2.48) | -0.64 | <0.001 | -0.76 | <0.001 | 2.07 (1.65, 2.48) | -0.21 | 0.099 | -0.16 | 0.344 |
| Mid-late | 1.80 (1.66, 1.93) | -0.56 | <0.001 | 1.90 (1.74, 2.06) | -0.48 | <0.001 | - | - | - | - | - | 1.54 (1.12, 1.96) | -0.74 | <0.001 | -0.27 | 0.056 |
| Postpartum | 2.35 (2.22, 2.49) | - | - | 2.38 (2.21, 2.56) | - | - | 2.55 (2.55, 3.08) | - | - | -0.16 | 0.343 | 2.28 (1.89, 2.67) | - | - | -0.04 | 0.773 |
| log Stumvoll, adjusted for log Matsuda | Early | 6.83 (6.75, 6.92) | 0.15 | <0.001 | 6.94 (6.86, 7.02) | 0.15 | 0.003 | 6.22 (5.65, 6.80) | 0.24 | 0.150 | -0.53 | <0.001 | 6.50 (6.30, 6.70) | 0.03 | 0.771 | -0.36 | <0.001 |
| Mid-late | 6.84 (6.75, 6.93) | 0.15 | 0.002 | 6.91 (6.83, 6.99) | 0.12 | 0.018 | - | - | - | - | - | 6.44 (6.21, 6.67) | -0.03 | 0.822 | -0.30 | <0.001 |
| Postpartum | 6.69 (6.59, 6.78) | - | - | 6.79 (6.70, 6.88) | - | - | 5.98 (5.39, 6.57) | - | - | -0.44 | 0.004 | 6.47 (6.25, 6.69) | - | - | -0.24 | 0.057 |
| Pregnancy Insulin Physiology (PIP) Index | Early | 2437 (2252, 2621) | 212 | 0.044 | 2653 (2447, 2859) | 286 | 0.033 | 1414 (778, 2050) | 118 | 0.586 | -1022 | <0.001 | 1768 (1444, 2091) | -16 | 0.950 | -811 | <0.001 |
| Mid-late | 2286 (2118, 2455) | 61 | 0.575 | 2426 (2244, 2608) | 59 | 0.656 | - | - | - | - | - | 1680 (1395, 1966) | -103 | 0.681 | -683 | <0.001 |
| Postpartum | 2225 (1993, 2457) | - | - | 2367 (2111, 2623) | - | - | 1296 (578, 2013) | - | - | -723 | 0.008 | 1784 (1209, 2358) | - | - | -485 | 0.071 |

Estimated mean values for each outcome are obtained from covariate-adjusted linear mixed effects models, fit among all participants together and then in separate models stratified by GDM subgroup. The estimated mean value at each visit corresponds to a participant who is the mean age of the participant population, has no family history of diabetes, is not Hispanic, is married, and has completed college. For log Matsuda-adjusted models, the reference individual also has the mean observed value of log Stumvoll at the visit being modeled. Under “Compared to postpartum”, the “diff” column refers to the estimated value of the outcome at a given visit minus the estimated value at postpartum, and the p-value corresponds to the statistical test of the null hypothesis that that difference is 0. The results under “Compared to No GDM” corresponds to a different set of linear regression models, where the observations included in each model are from participants in all 3 subgroups at a single visit. The “diff” column refers to the difference between the value for a given subgroup minus the value for the no GDM group at the given visit, and the p-value corresponds to the null hypothesis that the difference is 0. These differences do not align exactly with the estimated mean values from the longitudinal models as they are obtained from a separate set of regression models that leverage data from all 3 subgroups simultaneously, but only at one visit at a time.

**Supplemental Table 8.** Covariate-adjusted differences obtained from longitudinal models for log Stumvoll, adjusted for log Matsuda, comparing visits and GDM subgroups, with additional adjustment for breastfeeding practice at postpartum

|  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **Model** | **Visit** | **All** | | | **No GDM** | | | **Early GDM** | | | | | **Classic GDM** | | | | |
| **Est. mean** | **Compared to postpartum** | | **Est. mean** | **Compared to postpartum** | | **Est. mean** | **Compared to postpartum** | | **Compared to No GDM** | | **Est. mean** | **Compared to postpartum** | | **Compared to No GDM** | |
| **Diff** | **p** | **Diff** | **p** | **Diff** | **p** | **Diff** | **p** | **Diff** | **p** | **Diff** | **p** |
| Log Stumvoll, adjusted for log Matsuda | Early | 6.86 (6.77, 6.96) | 0.16 | <0.001 | 6.96 (6.87, 7.05) | 0.16 | 0.002 | 6.13 (5.49, 6.77) | 0.25 | 0.147 | -0.52 | <0.001 | 6.58 (6.34, 6.81) | 0.03 | 0.747 | -0.35 | <0.001 |
| Mid-late | 6.86 (6.76, 6.96) | 0.15 | 0.001 | 6.93 (6.83, 7.02) | 0.13 | 0.013 | - | - | - | - | - | 6.51 (6.25, 6.78) | -0.03 | 0.831 | -0.29 | <0.001 |
| Postpartum | 6.71 (6.60, 6.81) | - | - | 6.80 (6.70, 6.90) | - | - | 5.88 (5.23, 6.52) | - | - | -0.42 | 0.006 | 6.54 (6.30, 6.78) | - | - | -0.21 | 0.124 |

Estimated mean values for each outcome are obtained from covariate-adjusted linear mixed effects models, fit among all participants together and then in separate models stratified by GDM subgroup. The estimated mean value at each visit corresponds to a participant who is the mean age of the participant population, has the mean BMI of all participants at the early pregnancy visit, has no family history of diabetes, is not Hispanic, is married, has completed college, and is not exclusively breastfeeding. Under “Compared to postpartum”, the “diff” column refers to the estimated value of the outcome at a given visit minus the estimated value at postpartum, and the p-value corresponds to the statistical test of the null hypothesis that that difference is 0. The results under “Compared to No GDM” corresponds to a different set of linear regression models, where the observations included in each model are from participants in all 3 subgroups at a single visit. The “diff” column refers to the difference between the value for a given subgroup minus the value for the no GDM group at the given visit, and the p-value corresponds to the null hypothesis that the difference is 0. These differences do not align exactly with the estimated mean values from the longitudinal models as they are obtained from a separate set of regression models that leverage data from all 3 subgroups simultaneously, but only at one visit at a time. Breastfeeding was adjusted for with a binary variable (exclusively breastfeeding vs. not). Participants were categorized as exclusively breastfeeding based on their data at the postpartum visit. 1 participant who attended the postpartum visit but had missing data for breastfeeding, and participants who did not attend the postpartum visit, were categorized as not exclusively breastfeeding.