# Supplemental Information

## MISSING DATA

Mothers who completed more waves (1 vs 2 vs 3 vs 4 waves) had lower depressive symptoms at 4 months ($F_{3,\ 4693}$ = 23.45; $P < .001$), were older ($F_{3,\ 4865}$ = 46.18; $P < .001$), had lower prepregnancy BMI ($F_{3,\ 4858}$ = 6.31; $P < .001$), had infants with higher gestational age ($F_{3,\ 4865}$ = 9.13; $P < .001$), and were more likely to have completed college ($\chi^3_2$ [$n$ = 4866] = 329.53; $P < .001$), be married ($\chi^2_3$ [$n$ = 4735] = 252.87; $P < .001$), and identify as non-Hispanic white ($\chi^2_3$[$n$ = 4866] = 135.39; $P < .001$). Mothers who completed more waves were also more likely to have had infertility treatment ($\chi^2_3$ [$n$ = 4866] = 92.54; $P < .001$) and be nulliparous ($\chi^2_3$ [$n$ = 4830] = 71.44; $P < .001$) but less likely to have given birth to multiples ($\chi^2_3$ [$n$ = 4866] = 36.23; $P < .001$) and smoked during pregnancy ($\chi^2_3$ [$n$ = 4865] = 160.23; $P < .001$). Dropout was unrelated to GDM, gestational hypertension, and infant sex ($\chi^2_3$ [$n$ = 4866] = 0.70–1.41; $P$ = .70–.87). These patterns of differential missingness make it particularly important to handle missing data in depressive symptoms over time, taking demographic and perinatal factors into account by using FIML.

## MODELING STRATEGY

A single-group linear latent growth model fit the data well ($\chi^2_3$ = 30.41; $P < .001$; comparative fit index [CFI] = .986; Tucker-Lewis fit index [TLI] = .983; root mean square error of approximation [RMSEA] = .032; 90% CI: .022 to .044; standardized root mean residual [SRMR] =.025). The intercept was $b$ = 2.68 ($P < .001$), and the slope was $b$ = −.08 ($P < .001$), suggesting a low average level of depressive symptoms at 4 months and a slight decline over the study period. The intercept and slope were inversely related ($r$ = −.52; $P < .001$), and the variances of both the intercept ($b$ = 4.58; $SE$ = .24; $P < .001$) and slope ($b$ = .13; $SE$ = .05; $P$ = .005) were significant, suggesting that there were individual differences around the intercept and slope that could be modeled with latent class growth analysis.

Next, the relative fit of latent class models from 2 to 7 classes was compared with the model with 1 less class (eg, 3 classes was compared with 2, and 7 classes was compared with 6). Model fit was compared by using decreasing Akaike's information criterion, Bayesian information criterion, and sample-size–adjusted Bayesian information criterion and with the Lo-Mendell-Rubin adjusted likelihood ratio test (LRT) and bootstrapped LRT. Priority was given to the bootstrapped LRT because it was found to perform the best under varied model conditions.[59] Theoretical considerations, entropy, and the sample sizes within classes were also considered when choosing the best number of classes. Finally, to confirm the number of classes and ensure that classes did not change markedly with the addition of demographic and perinatal predictors, a conditional model regressing the intercept, slope, and latent class membership variable on the demographic and perinatal predictors was estimated, following the methods of Jung and Wickrama.[60]

Supplemental Table 4 presents the fit statistics and model comparisons for the first step of the 3-step model. In these models, the variances of the intercept and slope were estimated, but were constrained across classes to achieve model convergence. The bootstrapped LRT and decreasing Akaike's information criterion, Bayesian information criterion, and sample-size–adjusted Bayesian information criterion indicated that each additional class improved model fit. However, in the 6- and 7-class solutions, the latent variable covariance matrix was not positive definite, suggesting that classes were not distinct, and at least 1 class contained <100 mothers (<2% of cases). The 4- and 5-class solutions were both acceptable, but entropy was slightly higher and the Lo-Mendell-Rubin test was significant only for the 4-class model, suggesting 4 classes may be the best solution. To confirm this choice, 4- and 5-class conditional models were computed with the predictors in the model (data are not shown). The class solution was similar for the 3-step and conditional models when estimating 4-classes but not 5-classes, indicating that the 4-class solution was more stable.

**SUPPLEMENTAL TABLE 4** Model Comparisons of Unconditional Growth Mixture Models

| Classes | AIC | BIC | SSA-BIC | Lo-Mendell-Rubin Adjusted LRT | Bootstrapped LRT | Entropy | Δ AIC | Δ BIC | Δ SSA-BIC |
|---|---|---|---|---|---|---|---|---|---|
| 2 | 53 496.76 | 53 568.15 | 53 533.20 | 1102.53*** | 1145.81*** | 0.869 | 1139.82 | 1120.34 | 1129.88 |
| 3 | 53 203.63 | 53 294.49 | 53 250.00 | 287.83** | 299.13*** | 0.826 | 293.13 | 273.66 | 283.19 |
| 4 | 52 970.74 | 53 081.07 | 53 027.05 | 229.86* | 238.89*** | 0.795 | 232.89 | 213.42 | 222.95 |
| 5 | 52 851.26 | 52 981.06 | 52 917.51 | 120.74 | 125.48*** | 0.784 | 119.48 | 100.01 | 109.55 |
| 6[a] | 52 786.04 | 52 935.31 | 52 862.22 | 68.53 | 71.23*** | 0.791 | 65.23 | 45.76 | 55.29 |
| 7[a] | 52 661.69 | 52 830.43 | 52 747.81 | 94.11 | 97.80*** | 0.748 | 124.35 | 104.88 | 114.41 |

AIC, Akaike's information criterion; BIC, Bayesian information criterion; LRT, likelihood ratio test; SSA-BIC, sample-size–adjusted Bayesian information criterion.

[a] The latent variable covariance matrix ($\psi$) was not positive definite in this model.

* $P < .05$.

** $P < .01$.

*** $P < .001$.