

Additional Table 4: 10-fold cross-validated AUC comparisons via Delong's test on the validation set

Prediction Method	Predictor Levels ¹	P-values
CART	1-2 vs. 1	0.0127
	1-3 vs. 1-2	<2.2e-16
	1-4 vs. 1-3	8.495E-08
LASSO regression	1-2 vs. 1	0.0041
	1-3 vs. 1-2	<2.2e-16
	1-4 vs. 1-3	<2.2e-16
Simple super learner ²	1-2 vs. 1	0.0103
	1-3 vs. 1-2	<2.2e-16
	1-4 vs. 1-3	<2.2e-16
Complex super learner ³	1-2 vs. 1	0.0008
	1-3 vs. 1-2	<2.2e-16
	1-4 vs. 1-3	<2.2e-16

AUC, area under the curve; CART, classification and regression tree; LASSO, least absolute shrinkage and selection operator; XGBoost, extreme gradient boosting

¹Level 1: one-year preconception to last menstrual period; level 2: last menstrual period to before diagnosis of gestational diabetes; level 3: at the time of diagnosis of gestational diabetes; level 4: one week following diagnosis of gestational diabetes.

²Candidate algorithms in simple super learner included response-mean, LASSO regression, and CART.

³Candidate algorithms in complex super learner included response-mean, LASSO regression, CART, random forest, and XGBoost algorithms.