Supplemental Table 4: Pipeline ensembles

Pipeline ensemble for each model developed using *AutoPrognosis 2.0*, for both multi-class and binary predictions.

Multi-class predictions	
Model	**Pipeline Ensemble**
AP1_mu	gain->data_cleanup->random_forest + softimpute->variance_threshold->feature_normalizer->data_cleanup->logistic_regression + ice->feature_agglomeration->maxabs_scaler->data_cleanup->lgbm --> gain->data_cleanup->random_forest
AP2_mu	0.624999999609375 * (mice->pca->minmax_scaler->data_cleanup->random_forest) + 0.1874999998828125 * (median->variance_threshold->feature_normalizer->data_cleanup->logistic_regression) + 0.1874999998828125 * (missforest->variance_threshold->scaler->data_cleanup->lgbm)
AP3_mu	0.3124999998046875 * (gain->data_cleanup->random_forest) + 0.24999999984375 * (gain->data_cleanup->lgbm) + 0.4374999997265625 * (gain->data_cleanup->xgboost)
AP4_mu	0.3157894735180055 * (median->nop->nop->data_cleanup->random_forest) + 0.3684210524376731 * (gain->data_cleanup->lgbm) + 0.3157894735180055 * (median->feature_agglomeration->minmax_scaler->data_cleanup->logistic_regression)
AP5_mu	average(median->nop->nop->data_cleanup->random_forest + gain->data_cleanup->lgbm + gain->data_cleanup->xgboost)
AP5_top5_mu	gain->data_cleanup->random_forest + missforest->data_cleanup->scaler->data_cleanup->logistic_regression + gain->data_cleanup->lgbm --> gain->data_cleanup->random_forest

Binary predictions	
Model	**Pipeline Ensemble**
AP1_bi	ice->nop->scaler->data_cleanup->random_forest + median->feature_agglomeration->scaler->data_cleanup->catboost + nop->data_cleanup->uniform_transform->data_cleanup->lgbm --> ice->nop->scaler->data_cleanup->random_forest
AP2_bi	0.5882352937716263 * (mice->pca->minmax_scaler->data_cleanup->random_forest) + 0.2352941175086505 * (median->feature_agglomeration->scaler->data_cleanup->catboost) + 0.1764705881314879 * (ice->nop->uniform_transform->data_cleanup->xgboost)
AP3_bi	0.5555555552469136 * (ice->data_cleanup->nop->data_cleanup->catboost) + 0.2777777776234568 * (softimpute->variance_threshold->maxabs_scaler->data_cleanup->xgboost) + 0.16666666657407406 * (gain->data_cleanup->random_forest)
AP4_bi	median->nop->nop->data_cleanup->random_forest + gain->data_cleanup->lgbm + softimpute->nop->minmax_scaler->data_cleanup->catboost --> median->nop->nop->data_cleanup->random_forest
AP5_bi	median->nop->nop->data_cleanup->random_forest + gain->data_cleanup->lgbm + softimpute->nop->minmax_scaler->data_cleanup->catboost --> median->nop->nop->data_cleanup->random_forest
AP5_top5_bi	average(median->pca->minmax_scaler->data_cleanup->random_forest + gain->nop->uniform_transform->data_cleanup->catboost + ice->pca->maxabs_scaler->data_cleanup->logistic_regression)