

Additional Table 5: Complex super learner output

Predictors level	Candidate algorithms	Risk¹	Coefficient
Level 1	Response-mean	0.501	0.027
	LASSO regression	0.331	0.301
	CART	0.393	0
	Random forest	0.351	0
	XGBoost	N/A ²	0.673
Levels 1-2	Response-mean	0.501	0.033
	LASSO regression	0.316	0.379
	CART	0.387	0
	Random forest	0.330	0.050
	XGBoost	N/A ²	0.539
Levels 1-3	Response-mean	0.501	0.027
	LASSO regression	0.214	0.223
	CART	0.264	0
	Random forest	0.215	0.080
	XGBoost	N/A ²	0.670
Levels 1-4	Response-mean	0.501	0.035
	LASSO regression	0.151	0.190
	CART	0.219	0
	Random forest	0.146	0.164
	XGBoost	N/A ²	0.611

CART, classification and regression trees; LASSO, least absolute shrinkage and selection operator; XGBoost, extreme gradient boosting

¹Risk is the mean of the loss measured by the negative binomial log-likelihood

²N/A for XGBoost due to multiple tuning parameters used in the candidate learners set. Risk over multiple XGBoost algorithms is not shown.