Table S1. The optimal hyperparameter values and prediction performance of the RF full model and reduced model trained by 9 combinations (3 datasets×3 resampling methods).

Resampling method	Predictor	Set	Hyperparameter		OOB data					Testing data				
			ntree	*mtry*	*Sens*	*Prec*	*F1*	*Acc*	*bAcc*	*Sens*	*Prec*	*F1*	*Acc*	*bAcc*
None	full	1	200	6	82.7%	85.3%	83.9%	88.3%	87.1%	85.3%	80.6%	82.9%	89.4%	88.2%
		2	400	6	83.5%	87.1%	85.3%	89.4%	88.2%	68.6%	82.8%	75.0%	85.8%	81.1%
		3	400	6	76.8%	86.9%	81.6%	87.5%	85.2%	83.8%	83.8%	83.8%	89.4%	87.9%
	reduced	1	200	1	89.5%	78.6%	83.7%	88.7%	88.9%	79.4%	71.1%	75.0%	84.1%	82.7%
		2	200	1	87.4%	78.4%	82.6%	87.9%	87.8%	68.6%	92.3%	78.7%	88.5%	83.0%
		3	200	1	84.7%	75.8%	80.0%	86.4%	86.0%	83.8%	86.1%	84.9%	90.3%	88.6%
SMOTE	full	1	600	3	91.8%	90.9%	91.4%	90.6%	90.5%	88.2%	71.4%	78.9%	85.8%	86.5%
		2	700	3	93.3%	92.3%	92.8%	92.3%	92.2%	74.3%	81.3%	77.6%	86.7%	83.3%
		3	150	6	88.9%	92.3%	90.6%	90.3%	90.4%	86.5%	82.1%	84.2%	89.4%	88.6%
	reduced	1	200	1	91.7%	89.8%	90.7%	90.1%	90.0%	88.2%	66.7%	75.9%	83.2%	84.6%
		2	200	1	93.2%	92.3%	92.7%	92.3%	92.2%	74.3%	81.3%	77.6%	86.7%	83.3%
		3	200	1	89.9%	88.9%	89.4%	88.9%	88.8%	86.5%	82.1%	84.2%	89.4%	88.6%
B-SMOTE1	full	1	100	3	89.4%	90.4%	89.9%	91.0%	90.8%	90.5%	88.1%	89.1%	91.8%	92.1%
		2	200	3	88.6%	93.7%	91.1%	92.4%	91.9%	73.8%	84.7%	79.0%	86.5%	83.7%
		3	300	3	73.5%	76.0%	74.7%	77.1%	76.8%	72.8%	72.2%	72.3%	79.4%	78.4%
	reduced	1	200	1	94.5%	90.8%	92.5%	93.6%	93.8%	91.1%	83.2%	86.7%	90.0%	90.8%
		2	200	1	87.5%	89.2%	88.3%	89.7%	89.5%	82.9%	90.1%	86.6%	90.9%	89.2%
		3	200	1	68.1%	68.8%	68.5%	71.4%	71.2%	65.3%	67.1%	66.2%	75.9%	73.6%