

Supplementary table 1. Final predictors for developing machine learning models selected by the Lasso procedure.

Lambda	No. of non-zero coefficient	Cross-validation mean deviance	Variables (A)dded, (R)emoved, or left (U)nchanged	
0.1874	1	1.3399	A	STK1p
0.1177	2	1.2365	A	FPSA
0.0673	3	1.1507	A	age
0.0423	4	1.1182	A	FTPSA
0.0291*	4	1.1076	U	

* Lambda selected by the minimum cross-validation mean deviance. STK1p = serum thymidine kinase 1 protein; FPSA = free prostate-specific antigen; FTPSA = free/total prostate-specific antigen ratio.

Supplementary table 2. Multicollinearity test following logistic model

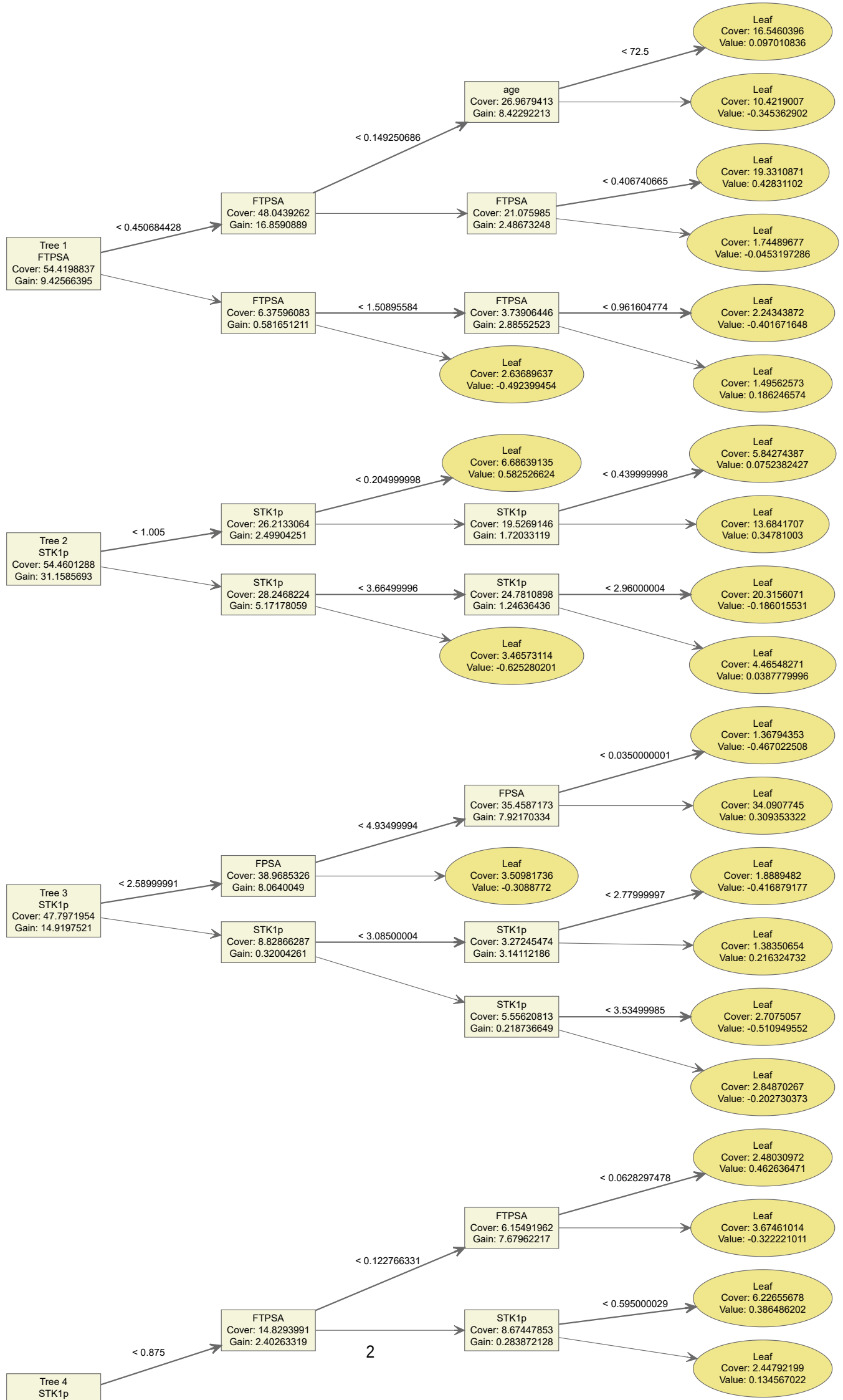
Variable	Variance inflation factor	Tolerance
STK1p (pmol/L)	1.01	0.99
Age (year)	1.01	0.99
TPSA (μg/L)	1.00	1.00

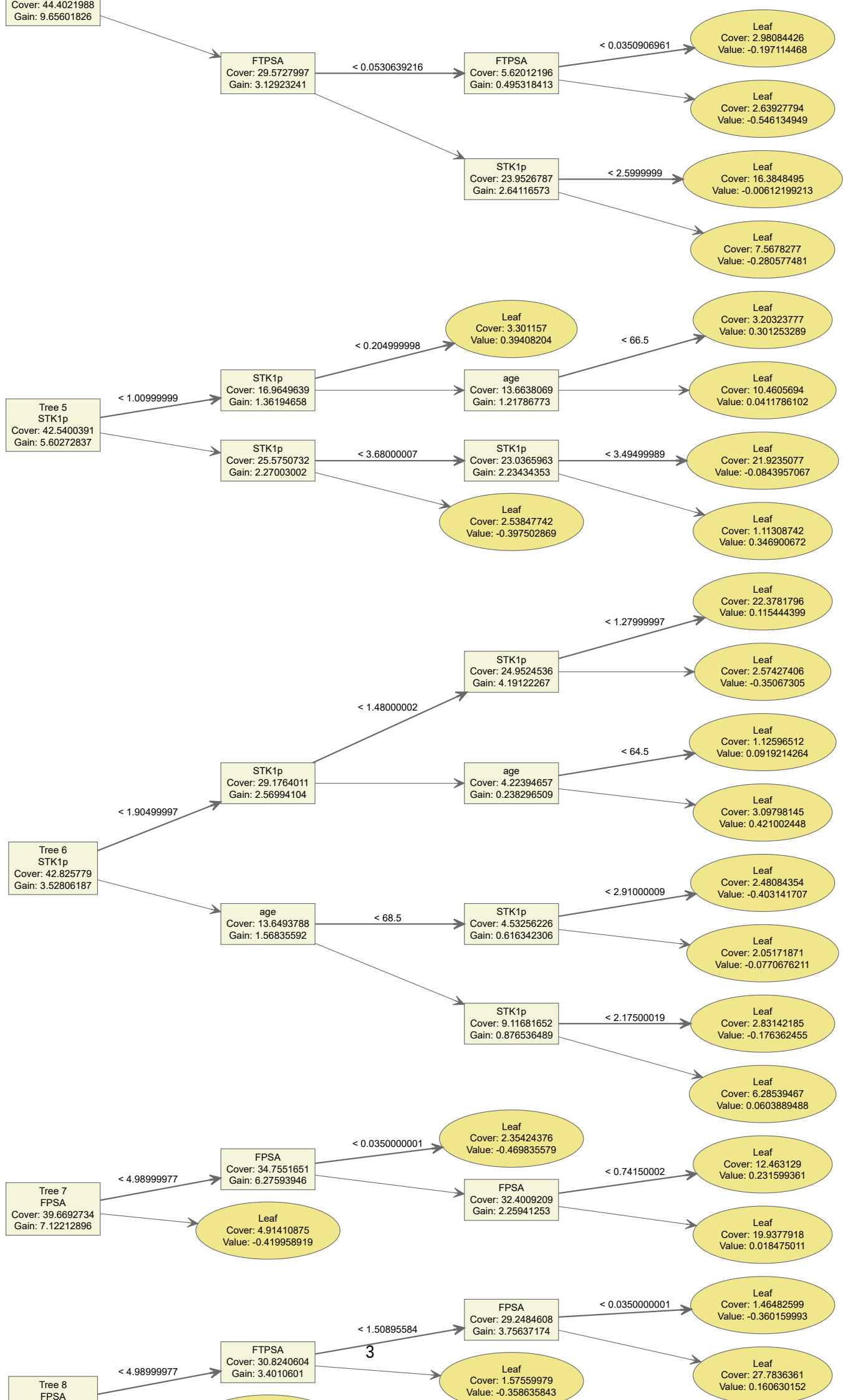
STK1p=serum thymidine kinase 1 protein; TPSA=total prostate-specific antigen

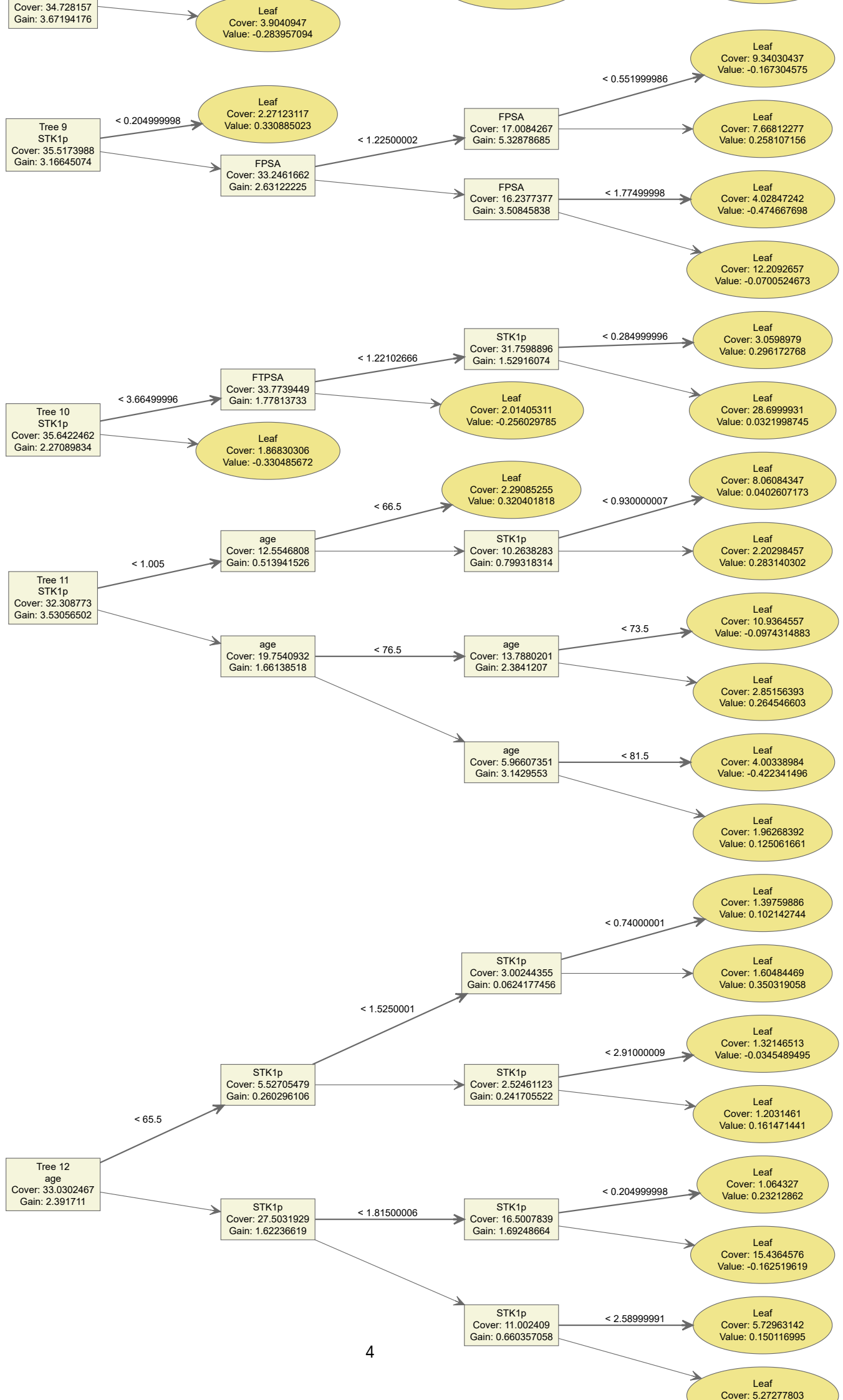
Supplementary table 3. Performance comparison between the XGBOOST and logistic model by subgroups.

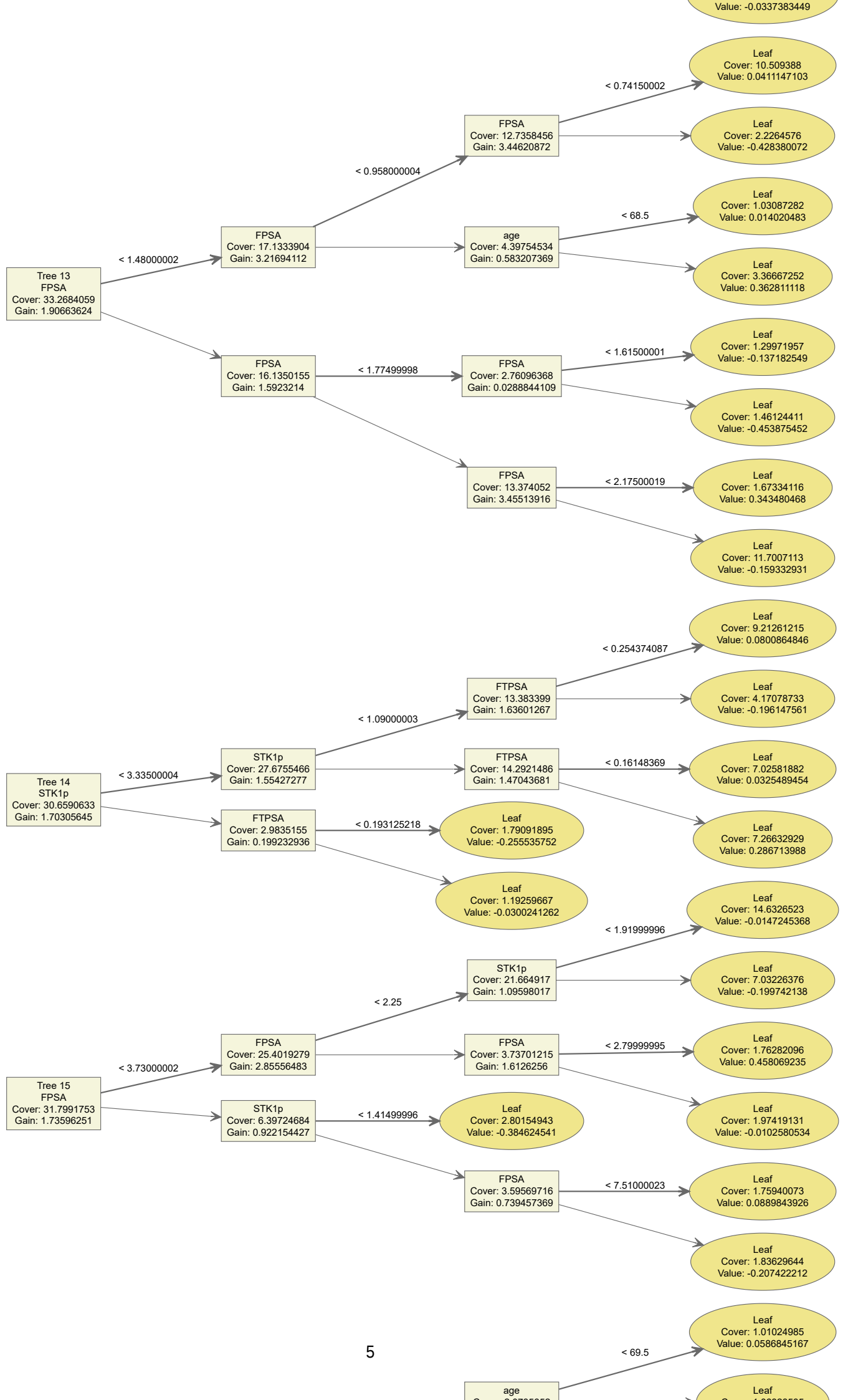
Subgroups	n	Model	AUC (95% confidence interval)	P
Overall	310	XGBOOST	0.965 (0.944, 0.987)	<0.001
		Logistic	0.813 (0.766, 0.861)	
Aged 30-60 years	29	XGBOOST	0.938 (0.815, 1.000)	0.807
		Logistic	0.929 (0.822, 1.000)	
Aged 60-69 years	93	XGBOOST	0.960 (0.913, 1.000)	0.031
		Logistic	0.858 (0.772, 0.944)	
Aged 70-79 years	137	XGBOOST	0.962 (0.928, 0.995)	<0.001
		Logistic	0.738 (0.655, 0.821)	
Aged ≥80 years	51	XGBOOST	0.978 (0.936, 1.000)	0.001
		Logistic	0.784 (0.659, 0.909)	
Center 1	212	XGBOOST	0.931 (0.886, 0.976)	<0.001
		Logistic	0.727 (0.649, 0.805)	
Center 2	98	XGBOOST	1.000 (1.000, 1.000)	0.022
		Logistic	0.938 (0.885, 0.991)	

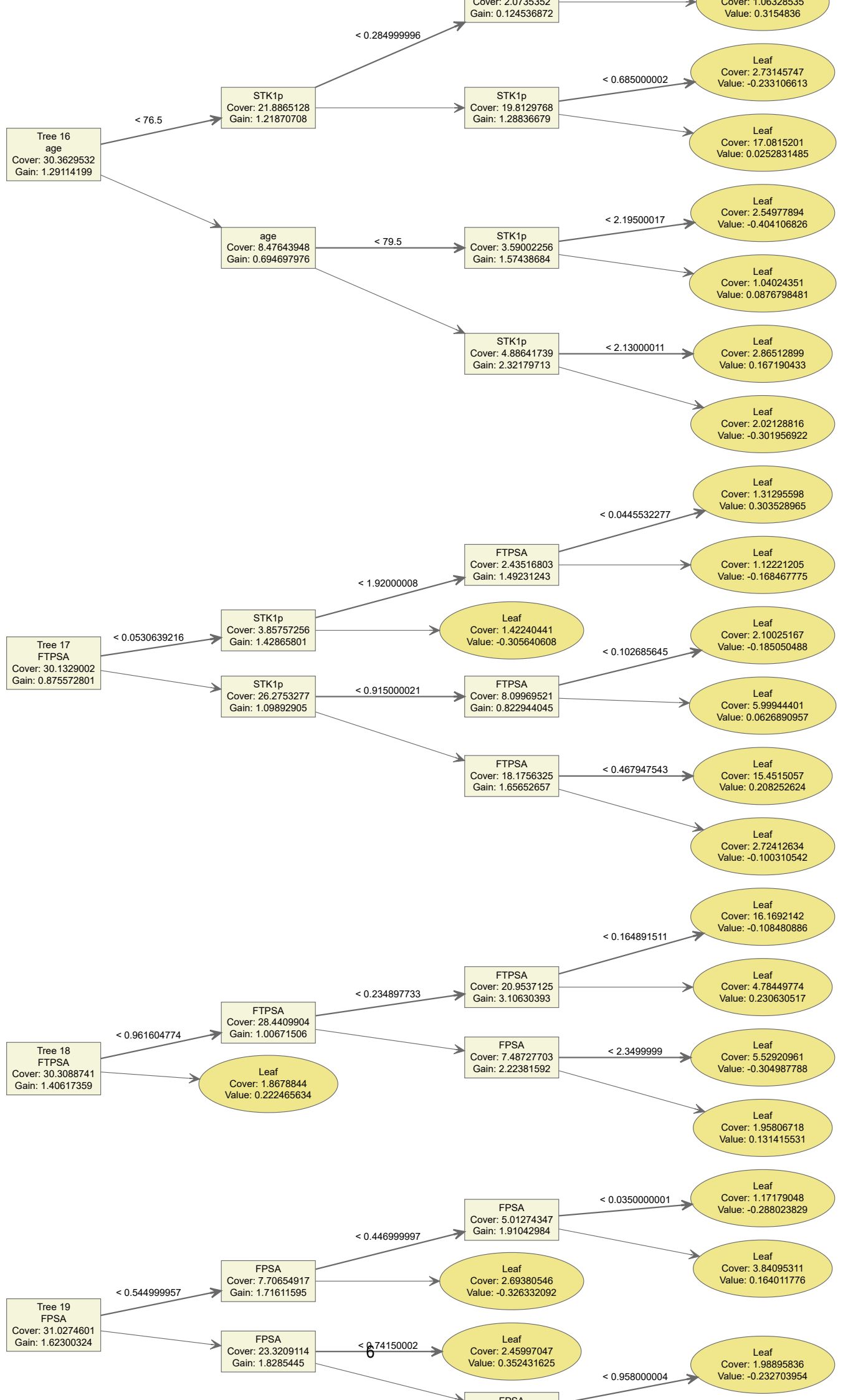
AUC = area under the receiver operating characteristic curve; P values were provided by the DeLong test for comparing AUCs. Center 1 refers to the First People's Hospital of Longquanyi District; center 2 refers to the Daping Hospital of the Third Military Medical University.

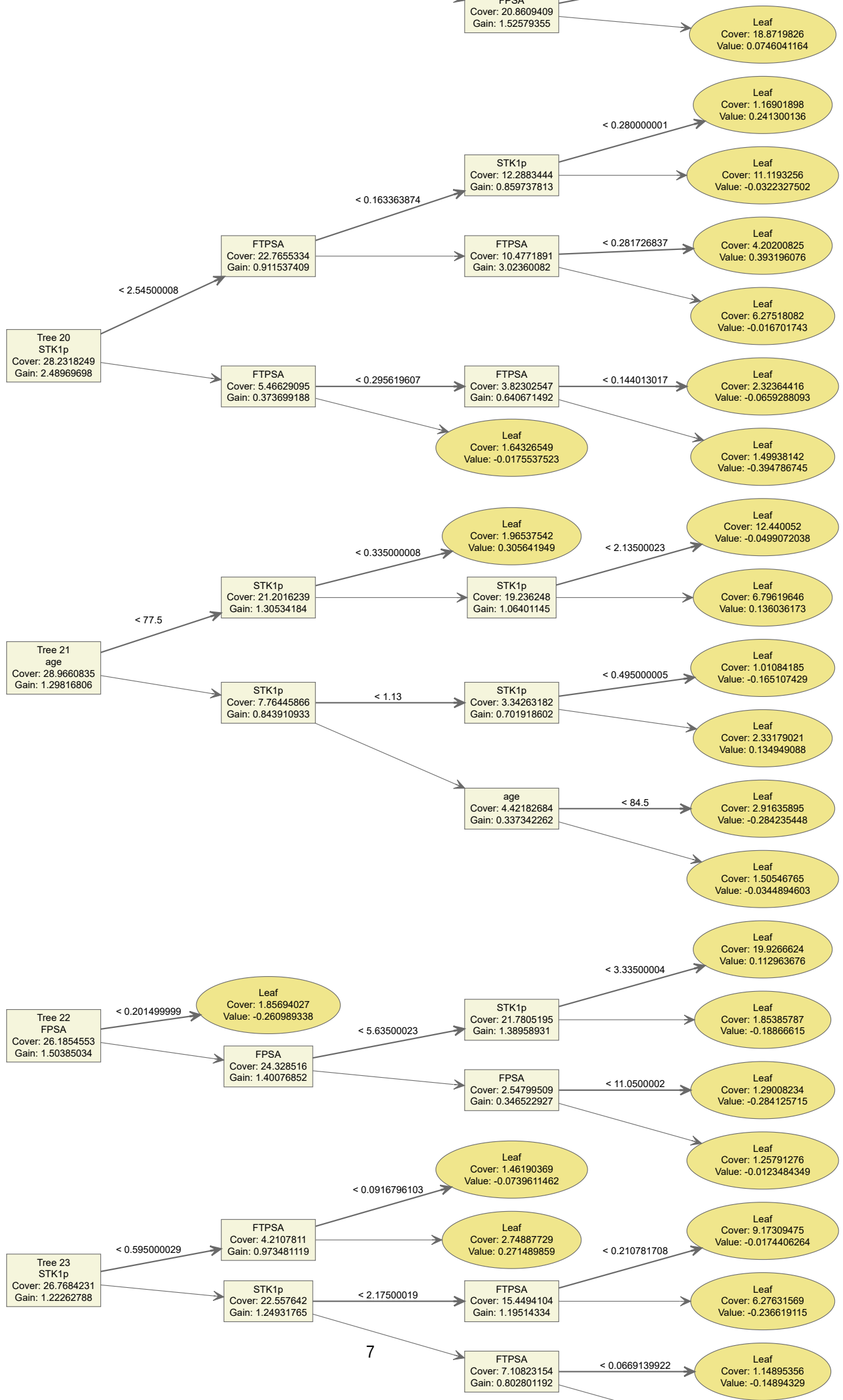


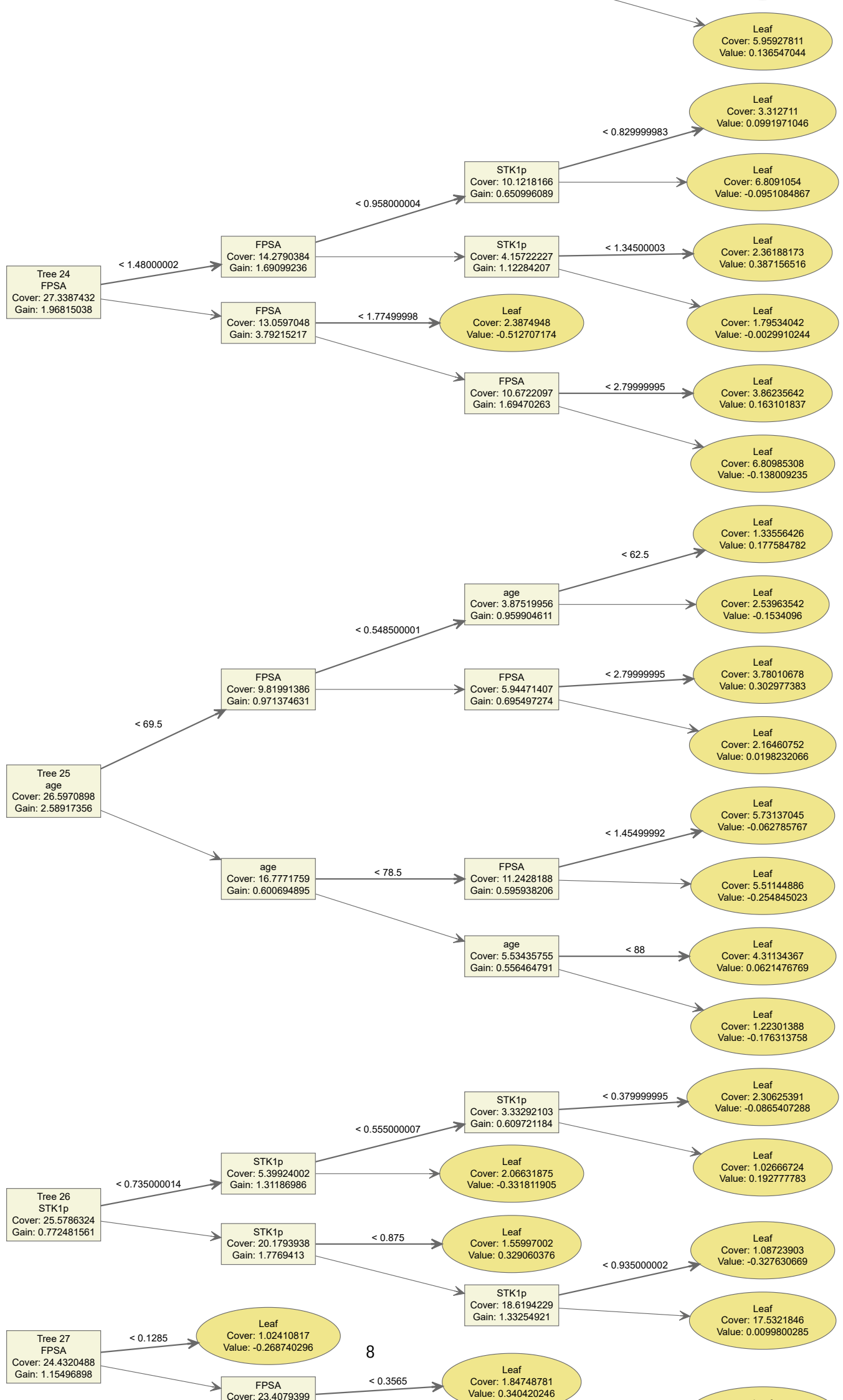


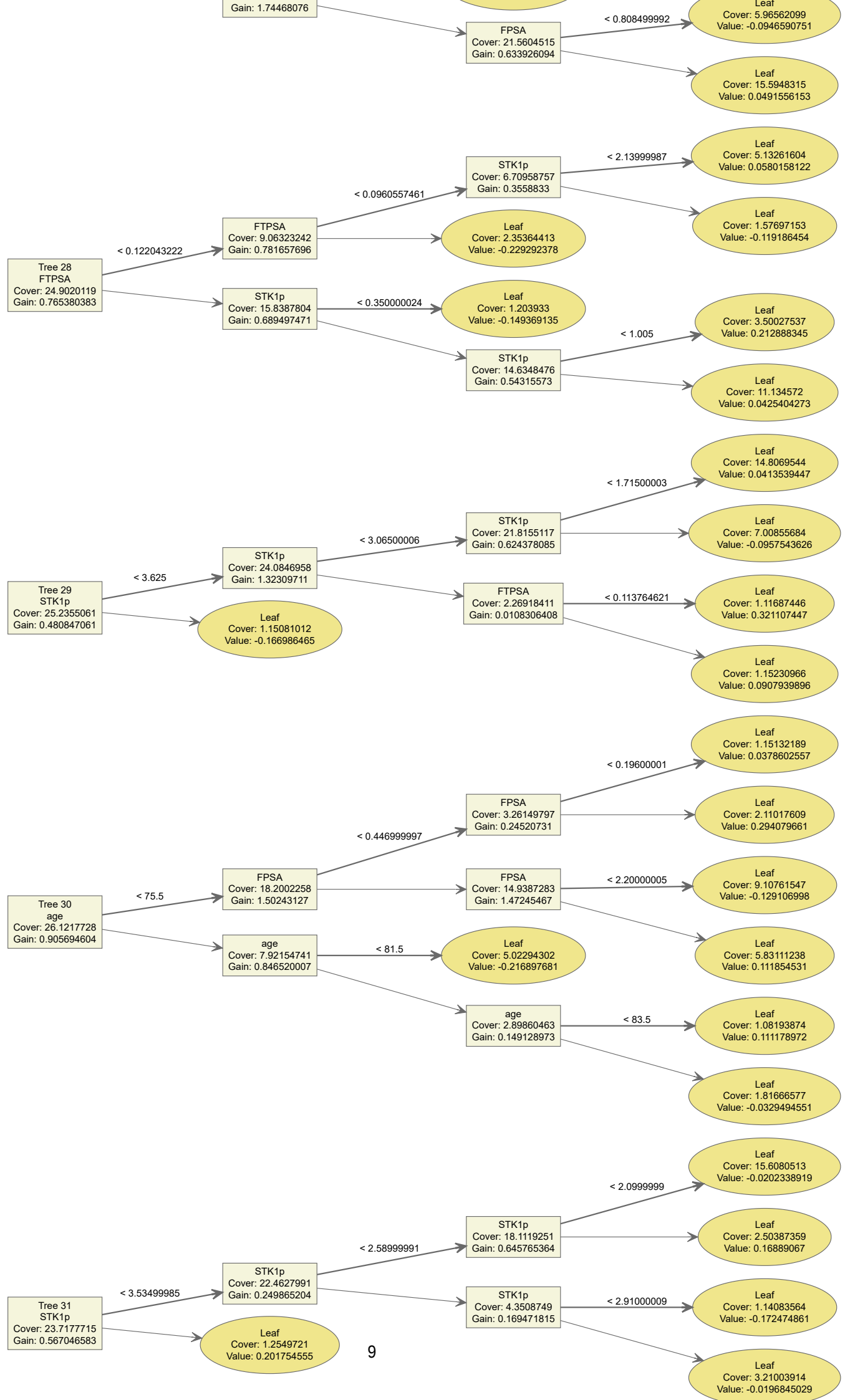


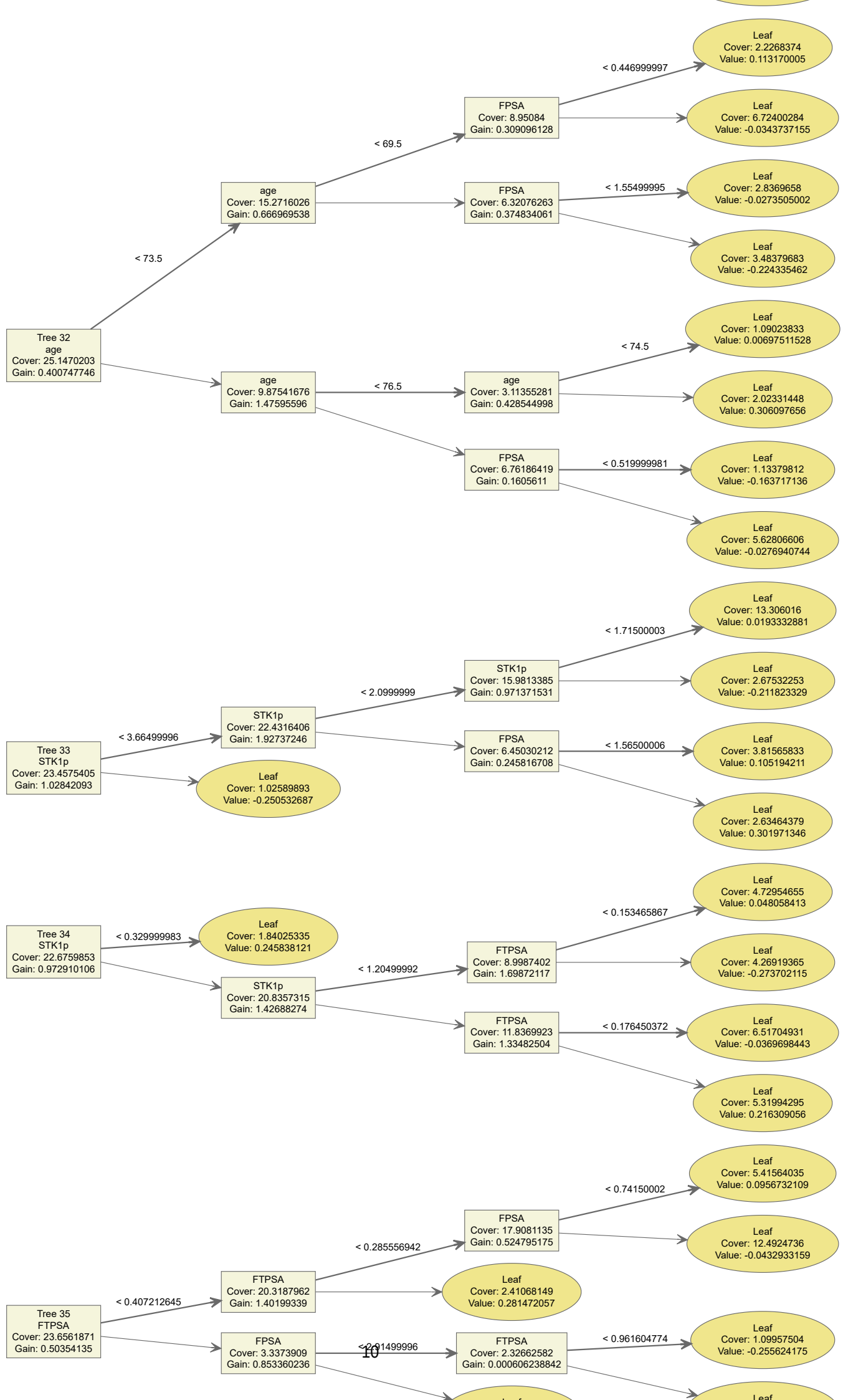


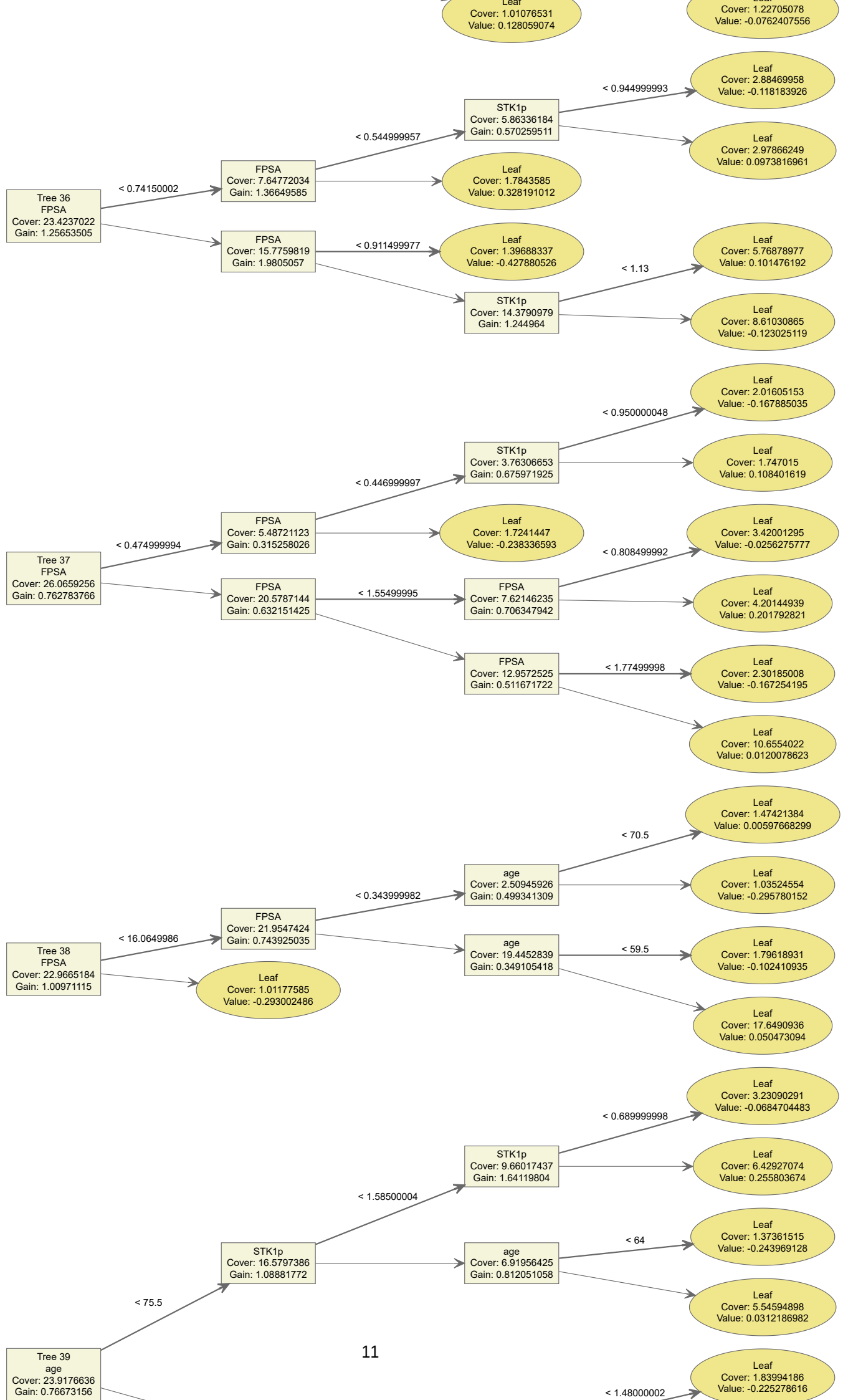


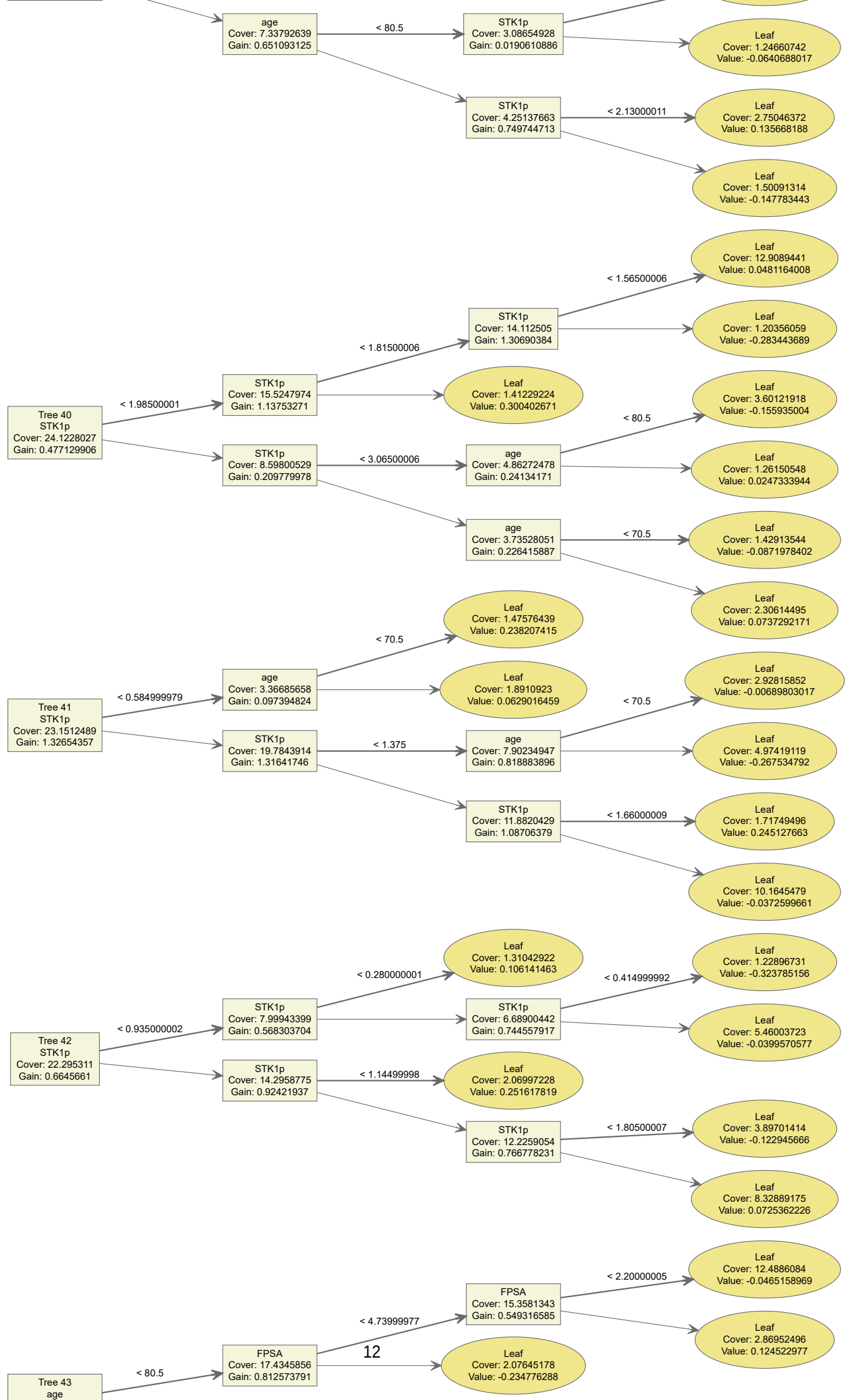


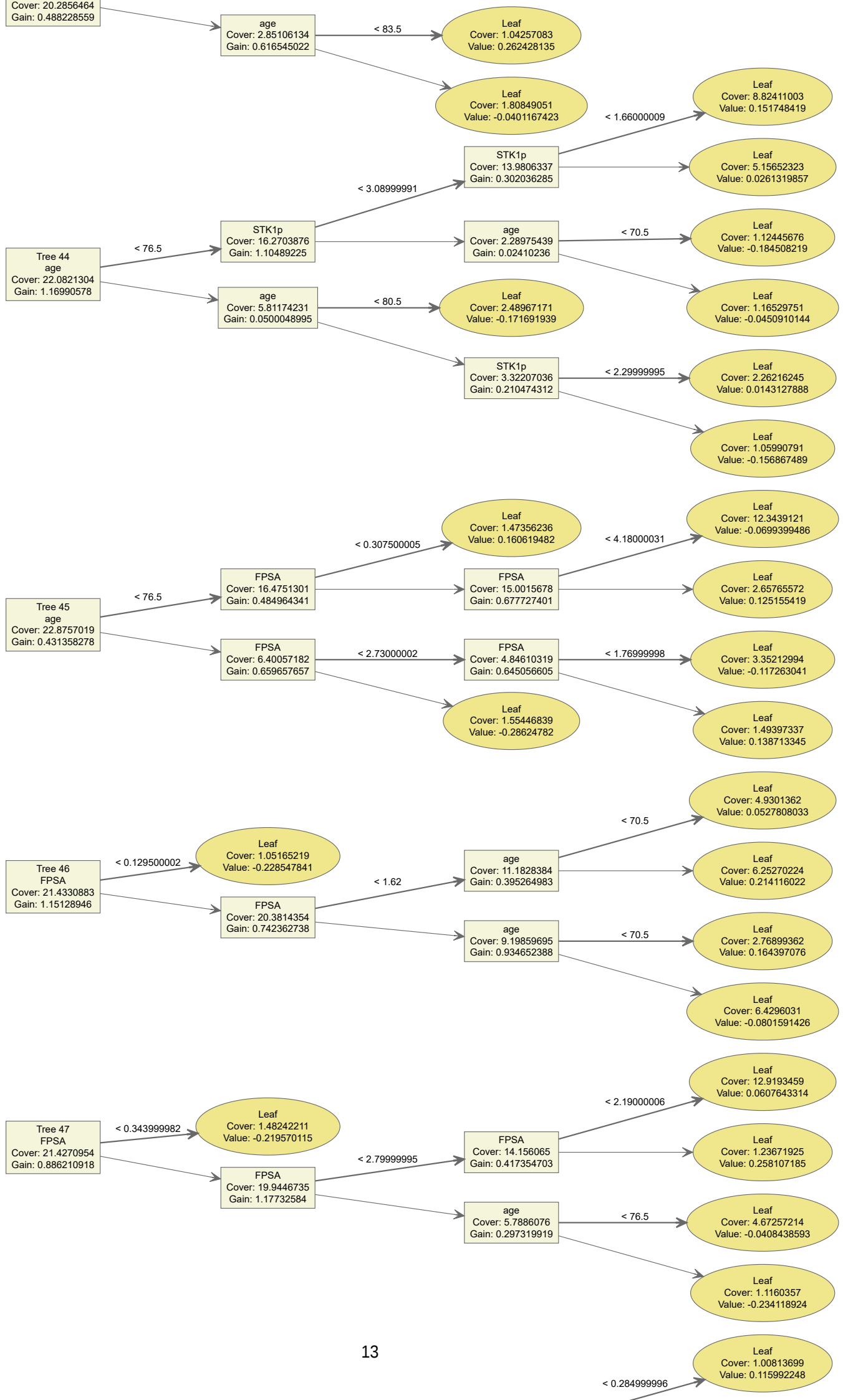


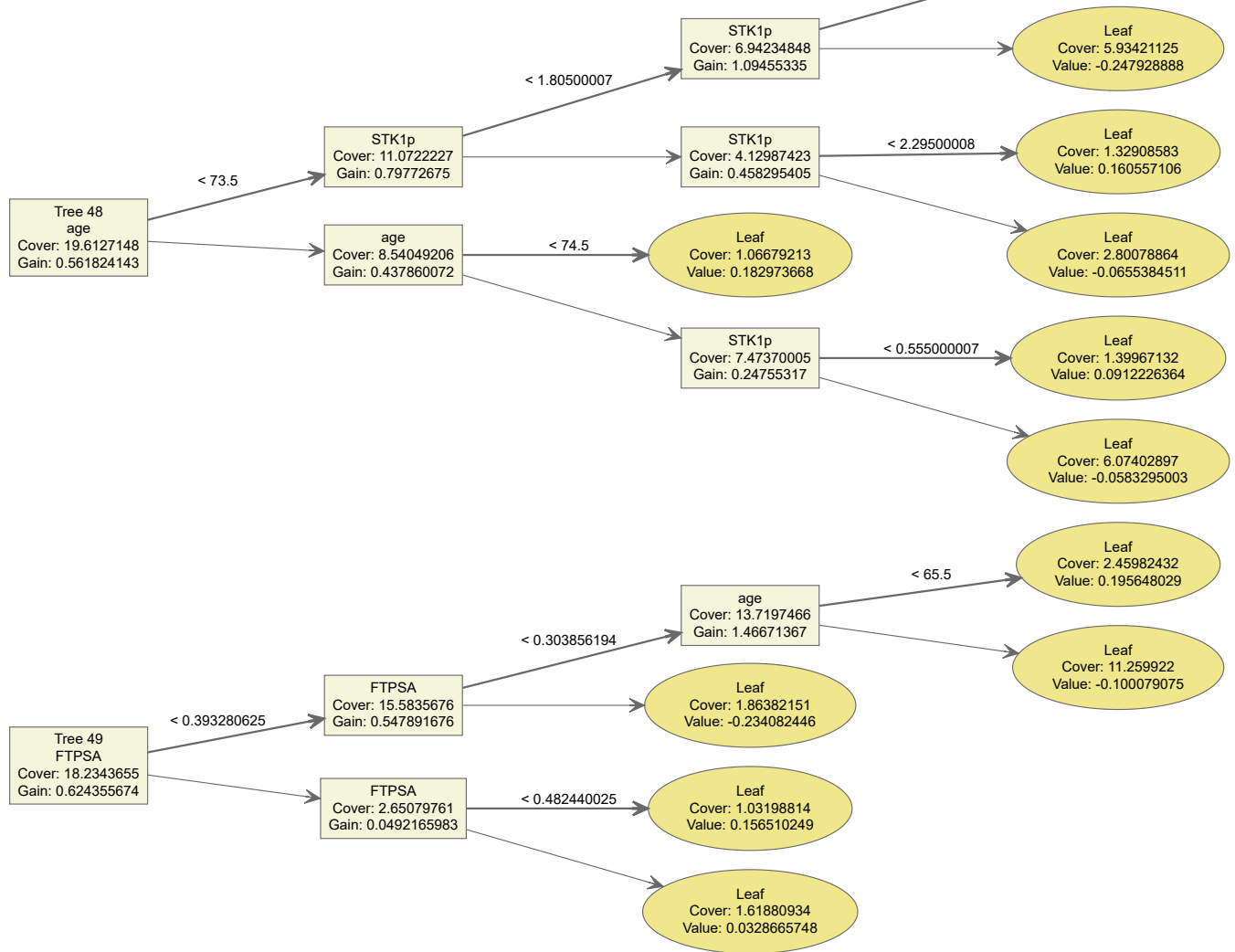












Supplementary Fig 1. All decision trees of the XGBOOST model.