

HOSEA Aim I – Results (MICE)

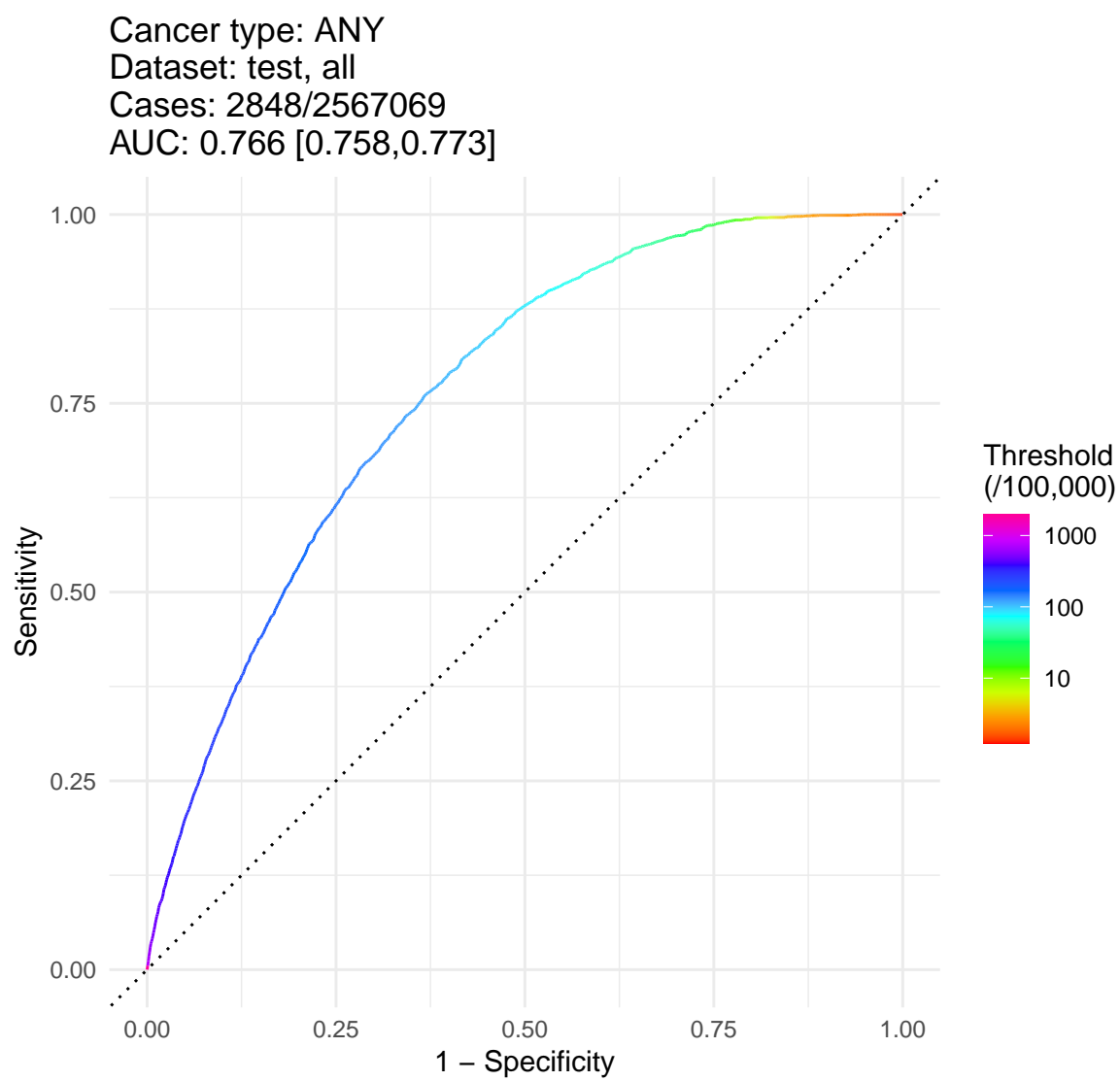
Simon Fontaine

October 31, 2022

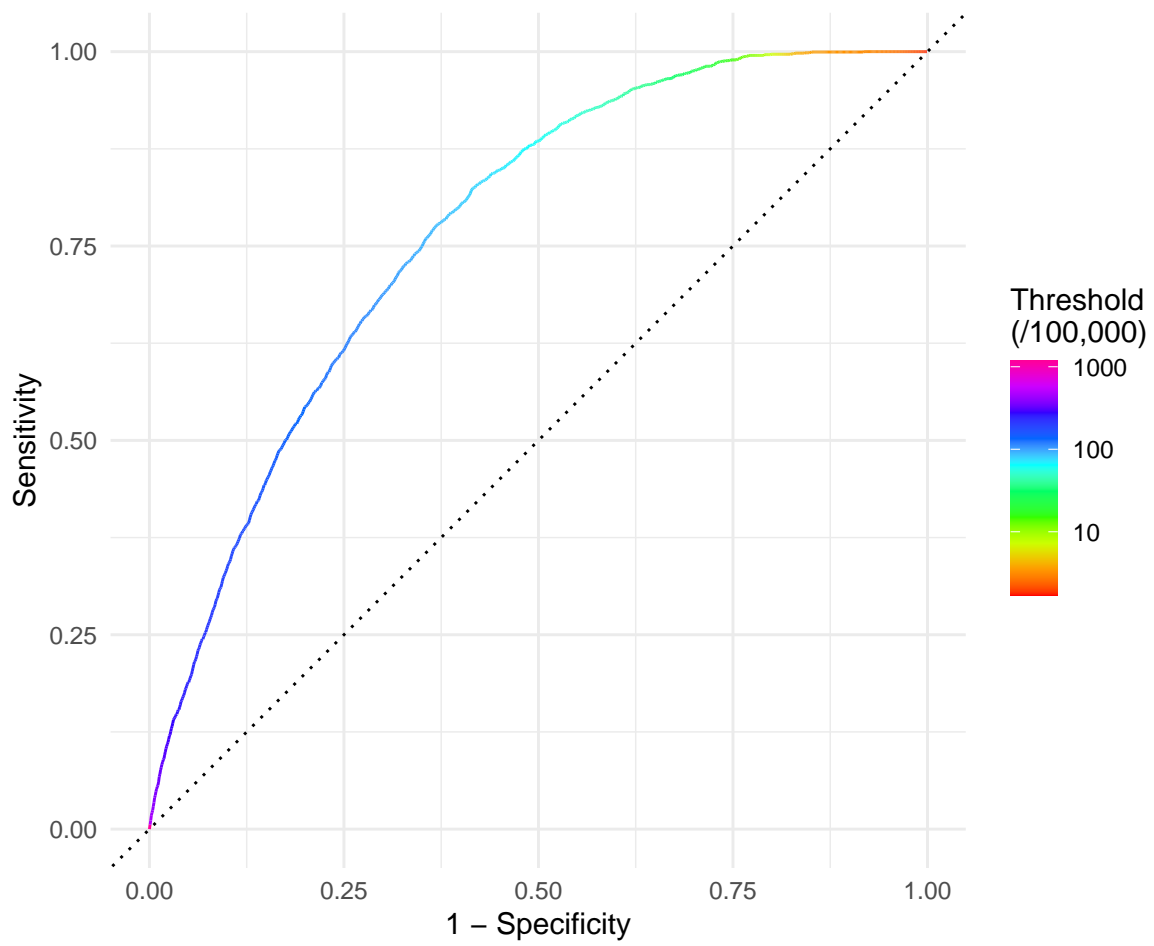
Contents

1	Overall performance	2
2	Comparison to HUNT, Kunzmann and Guidelines	5
2.1	Full test data	5
2.2	Complete test patients	8
2.3	Representative sample	11
3	Calibration	14
4	Thresholds	17
5	Stratification by identity groups	21
5.1	Age	21
5.2	Sex	24
5.3	Race	27
6	Cancer stage	30
7	Variable importance	31
7.1	Gain VI	31
7.2	SHAP VI	33
8	Years prior	35

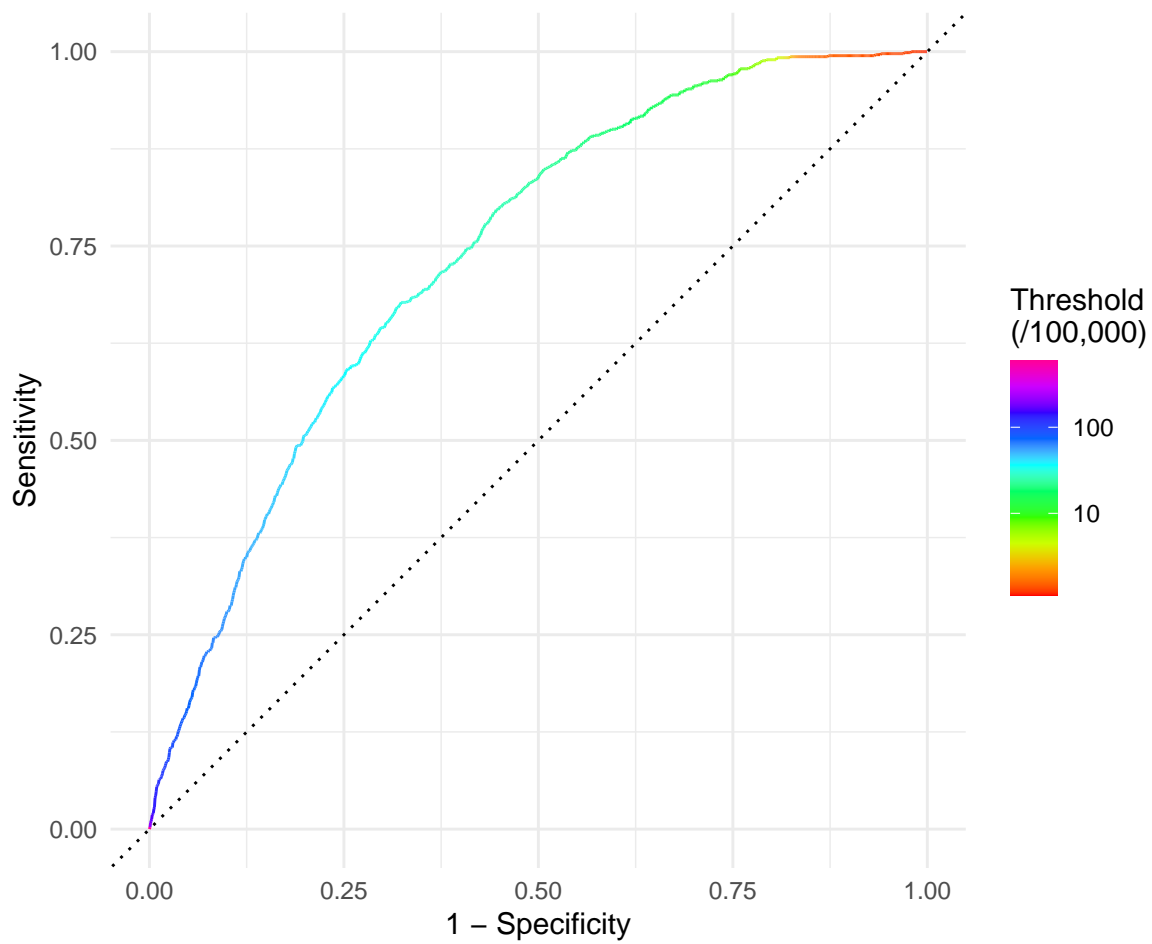
1 Overall performance



Cancer type: EAC
Dataset: test, all
Cases: 2076/2567069
AUC: 0.771 [0.762,0.779]

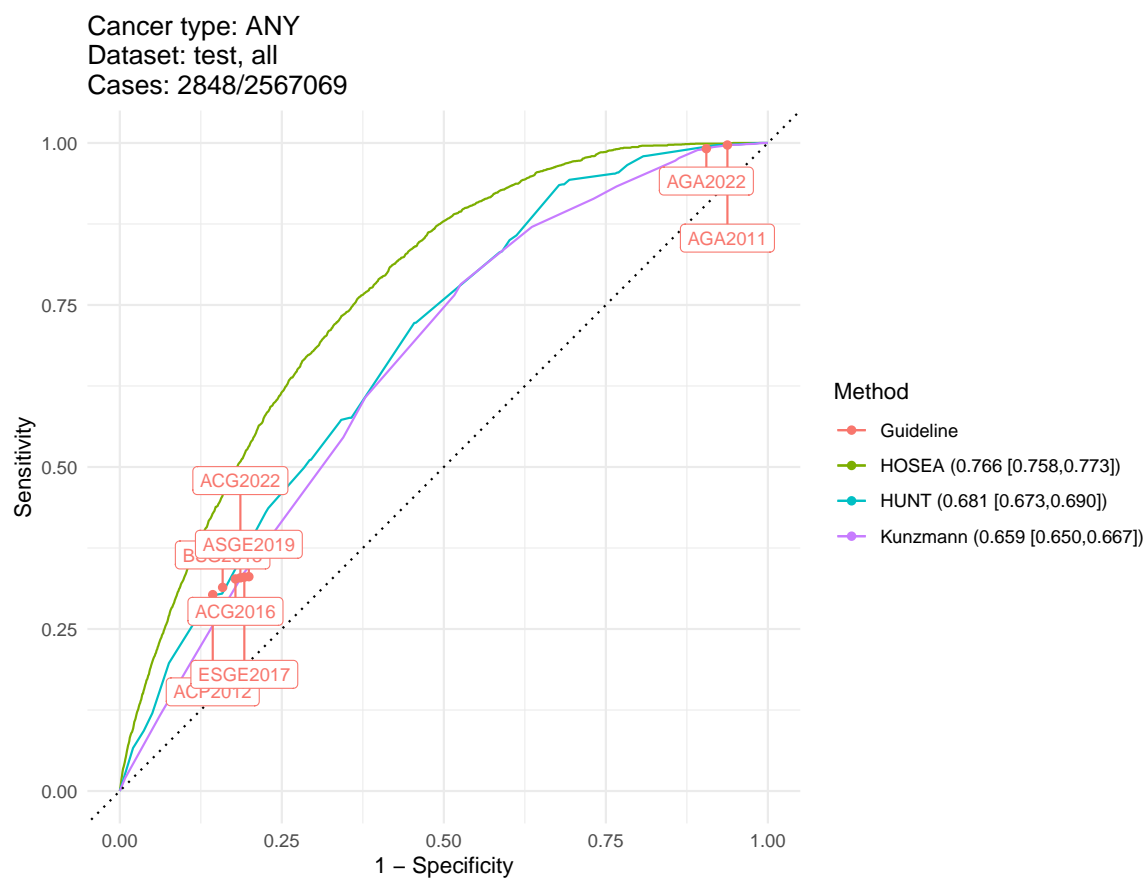


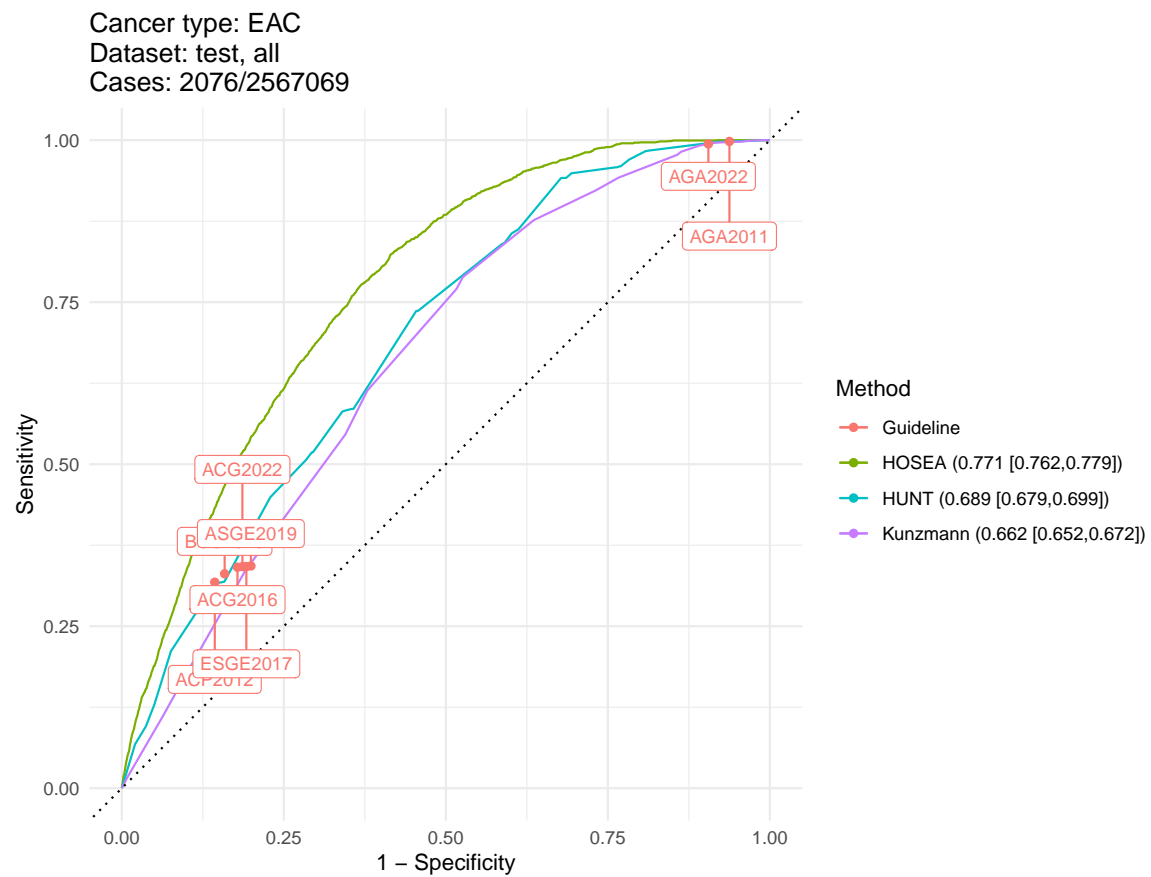
Cancer type: EGJAC
Dataset: test, all
Cases: 772/2567069
AUC: 0.739 [0.724,0.754]



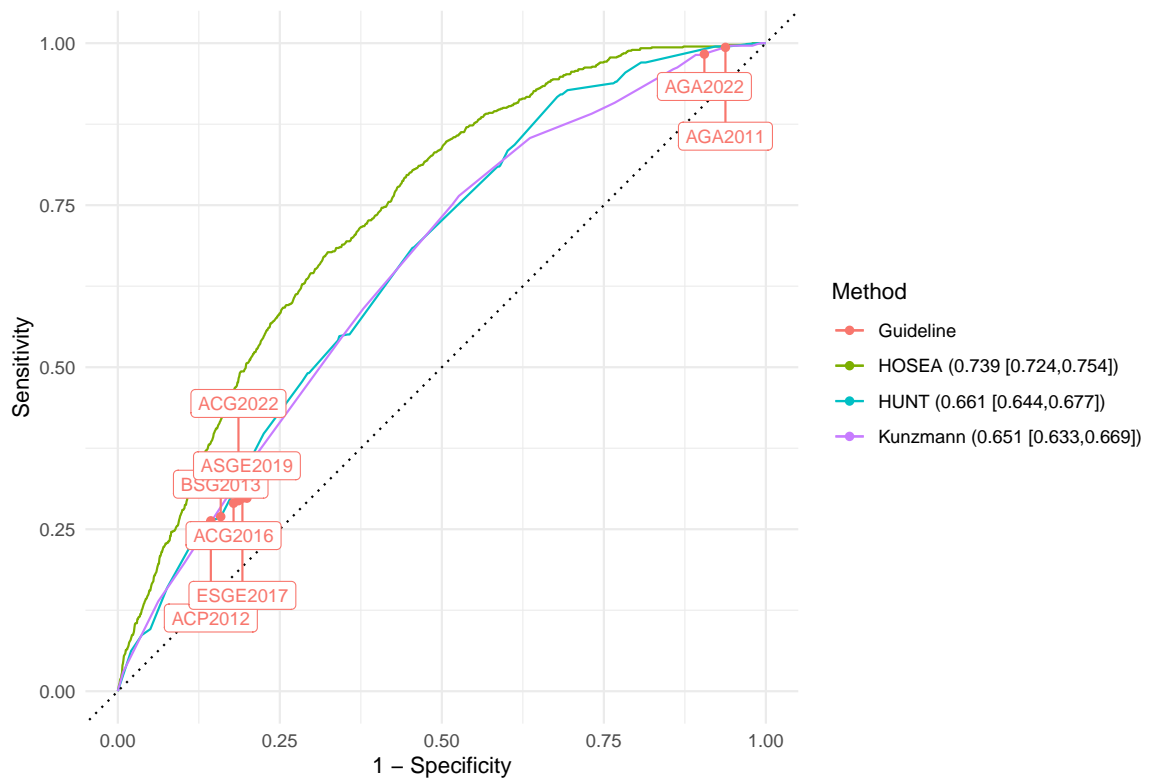
2 Comparison to HUNT, Kunzmann and Guidelines

2.1 Full test data



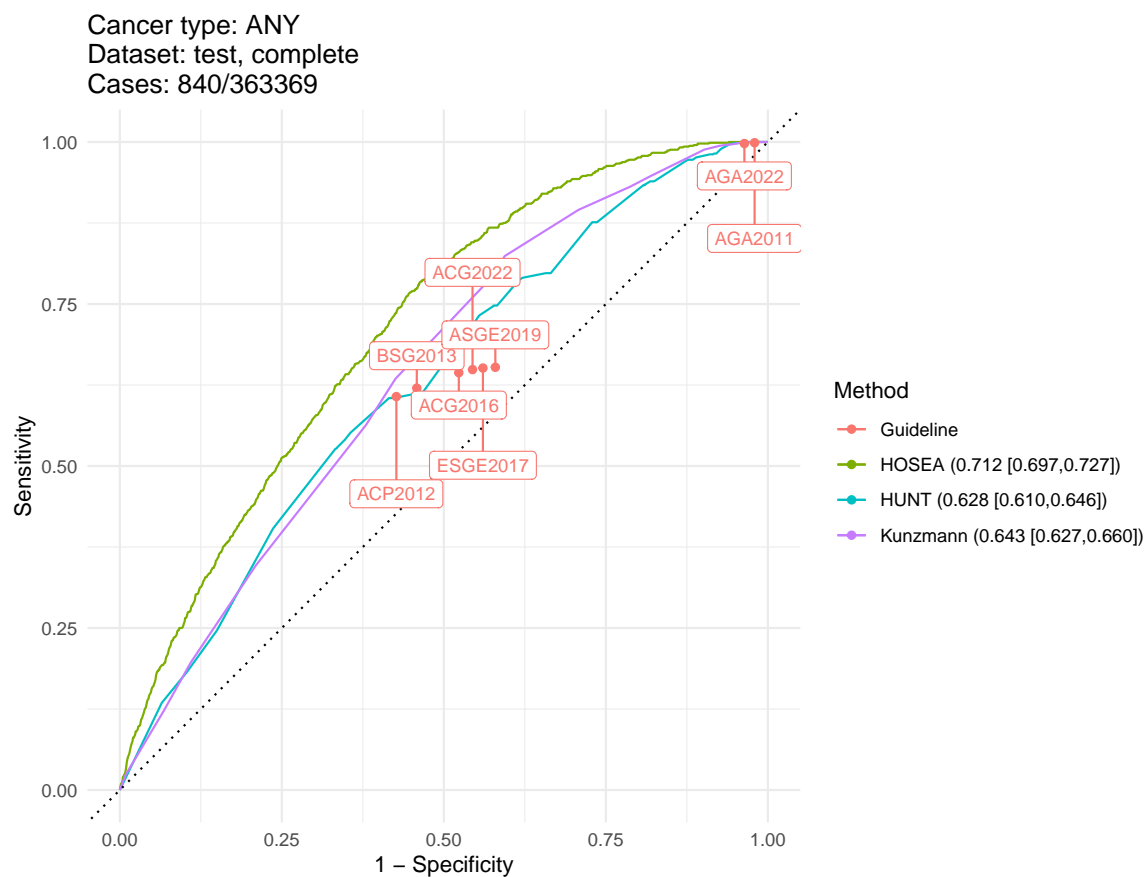


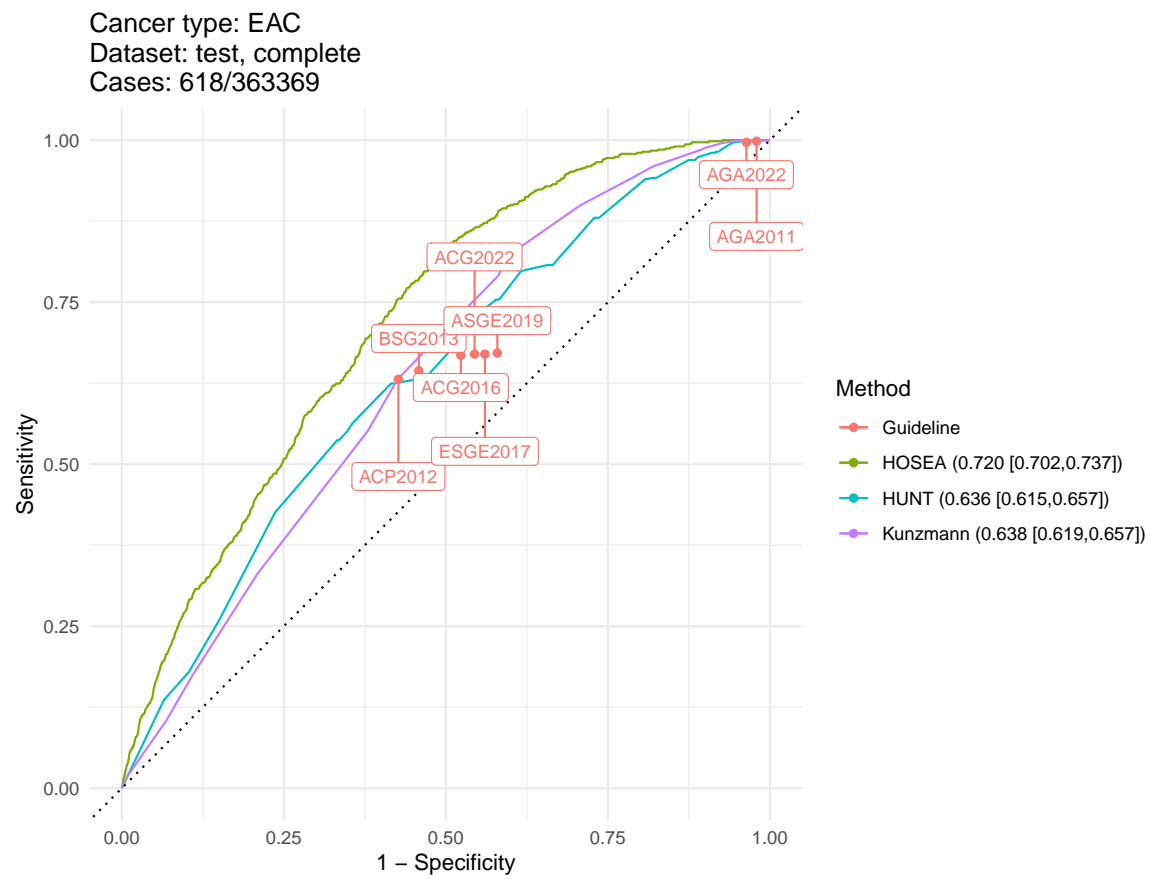
Cancer type: EGJAC
 Dataset: test, all
 Cases: 772/2567069



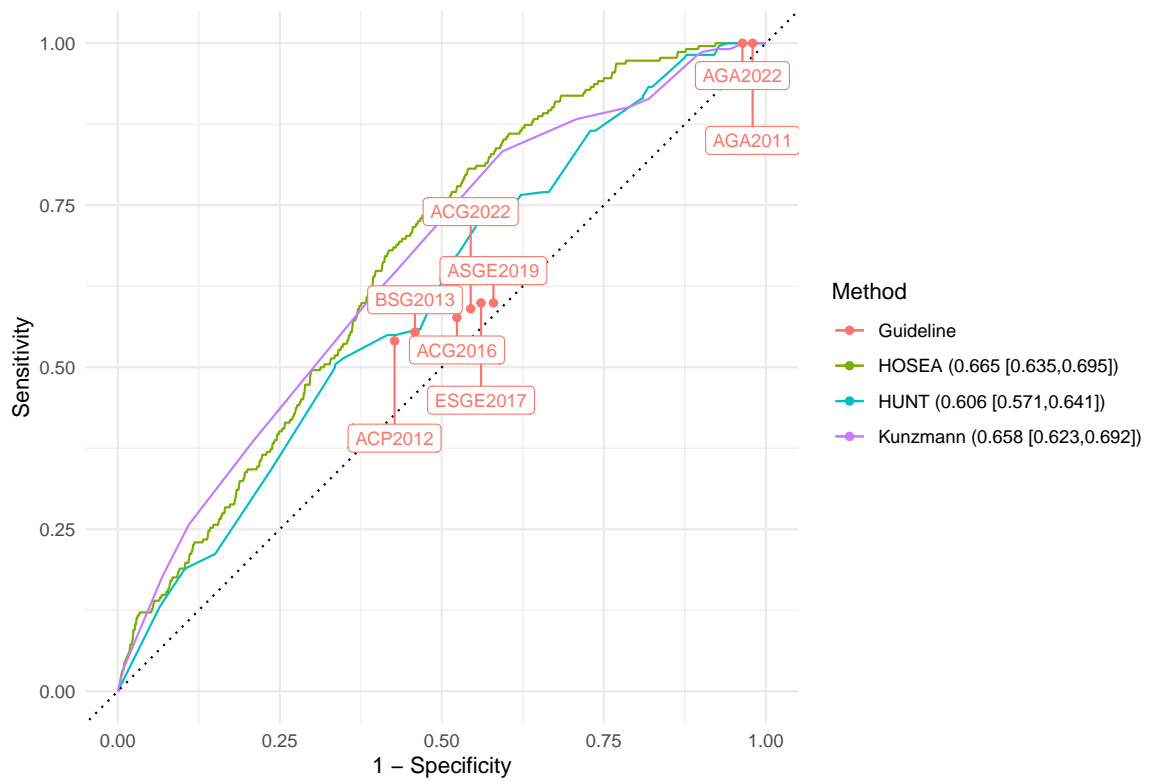
2.2 Complete test patients

w.r.t. *HUNT*, *Kunzmann*, & *Guidelines* (i.e., requires age, sex, GERD, bmi, smoking, etc.)



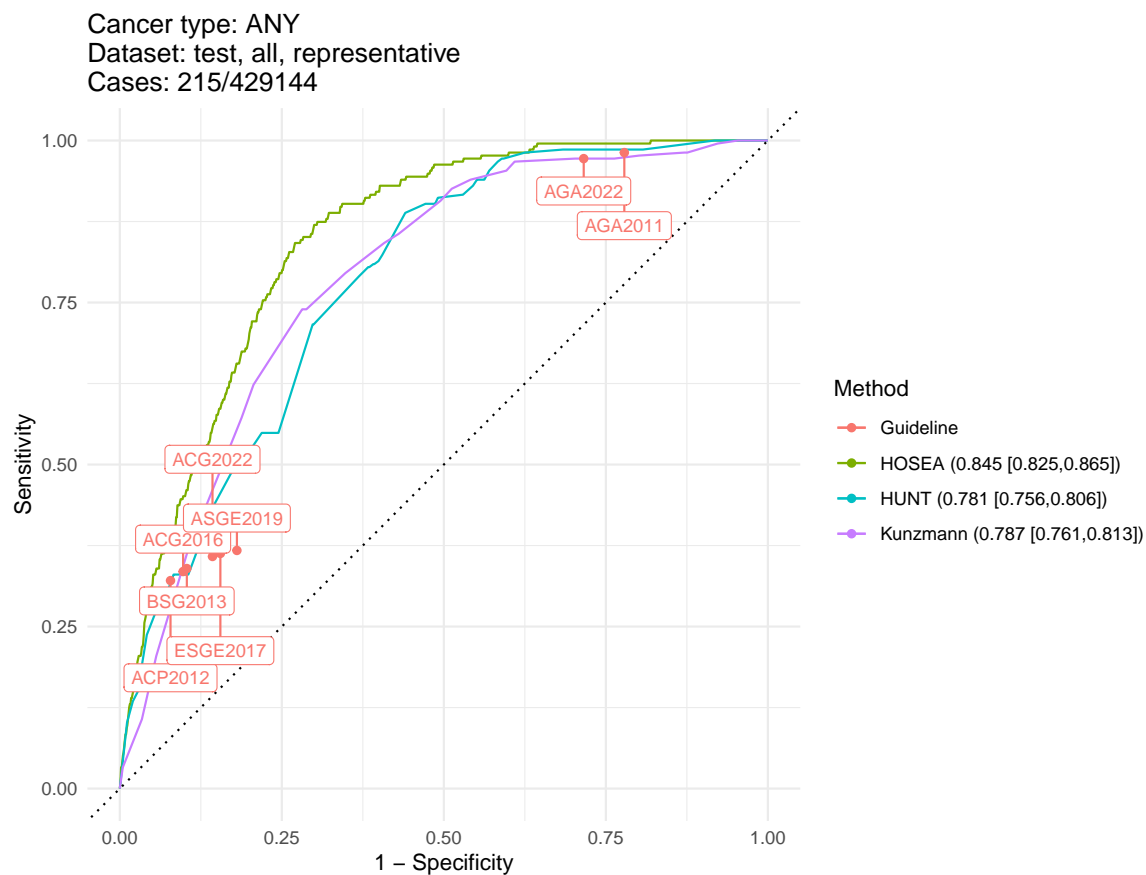


Cancer type: EGJAC
 Dataset: test, complete
 Cases: 222/363369

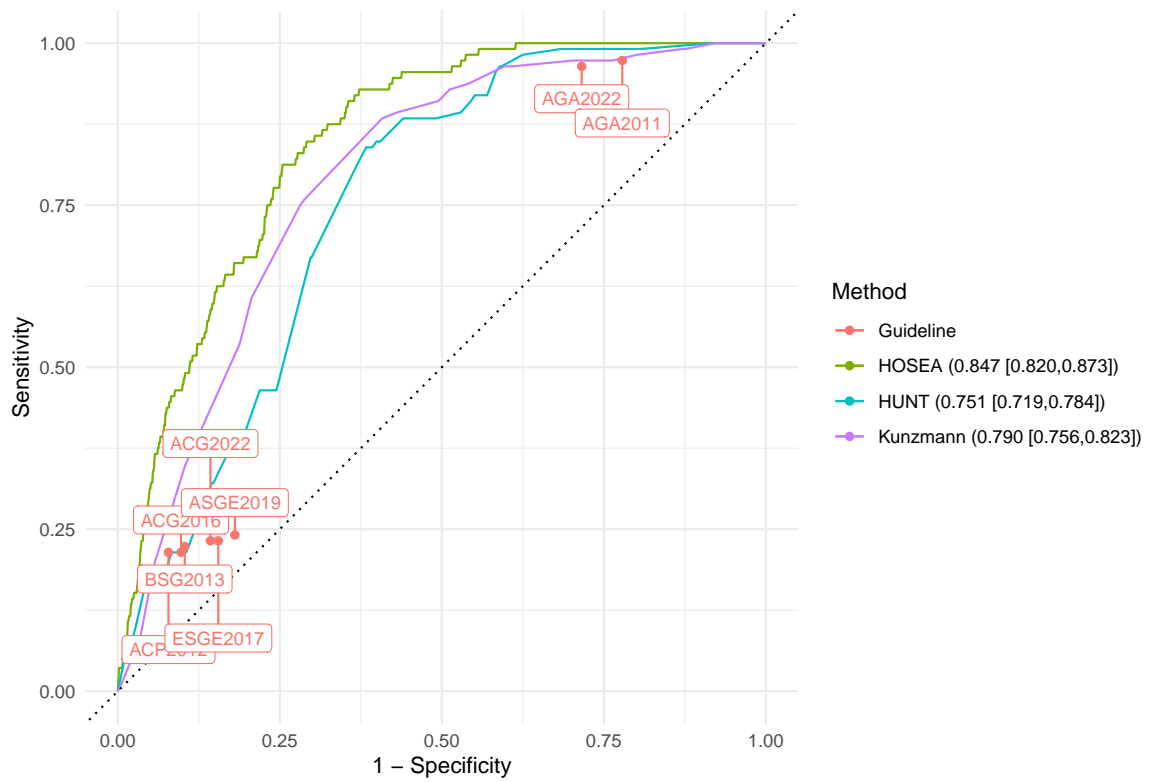


2.3 Representative sample

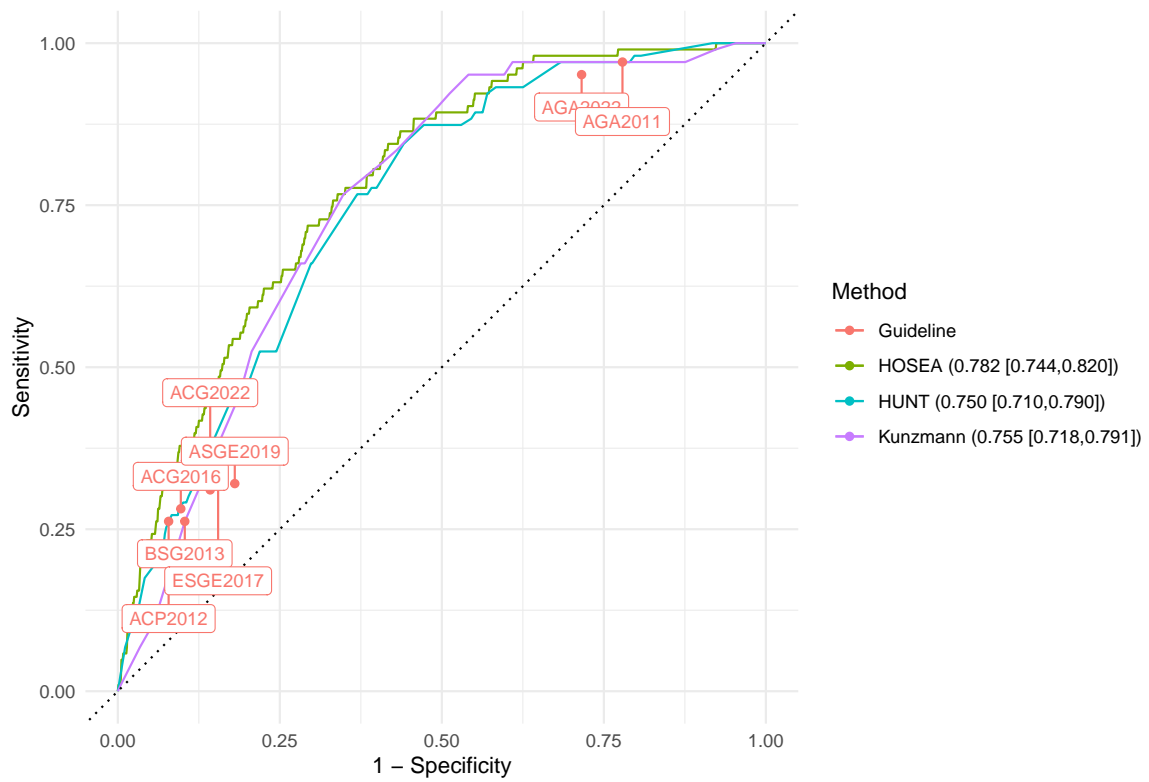
w.r.t. sex and prevalence ratio by sex



Cancer type: EAC
 Dataset: test, all, representative
 Cases: 112/429144

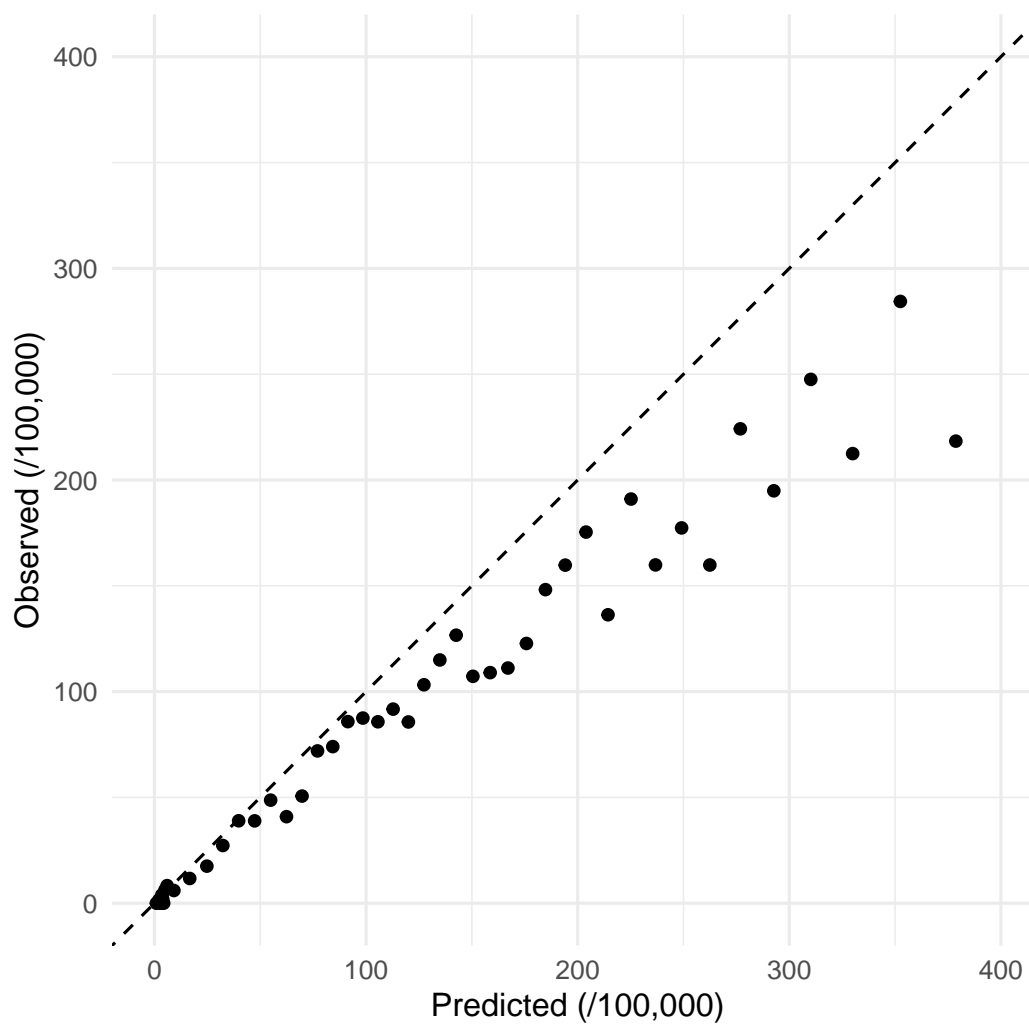


Cancer type: EGJAC
 Dataset: test, all, representative
 Cases: 103/429144

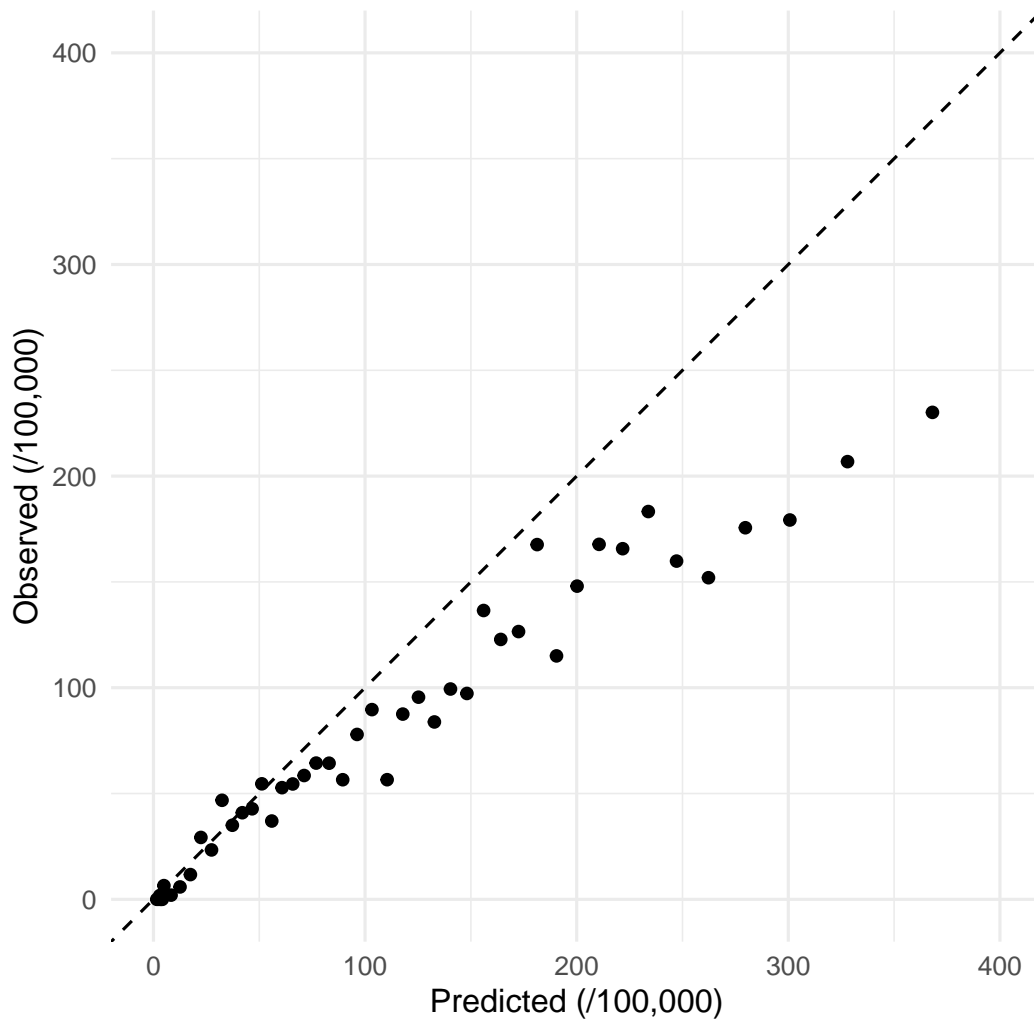


3 Calibration

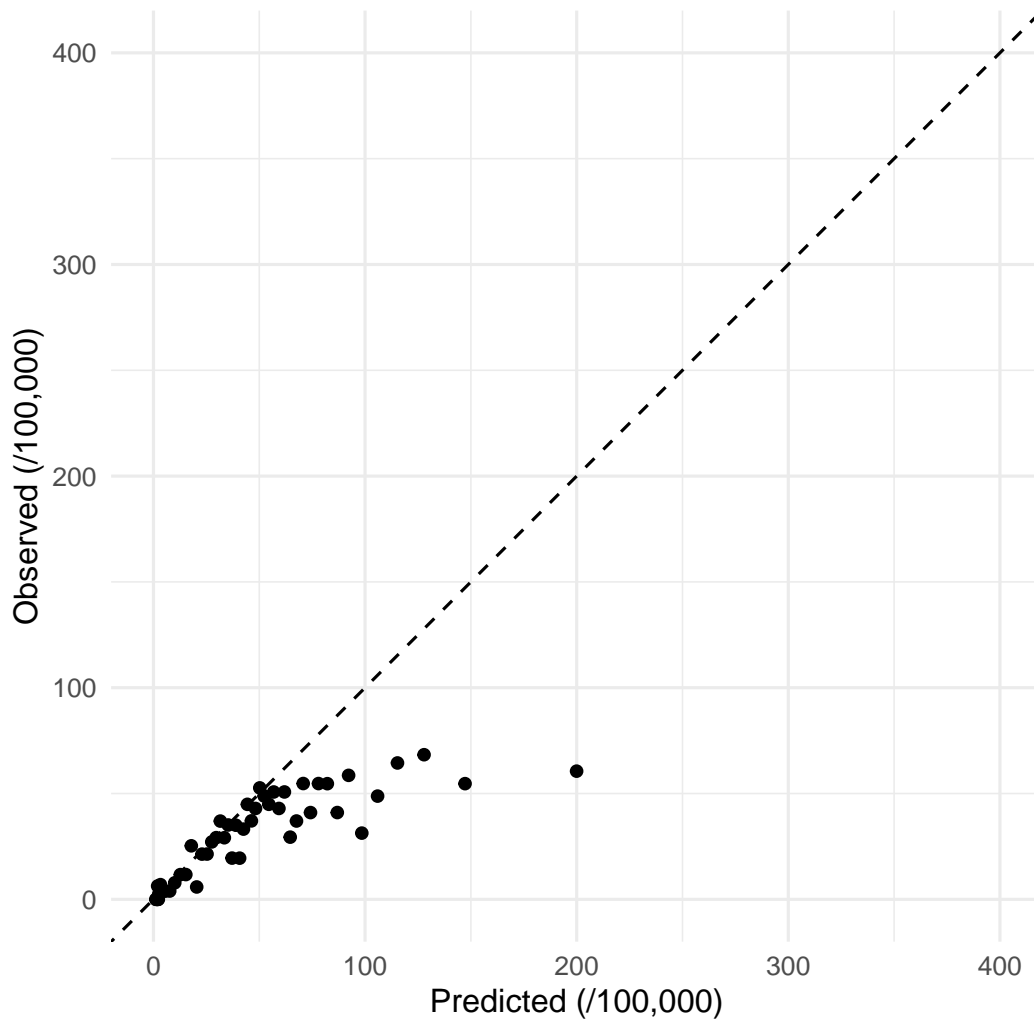
Cancer type: ANY
Dataset: test, all
Cases: 2848/2567069
HL: H=723.69, df=50, p=0



Cancer type: EAC
Dataset: test, all
Cases: 2076/2567069
HL: $H=334.91$, $df=50$, $p=0$



Cancer type: EGJAC
Dataset: test, all
Cases: 772/2567069
HL: $H=237.42$, $df=50$, $p=0$



4 Thresholds

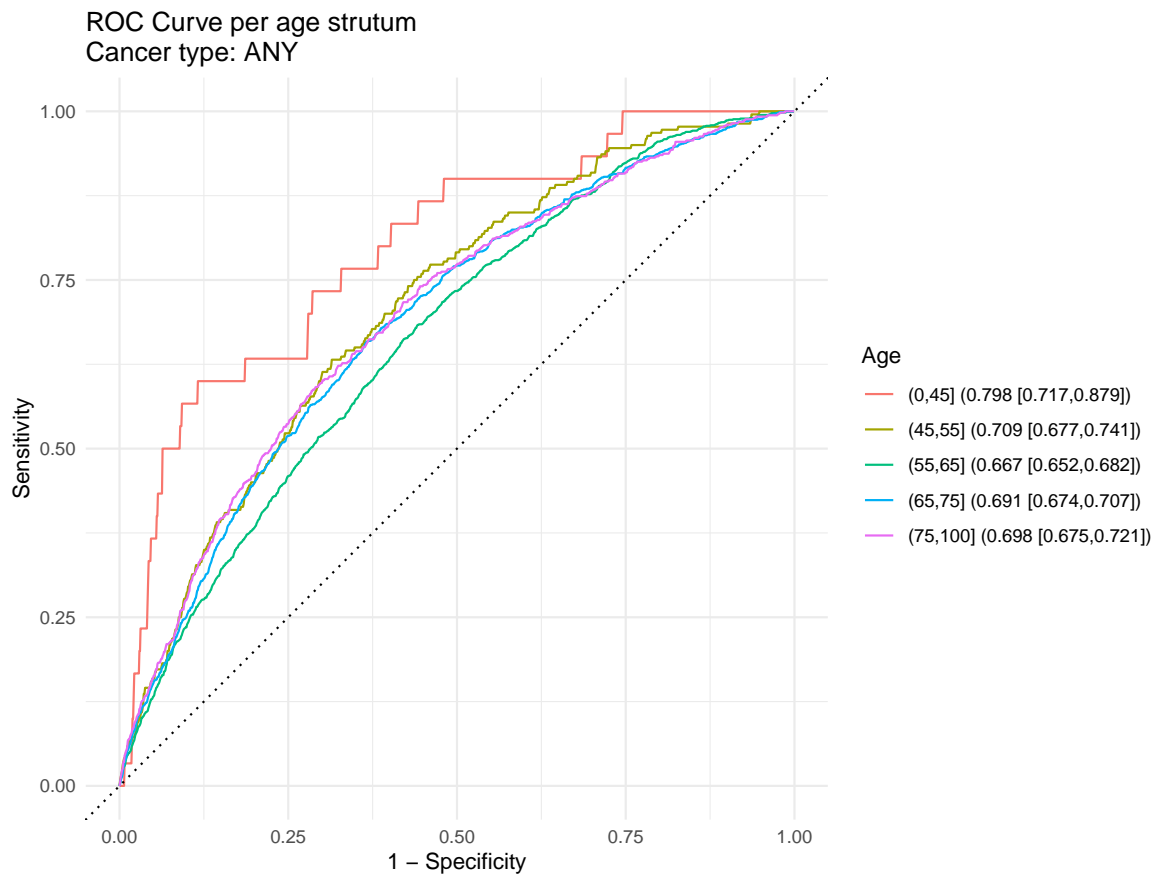
Threshold	TPR	PPV	DetPrevalence
0	100.00	0.11	100.00
5	99.82	0.13	86.56
10	99.58	0.13	82.62
15	99.47	0.14	81.37
20	99.33	0.14	80.19
25	99.23	0.14	78.92
30	98.95	0.14	77.60
35	98.56	0.14	76.25
40	98.17	0.15	74.92
45	97.58	0.15	73.59
50	97.12	0.15	72.28
55	96.73	0.15	70.95
60	96.03	0.15	69.62
65	95.54	0.16	68.27
70	94.87	0.16	66.90
75	94.24	0.16	65.54
80	93.40	0.16	64.17
85	92.45	0.16	62.77
90	91.43	0.17	61.36
95	90.38	0.17	59.96
100	89.29	0.17	58.54
105	88.06	0.17	57.13
110	86.97	0.17	55.75
115	85.74	0.18	54.35
120	84.83	0.18	52.97
125	83.67	0.18	51.61
130	82.44	0.18	50.27
135	81.21	0.18	48.94
140	79.56	0.19	47.63
145	78.34	0.19	46.36
150	76.97	0.19	45.09
155	75.88	0.19	43.86
160	74.72	0.19	42.65
165	73.38	0.20	41.45
170	72.12	0.20	40.28
175	71.17	0.20	39.15
180	69.77	0.20	38.03
185	68.47	0.21	36.93
190	66.89	0.21	35.84
195	65.59	0.21	34.79
200	63.87	0.21	33.77
220	58.53	0.22	29.92
240	53.02	0.22	26.44
260	47.93	0.23	23.34
280	43.19	0.23	20.57
300	38.80	0.24	18.11
320	34.16	0.24	15.96
340	30.51	0.24	14.05
360	26.26	0.24	12.38
380	22.96	0.23	10.88
400	20.44	0.24	9.58
420	18.36	0.24	8.43
440	16.71	0.25	7.42
460	14.99	0.24	6.54

Threshold	TPR	PPV	DetPrevalence
0	100.00	0.08	100.00
5	99.76	0.10	84.41
10	99.61	0.10	80.08
15	99.47	0.10	77.95
20	99.18	0.11	75.93
25	98.46	0.11	73.93
30	97.88	0.11	71.93
35	96.72	0.11	69.90
40	95.86	0.11	67.81
45	94.70	0.12	65.66
50	93.55	0.12	63.49
55	92.10	0.12	61.34
60	91.09	0.12	59.26
65	89.69	0.13	57.27
70	88.44	0.13	55.40
75	87.24	0.13	53.61
80	85.79	0.13	51.93
85	84.54	0.14	50.33
90	83.33	0.14	48.80
95	82.03	0.14	47.32
100	80.88	0.14	45.89
105	79.19	0.14	44.47
110	77.89	0.15	43.09
115	76.97	0.15	41.73
120	75.53	0.15	40.39
125	74.04	0.15	39.04
130	72.35	0.16	37.71
135	71.19	0.16	36.39
140	69.46	0.16	35.07
145	68.06	0.16	33.78
150	66.33	0.17	32.50
155	64.69	0.17	31.23
160	62.48	0.17	29.98
165	60.79	0.17	28.77
170	58.96	0.17	27.57
175	57.18	0.18	26.40
180	54.82	0.18	25.27
185	52.75	0.18	24.16
190	50.48	0.18	23.07
195	49.37	0.18	22.02
200	47.50	0.18	21.02
220	39.98	0.19	17.28
240	33.04	0.19	14.03
260	27.36	0.20	11.26
280	23.07	0.21	8.94
300	18.88	0.22	7.03
320	15.22	0.22	5.49
340	12.52	0.24	4.26
360	9.20	0.23	3.28
380	6.94	0.22	2.51
400	5.30	0.22	1.91
420	3.81	0.21	1.45
440	2.94	0.22	1.10
460	2.33	0.23	0.83

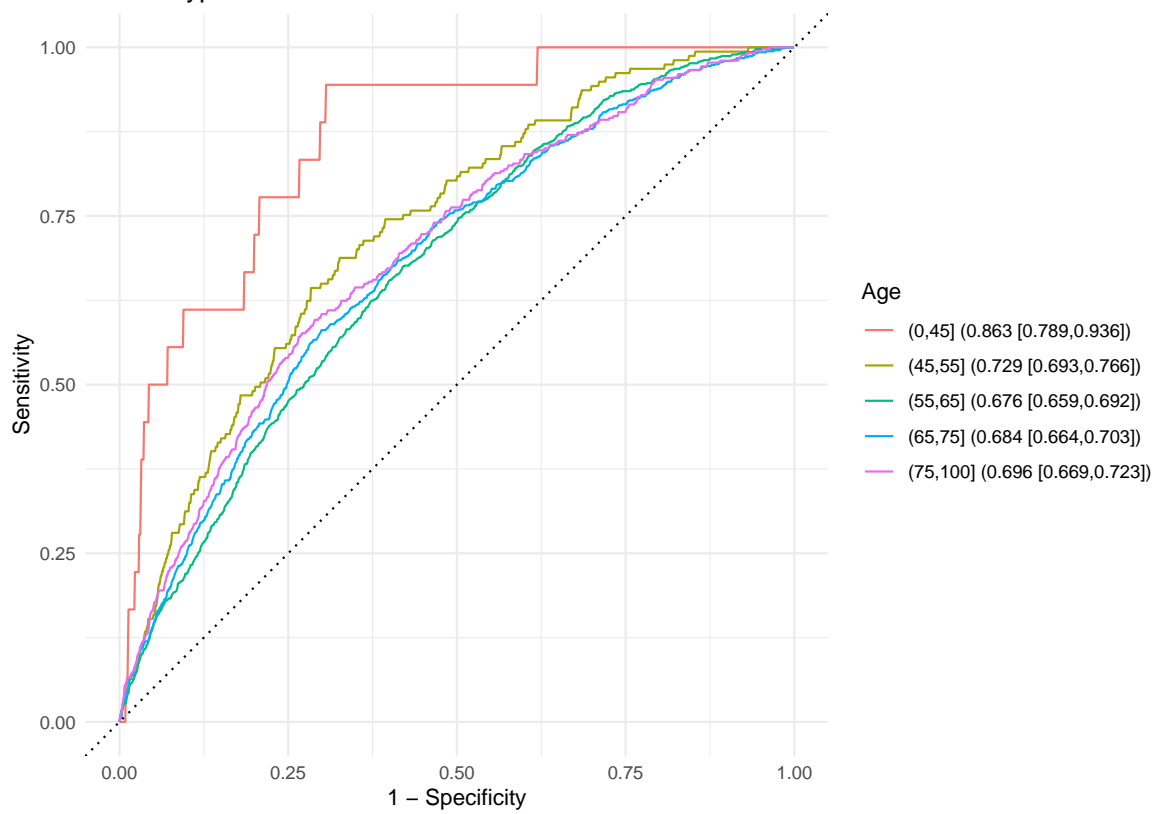
Threshold	TPR	PPV	DetPrevalence
0	100.00	0.03	100.00
5	99.09	0.04	81.40
10	98.19	0.04	76.96
15	96.89	0.04	73.16
20	94.56	0.04	69.26
25	92.62	0.04	65.19
30	88.73	0.04	60.53
35	82.90	0.05	55.25
40	77.46	0.05	49.69
45	71.63	0.05	44.23
50	64.25	0.05	39.13
55	57.38	0.05	34.49
60	50.39	0.05	30.36
65	44.82	0.05	26.67
70	41.19	0.05	23.40
75	36.27	0.05	20.51
80	31.99	0.05	17.95
85	27.72	0.05	15.72
90	25.52	0.06	13.75
95	22.02	0.06	12.03
100	20.08	0.06	10.51
105	18.39	0.06	9.18
110	16.71	0.06	8.03
115	14.25	0.06	7.01
120	12.44	0.06	6.13
125	10.75	0.06	5.36
130	9.33	0.06	4.67
135	7.77	0.06	4.09
140	6.61	0.06	3.57
145	5.83	0.06	3.12
150	4.92	0.05	2.72
155	4.53	0.06	2.38
160	4.27	0.06	2.08
165	3.63	0.06	1.82
170	2.85	0.05	1.59
175	2.59	0.06	1.39
180	2.59	0.06	1.22
185	2.20	0.06	1.07
190	1.94	0.06	0.93
195	1.81	0.07	0.82
200	1.42	0.06	0.72
220	0.52	0.04	0.42
240	0.26	0.03	0.26
260	0.13	0.02	0.16
280	0.13	0.04	0.10
300	0.13	0.06	0.06
320	0.13	0.10	0.04
340	0.00	0.00	0.02
360	0.00	0.00	0.02
380	0.00	0.00	0.01
400	0.00	0.00	0.01
420	0.00	0.00	0.00
440	0.00	0.00	0.00
460	0.00	0.00	0.00

5 Stratification by identity groups

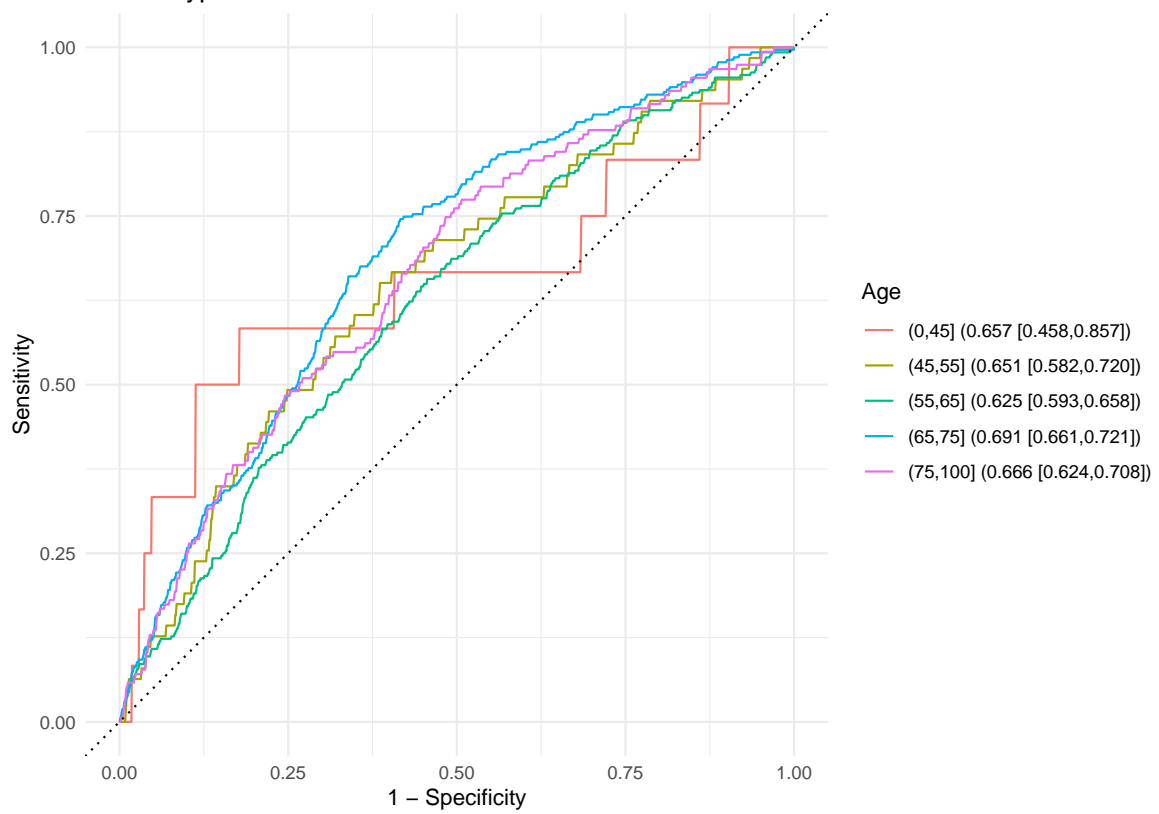
5.1 Age



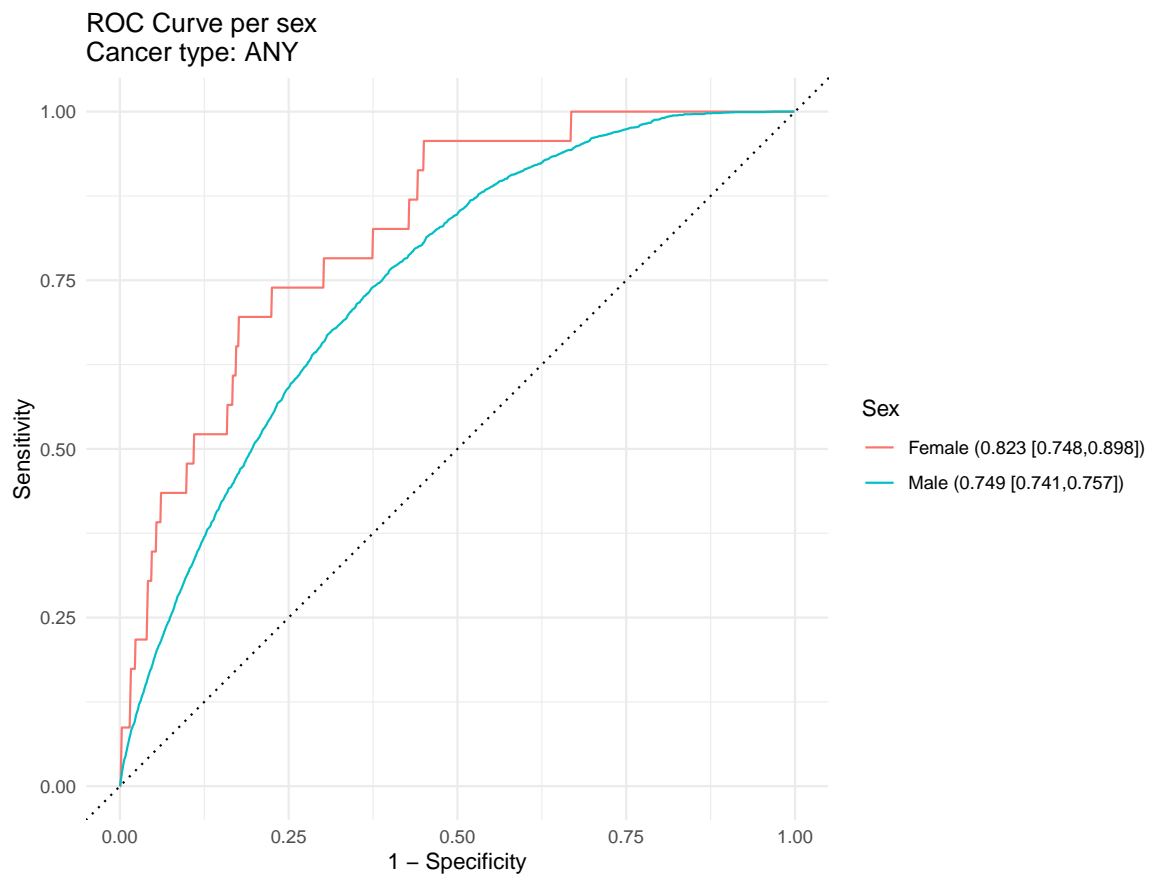
ROC Curve per age stratum
Cancer type: EAC

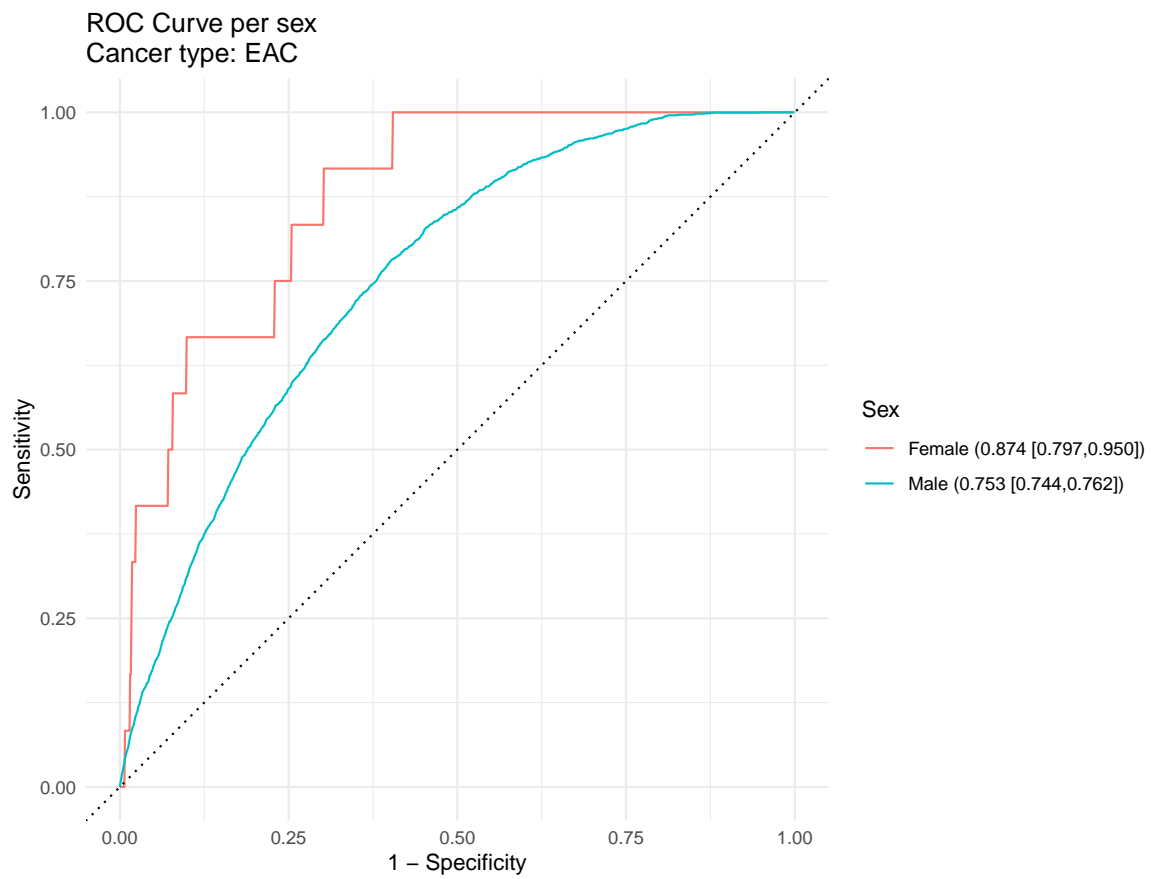


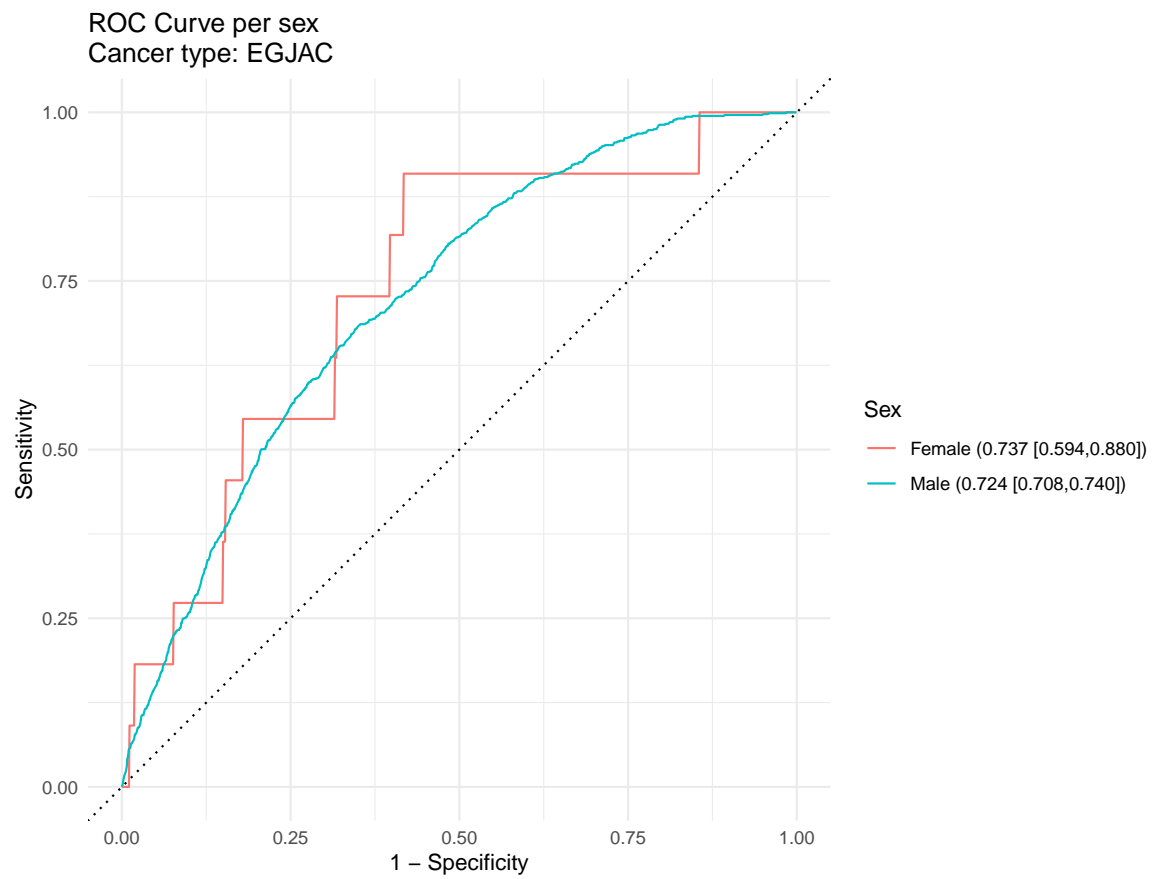
ROC Curve per age stratum
Cancer type: EGJAC



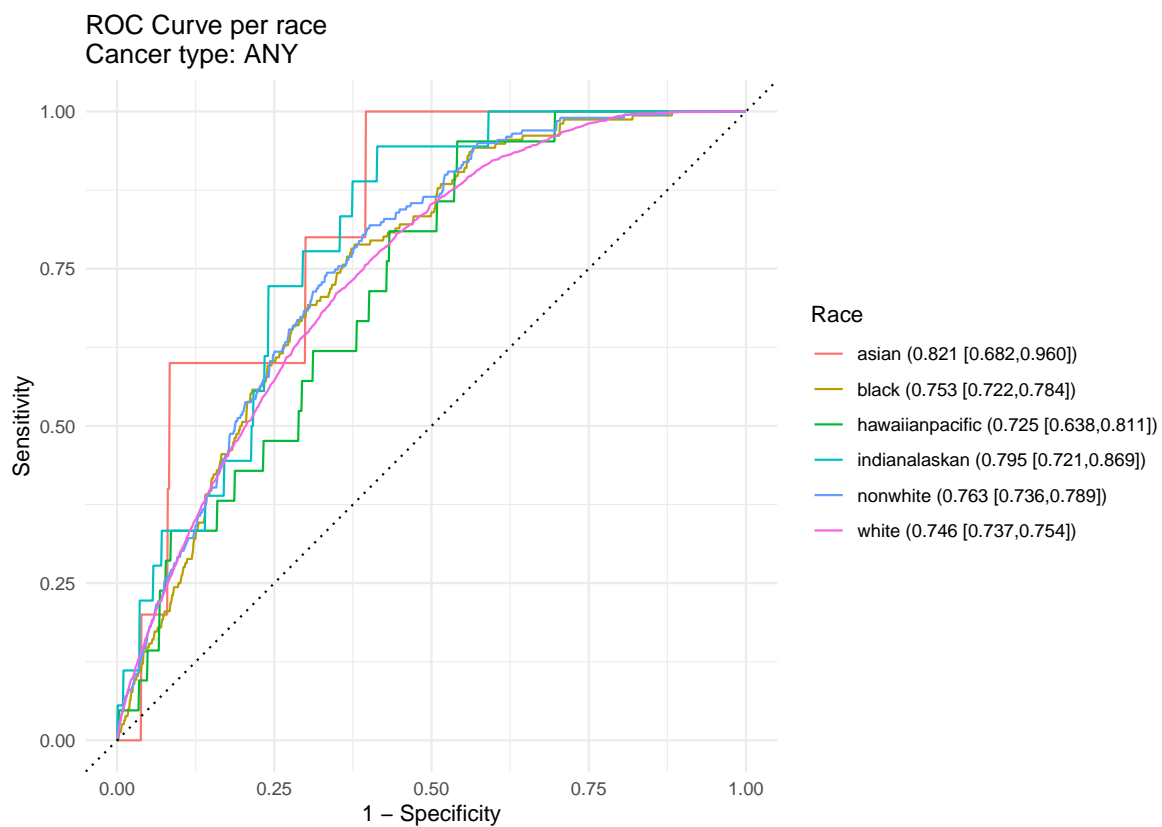
5.2 Sex

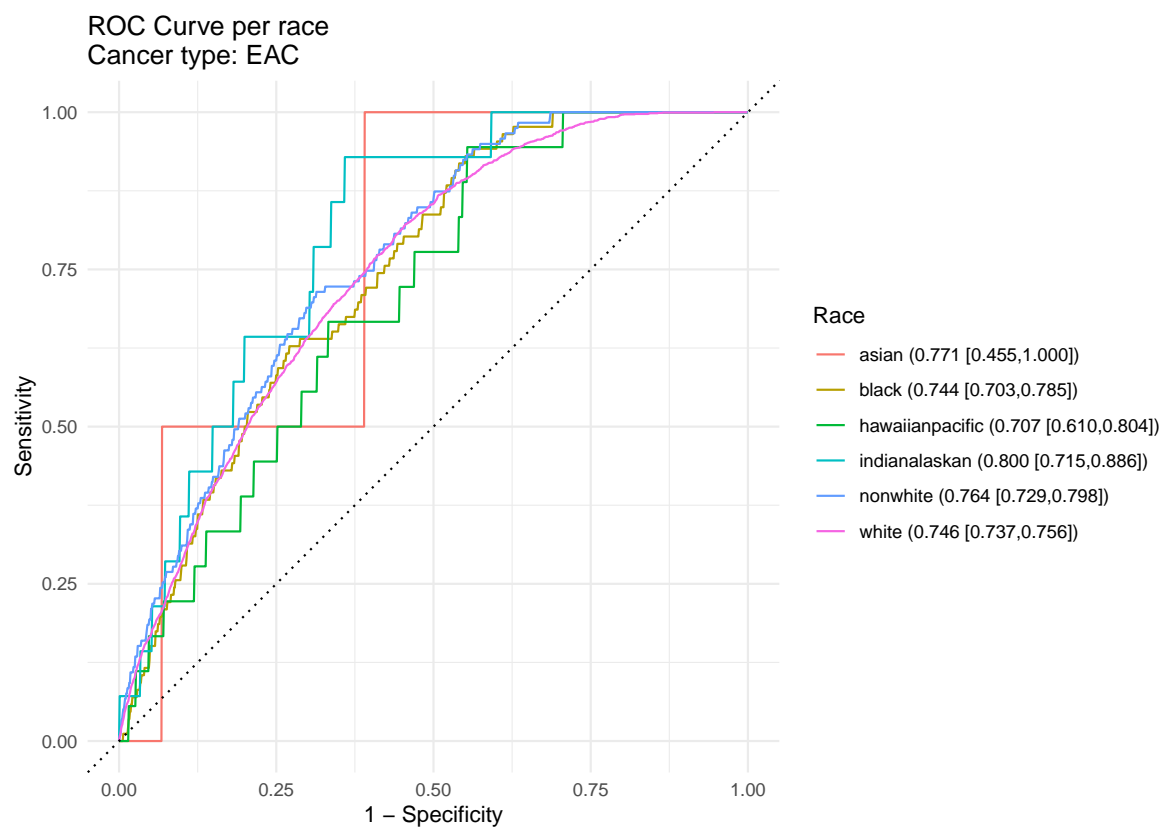


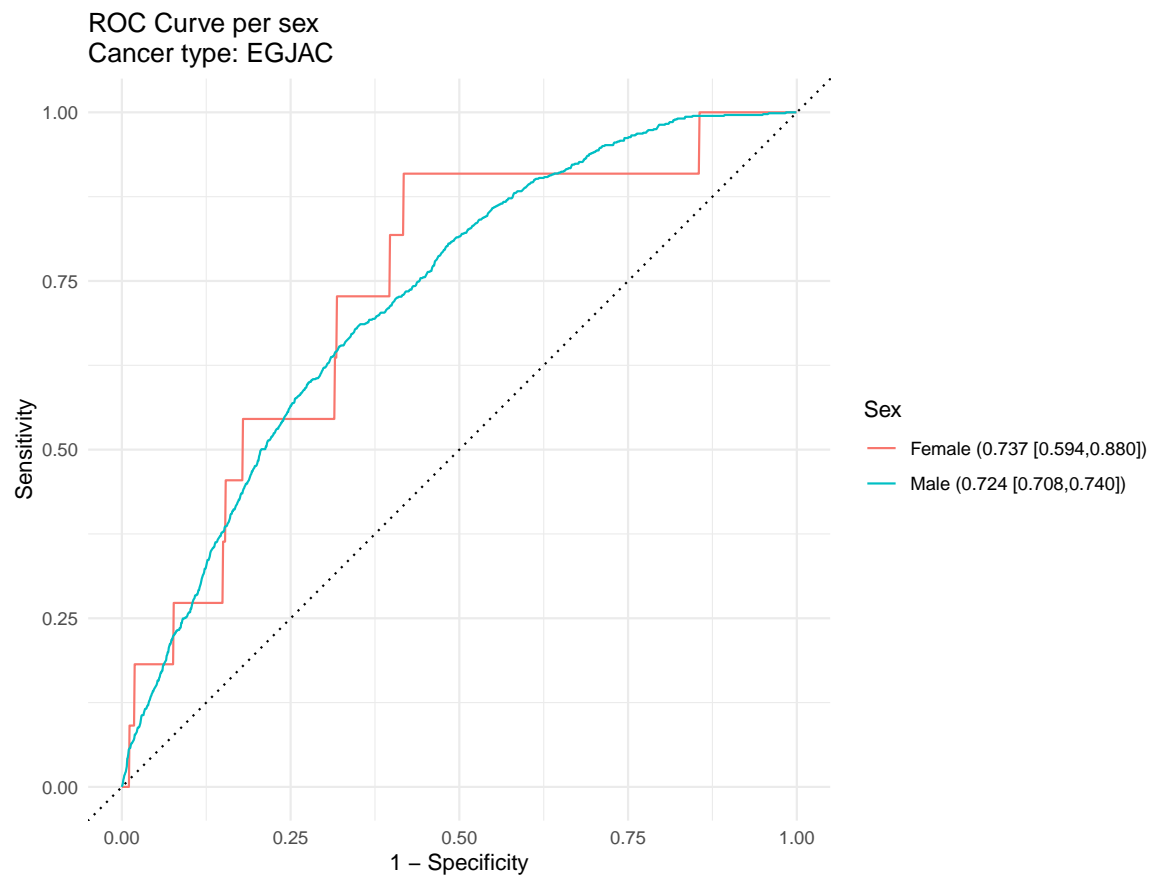




5.3 Race







6 Cancer stage

NB: the discrepancy in number of cases between I+ and Any comes from the fact that some were classified as unknown stage according to the provided staging.

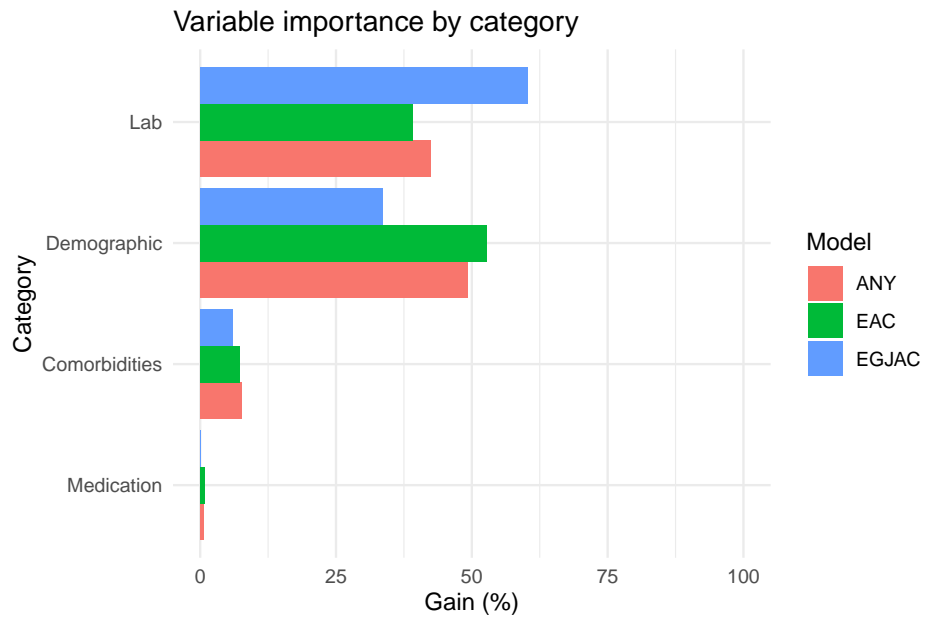
Stage	Test. AUC	Nb. cases
Any	0.765 [0.758,0.773]	2818
I	0.799 [0.782,0.817]	495
II	0.776 [0.757,0.795]	435
III	0.769 [0.756,0.782]	878
IV	0.759 [0.747,0.770]	1244
I+	0.768 [0.760,0.776]	2448
II+	0.762 [0.754,0.770]	2169
III+	0.760 [0.752,0.769]	1950
IV+	0.759 [0.747,0.770]	1244

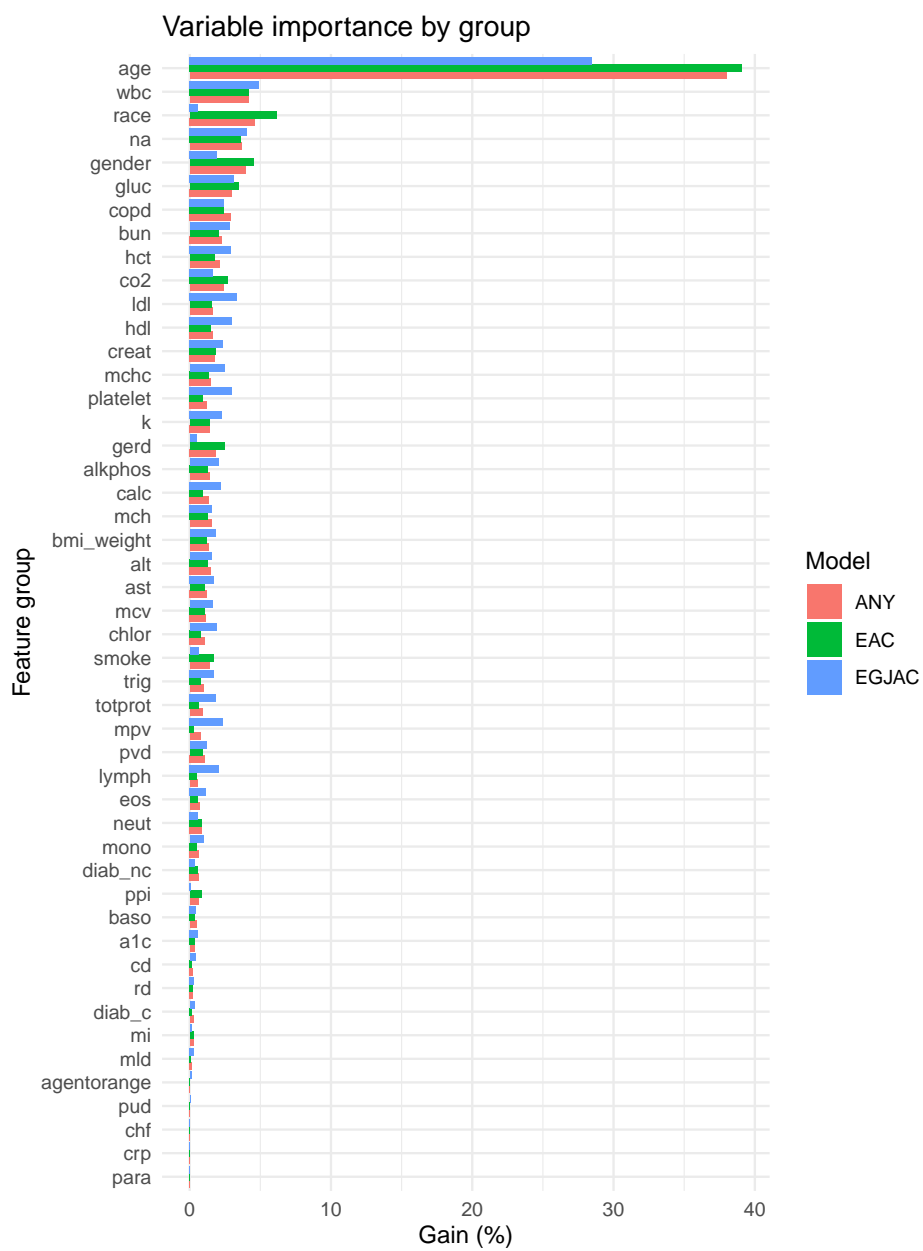
Stage	Test. AUC	Nb. cases
Any	0.771 [0.762,0.779]	2054
I	0.810 [0.791,0.828]	350
II	0.784 [0.763,0.806]	302
III	0.770 [0.756,0.785]	650
IV	0.763 [0.750,0.775]	913
I+	0.771 [0.762,0.780]	1813
II+	0.765 [0.755,0.774]	1608
III+	0.763 [0.753,0.773]	1451
IV+	0.763 [0.750,0.775]	913

Stage	Test. AUC	Nb. cases
Any	0.738 [0.723,0.753]	764
I	0.765 [0.729,0.801]	145
II	0.748 [0.710,0.786]	133
III	0.751 [0.723,0.778]	228
IV	0.727 [0.703,0.750]	331
I+	0.742 [0.725,0.758]	635
II+	0.736 [0.718,0.754]	561
III+	0.733 [0.714,0.752]	499
IV+	0.727 [0.703,0.750]	331

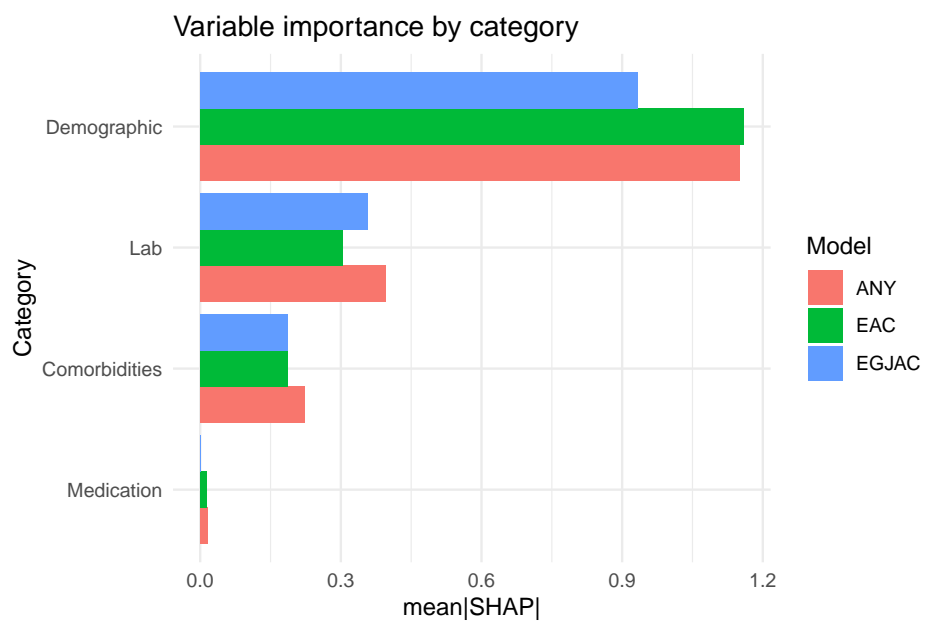
7 Variable importance

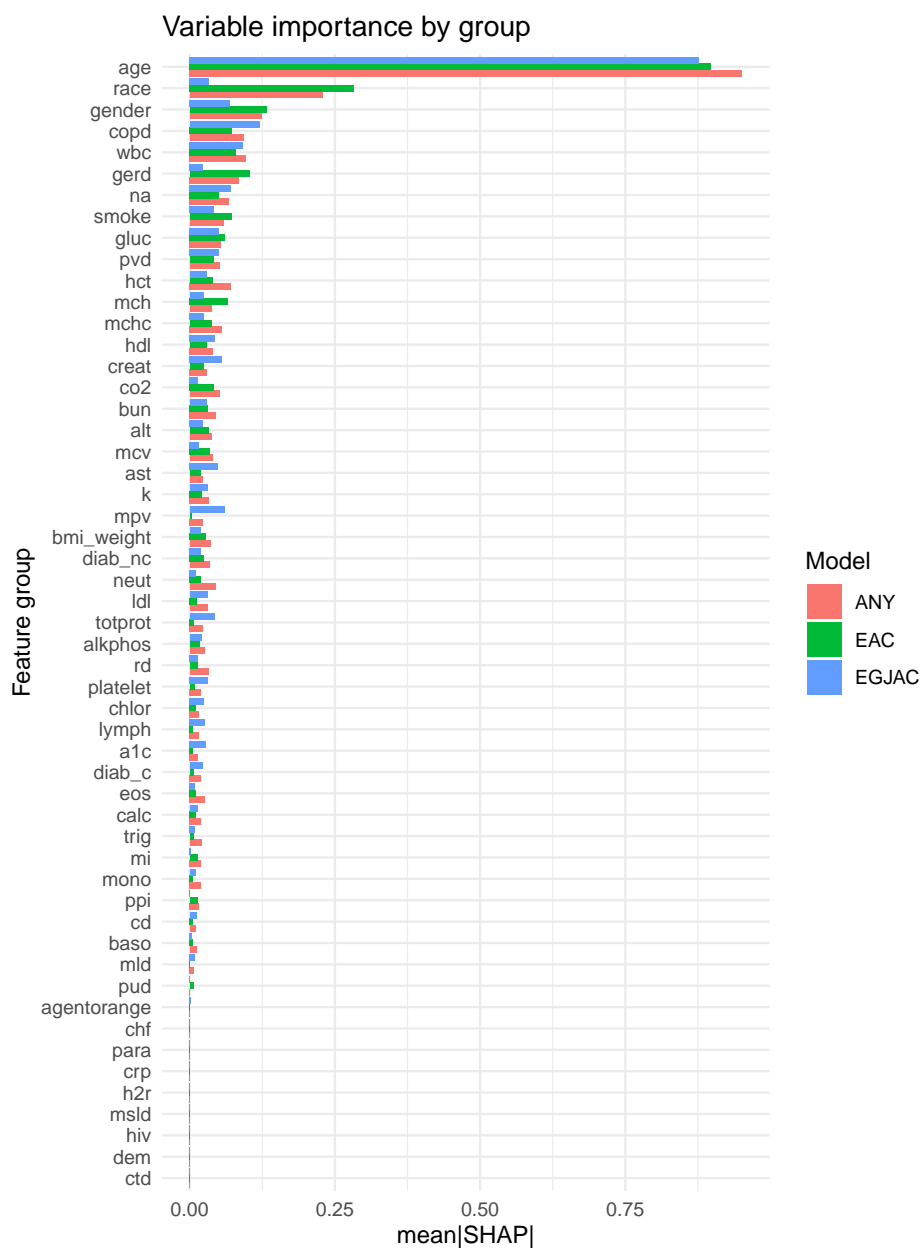
7.1 Gain VI





7.2 SHAP VI





8 Years prior

