# A comparative study of survival models for breast cancer prognostication based on microarray data

B. Haibe-Kains[1,2]

[1]Functional Genomics Unit, Institut Jules Bordet
[2]Machine Learning Group, Université Libre de Bruxelles

April 21, 2008

# Contents

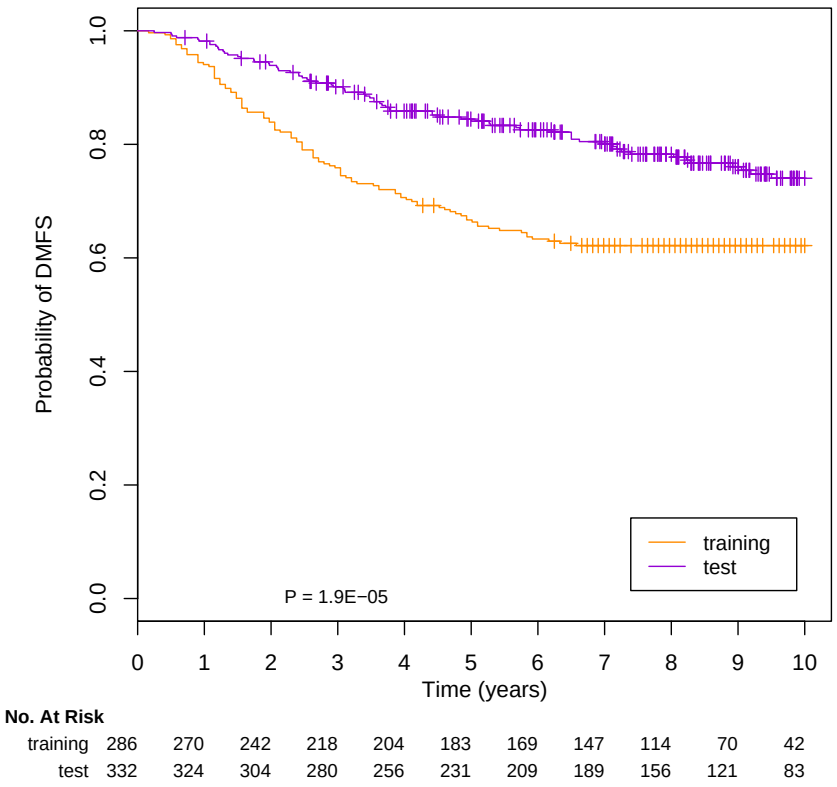

**1 From TRAINING to TEST** **3**

1.1 Survival Curves for TRAINING and TEST Sets . . . 3
1.2 AURKA Alone . . . 3
1.3 BD.COMBUNIV.WILCOXON.HG . . . 6
1.4 BD.COMBUNIV.COX.SURV . . . 9
1.5 BD.MULTIV.LM.TOE . . . 12
1.6 BD.MULTIV.COX.SURV . . . 15
1.7 GW.RANK.COMBUNIV.WILCOXON.HG . . . 18
1.8 GW.RANKCV.COMBUNIV.WILCOXON.HG . . . 22
1.9 GW.RANK.COMBUNIV.COX.SURV . . . 26
1.10 GW.RANKCV.COMBUNIV.COX.SURV . . . 30
1.11 GW.RANK.MULTIV.RCOX.SURV . . . 34
1.12 GW.RANKCV.MULTIV.RCOX.SURV . . . 38
1.13 GW.PCA.COMBUNIV.WILCOXON.HG . . . 42
1.14 GW.PCACV.COMBUNIV.WILCOXON.HG . . . 45
1.15 GW.PCA.COMBUNIV.COX.SURV . . . 49
1.16 GW.PCACV.COMBUNIV.COX.SURV . . . 52
1.17 GW.PCA.MULTIV.RCOX.SURV . . . 56
1.18 GW.PCACV.MULTIV.RCOX.SURV . . . 59
1.19 GENE76 . . . 63
1.20 GGI . . . 66
1.21 Model Comparison . . . 69
1.21.1 Training Set . . . 69
1.21.2 Test Set . . . 90

TRAINING set = VDX and TEST set = TBG.

# 1 From TRAINING to TEST

## 1.1 Survival Curves for TRAINING and TEST Sets



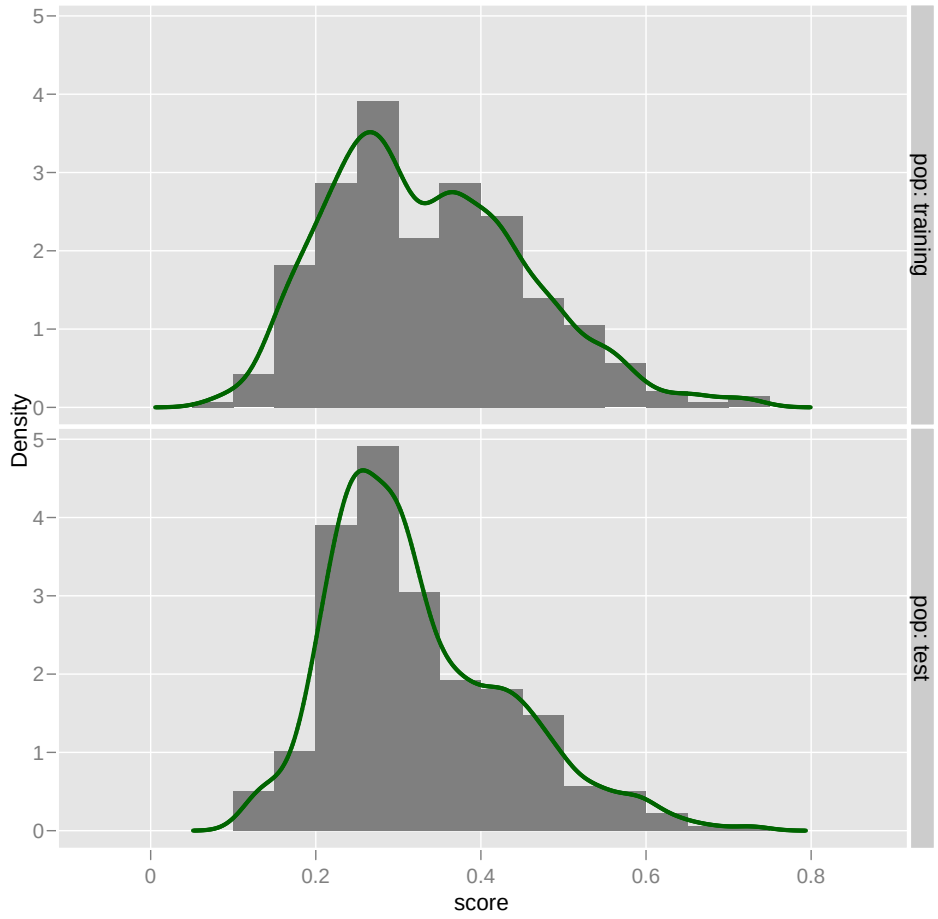

**No. At Risk**

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| training | 286 | 270 | 242 | 218 | 204 | 183 | 169 | 147 | 114 | 70 | 42 |
| test | 332 | 324 | 304 | 280 | 256 | 231 | 209 | 189 | 156 | 121 | 83 |

As we can see, there is a significant difference between the two datasets. So we expect a bias in the prediction of survival probabilities from the TRAINING and the TEST sets. This bias will results in large prediction errors as defined by the Brier score. However, this should not have a large impact on the concordance index, the time-dependent ROC curve and the hazard ratio estimates.

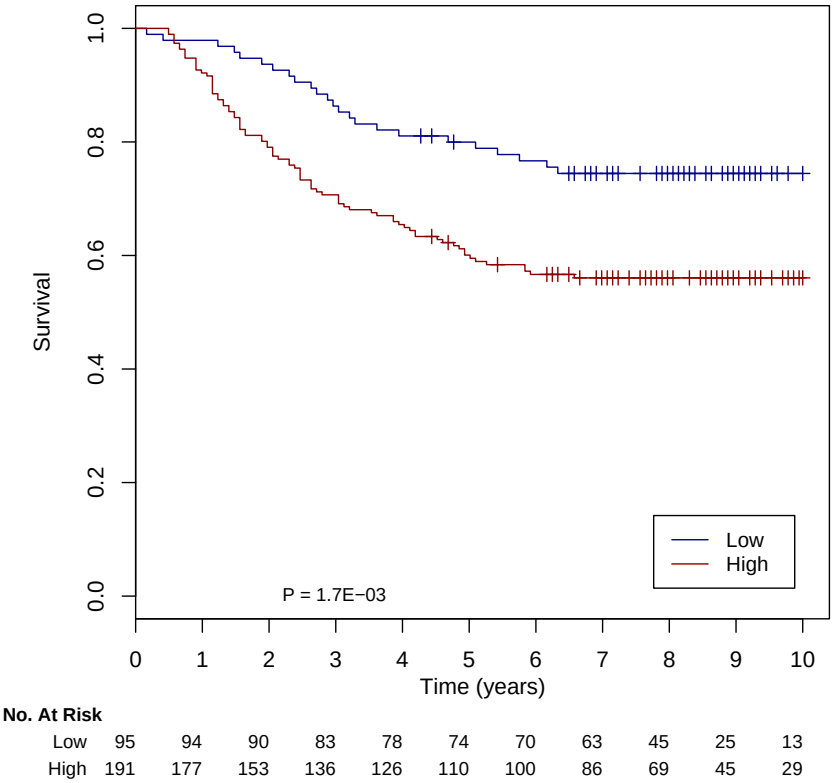
## 1.2 AURKA Alone

The following figure shows the distribution of the ggi :




**Risk Score** On TRAINING, the risk score exhibits a Cox p-value of 1.05E-05, a concordance index of 0.636 95CI[0.583,0.69] (p-value of 3.03E-07) and an integrated Brier score of 0.178.

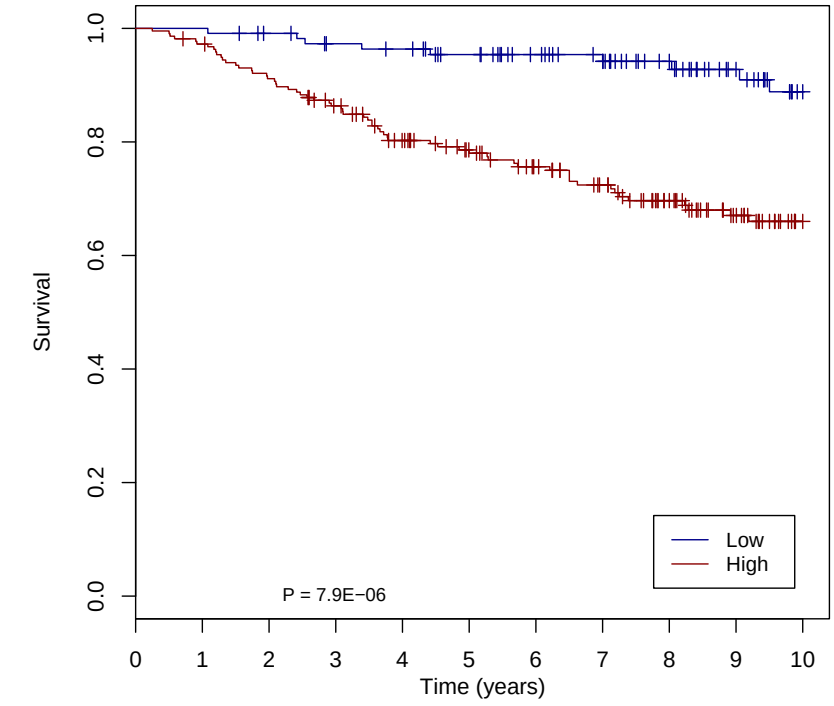
**Risk Group** On TRAINING, the binary classification computed from the risk score exhibits a Cox p-value of 1.10E-03, a concordance index of 0.685 95CI[0.587,0.783] (p-value of 1.13E-04) and an integrated Brier score of 0.182. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.85 [0.78,0.93] | 0.79 [0.71,0.88] | 0.74 [0.66,0.84] |
| High | 0.69 [0.63,0.76] | 0.60 [0.53,0.67] | 0.56 [0.49,0.64] |

**Risk Score** On TEST, the risk score exhibits a Cox p-value of 1.47E-05, a concordance index of 0.683 95CI[0.626,0.74] (p-value of 1.79E-10) and an integrated Brier score of 0.132.

**Risk Group** On TEST, the binary classification computed from the risk score exhibits a Cox p-value of 3.78E-07, a concordance index of 0.834 95CI[0.732,0.937] (p-value of 7.36E-11) and an integrated Brier score of 0.133. The following figure shows the Kaplan-Meier survival curves for the two groups :



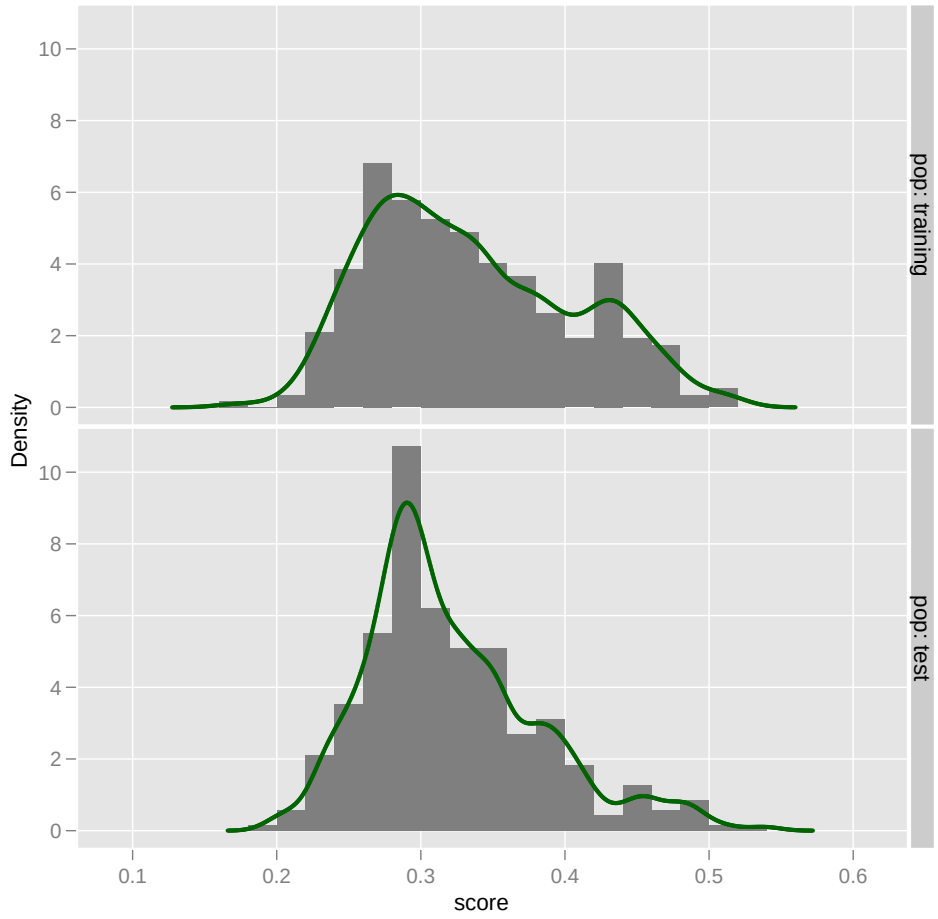

**No. At Risk**

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Low | 114 | 114 | 111 | 105 | 103 | 96 | 87 | 80 | 67 | 52 | 35 |
| High | 218 | 210 | 194 | 176 | 154 | 136 | 123 | 110 | 89 | 70 | 48 |

| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.96 [0.93,1.00] | 0.95 [0.92,0.99] | 0.89 [0.82,0.97] |
| High | 0.86 [0.82,0.91] | 0.78 [0.73,0.84] | 0.66 [0.59,0.74] |

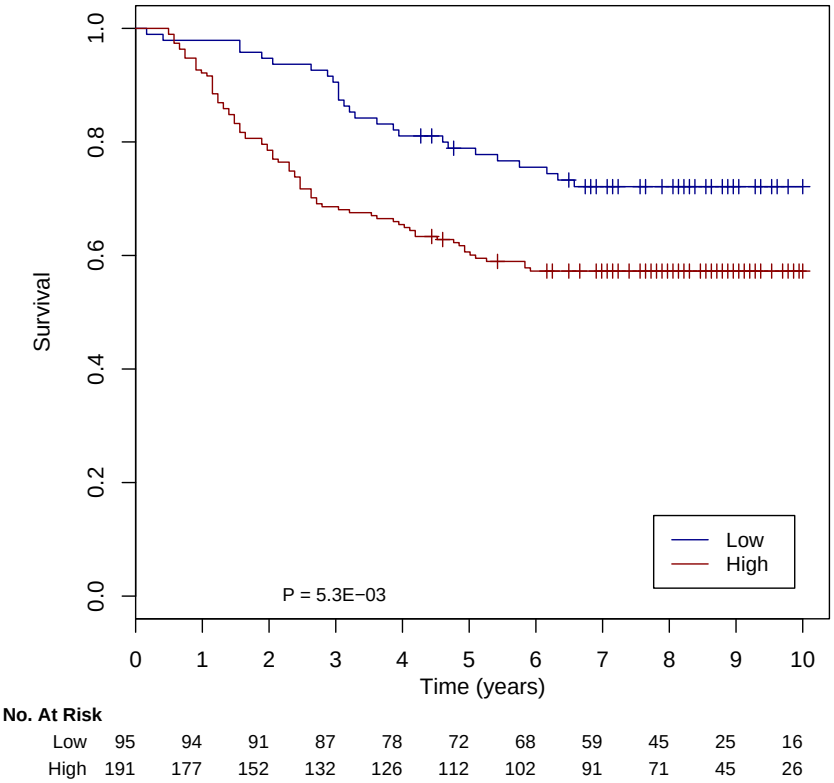
## 1.3 BD.COMBUNIV.WILCOXON.HG

The following figure shows the distribution of the ggi :




**Risk Score** On TRAINING, the risk score exhibits a Cox p-value of 1.16E-02, a concordance index of 0.606 95CI[0.552,0.659] (p-value of 5.55E-05) and an integrated Brier score of 0.185.

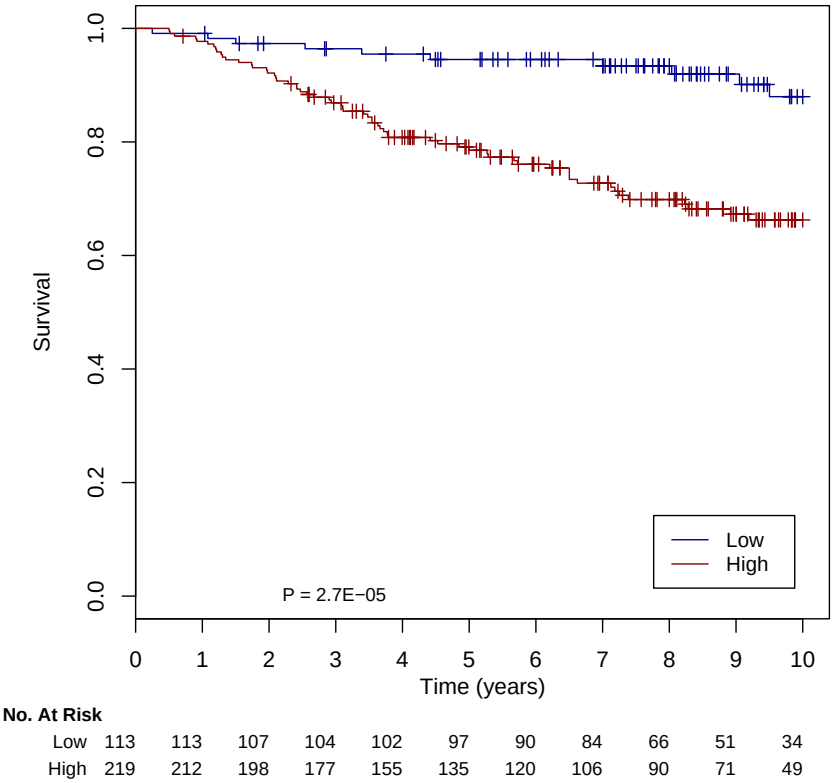
**Risk Group** On TRAINING, the binary classification computed from the risk score exhibits a Cox p-value of 3.99E-03, a concordance index of 0.675 95CI[0.579,0.772] (p-value of 1.86E-04) and an integrated Brier score of 0.184. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.87 [0.81,0.94] | 0.78 [0.70,0.87] | 0.72 [0.64,0.82] |
| High | 0.68 [0.62,0.75] | 0.60 [0.53,0.67] | 0.57 [0.51,0.65] |

**Risk Score** On TEST, the risk score exhibits a Cox p-value of 1.19E-06, a concordance index of 0.687 95CI[0.626,0.747] (p-value of 5.85E-10) and an integrated Brier score of 0.131.

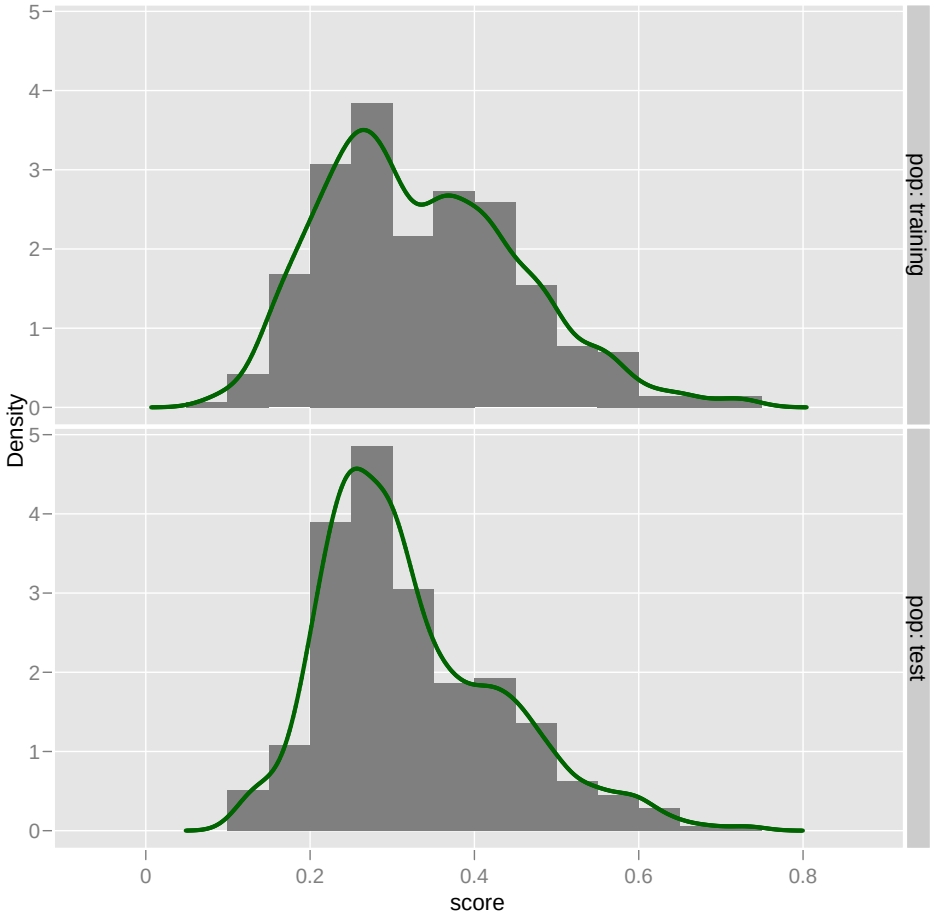
**Risk Group** On TEST, the binary classification computed from the risk score exhibits a Cox p-value of 2.35E-06, a concordance index of 0.804 95CI[0.692,0.916] (p-value of 5.49E-08) and an integrated Brier score of 0.134. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.95 [0.92,0.99] | 0.95 [0.90,0.99] | 0.88 [0.81,0.96] |
| High | 0.87 [0.82,0.92] | 0.79 [0.73,0.84] | 0.66 [0.59,0.74] |

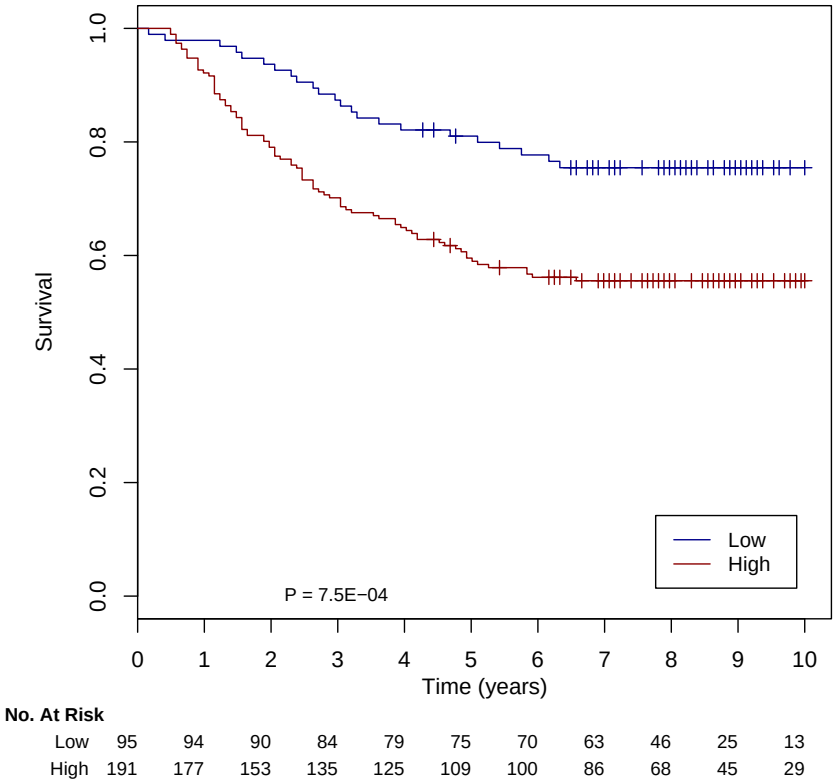
## 1.4 BD.COMBUNIV.COX.SURV

The following figure shows the distribution of the ggi :




**Risk Score** On TRAINING, the risk score exhibits a Cox p-value of 1.01E-05, a concordance index of 0.638 95CI[0.584,0.691] (p-value of 2.30E-07) and an integrated Brier score of 0.178.

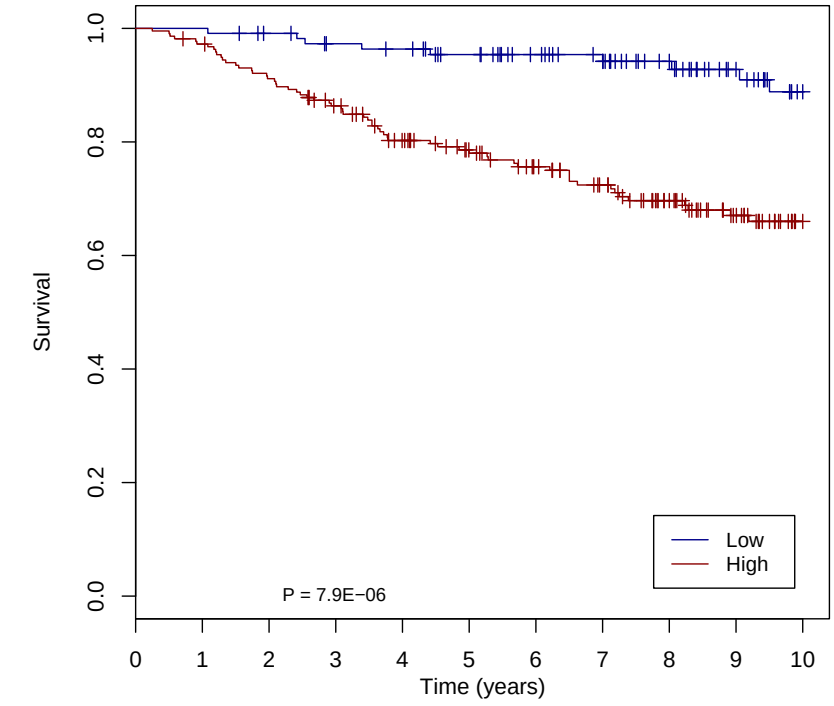
**Risk Group** On TRAINING, the binary classification computed from the risk score exhibits a Cox p-value of 4.40E-04, a concordance index of 0.698 95CI[0.6,0.795] (p-value of 3.65E-05) and an integrated Brier score of 0.181. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.86 [0.80,0.94] | 0.80 [0.72,0.88] | 0.75 [0.67,0.85] |
| High | 0.69 [0.62,0.75] | 0.59 [0.52,0.66] | 0.56 [0.49,0.63] |

**Risk Score** On TEST, the risk score exhibits a Cox p-value of 1.35E-05, a concordance index of 0.684 95CI[0.627,0.742] (p-value of 1.43E-10) and an integrated Brier score of 0.131.

**Risk Group** On TEST, the binary classification computed from the risk score exhibits a Cox p-value of 4.75E-07, a concordance index of 0.834 95CI[0.732,0.937] (p-value of 7.36E-11) and an integrated Brier score of 0.133. The following figure shows the Kaplan-Meier survival curves for the two groups :



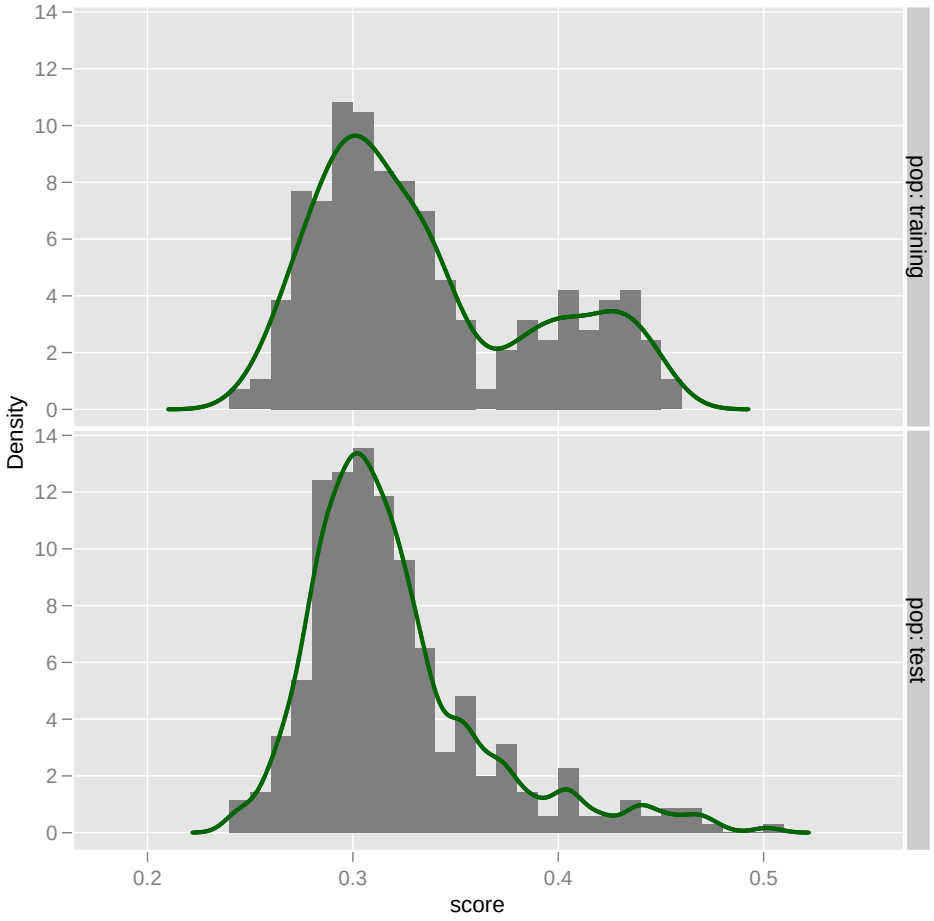

**No. At Risk**

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Low | 114 | 114 | 111 | 105 | 103 | 96 | 87 | 80 | 67 | 52 | 36 |
| High | 218 | 210 | 194 | 176 | 154 | 136 | 123 | 110 | 89 | 70 | 47 |

| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.96 [0.93,1.00] | 0.95 [0.92,0.99] | 0.89 [0.82,0.97] |
| High | 0.86 [0.82,0.91] | 0.78 [0.73,0.84] | 0.66 [0.59,0.74] |

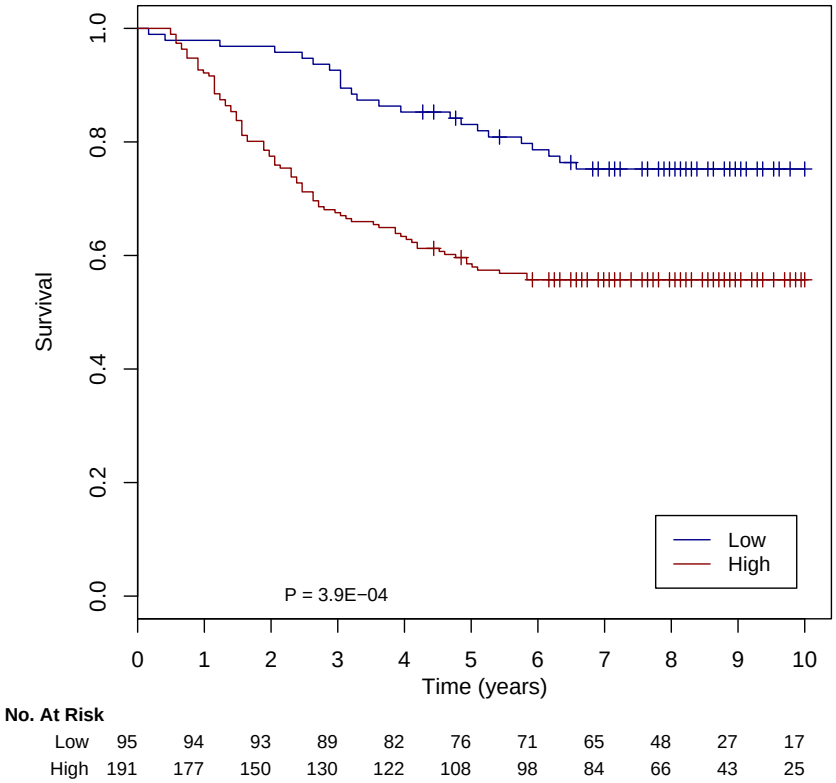
## 1.5 BD.MULTIV.LM.TOE

The following figure shows the distribution of the ggi in the two populations :




**Risk Score** On TRAINING, the risk score exhibits a Cox p-value of 5.80E-02, a concordance index of 0.601 95CI[0.55,0.653] (p-value of 5.95E-05) and an integrated Brier score of 0.186.

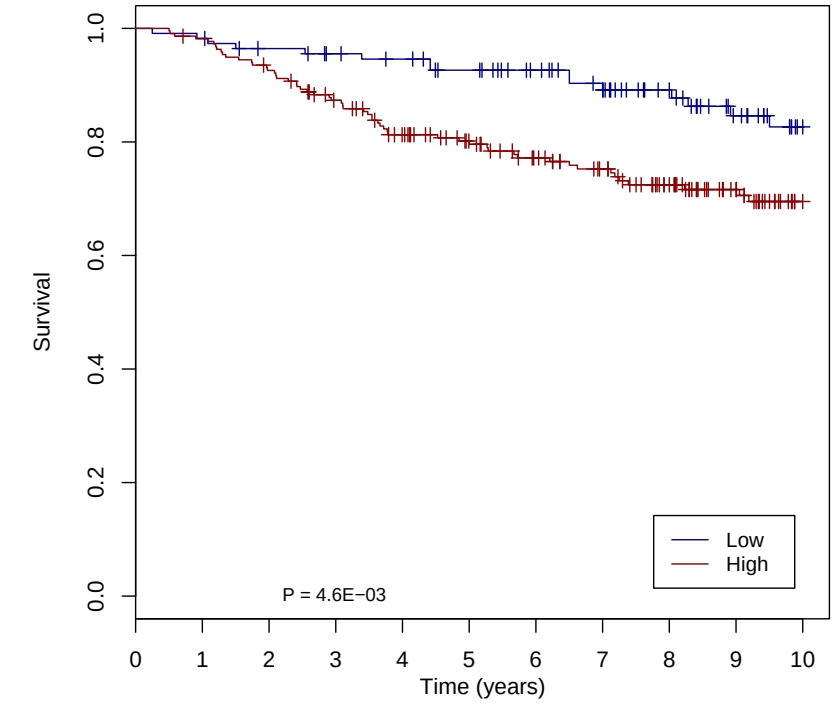
**Risk Group** On TRAINING, the binary classification computed from the risk score exhibits a Cox p-value of 2.15E-04, a concordance index of 0.721 95CI[0.628,0.814] (p-value of 1.68E-06) and an integrated Brier score of 0.18. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.89 [0.84,0.96] | 0.82 [0.75,0.90] | 0.75 [0.67,0.85] |
| High | 0.67 [0.61,0.74] | 0.58 [0.51,0.65] | 0.56 [0.49,0.63] |

**Risk Score** On TEST, the risk score exhibits a Cox p-value of 2.43E-06, a concordance index of 0.683 95CI[0.617,0.749] (p-value of 2.47E-08) and an integrated Brier score of 0.132.

**Risk Group** On TEST, the binary classification computed from the risk score exhibits a Cox p-value of 7.08E-04, a concordance index of 0.716 95CI[0.594,0.837] (p-value of 2.63E-04) and an integrated Brier score of 0.138. The following figure shows the Kaplan-Meier survival curves for the two groups :



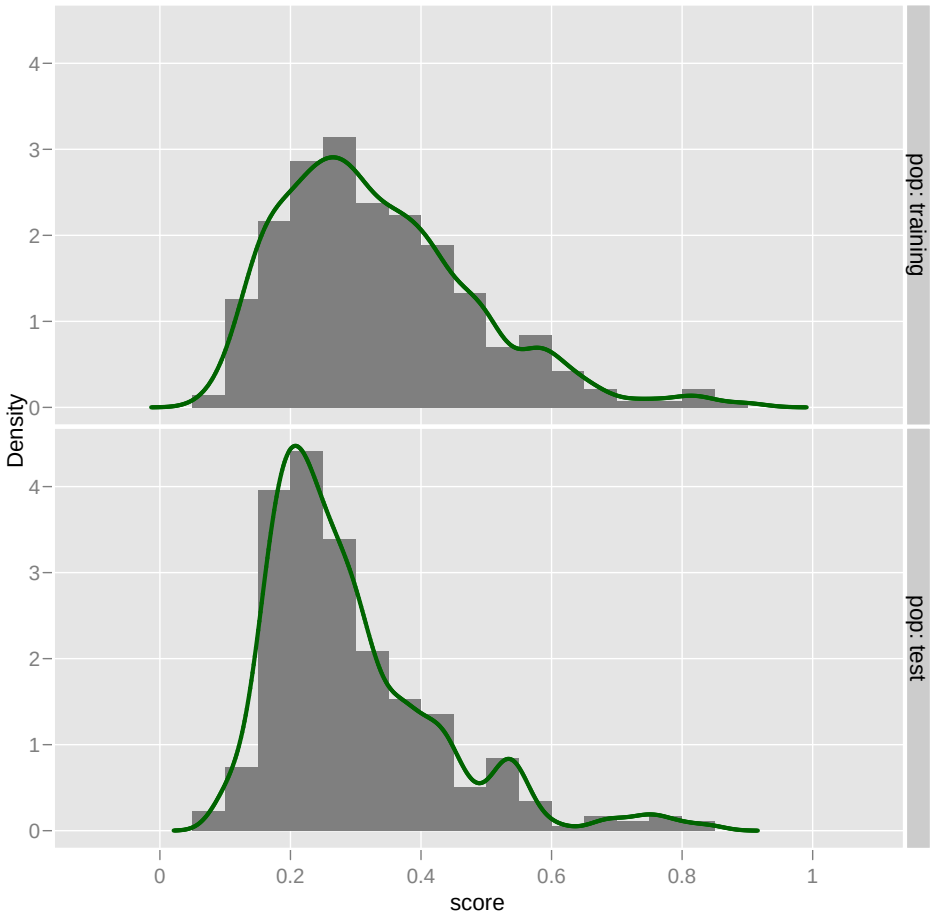

**No. At Risk**

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Low | 113 | 112 | 107 | 103 | 100 | 93 | 85 | 77 | 64 | 51 | 35 |
| High | 219 | 214 | 198 | 178 | 157 | 139 | 125 | 113 | 92 | 71 | 48 |

| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.96 [0.92,0.99] | 0.93 [0.88,0.98] | 0.83 [0.75,0.91] |
| High | 0.87 [0.82,0.92] | 0.80 [0.74,0.85] | 0.70 [0.63,0.77] |

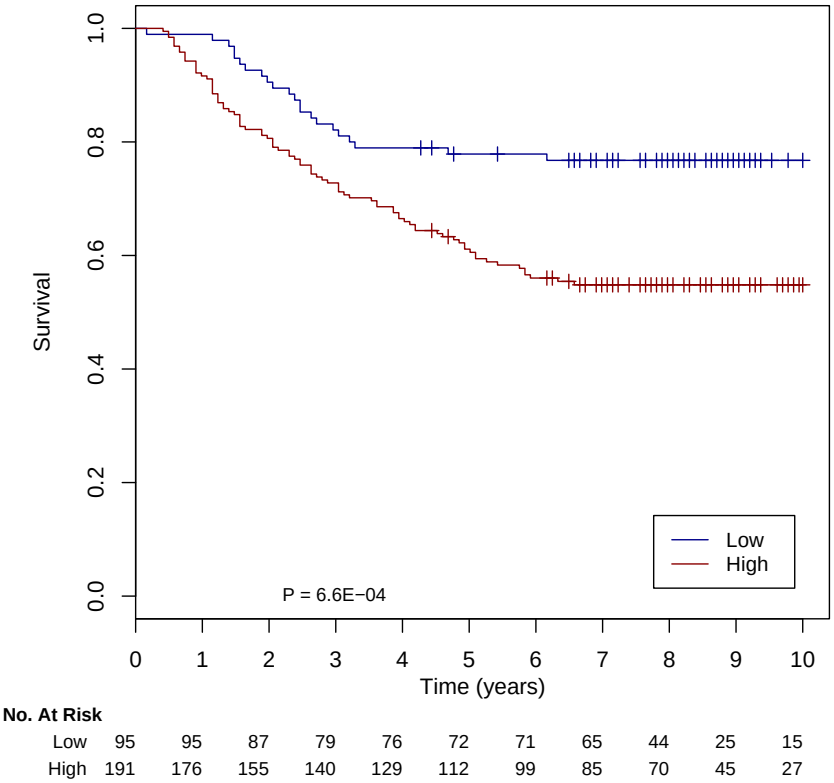
## 1.6 BD.MULTIV.COX.SURV

The following figure shows the distribution of the ggi in the two populations :




**Risk Score** On TRAINING, the risk score exhibits a Cox p-value of 3.30E-08, a concordance index of 0.649 95CI[0.594,0.704] (p-value of 5.59E-08) and an integrated Brier score of 0.172.

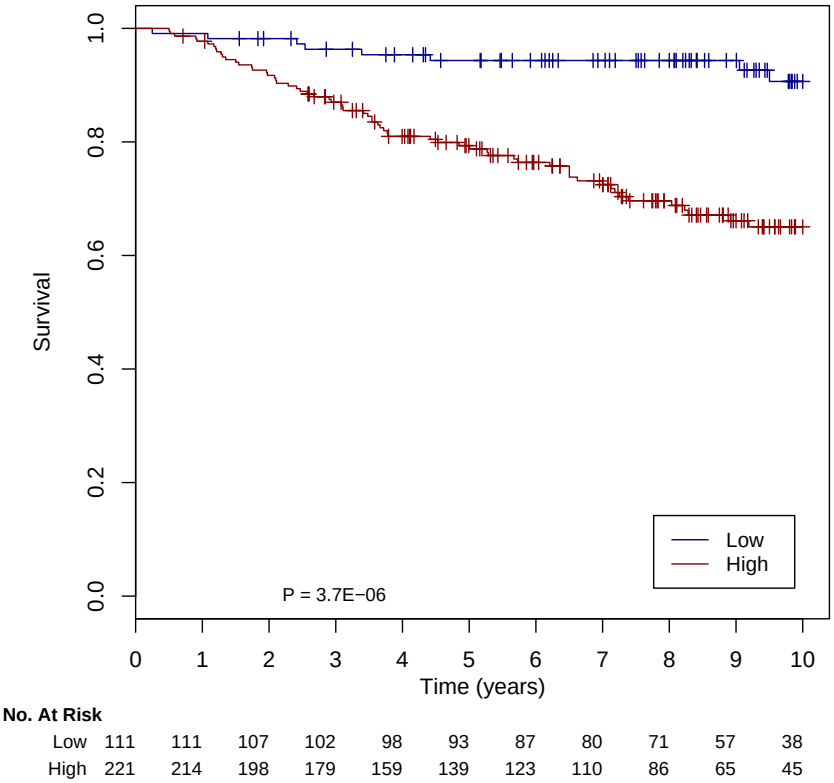
**Risk Group** On TRAINING, the binary classification computed from the risk score exhibits a Cox p-value of 3.72E-04, a concordance index of 0.685 95CI[0.583,0.787] (p-value of 1.82E-04) and an integrated Brier score of 0.182. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.81 [0.74,0.89] | 0.78 [0.70,0.87] | 0.77 [0.69,0.86] |
| High | 0.71 [0.65,0.78] | 0.61 [0.54,0.68] | 0.55 [0.48,0.62] |

**Risk Score** On TEST, the risk score exhibits a Cox p-value of 1.59E-03, a concordance index of 0.657 95CI[0.6,0.713] (p-value of 2.70E-08) and an integrated Brier score of 0.132.

**Risk Group** On TEST, the binary classification computed from the risk score exhibits a Cox p-value of 3.79E-07, a concordance index of 0.828 95CI[0.716,0.939] (p-value of 4.43E-09) and an integrated Brier score of 0.132. The following figure shows the Kaplan-Meier survival curves for the two groups :



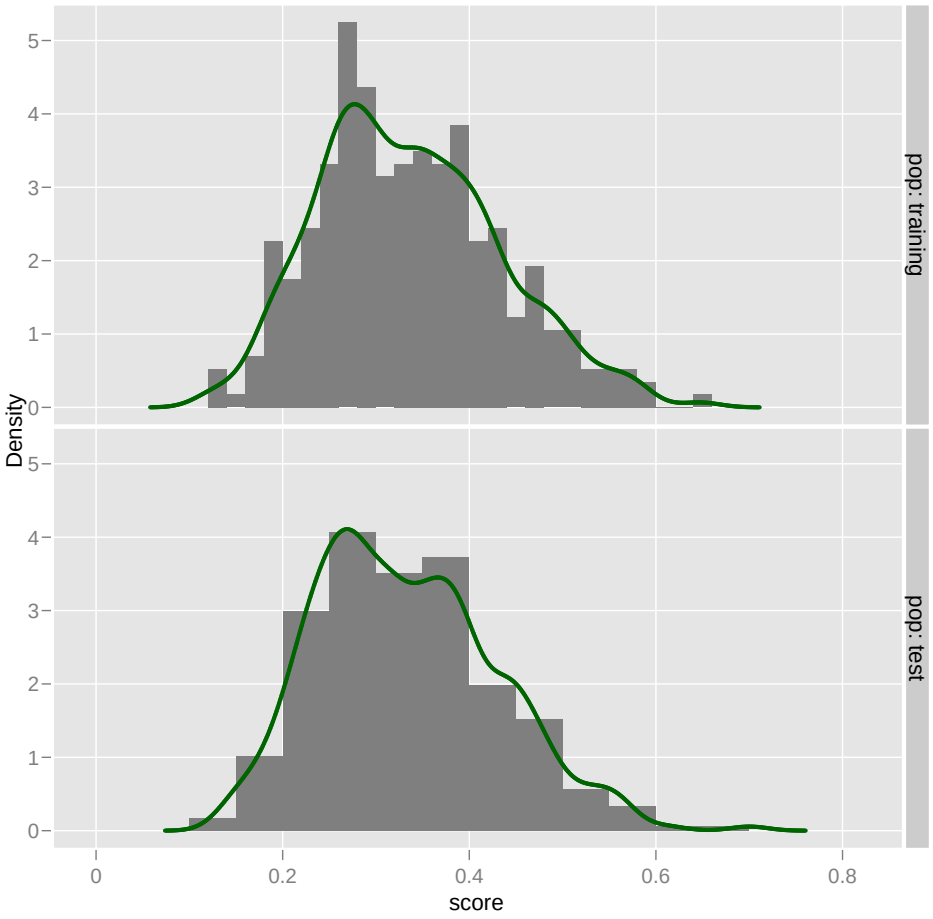

| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.96 [0.93,1.00] | 0.94 [0.90,0.99] | 0.91 [0.84,0.97] |
| High | 0.87 [0.83,0.92] | 0.79 [0.73,0.85] | 0.65 [0.58,0.73] |

## 1.7 GW.RANK.COMBUNIV.WILCOXON.HG

The following table shows the selected features :

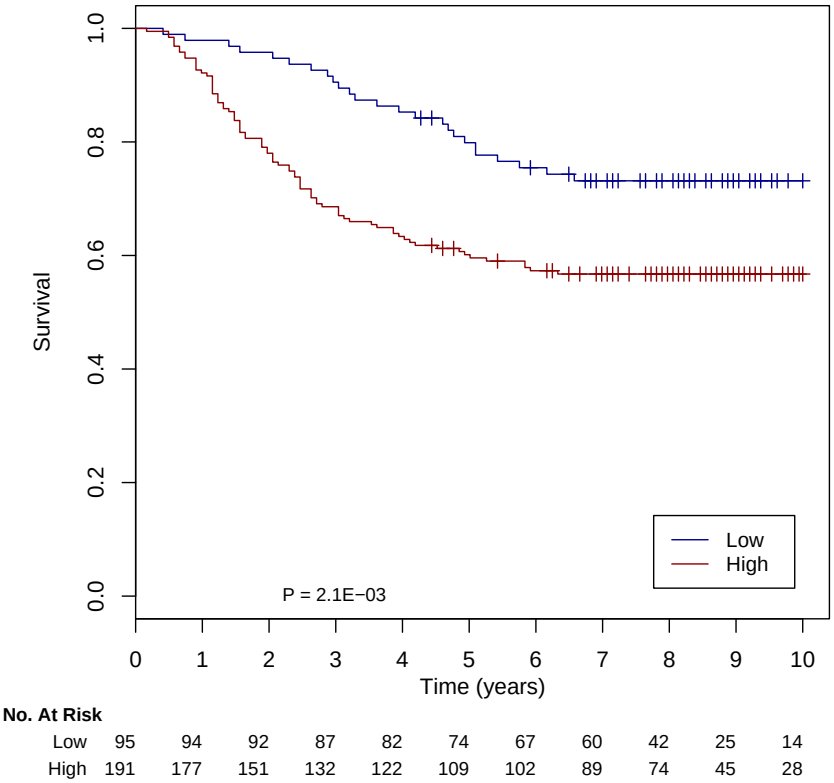
| | diff | p.value | c.index | p.value | NCBI.gene.symbol | EntrezGene.ID |
|---|---|---|---|---|---|---|
| 206887.at | −1.000 | 7.2E−05 | 0.473 | 1.6E−01 | CCBP2 | 1238 |
| 222077.s.at | 1.000 | 8.7E−05 | 0.643 | 8.1E−09 | RACGAP1 | 29127 |
| 218728.s.at | 1.000 | 9.3E−05 | 0.564 | 1.1E−02 | CNIH4 | 29097 |
| 212250.at | 1.000 | 1.0E−04 | 0.575 | 3.7E−03 | MTDH | 92140 |
| 201897.s.at | 1.000 | 1.0E−04 | 0.563 | 1.5E−02 | CKS1B | 1163 |
| 215850.s.at | −1.000 | 1.1E−04 | 0.501 | 4.9E−01 | | |
| 219494.at | 1.000 | 1.2E−04 | 0.595 | 2.8E−04 | RAD54B | 25788 |
| 206791.s.at | −1.000 | 1.2E−04 | 0.449 | 3.7E−02 | PDE4C | 5143 |
| 208079.s.at | 1.000 | 1.2E−04 | 0.636 | 3.0E−07 | AURKA | 6790 |
| 215021.s.at | −1.000 | 1.3E−04 | 0.411 | 7.7E−04 | NRXN3 | 9369 |
| 203213.at | 1.000 | 1.4E−04 | 0.593 | 2.7E−04 | | |
| 202188.at | 1.000 | 1.4E−04 | 0.563 | 1.1E−02 | NUP93 | 9688 |
| 221676.s.at | 1.000 | 1.4E−04 | 0.573 | 3.6E−03 | CORO1C | 23603 |
| 204886.at | 1.000 | 1.4E−04 | 0.594 | 1.9E−04 | PLK4 | 10733 |
| 220245.at | −1.000 | 1.7E−04 | 0.461 | 7.2E−02 | SLC45A2 | 51151 |
| 219010.at | 1.000 | 1.9E−04 | 0.543 | 6.5E−02 | C1orf106 | 55765 |
| 215999.at | −1.000 | 2.2E−04 | 0.477 | 2.0E−01 | | |
| 203764.at | 1.000 | 2.2E−04 | 0.614 | 1.5E−05 | DLG7 | 9787 |
| 204641.at | 1.000 | 2.2E−04 | 0.646 | 4.4E−09 | NEK2 | 4751 |
| 210691.s.at | 1.000 | 2.5E−04 | 0.566 | 1.1E−02 | CACYBP | 27101 |
| 201834.at | −1.000 | 2.6E−04 | 0.472 | 1.6E−01 | PRKAB1 | 5564 |
| 206515.at | −1.000 | 2.6E−04 | 0.495 | 4.3E−01 | | |
| 202107.s.at | 1.000 | 2.7E−04 | 0.567 | 9.9E−03 | MCM2 | 4171 |
| 218732.at | 1.000 | 2.9E−04 | 0.547 | 4.8E−02 | PTRH2 | 51651 |
| 216841.s.at | 1.000 | 3.0E−04 | 0.533 | 1.2E−01 | SOD2 | 6648 |
| 48580.at | −1.000 | 3.1E−04 | 0.425 | 4.2E−03 | CXXC1 | 30827 |
| 218009.s.at | 1.000 | 3.3E−04 | 0.605 | 7.0E−05 | PRC1 | 9055 |
| 200940.s.at | −1.000 | 3.3E−04 | 0.433 | 8.1E−03 | RERE | 473 |
| 201557.at | −1.000 | 3.4E−04 | 0.455 | 5.4E−02 | VAMP2 | 6844 |
| 219037.at | 1.000 | 3.4E−04 | 0.605 | 4.3E−05 | CGI-115 | 51018 |

The following figure shows the distribution of the ggi :




**Risk Score** On TRAINING, the risk score exhibits a Cox p-value of 3.72E-04, a concordance index of 0.619 95CI[0.568,0.67] (p-value of 2.16E-06) and an integrated Brier score of 0.182.

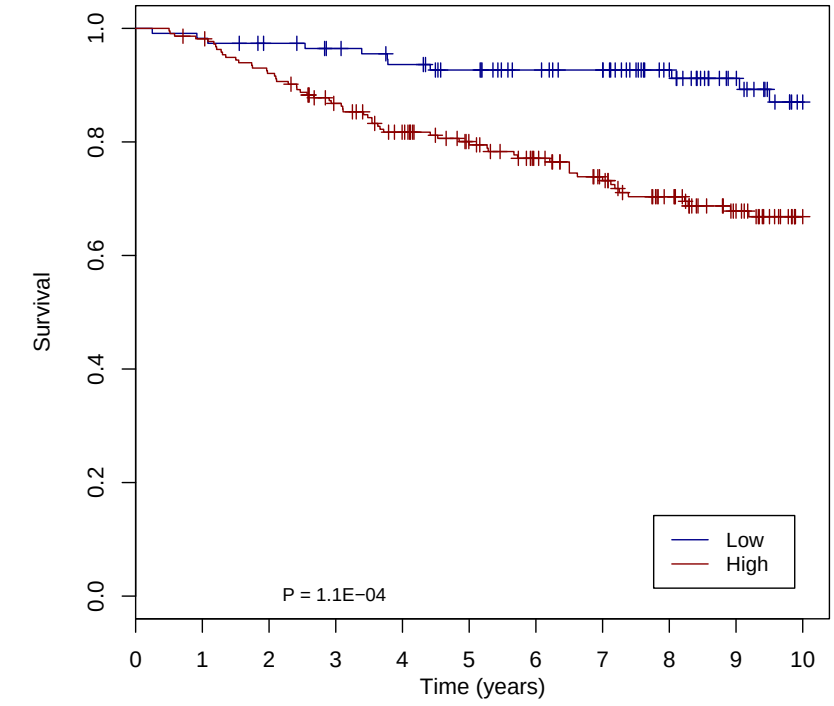
**Risk Group** On TRAINING, the binary classification computed from the risk score exhibits a Cox p-value of 1.46E-03, a concordance index of 0.694 95CI[0.599,0.789] (p-value of 3.03E-05) and an integrated Brier score of 0.182. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.89 [0.84,0.96] | 0.78 [0.70,0.87] | 0.73 [0.65,0.83] |
| High | 0.67 [0.61,0.74] | 0.60 [0.53,0.67] | 0.57 [0.50,0.64] |

**Risk Score** On TEST, the risk score exhibits a Cox p-value of 1.44E-06, a concordance index of 0.691 95CI[0.629,0.752] (p-value of 6.30E-10) and an integrated Brier score of 0.131.

**Risk Group** On TEST, the binary classification computed from the risk score exhibits a Cox p-value of 2.19E-05, a concordance index of 0.775 95CI[0.657,0.893] (p-value of 2.30E-06) and an integrated Brier score of 0.136. The following figure shows the Kaplan-Meier survival curves for the two groups :



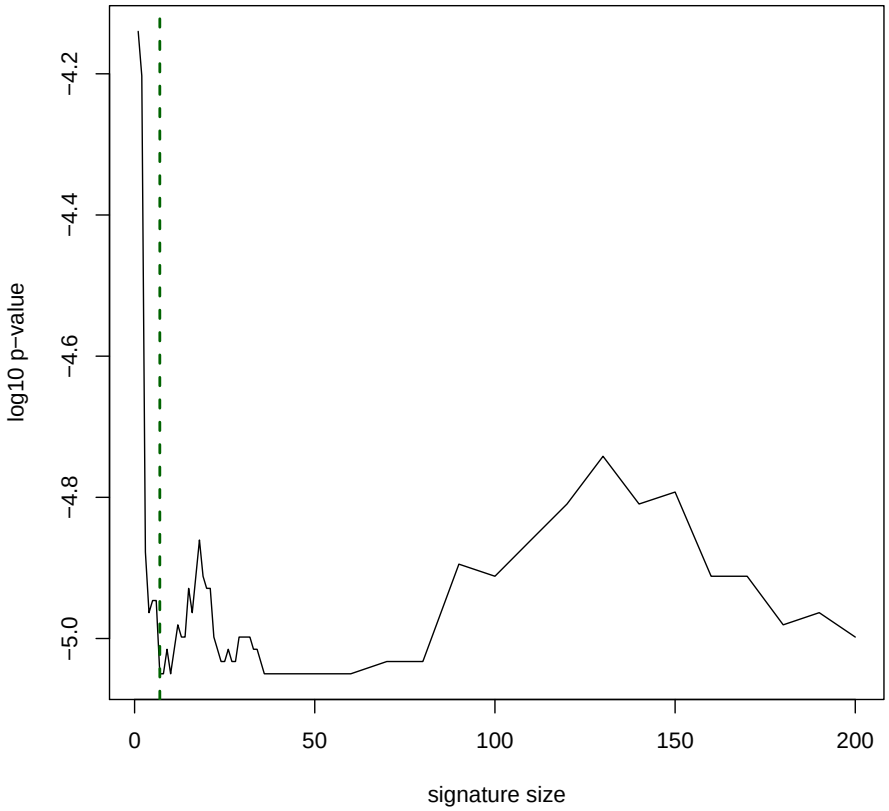

**No. At Risk**

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Low | 114 | 113 | 109 | 105 | 100 | 93 | 85 | 81 | 65 | 49 | 32 |
| High | 218 | 213 | 196 | 176 | 157 | 139 | 125 | 109 | 92 | 73 | 51 |

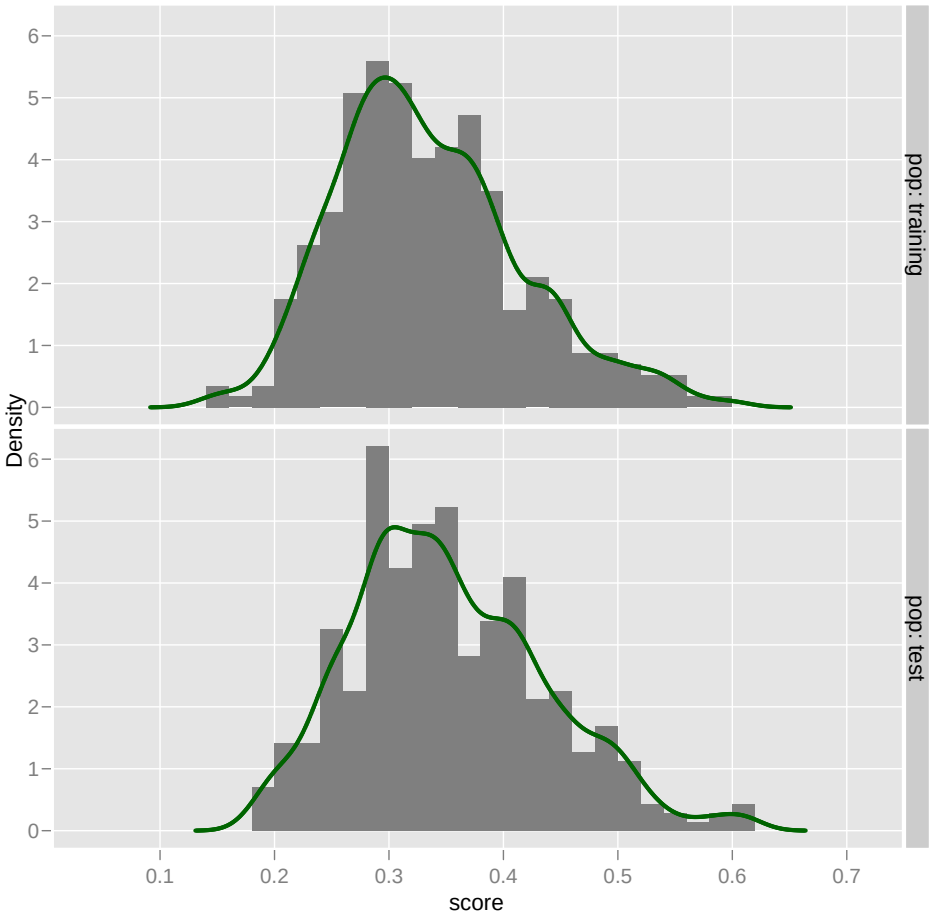
| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.96 [0.93,1.00] | 0.93 [0.88,0.98] | 0.87 [0.80,0.95] |
| High | 0.86 [0.82,0.91] | 0.79 [0.74,0.85] | 0.67 [0.60,0.74] |

## 1.8 GW.RANKCV.COMBUNIV.WILCOXON.HG

The following figure shows the evolution of the performance w.r.t. the signature size (best signature size is 7):

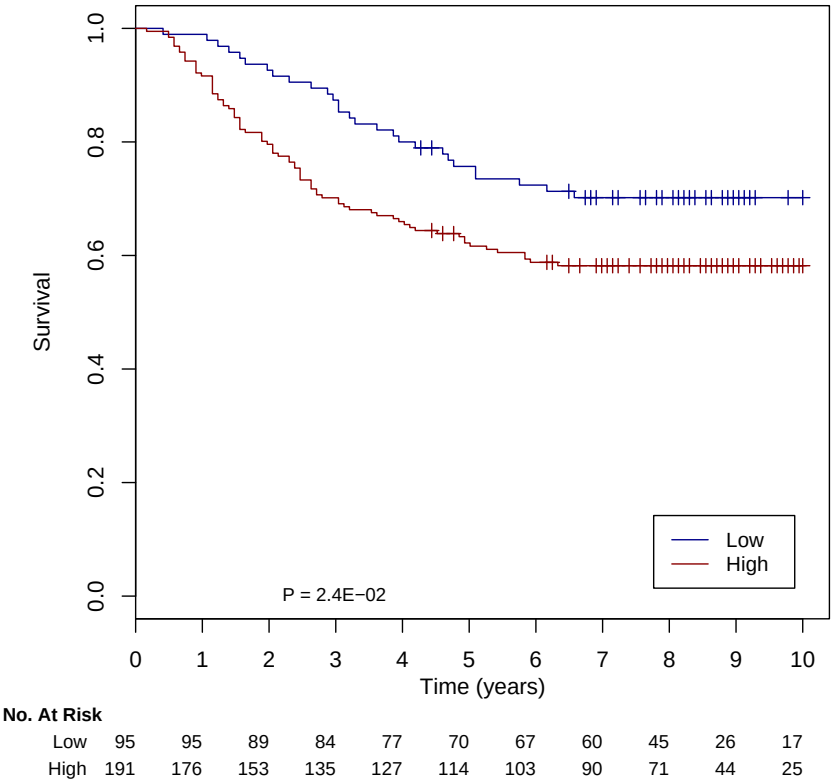



The following figure shows the distribution of the ggi :




**Risk Score** On TRAINING, the risk score exhibits a Cox p-value of 2.71E-03, a concordance index of 0.595 95CI[0.543,0.647] (p-value of 1.72E-04) and an integrated Brier score of 0.184.

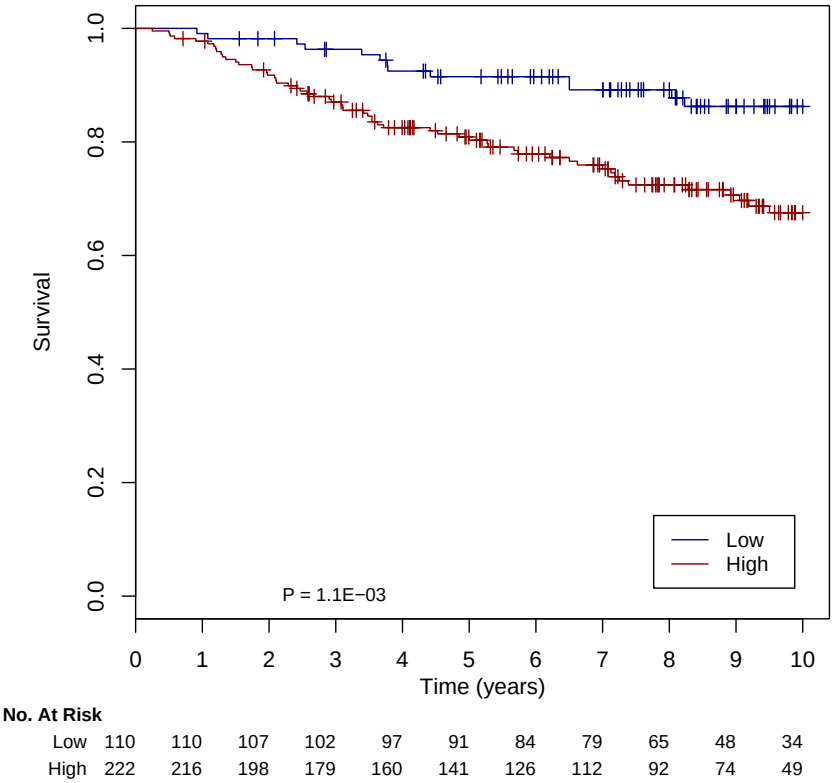
**Risk Group** On TRAINING, the binary classification computed from the risk score exhibits a Cox p-value of 1.99E-02, a concordance index of 0.642 95CI[0.544,0.741] (p-value of 2.28E-03) and an integrated Brier score of 0.185. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.85 [0.78,0.93] | 0.73 [0.65,0.83] | 0.70 [0.62,0.80] |
| High | 0.69 [0.63,0.76] | 0.62 [0.55,0.69] | 0.58 [0.52,0.66] |

**Risk Score** On TEST, the risk score exhibits a Cox p-value of 9.49E-05, a concordance index of 0.643 95CI[0.58,0.706] (p-value of 4.22E-06) and an integrated Brier score of 0.139.

**Risk Group** On TEST, the binary classification computed from the risk score exhibits a Cox p-value of 3.85E-04, a concordance index of 0.728 95CI[0.607,0.85] (p-value of 1.17E-04) and an integrated Brier score of 0.137. The following figure shows the Kaplan-Meier survival curves for the two groups :



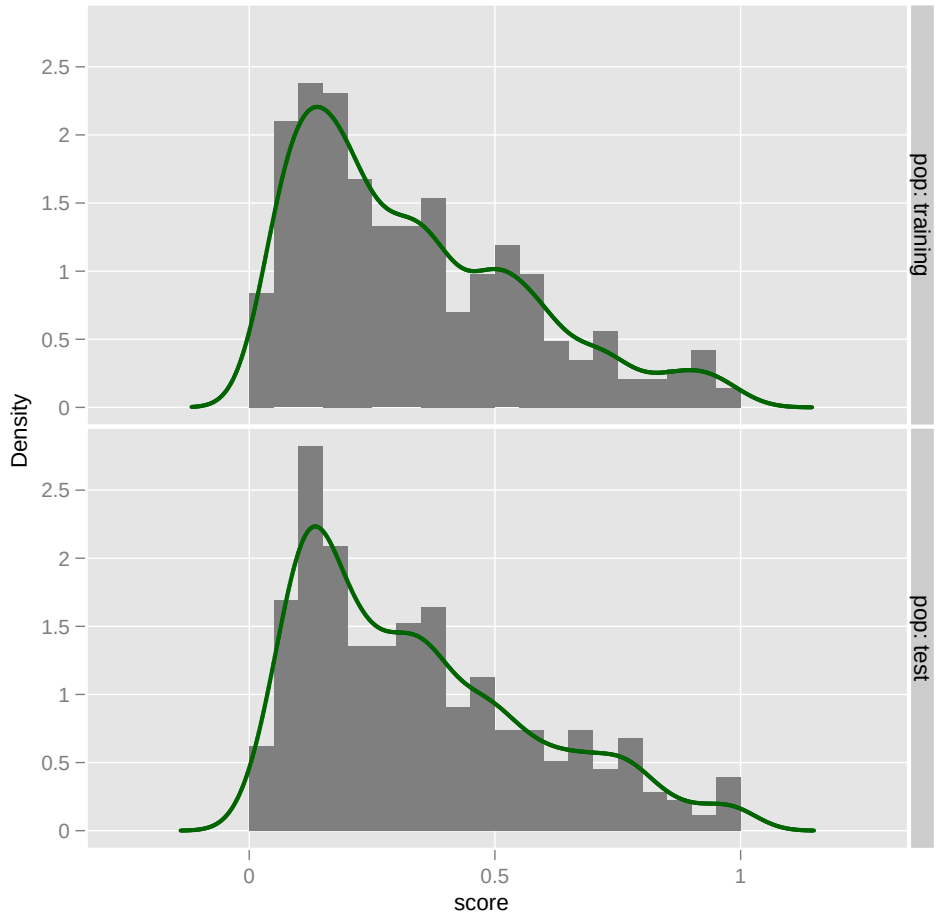

| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.95 [0.91,0.99] | 0.91 [0.86,0.97] | 0.86 [0.79,0.94] |
| High | 0.87 [0.83,0.92] | 0.80 [0.75,0.86] | 0.68 [0.61,0.75] |

## 1.9 GW.RANK.COMBUNIV.COX.SURV

The following table shows the selected features :

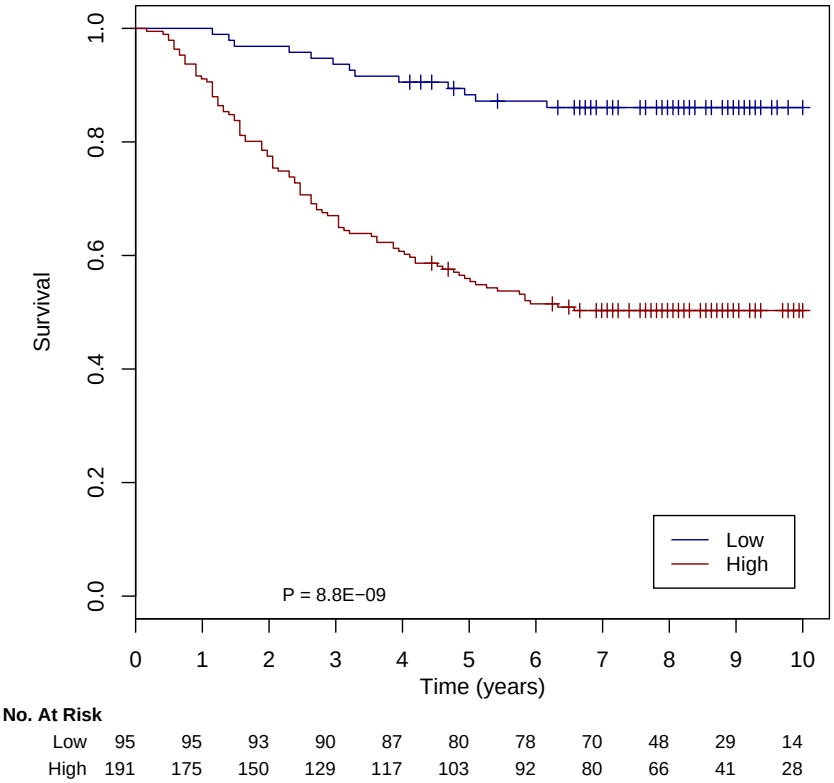
| | hr | p.value | c.index | p.value | NCBI.gene.symbol | EntrezGene.ID |
|---|---|---|---|---|---|---|
| 204641.at | 1.565 | 6.5E−07 | 0.646 | 4.4E−09 | NEK2 | 4751 |
| 222077.s.at | 1.919 | 7.8E−07 | 0.643 | 8.1E−09 | RACGAP1 | 29127 |
| 219478.at | 1.551 | 8.5E−07 | 0.641 | 1.7E−07 | WFDC1 | 58189 |
| 201769.at | 2.529 | 3.3E−06 | 0.643 | 3.5E−08 | CLINT1 | 9685 |
| 202324.s.at | 2.416 | 3.3E−06 | 0.625 | 1.7E−06 | ACBD3 | 64746 |
| 218252.at | 2.343 | 4.2E−06 | 0.627 | 1.7E−06 | CKAP2 | 26586 |
| 212898.at | 1.887 | 4.7E−06 | 0.628 | 9.2E−07 | KIAA0406 | 9675 |
| 208079.s.at | 1.540 | 5.2E−06 | 0.636 | 3.0E−07 | AURKA | 6790 |
| 202824.s.at | 2.042 | 6.6E−06 | 0.625 | 3.8E−06 | TCEB1 | 6921 |
| 209380.s.at | 1.777 | 6.7E−06 | 0.606 | 1.3E−04 | ABCC5 | 10057 |
| 213226.at | 1.657 | 7.7E−06 | 0.637 | 7.7E−08 | | |
| 201076.at | 0.349 | 8.3E−06 | 0.375 | 1.5E−06 | NHP2L1 | 4809 |
| 201664.at | 2.082 | 9.4E−06 | 0.621 | 3.2E−06 | SMC4 | 10051 |
| 202969.at | 1.894 | 1.0E−05 | 0.628 | 1.4E−06 | | |
| 212900.at | 2.439 | 1.0E−05 | 0.625 | 2.5E−06 | | |
| 218701.at | 1.783 | 1.2E−05 | 0.623 | 2.3E−06 | LACTB2 | 51110 |
| 218478.s.at | 2.707 | 1.3E−05 | 0.621 | 7.3E−06 | ZCCHC8 | 55596 |
| 217235.x.at | 0.753 | 1.3E−05 | 0.394 | 6.1E−05 | | |
| 214853.s.at | 2.447 | 1.4E−05 | 0.618 | 4.9E−06 | SHC1 | 6464 |
| 212149.at | 1.990 | 1.4E−05 | 0.623 | 2.5E−06 | KIAA0143 | 23167 |
| 210396.s.at | 1.922 | 1.6E−05 | 0.625 | 5.6E−07 | BOLA2 | 552900 |
| 202620.s.at | 1.534 | 1.6E−05 | 0.632 | 2.6E−07 | PLOD2 | 5352 |
| 207165.at | 1.613 | 1.6E−05 | 0.613 | 2.0E−05 | HMMR | 3161 |
| 212687.at | 2.135 | 1.7E−05 | 0.626 | 1.1E−06 | | |
| 209835.x.at | 0.604 | 1.8E−05 | 0.404 | 2.2E−04 | CD44 | 960 |
| 215205.x.at | 0.629 | 2.0E−05 | 0.368 | 1.8E−07 | NCOR2 | 9612 |
| 209276.s.at | 0.575 | 2.2E−05 | 0.373 | 5.5E−07 | GLRX | 2745 |
| 201369.s.at | 0.631 | 2.2E−05 | 0.378 | 1.2E−06 | ZFP36L2 | 678 |
| 217157.x.at | 0.776 | 2.3E−05 | 0.395 | 5.6E−05 | | |
| 212014.x.at | 0.640 | 2.4E−05 | 0.403 | 1.6E−04 | CD44 | 960 |

The following figure shows the distribution of the ggi :




**Risk Score** On TRAINING, the risk score exhibits a Cox p-value of 3.39E-19, a concordance index of 0.742 95CI[0.699,0.785] (p-value of 2.86E-28) and an integrated Brier score of 0.148.

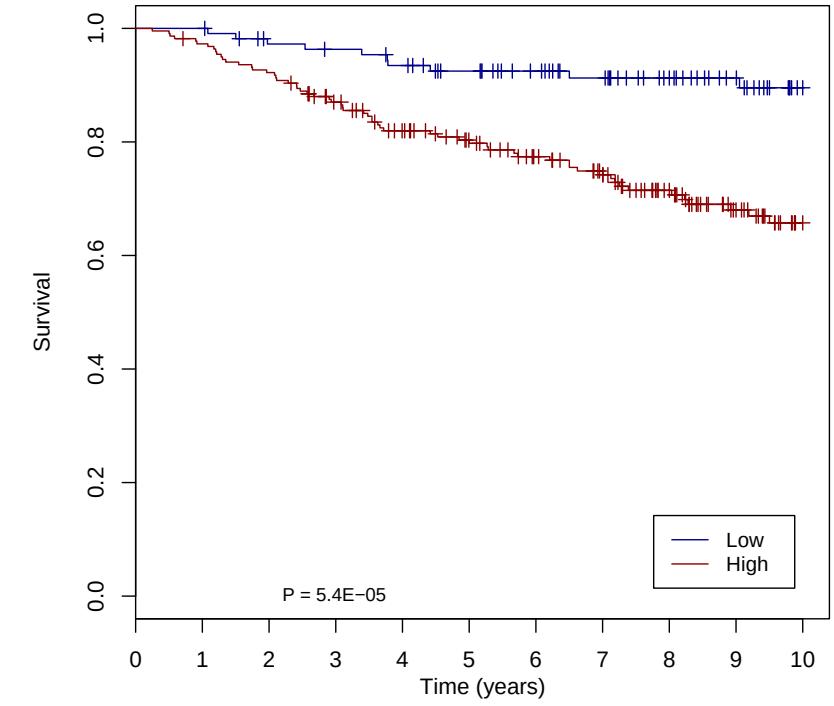
**Risk Group** On TRAINING, the binary classification computed from the risk score exhibits a Cox p-value of 4.31E-10, a concordance index of 0.836 95CI[0.756,0.916] (p-value of 8.15E-17) and an integrated Brier score of 0.168. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.93 [0.88,0.98] | 0.87 [0.81,0.94] | 0.86 [0.79,0.93] |
| High | 0.65 [0.58,0.72] | 0.55 [0.49,0.63] | 0.50 [0.44,0.58] |

**Risk Score** On TEST, the risk score exhibits a Cox p-value of 8.31E-04, a concordance index of 0.65 95CI[0.592,0.708] (p-value of 1.84E-07) and an integrated Brier score of 0.158.

**Risk Group** On TEST, the binary classification computed from the risk score exhibits a Cox p-value of 2.36E-05, a concordance index of 0.778 95CI[0.66,0.896] (p-value of 2.10E-06) and an integrated Brier score of 0.139. The following figure shows the Kaplan-Meier survival curves for the two groups :



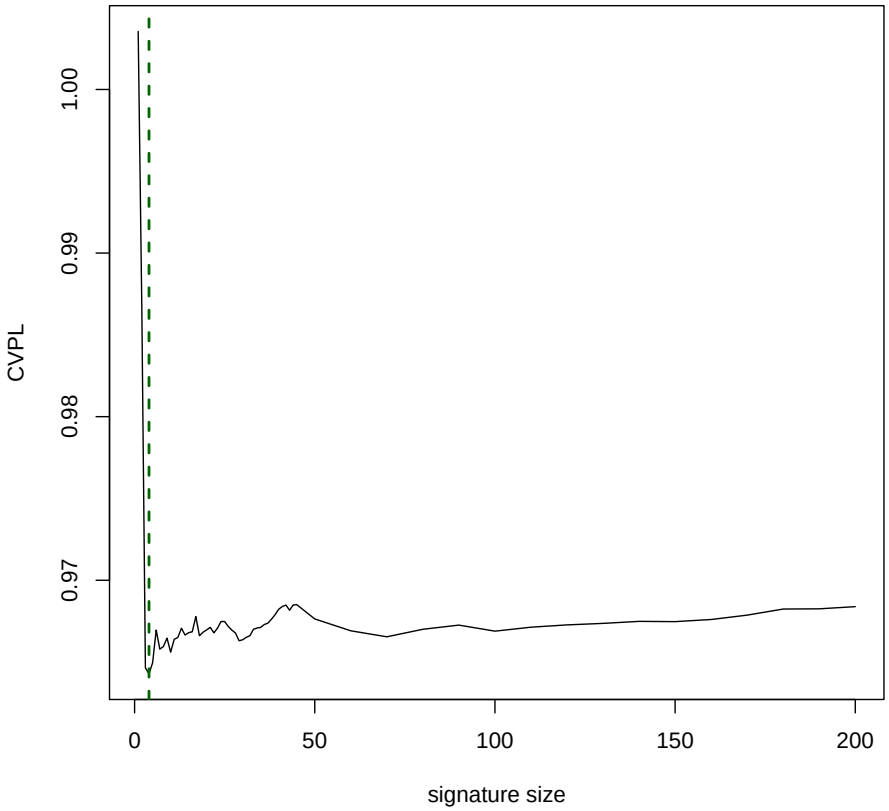

**No. At Risk**

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Low | 111 | 111 | 105 | 103 | 99 | 91 | 83 | 76 | 65 | 55 | 39 |
| High | 221 | 213 | 200 | 178 | 158 | 141 | 127 | 114 | 91 | 67 | 44 |

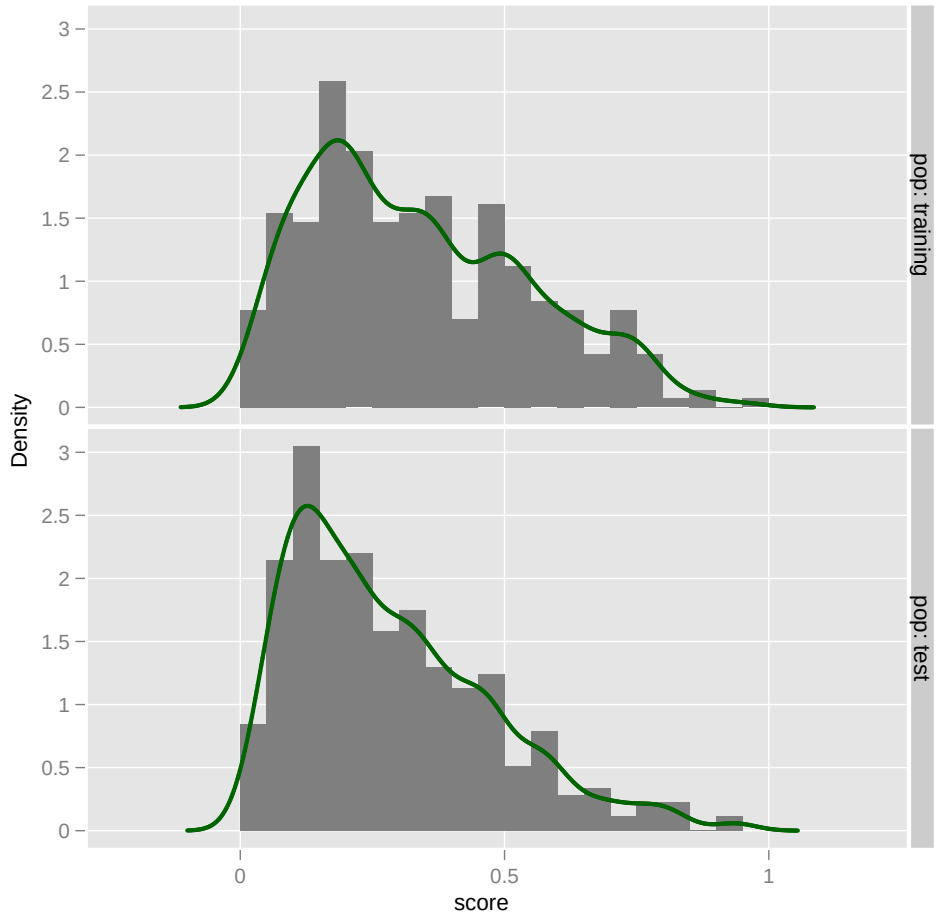
| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.95 [0.91,0.99] | 0.92 [0.88,0.98] | 0.90 [0.83,0.96] |
| High | 0.87 [0.83,0.92] | 0.80 [0.74,0.85] | 0.66 [0.59,0.74] |

## 1.10 GW.RANKCV.COMBUNIV.COX.SURV

The following figure shows the evolution of the performance w.r.t. the signature size (best signature size is 4):

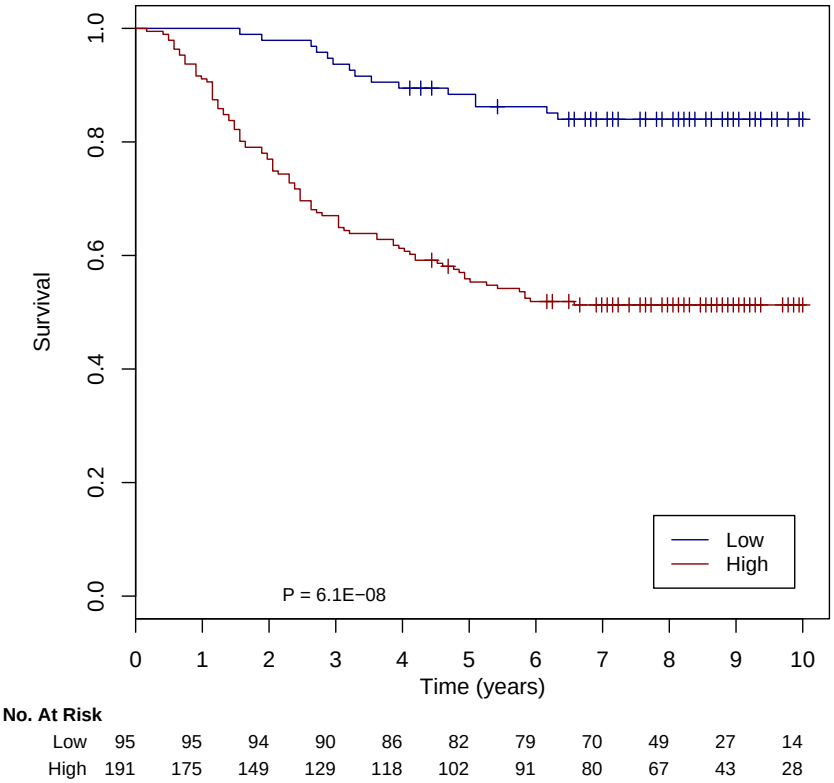



The following figure shows the distribution of the ggi :




**Risk Score** On TRAINING, the risk score exhibits a Cox p-value of 3.95E-15, a concordance index of 0.717 95CI[0.671,0.762] (p-value of 7.07E-21) and an integrated Brier score of 0.156.

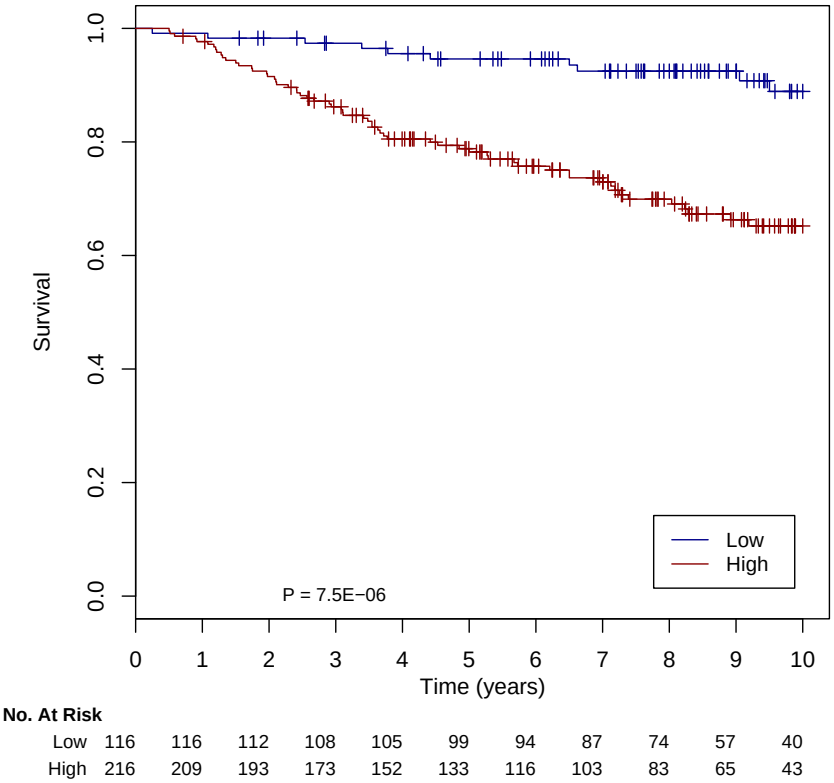
**Risk Group** On TRAINING, the binary classification computed from the risk score exhibits a Cox p-value of 5.82E-09, a concordance index of 0.824 95CI[0.745,0.903] (p-value of 3.60E-16) and an integrated Brier score of 0.17. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.93 [0.88,0.98] | 0.86 [0.80,0.93] | 0.84 [0.77,0.92] |
| High | 0.65 [0.58,0.72] | 0.55 [0.49,0.63] | 0.51 [0.45,0.59] |

**Risk Score** On TEST, the risk score exhibits a Cox p-value of 6.72E-06, a concordance index of 0.676 95CI[0.62,0.732] (p-value of 4.08E-10) and an integrated Brier score of 0.131.

**Risk Group** On TEST, the binary classification computed from the risk score exhibits a Cox p-value of 8.10E-07, a concordance index of 0.816 95CI[0.71,0.922] (p-value of 2.70E-09) and an integrated Brier score of 0.133. The following figure shows the Kaplan-Meier survival curves for the two groups :



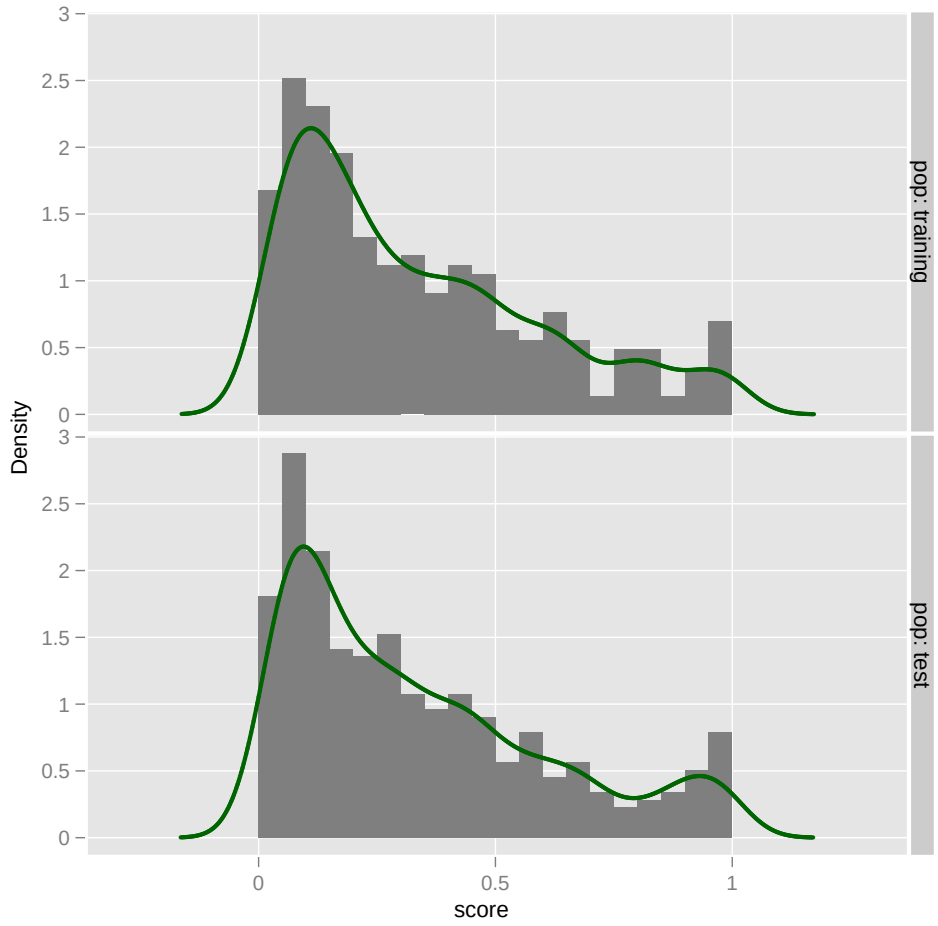

| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.96 [0.93,1.00] | 0.95 [0.91,0.99] | 0.89 [0.82,0.96] |
| High | 0.86 [0.82,0.91] | 0.78 [0.73,0.84] | 0.65 [0.58,0.73] |

## 1.11 GW.RANK.MULTIV.RCOX.SURV

The following table shows the selected features :

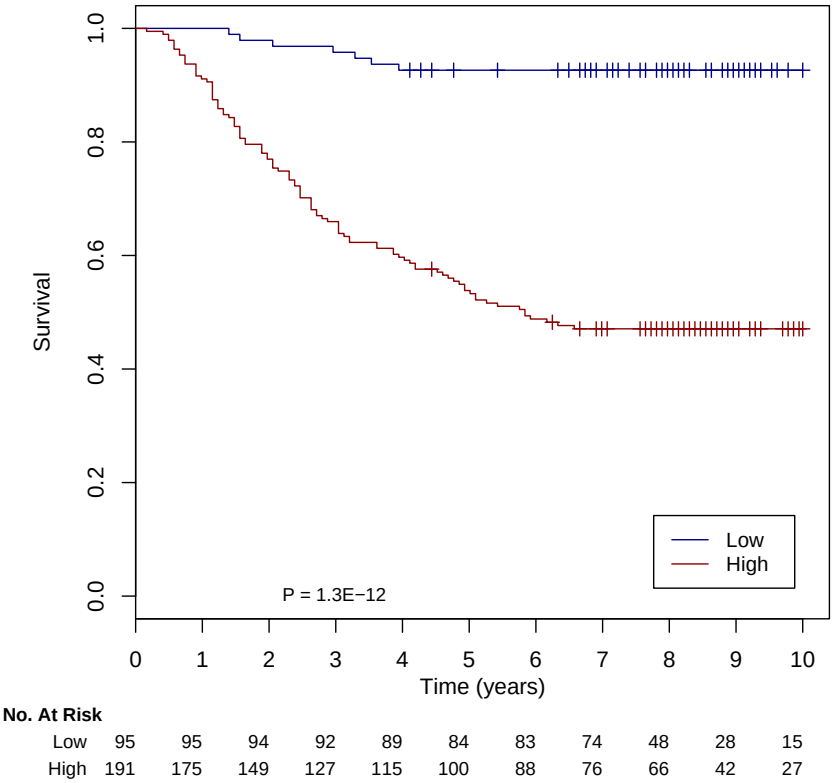
| | hr | p.value | c.index | p.value | cox.coef | NCBI.gene.symbol | EntrezGene.ID |
|---|---|---|---|---|---|---|---|
| 201076.at | 0.349 | 8.3E−06 | 0.375 | 1.5E−06 | −0.415 | NHP2L1 | 4809 |
| 219478.at | 1.551 | 8.5E−07 | 0.641 | 1.7E−07 | 0.330 | WFDC1 | 58189 |
| 204641.at | 1.565 | 6.5E−07 | 0.646 | 4.4E−09 | 0.219 | NEK2 | 4751 |
| 201769.at | 2.529 | 3.3E−06 | 0.643 | 3.5E−08 | 0.216 | CLINT1 | 9685 |
| 218252.at | 2.343 | 4.2E−06 | 0.627 | 1.7E−06 | 0.167 | CKAP2 | 26586 |
| 209380.s.at | 1.777 | 6.7E−06 | 0.606 | 1.3E−04 | 0.162 | ABCC5 | 10057 |
| 215205.x.at | 0.629 | 2.0E−05 | 0.368 | 1.8E−07 | −0.153 | NCOR2 | 9612 |
| 202969.at | 1.894 | 1.0E−05 | 0.628 | 1.4E−06 | 0.133 | | |
| 217235.x.at | 0.753 | 1.3E−05 | 0.394 | 6.1E−05 | −0.130 | | |
| 212898.at | 1.887 | 4.7E−06 | 0.628 | 9.2E−07 | 0.119 | KIAA0406 | 9675 |
| 209835.x.at | 0.604 | 1.8E−05 | 0.404 | 2.2E−04 | −0.102 | CD44 | 960 |
| 202620.s.at | 1.534 | 1.6E−05 | 0.632 | 2.6E−07 | 0.094 | PLOD2 | 5352 |
| 201369.s.at | 0.631 | 2.2E−05 | 0.378 | 1.2E−06 | −0.030 | ZFP36L2 | 678 |
| 218701.at | 1.783 | 1.2E−05 | 0.623 | 2.3E−06 | 0.022 | LACTB2 | 51110 |
| 217157.x.at | 0.776 | 2.3E−05 | 0.395 | 5.6E−05 | −0.021 | | |
| 212900.at | 2.439 | 1.0E−05 | 0.625 | 2.5E−06 | 0.021 | | |
| 209276.s.at | 0.575 | 2.2E−05 | 0.373 | 5.5E−07 | −0.013 | GLRX | 2745 |
| 201664.at | 2.082 | 9.4E−06 | 0.621 | 3.2E−06 | 0.000 | SMC4 | 10051 |

The following figure shows the distribution of the ggi :




**Risk Score** On TRAINING, the risk score exhibits a Cox p-value of 1.41E-25, a concordance index of 0.774 95CI[0.733,0.815] (p-value of 4.21E-39) and an integrated Brier score of 0.136.

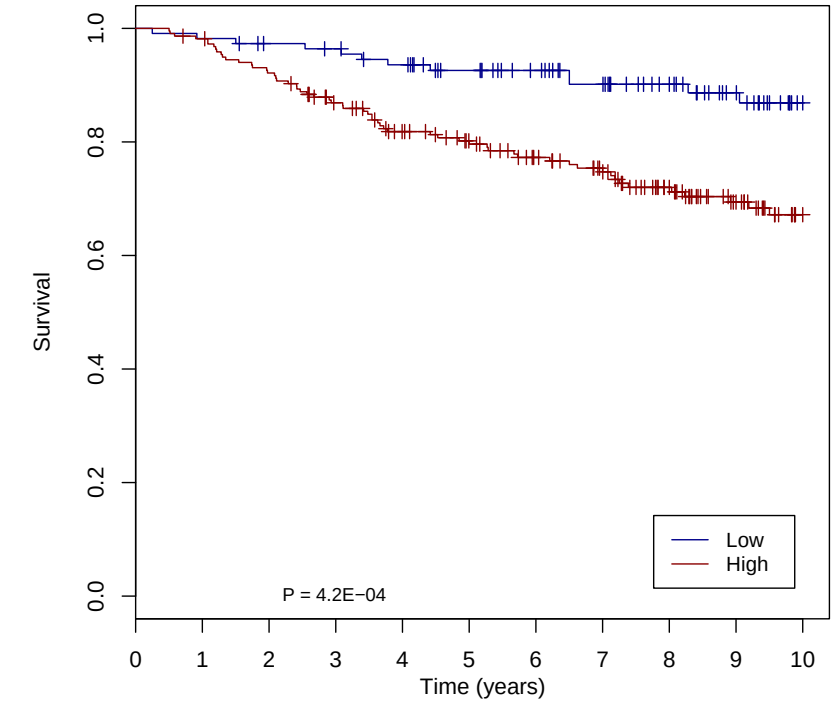
**Risk Group** On TRAINING, the binary classification computed from the risk score exhibits a Cox p-value of 1.51E-15, a concordance index of 0.906 95CI[0.841,0.972] (p-value of 9.74E-35) and an integrated Brier score of 0.159. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.95 [0.90,0.99] | 0.93 [0.88,0.98] | 0.93 [0.88,0.98] |
| High | 0.64 [0.57,0.71] | 0.53 [0.47,0.61] | 0.47 [0.40,0.55] |

**Risk Score** On TEST, the risk score exhibits a Cox p-value of 1.65E-03, a concordance index of 0.638 95CI[0.579,0.697] (p-value of 2.10E-06) and an integrated Brier score of 0.175.

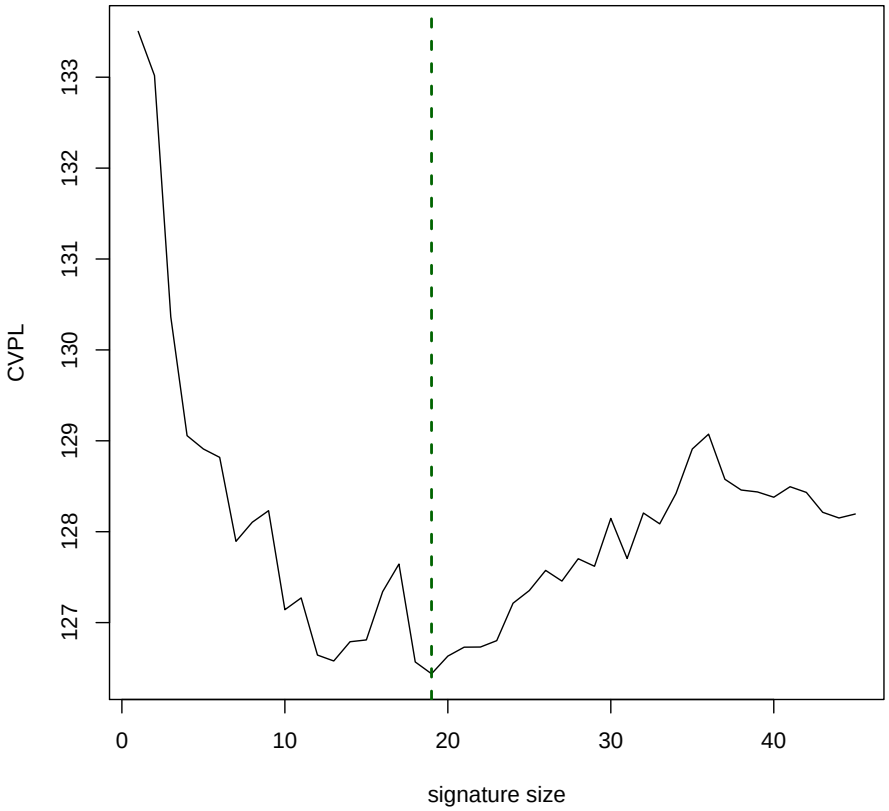
**Risk Group** On TEST, the binary classification computed from the risk score exhibits a Cox p-value of 1.13E-04, a concordance index of 0.749 95CI[0.628,0.87] (p-value of 2.78E-05) and an integrated Brier score of 0.147. The following figure shows the Kaplan-Meier survival curves for the two groups :



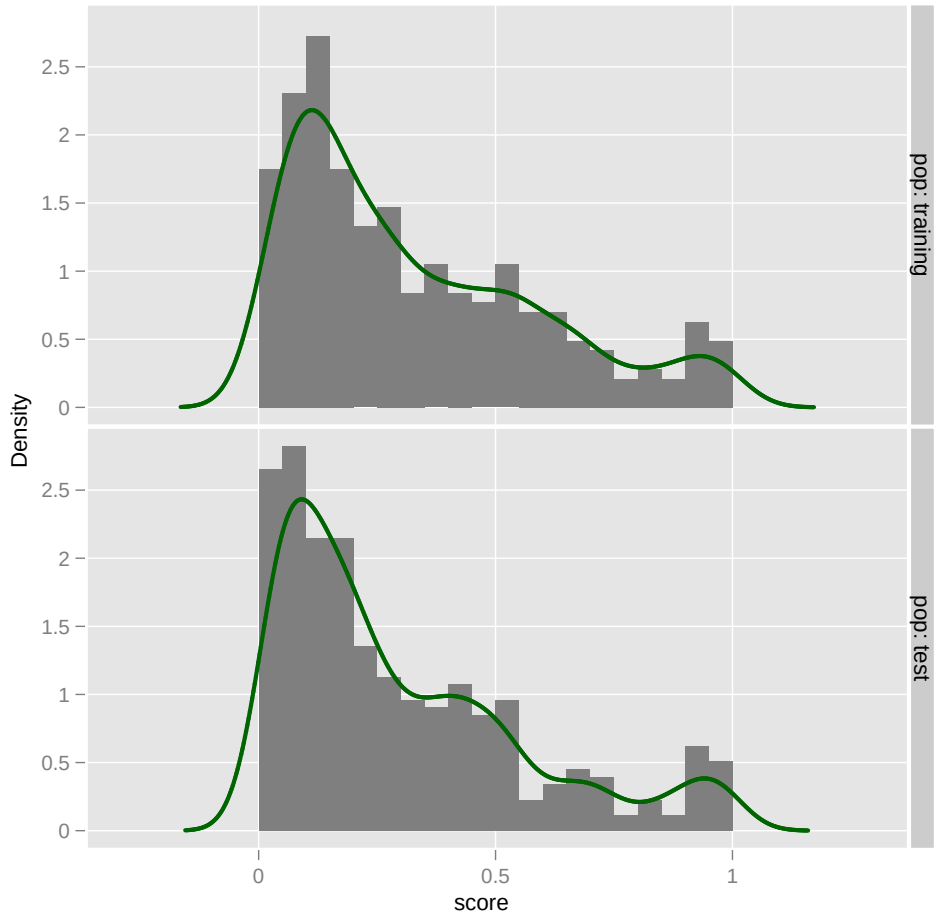

| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.96 [0.93,1.00] | 0.93 [0.88,0.98] | 0.87 [0.80,0.94] |
| High | 0.86 [0.82,0.91] | 0.80 [0.74,0.85] | 0.67 [0.60,0.75] |

## 1.12 GW.RANKCV.MULTIV.RCOX.SURV

The following figure shows the evolution of the performance w.r.t. the signature size (best signature size is 19):

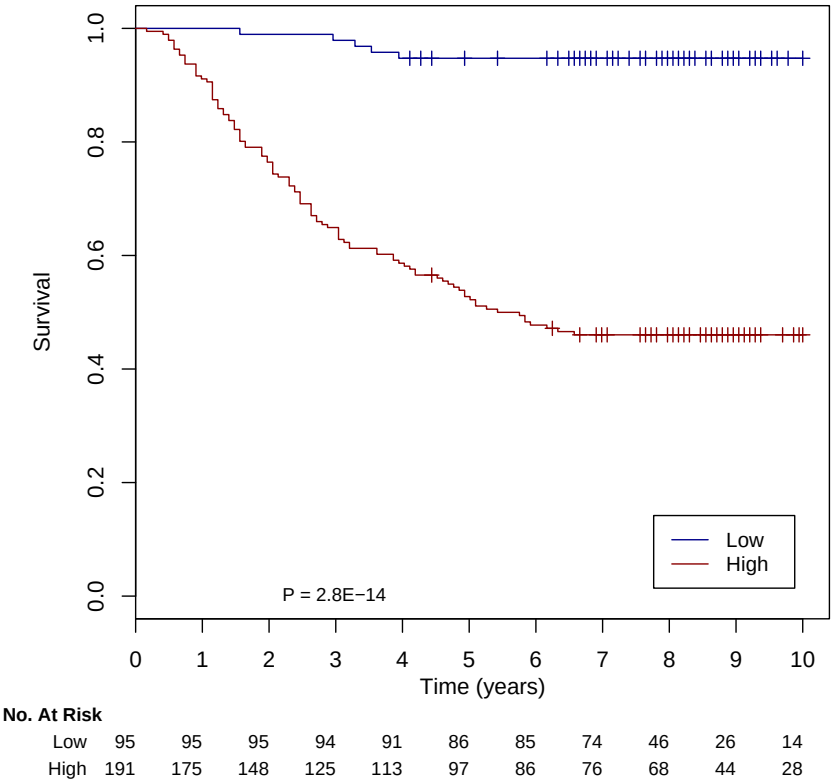



The following figure shows the distribution of the ggi :




**Risk Score** On TRAINING, the risk score exhibits a Cox p-value of 5.60E-25, a concordance index of 0.777 95CI[0.737,0.817] (p-value of 9.72E-42) and an integrated Brier score of 0.138.

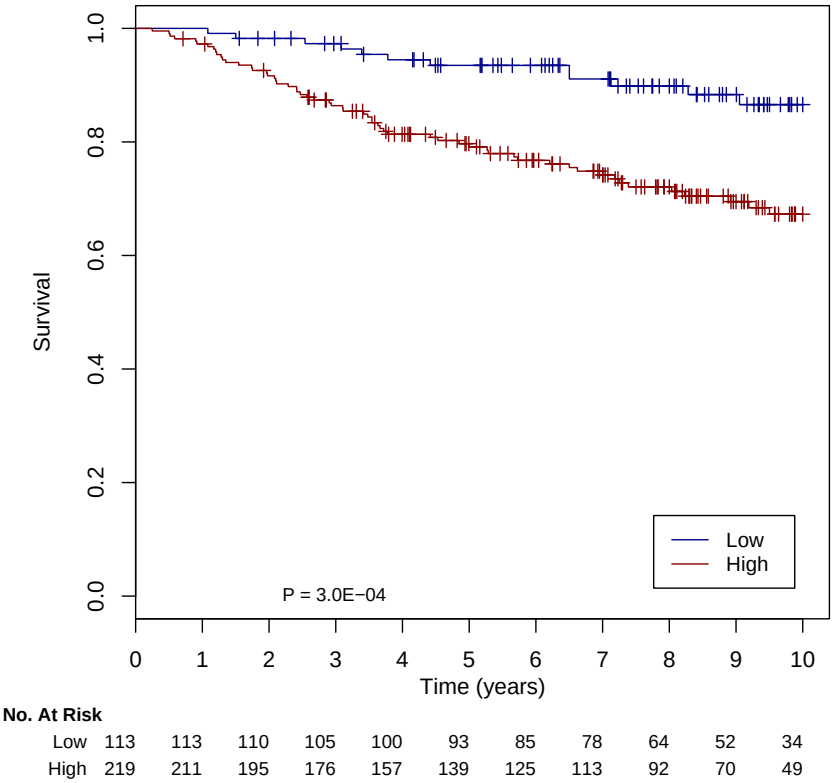
**Risk Group** On TRAINING, the binary classification computed from the risk score exhibits a Cox p-value of 3.86E-18, a concordance index of 0.938 95CI[0.885,0.99] (p-value of 9.53E-61) and an integrated Brier score of 0.155. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.97 [0.93,1.00] | 0.95 [0.90,0.99] | 0.95 [0.90,0.99] |
| High | 0.63 [0.56,0.70] | 0.52 [0.46,0.60] | 0.46 [0.39,0.54] |

**Risk Score** On TEST, the risk score exhibits a Cox p-value of 3.39E-03, a concordance index of 0.63 95CI[0.57,0.691] (p-value of 1.20E-05) and an integrated Brier score of 0.161.

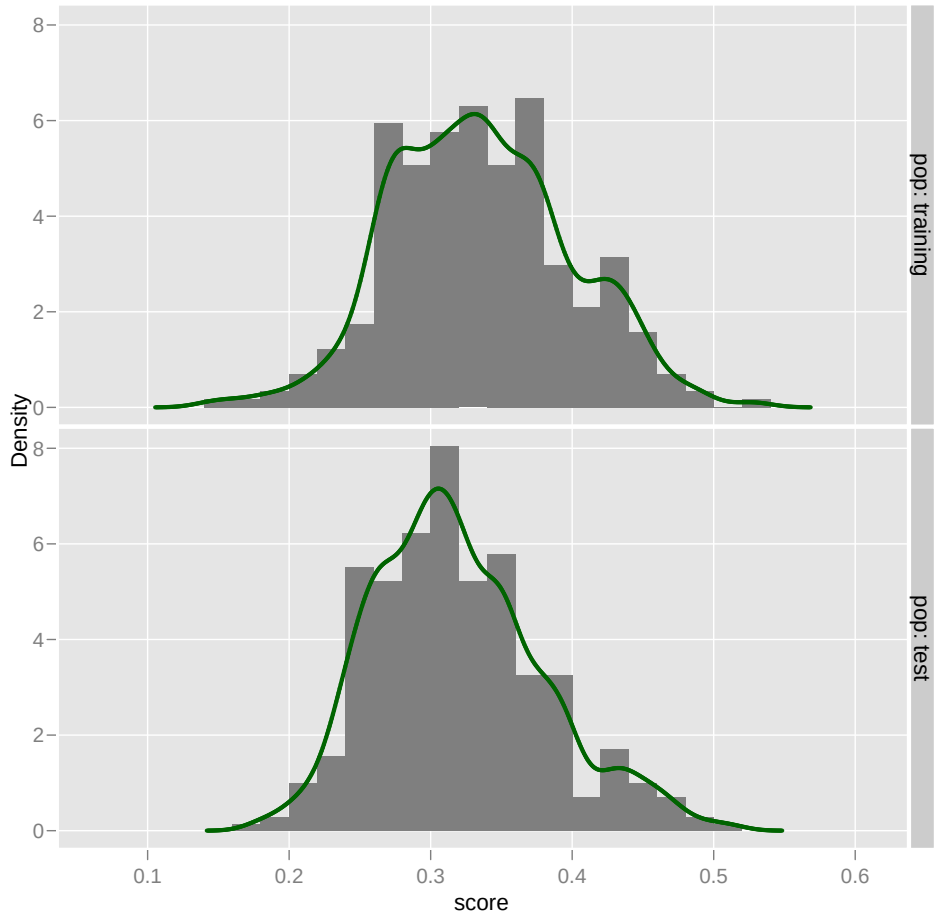
**Risk Group** On TEST, the binary classification computed from the risk score exhibits a Cox p-value of 8.71E-05, a concordance index of 0.766 95CI[0.651,0.88] (p-value of 2.90E-06) and an integrated Brier score of 0.148. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.97 [0.94,1.00] | 0.93 [0.89,0.98] | 0.87 [0.80,0.94] |
| High | 0.86 [0.81,0.91] | 0.79 [0.74,0.85] | 0.67 [0.60,0.75] |

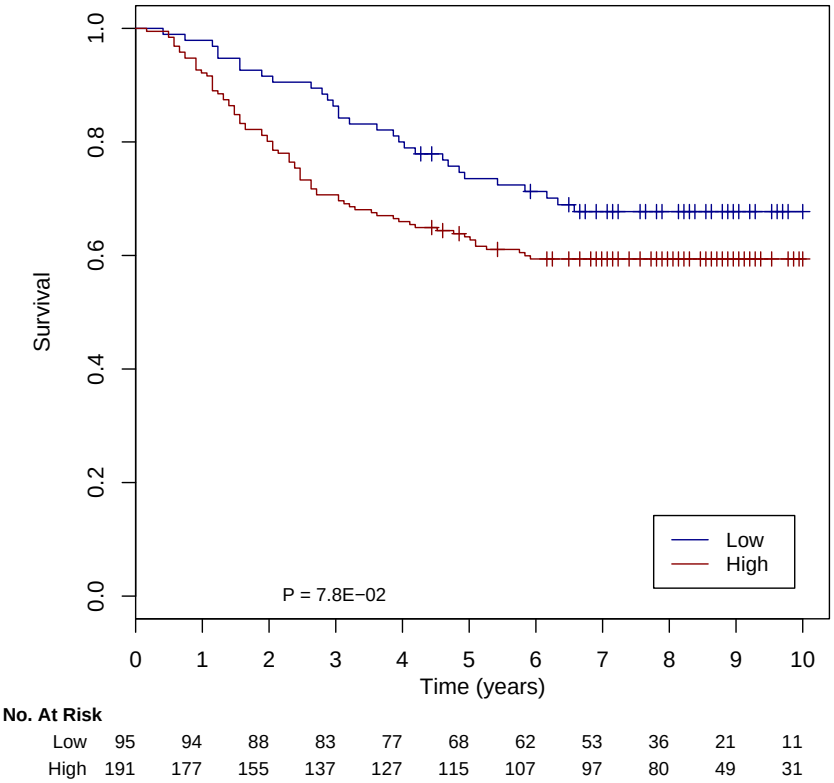
## 1.13 GW.PCA.COMBUNIV.WILCOXON.HG

The following figure shows the distribution of the ggi in the two populations :




**Risk Score** On TRAINING, the risk score exhibits a Cox p-value of 1.88E-02, a concordance index of 0.587 95CI[0.533,0.64] (p-value of 7.49E-04) and an integrated Brier score of 0.186.

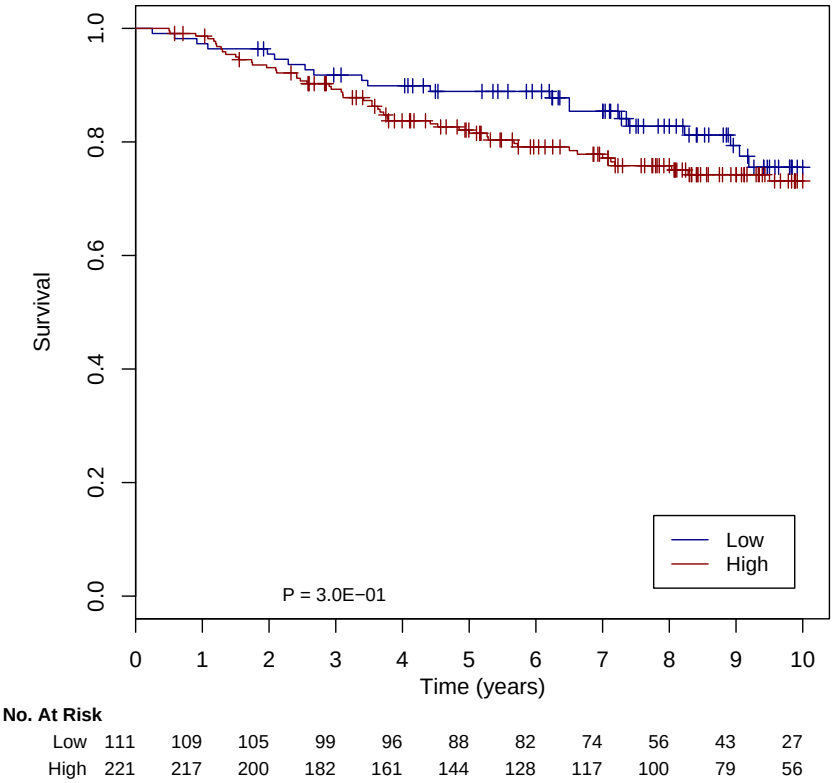
**Risk Group** On TRAINING, the binary classification computed from the risk score exhibits a Cox p-value of 7.19E-02, a concordance index of 0.616 95CI[0.517,0.715] (p-value of 1.08E-02) and an integrated Brier score of 0.187. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.84 [0.77,0.92] | 0.72 [0.64,0.82] | 0.68 [0.59,0.78] |
| High | 0.70 [0.63,0.76] | 0.63 [0.56,0.70] | 0.59 [0.53,0.67] |

**Risk Score** On TEST, the risk score exhibits a Cox p-value of 6.71E-02, a concordance index of 0.566 95CI[0.492,0.64] (p-value of 3.98E-02) and an integrated Brier score of 0.136.

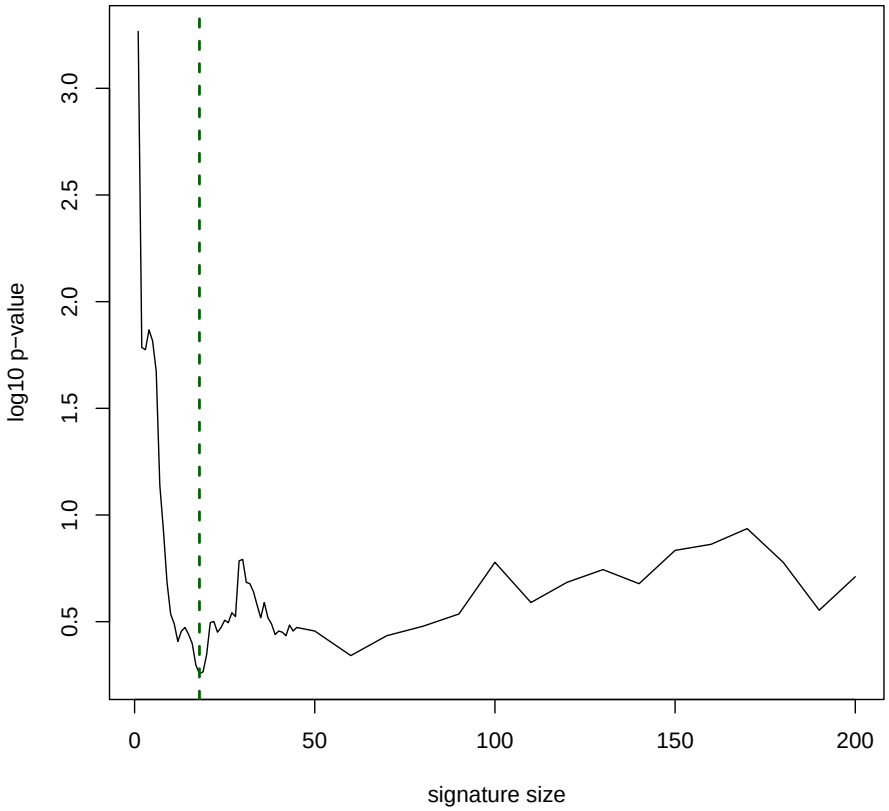
**Risk Group** On TEST, the binary classification computed from the risk score exhibits a Cox p-value of 2.99E-01, a concordance index of 0.589 95CI[0.461,0.716] (p-value of 8.74E-02) and an integrated Brier score of 0.14. The following figure shows the Kaplan-Meier survival curves for the two groups :



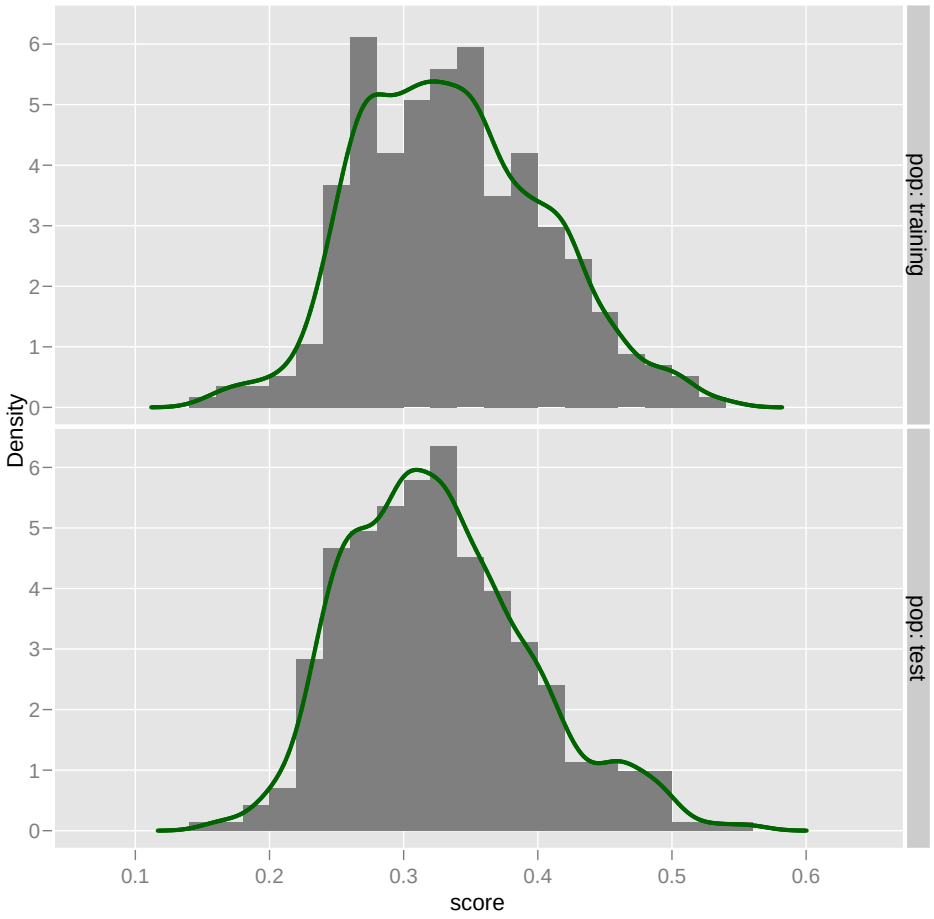

| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.92 [0.87,0.97] | 0.89 [0.83,0.95] | 0.76 [0.66,0.86] |
| High | 0.89 [0.85,0.93] | 0.82 [0.76,0.87] | 0.73 [0.67,0.80] |

## 1.14 GW.PCACV.COMBUNIV.WILCOXON.HG

The following figure shows the evolution of the performance w.r.t. the signature size (best signature size is 18):

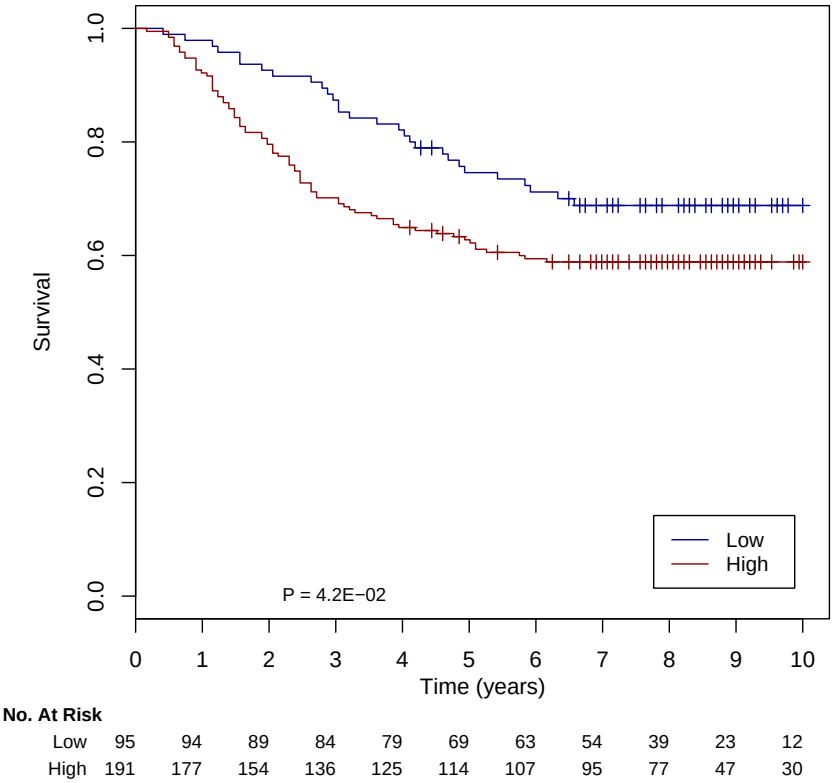



The following figure shows the distribution of the ggi in the two populations :




**Risk Score** On TRAINING, the risk score exhibits a Cox p-value of 9.69E-03, a concordance index of 0.595 95CI[0.542,0.649] (p-value of 2.26E-04) and an integrated Brier score of 0.185.

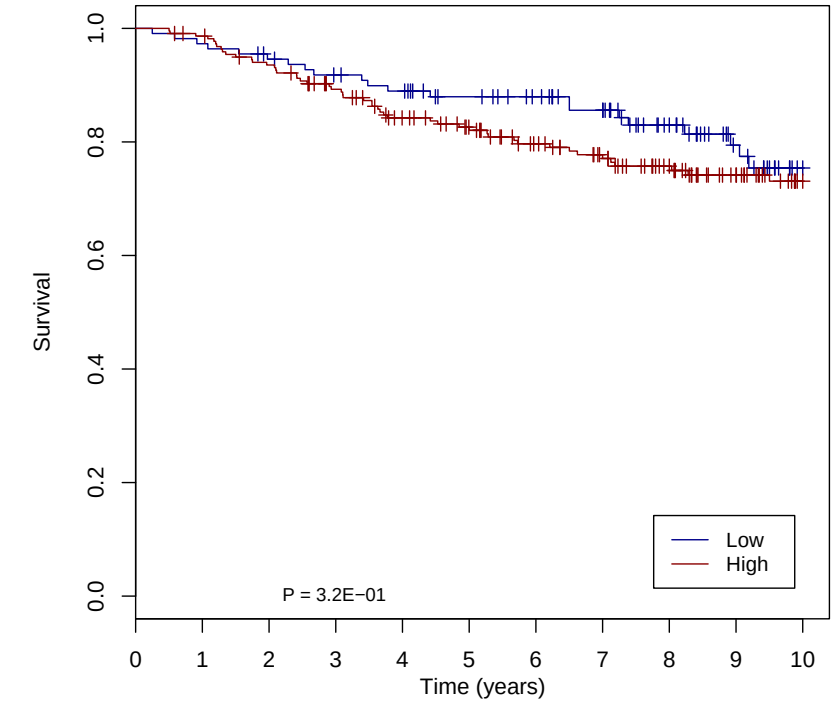
**Risk Group** On TRAINING, the binary classification computed from the risk score exhibits a Cox p-value of 3.68E-02, a concordance index of 0.633 95CI[0.535,0.731] (p-value of 3.97E-03) and an integrated Brier score of 0.186. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.85 [0.78,0.93] | 0.73 [0.65,0.83] | 0.69 [0.60,0.79] |
| High | 0.69 [0.63,0.76] | 0.62 [0.56,0.69] | 0.59 [0.52,0.66] |

**Risk Score** On TEST, the risk score exhibits a Cox p-value of 4.49E-02, a concordance index of 0.575 95CI[0.502,0.648] (p-value of 2.22E-02) and an integrated Brier score of 0.137.

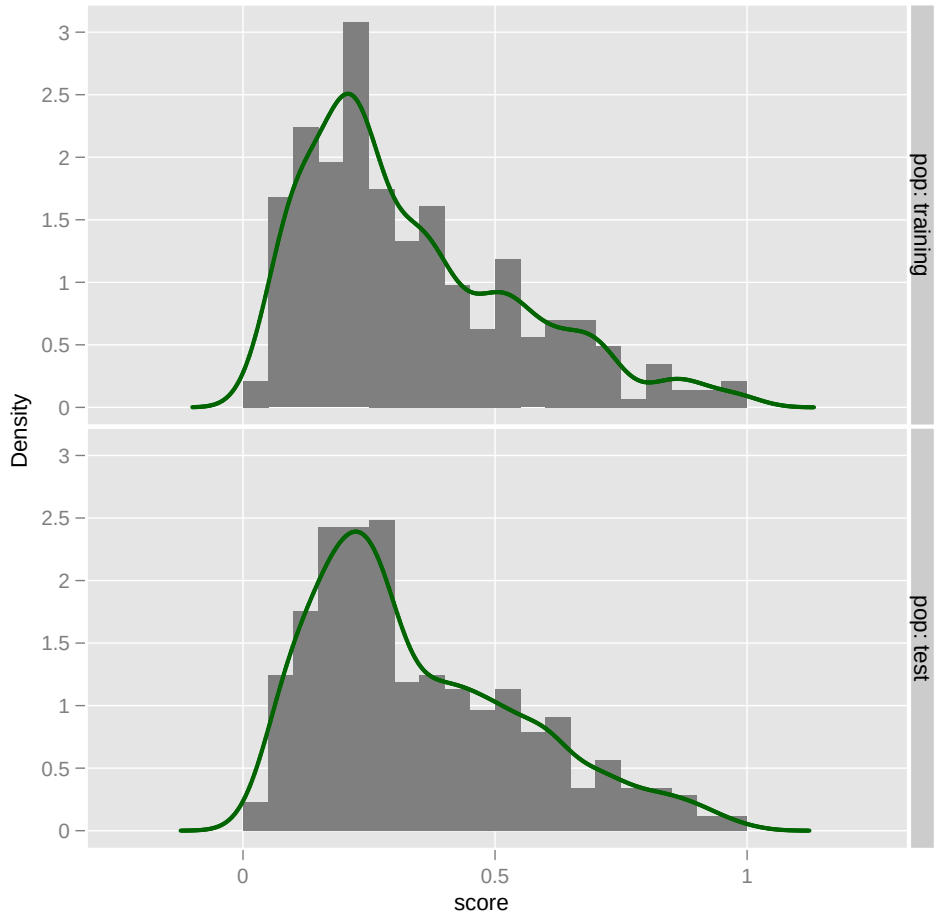
**Risk Group** On TEST, the binary classification computed from the risk score exhibits a Cox p-value of 3.78E-01, a concordance index of 0.582 95CI[0.454,0.711] (p-value of 1.05E-01) and an integrated Brier score of 0.141. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.92 [0.87,0.97] | 0.88 [0.82,0.94] | 0.75 [0.66,0.86] |
| High | 0.89 [0.85,0.93] | 0.82 [0.77,0.87] | 0.73 [0.67,0.80] |

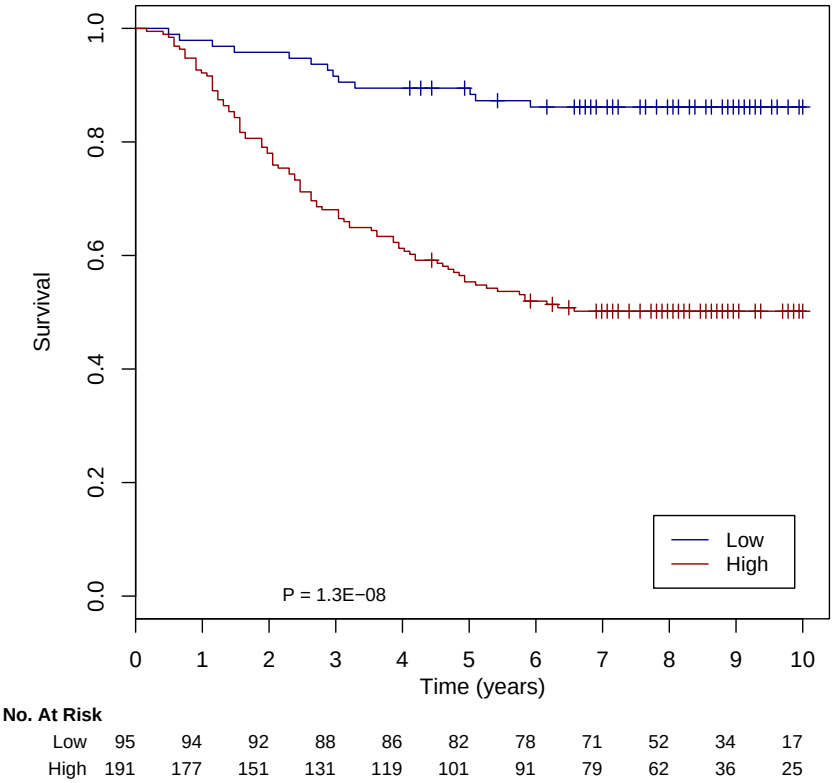
## 1.15 GW.PCA.COMBUNIV.COX.SURV

The following figure shows the distribution of the ggi in the two populations :




**Risk Score** On TRAINING, the risk score exhibits a Cox p-value of 2.42E-16, a concordance index of 0.727 95CI[0.68,0.774] (p-value of 2.68E-21) and an integrated Brier score of 0.154.

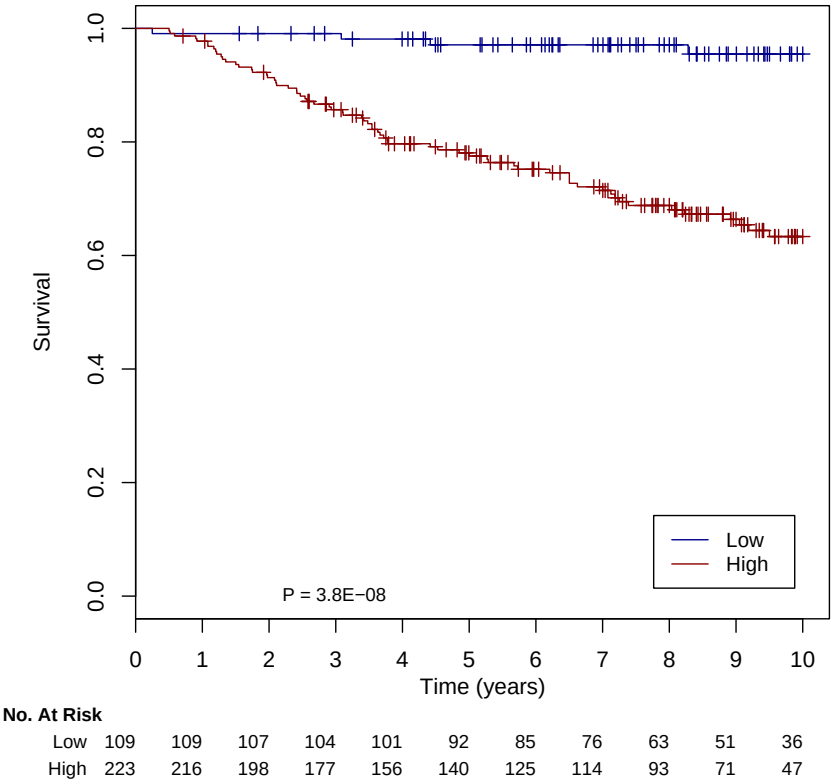
**Risk Group** On TRAINING, the binary classification computed from the risk score exhibits a Cox p-value of 6.86E-10, a concordance index of 0.824 95CI[0.738,0.909] (p-value of 6.68E-14) and an integrated Brier score of 0.168. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.91 [0.85,0.97] | 0.88 [0.82,0.95] | 0.86 [0.79,0.93] |
| High | 0.66 [0.60,0.74] | 0.55 [0.48,0.62] | 0.50 [0.44,0.58] |

**Risk Score** On TEST, the risk score exhibits a Cox p-value of 9.14E-06, a concordance index of 0.695 95CI[0.639,0.751] (p-value of 4.18E-12) and an integrated Brier score of 0.153.

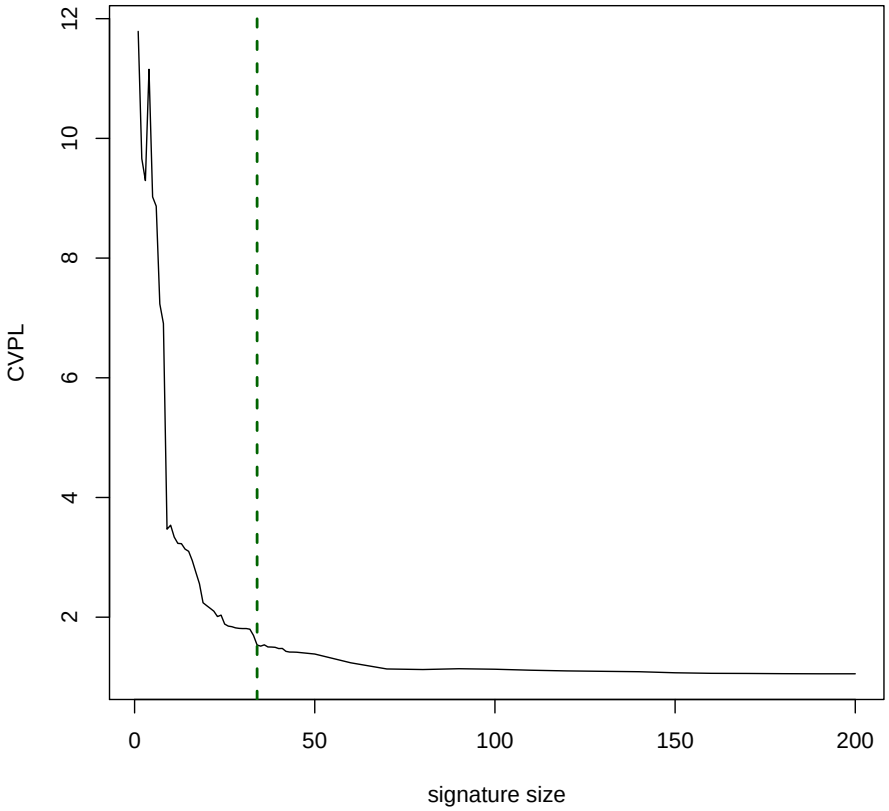
**Risk Group** On TEST, the binary classification computed from the risk score exhibits a Cox p-value of 1.02E-09, a concordance index of 0.909 95CI[0.822,0.996] (p-value of 2.00E-20) and an integrated Brier score of 0.133. The following figure shows the Kaplan-Meier survival curves for the two groups :



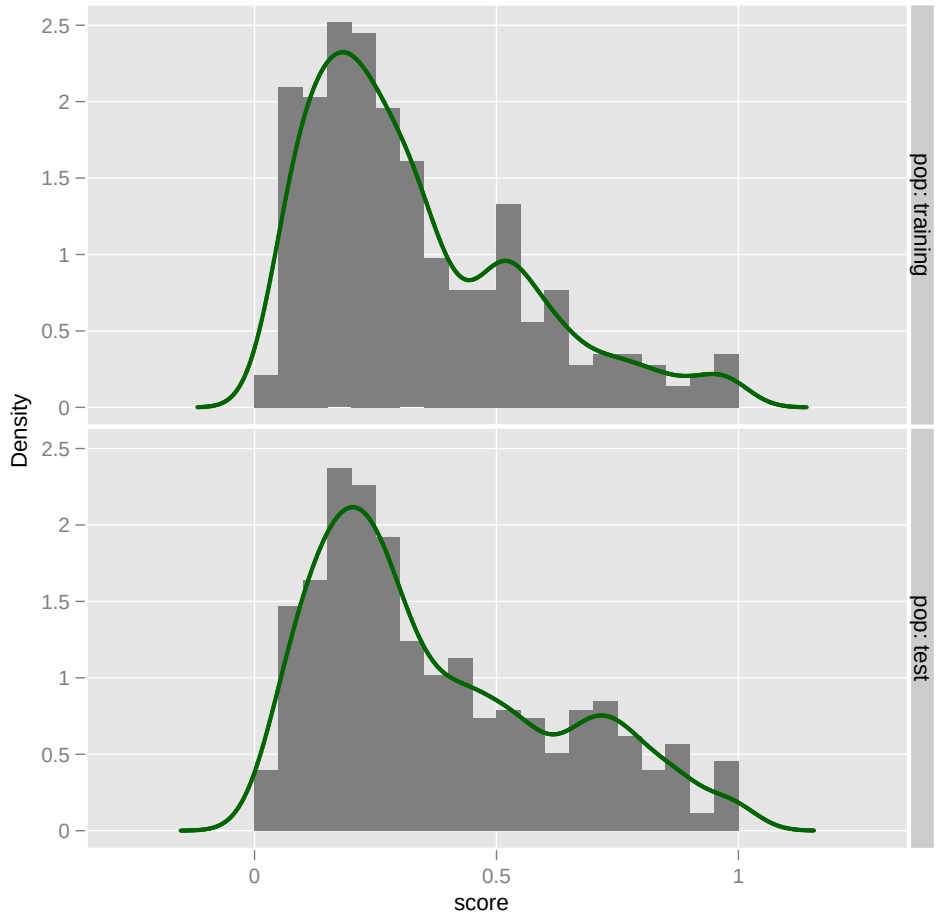

| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.98 [0.96,1.00] | 0.97 [0.94,1.00] | 0.95 [0.91,1.00] |
| High | 0.86 [0.81,0.90] | 0.78 [0.72,0.83] | 0.63 [0.56,0.71] |

## 1.16 GW.PCACV.COMBUNIV.COX.SURV

The following figure shows the evolution of the performance w.r.t. the signature size (best signature size is 34):

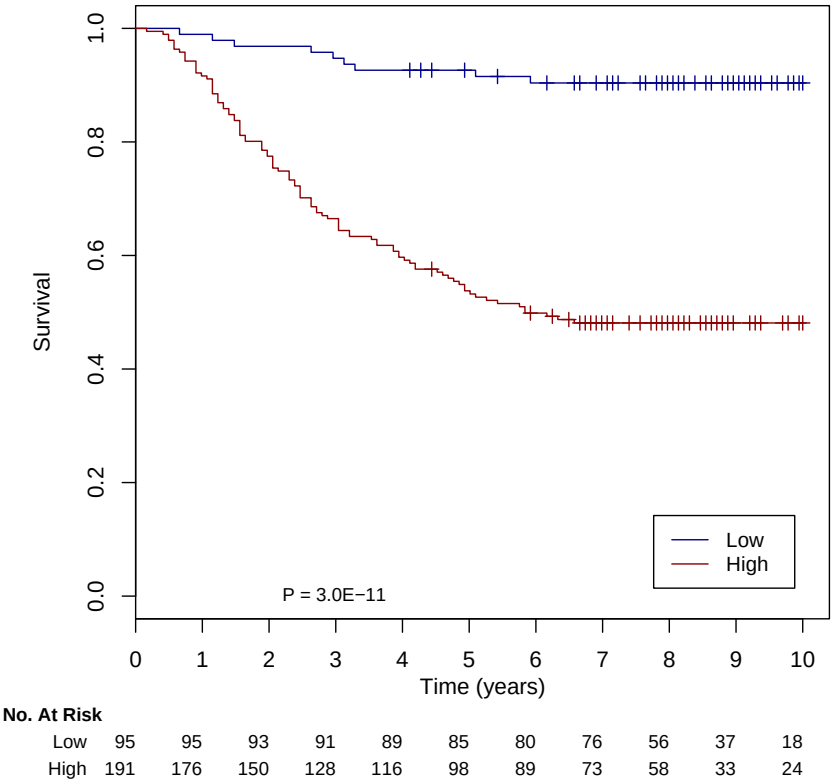



The following figure shows the distribution of the ggi in the two populations :




**Risk Score** On TRAINING, the risk score exhibits a Cox p-value of 1.15E-18, a concordance index of 0.748 95CI[0.702,0.793] (p-value of 6.86E-27) and an integrated Brier score of 0.149.

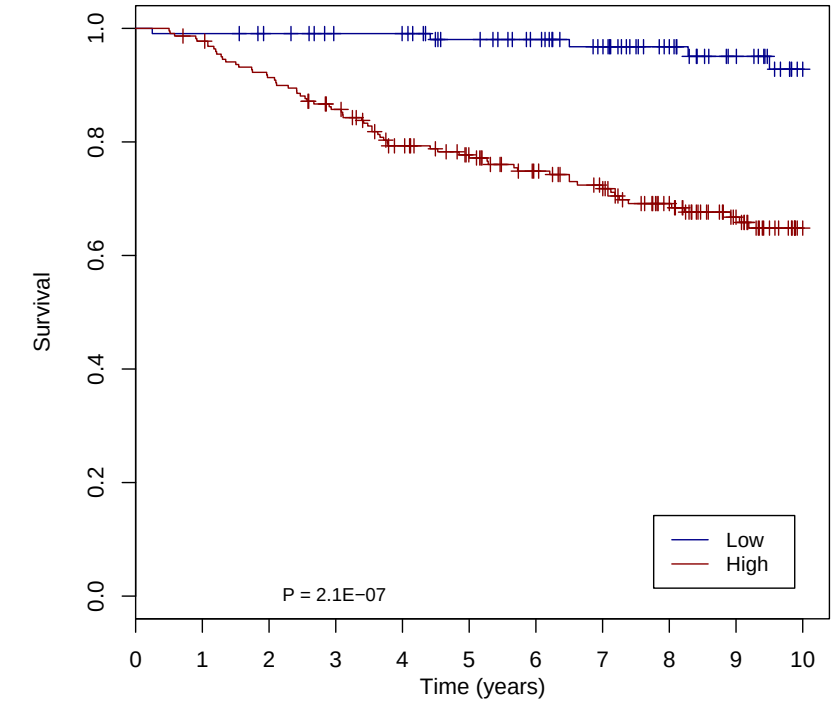
**Risk Group** On TRAINING, the binary classification computed from the risk score exhibits a Cox p-value of 1.68E-13, a concordance index of 0.881 95CI[0.808,0.954] (p-value of 8.52E-25) and an integrated Brier score of 0.162. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.94 [0.89,0.99] | 0.92 [0.86,0.97] | 0.90 [0.85,0.97] |
| High | 0.64 [0.58,0.72] | 0.53 [0.47,0.61] | 0.48 [0.41,0.56] |

**Risk Score** On TEST, the risk score exhibits a Cox p-value of 5.90E-06, a concordance index of 0.698 95CI[0.644,0.753] (p-value of 4.76E-13) and an integrated Brier score of 0.17.

**Risk Group** On TEST, the binary classification computed from the risk score exhibits a Cox p-value of 7.69E-09, a concordance index of 0.907 95CI[0.824,0.99] (p-value of 3.55E-22) and an integrated Brier score of 0.137. The following figure shows the Kaplan-Meier survival curves for the two groups :



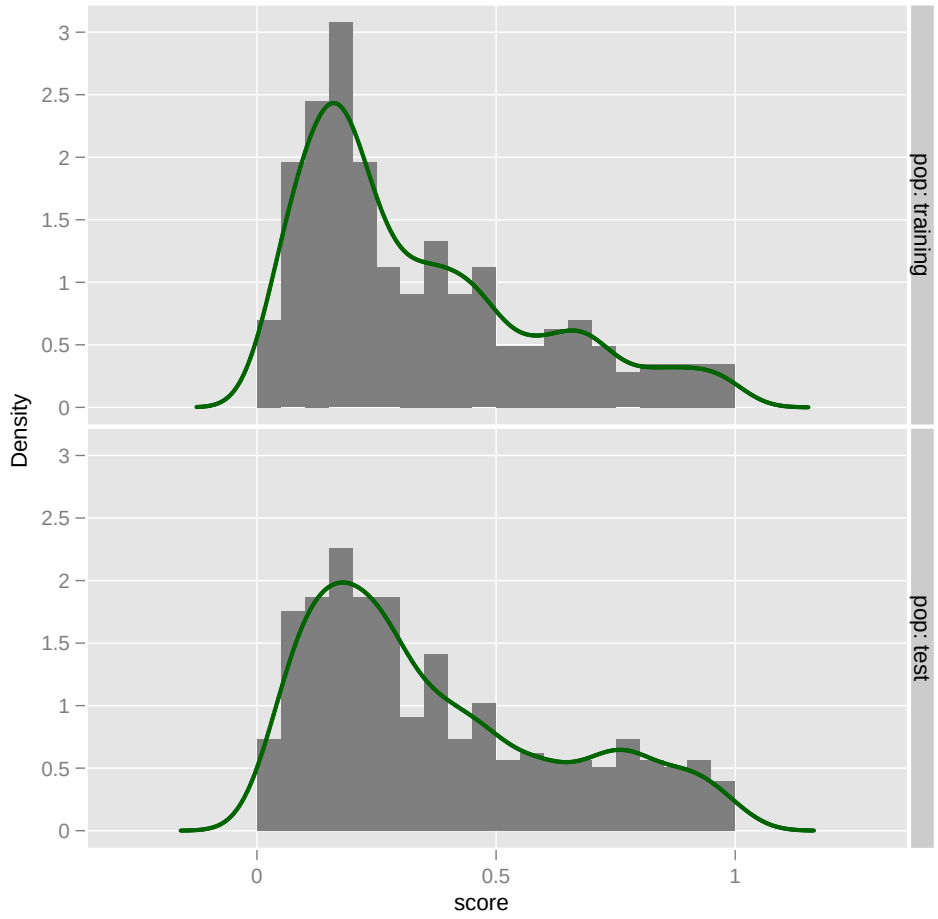

**No. At Risk**

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Low | 109 | 109 | 106 | 101 | 100 | 91 | 84 | 75 | 61 | 49 | 34 |
| High | 223 | 216 | 199 | 180 | 157 | 141 | 126 | 115 | 95 | 73 | 49 |

| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.99 [0.97,1.00] | 0.98 [0.95,1.00] | 0.93 [0.87,0.99] |
| High | 0.86 [0.81,0.91] | 0.77 [0.72,0.83] | 0.65 [0.58,0.72] |

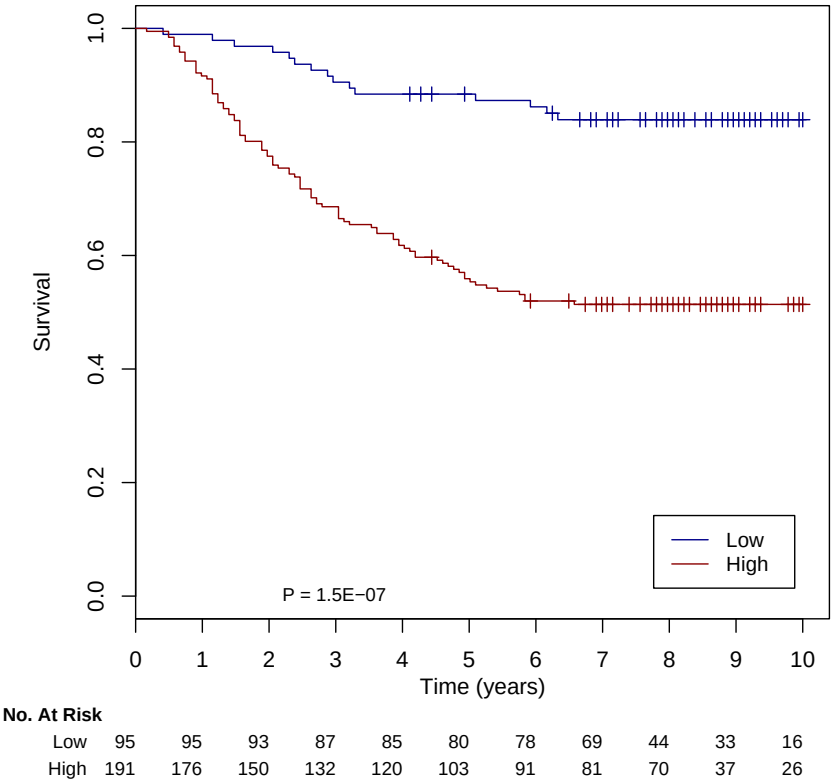
## 1.17 GW.PCA.MULTIV.RCOX.SURV

The following figure shows the distribution of the ggi :




**Risk Score** On TRAINING, the risk score exhibits a Cox p-value of 3.89E-21, a concordance index of 0.75 95CI[0.706,0.794] (p-value of 4.71E-29) and an integrated Brier score of 0.143.

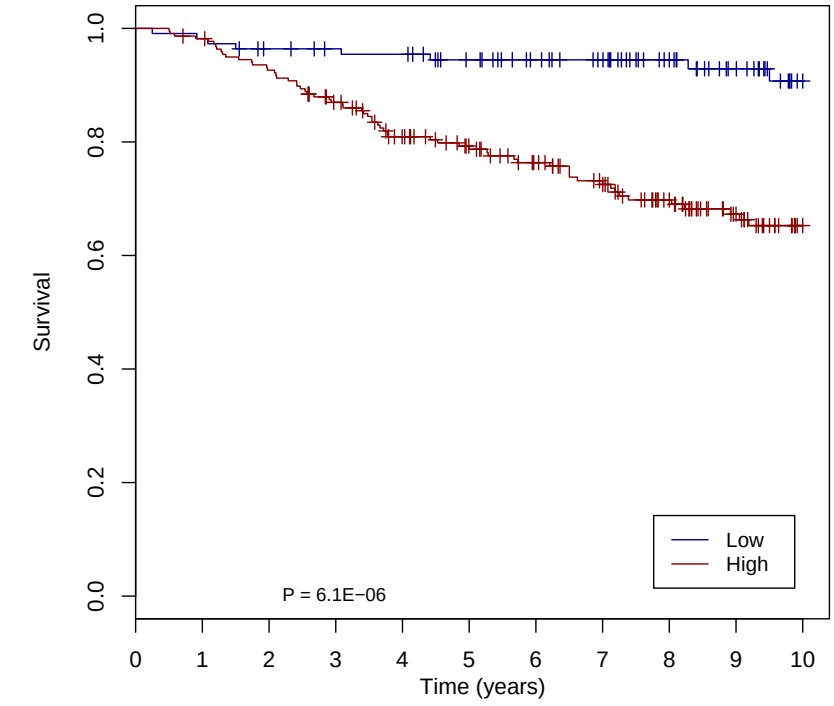
**Risk Group** On TRAINING, the binary classification computed from the risk score exhibits a Cox p-value of 1.57E-08, a concordance index of 0.806 95CI[0.719,0.892] (p-value of 2.21E-12) and an integrated Brier score of 0.171. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.89 [0.84,0.96] | 0.87 [0.81,0.94] | 0.84 [0.77,0.92] |
| High | 0.66 [0.60,0.74] | 0.55 [0.49,0.63] | 0.51 [0.45,0.59] |

**Risk Score** On TEST, the risk score exhibits a Cox p-value of 7.76E-06, a concordance index of 0.69 95CI[0.631,0.748] (p-value of 1.08E-10) and an integrated Brier score of 0.171.

**Risk Group** On TEST, the binary classification computed from the risk score exhibits a Cox p-value of 8.87E-07, a concordance index of 0.818 95CI[0.702,0.934] (p-value of 4.09E-08) and an integrated Brier score of 0.136. The following figure shows the Kaplan-Meier survival curves for the two groups :



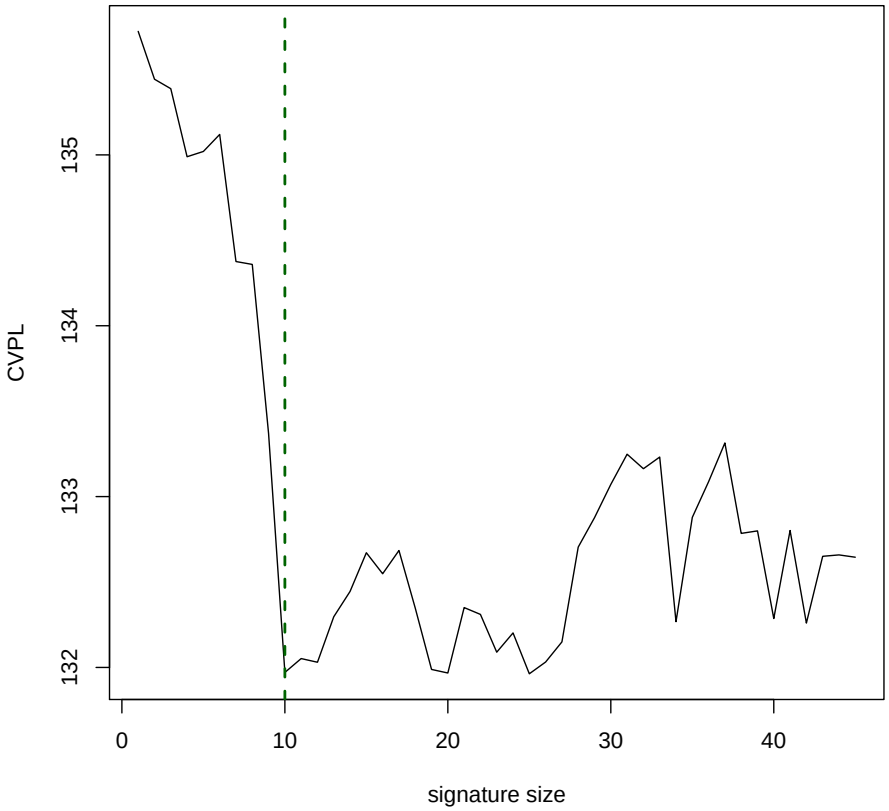

**No. At Risk**

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Low | 111 | 110 | 105 | 102 | 101 | 92 | 84 | 78 | 64 | 52 | 35 |
| High | 221 | 216 | 200 | 179 | 156 | 140 | 126 | 112 | 92 | 70 | 48 |

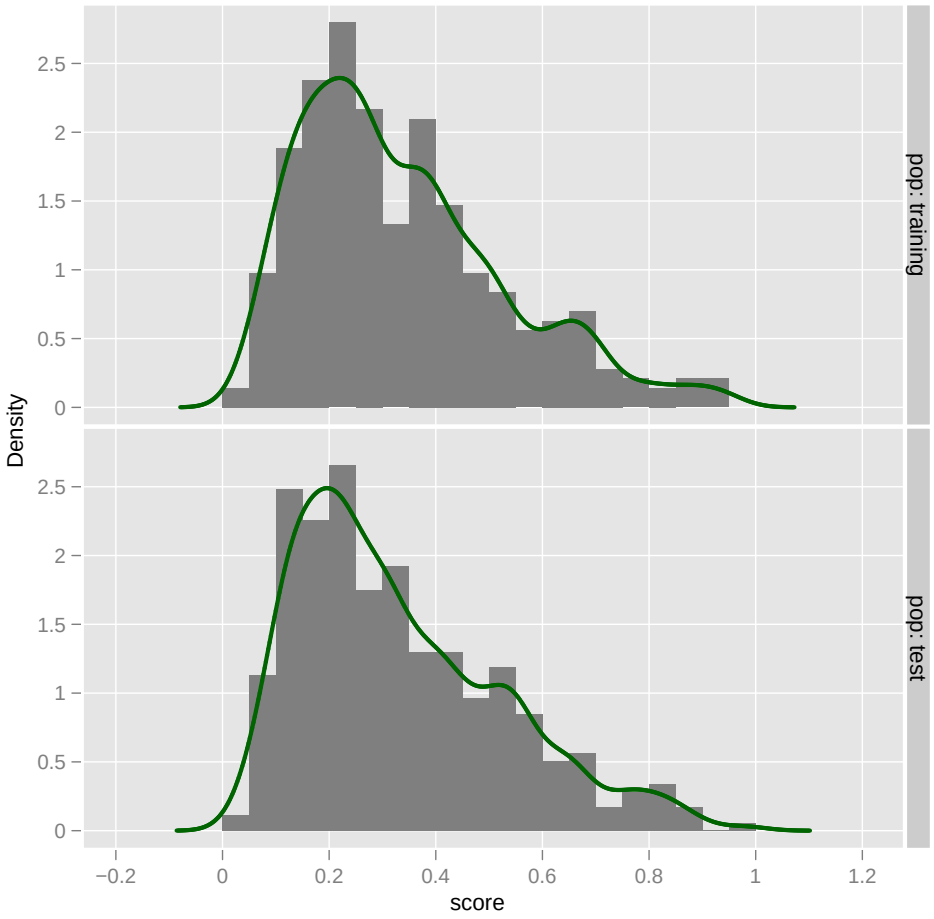
| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.95 [0.92,0.99] | 0.94 [0.90,0.99] | 0.91 [0.84,0.98] |
| High | 0.87 [0.83,0.92] | 0.79 [0.73,0.85] | 0.65 [0.58,0.73] |

## 1.18 GW.PCACV.MULTIV.RCOX.SURV

The following figure shows the evolution of the performance w.r.t. the signature size (best signature size is 10):

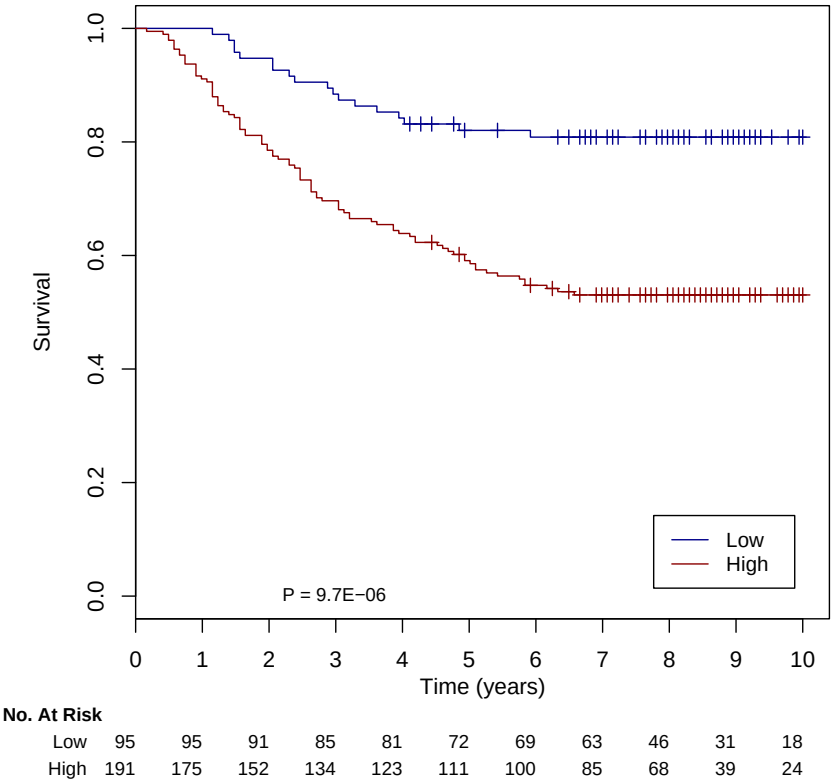



The following figure shows the distribution of the ggi :




**Risk Score** On TRAINING, the risk score exhibits a Cox p-value of 4.11E-13, a concordance index of 0.686 95CI[0.635,0.736] (p-value of 2.42E-13) and an integrated Brier score of 0.161.

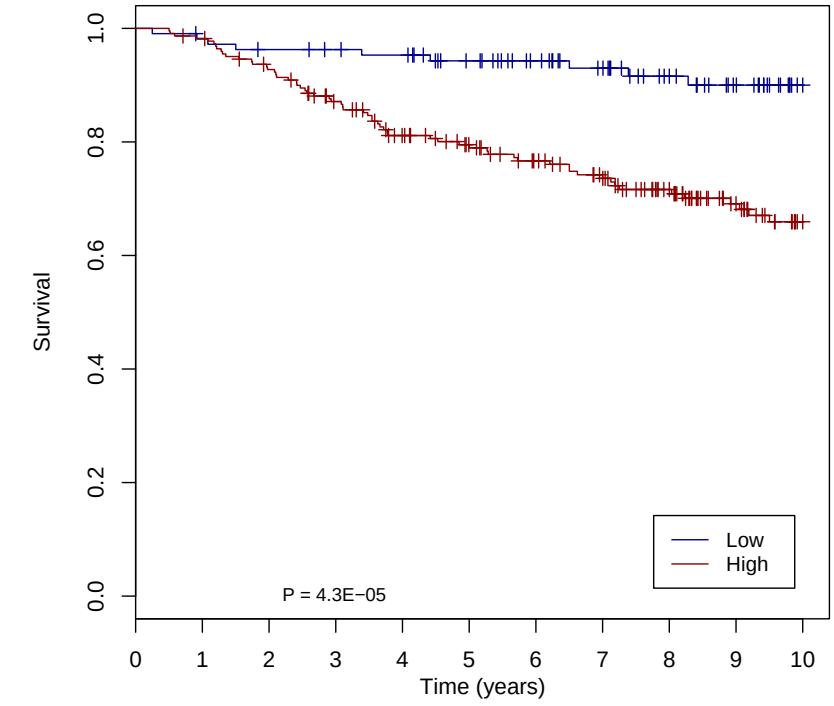
**Risk Group** On TRAINING, the binary classification computed from the risk score exhibits a Cox p-value of 2.65E-06, a concordance index of 0.754 95CI[0.66,0.848] (p-value of 6.18E-08) and an integrated Brier score of 0.178. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.87 [0.81,0.94] | 0.82 [0.75,0.90] | 0.81 [0.73,0.89] |
| High | 0.68 [0.62,0.75] | 0.59 [0.52,0.66] | 0.53 [0.46,0.61] |

**Risk Score** On TEST, the risk score exhibits a Cox p-value of 2.32E-04, a concordance index of 0.657 95CI[0.594,0.719] (p-value of 4.85E-07) and an integrated Brier score of 0.148.

**Risk Group** On TEST, the binary classification computed from the risk score exhibits a Cox p-value of 2.13E-05, a concordance index of 0.789 95CI[0.668,0.91] (p-value of 1.42E-06) and an integrated Brier score of 0.136. The following figure shows the Kaplan-Meier survival curves for the two groups :



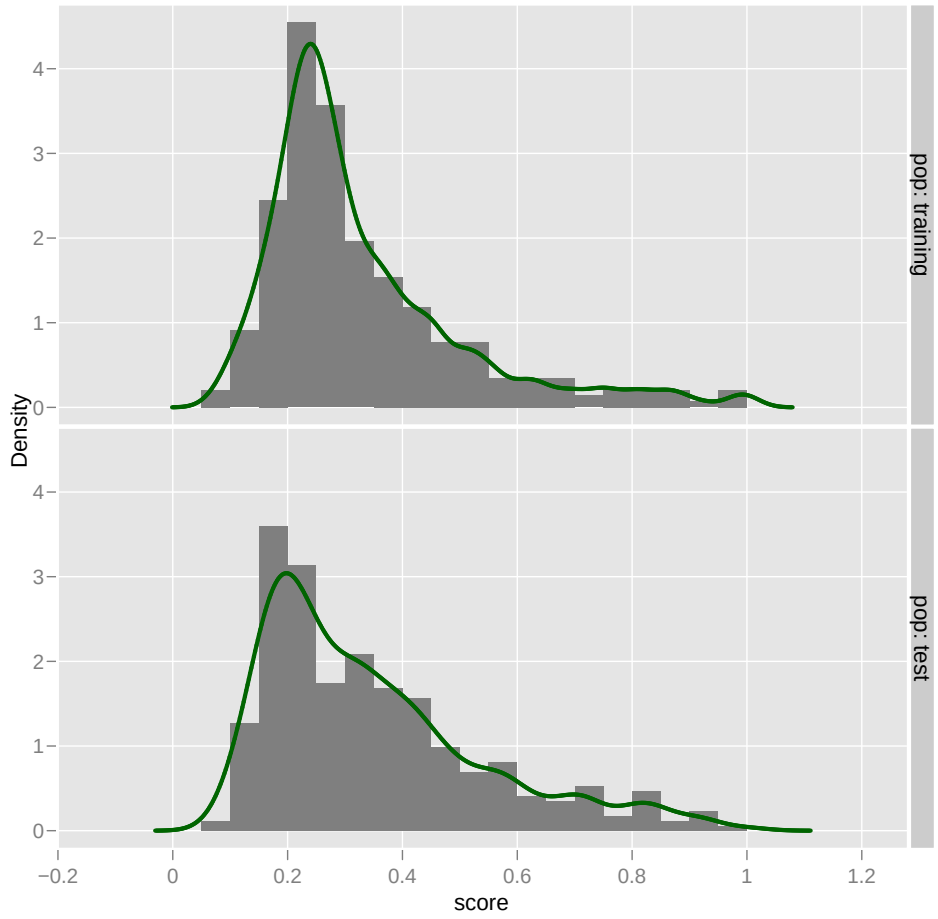

**No. At Risk**

| | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| Low | 108 | 106 | 103 | 101 | 99 | 89 | 80 | 72 | 60 | 51 | 34 |
| High | 224 | 219 | 202 | 180 | 158 | 143 | 130 | 118 | 96 | 71 | 49 |

| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.96 [0.93,1.00] | 0.94 [0.90,0.99] | 0.90 [0.84,0.97] |
| High | 0.87 [0.82,0.91] | 0.79 [0.74,0.85] | 0.66 [0.59,0.74] |

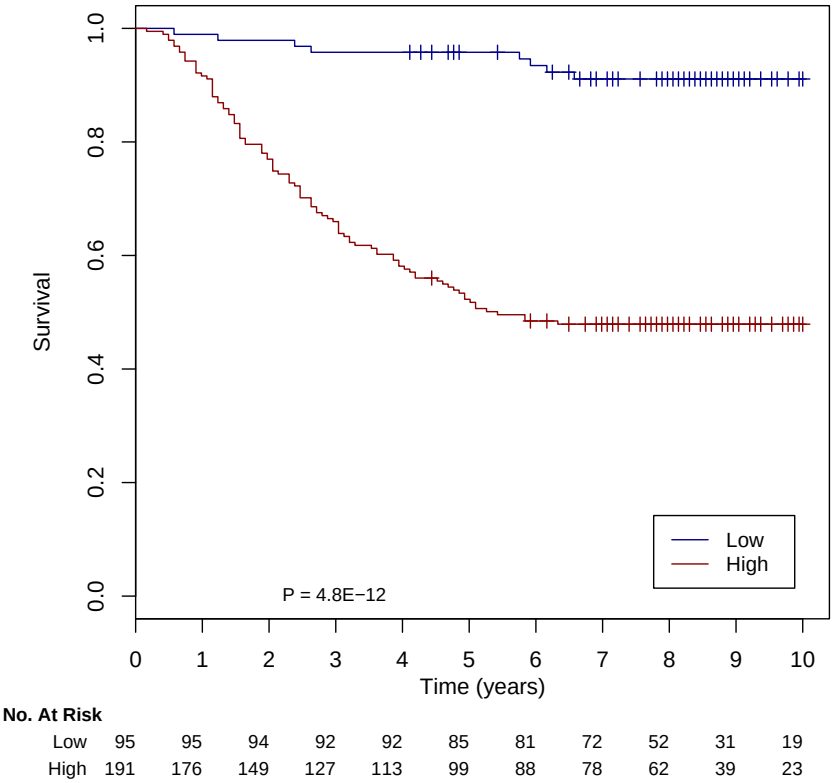
## 1.19 GENE76

The following figure shows the distribution of the ggi :




**Risk Score** On TRAINING, the risk score exhibits a Cox p-value of 2.63E-14, a concordance index of 0.754 95CI[0.713,0.795] (p-value of 2.22E-34) and an integrated Brier score of 0.158.

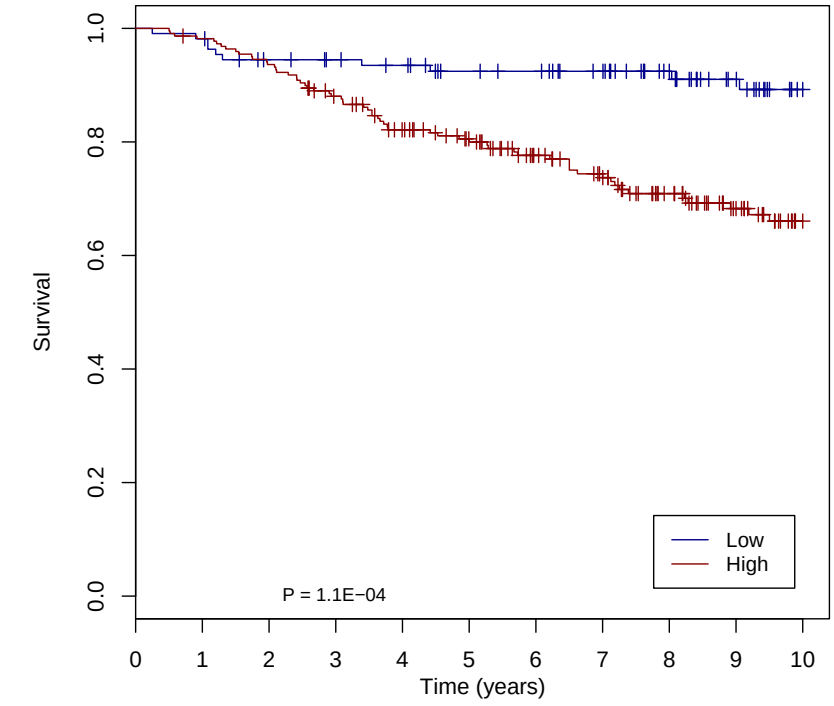
**Risk Group** On TRAINING, the binary classification computed from the risk score exhibits a Cox p-value of 1.21E-14, a concordance index of 0.903 95CI[0.838,0.969] (p-value of 4.23E-34) and an integrated Brier score of 0.16. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.96 [0.92,1.00] | 0.96 [0.92,1.00] | 0.91 [0.85,0.97] |
| High | 0.64 [0.57,0.71] | 0.52 [0.45,0.59] | 0.48 [0.41,0.56] |

**Risk Score** On TEST, the risk score exhibits a Cox p-value of 7.23E-06, a concordance index of 0.667 95CI[0.607,0.727] (p-value of 2.94E-08) and an integrated Brier score of 0.149.

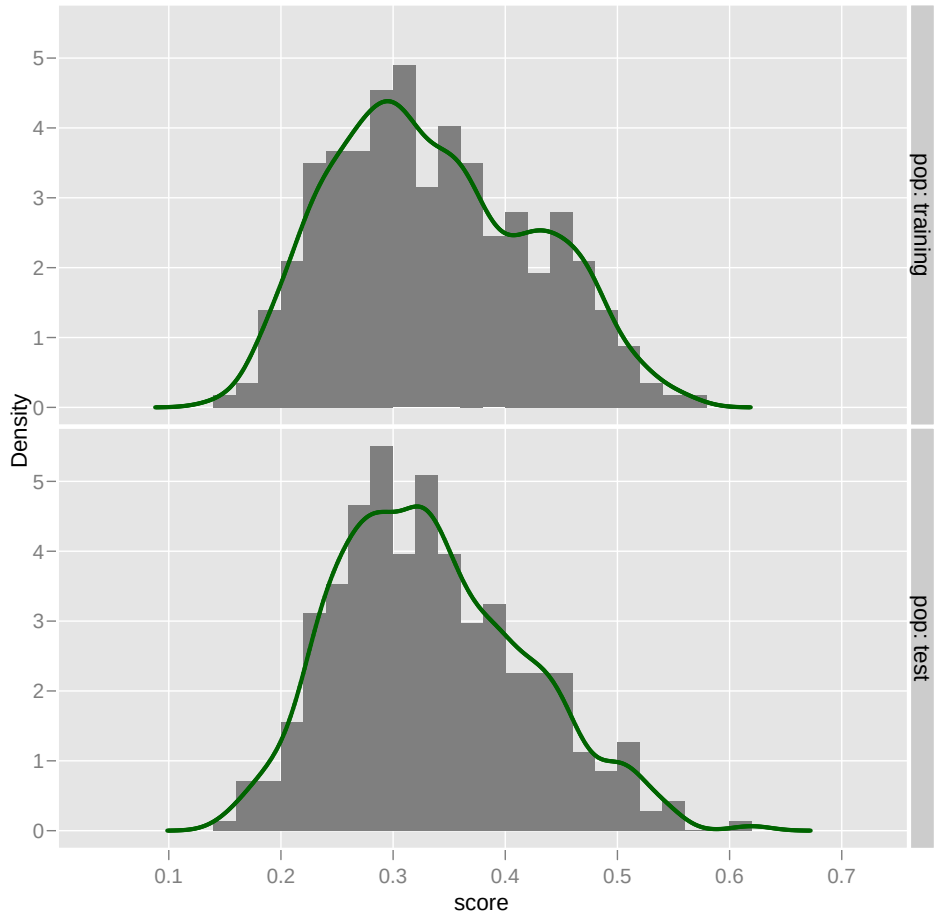
**Risk Group** On TEST, the binary classification computed from the risk score exhibits a Cox p-value of 2.71E-05, a concordance index of 0.754 95CI[0.624,0.884] (p-value of 6.38E-05) and an integrated Brier score of 0.145. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.94 [0.90,0.99] | 0.92 [0.88,0.98] | 0.89 [0.83,0.96] |
| High | 0.88 [0.83,0.92] | 0.80 [0.75,0.86] | 0.66 [0.59,0.74] |

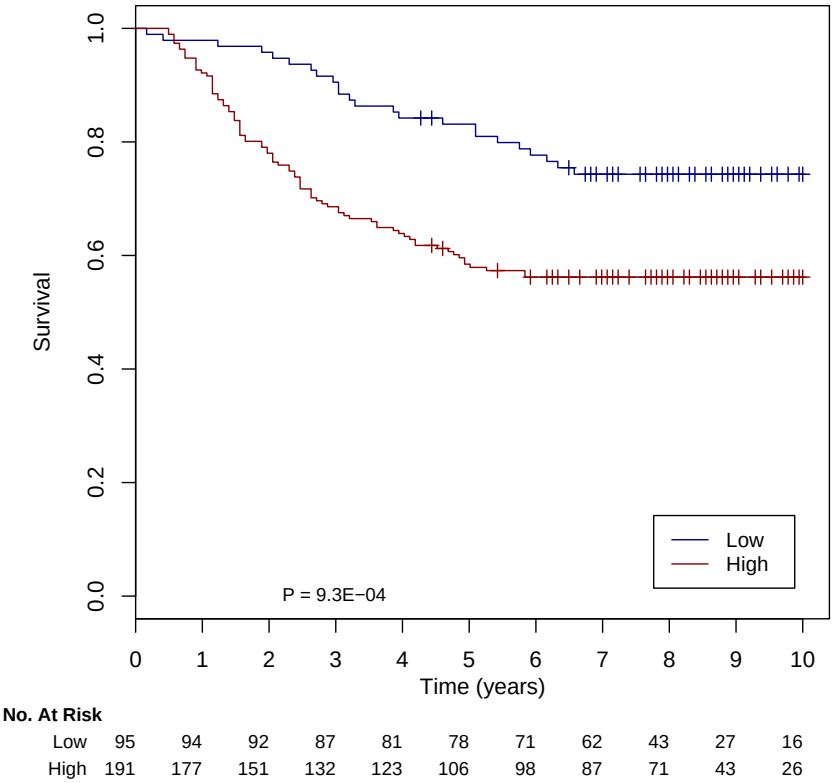
## 1.20 GGI

The following figure shows the distribution of the ggi :




**Risk Score** On TRAINING, the risk score exhibits a Cox p-value of 1.44E-03, a concordance index of 0.613 95CI[0.559,0.667] (p-value of 2.15E-05) and an integrated Brier score of 0.183.

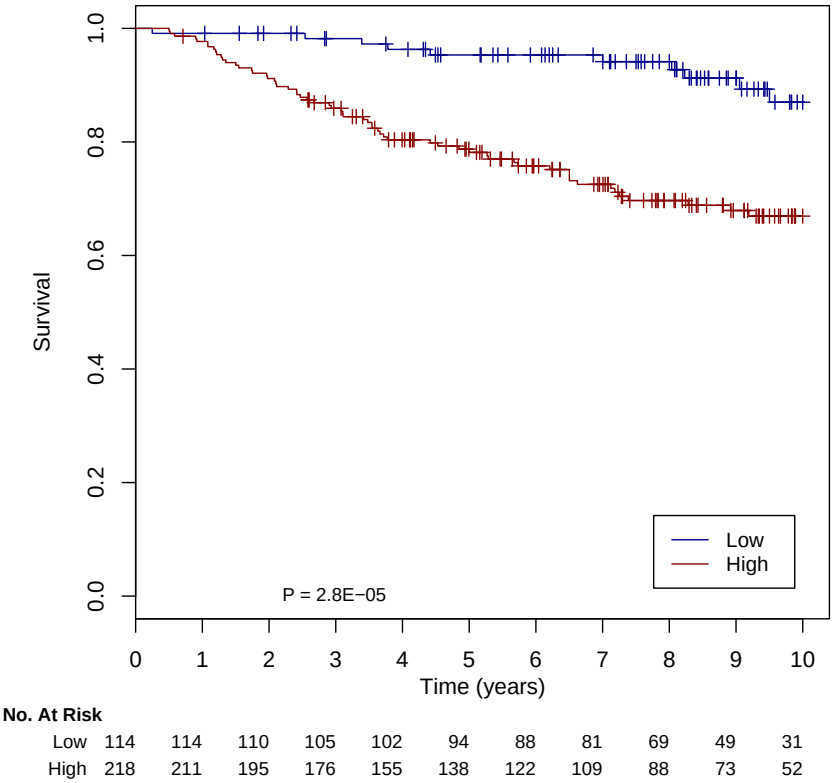
**Risk Group** On TRAINING, the binary classification computed from the risk score exhibits a Cox p-value of 5.72E-04, a concordance index of 0.706 95CI[0.611,0.8] (p-value of 9.82E-06) and an integrated Brier score of 0.181. The following figure shows the Kaplan-Meier survival curves for the two groups :




| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.88 [0.82,0.95] | 0.81 [0.73,0.89] | 0.74 [0.66,0.84] |
| High | 0.68 [0.61,0.75] | 0.58 [0.51,0.65] | 0.56 [0.50,0.64] |

**Risk Score** On TEST, the risk score exhibits a Cox p-value of 1.65E-09, a concordance index of 0.718 95CI[0.663,0.774] (p-value of 5.91E-15) and an integrated Brier score of 0.13.

**Risk Group** On TEST, the binary classification computed from the risk score exhibits a Cox p-value of 2.51E-06, a concordance index of 0.824 95CI[0.722,0.927] (p-value of 2.77E-10) and an integrated Brier score of 0.134. The following figure shows the Kaplan-Meier survival curves for the two groups :

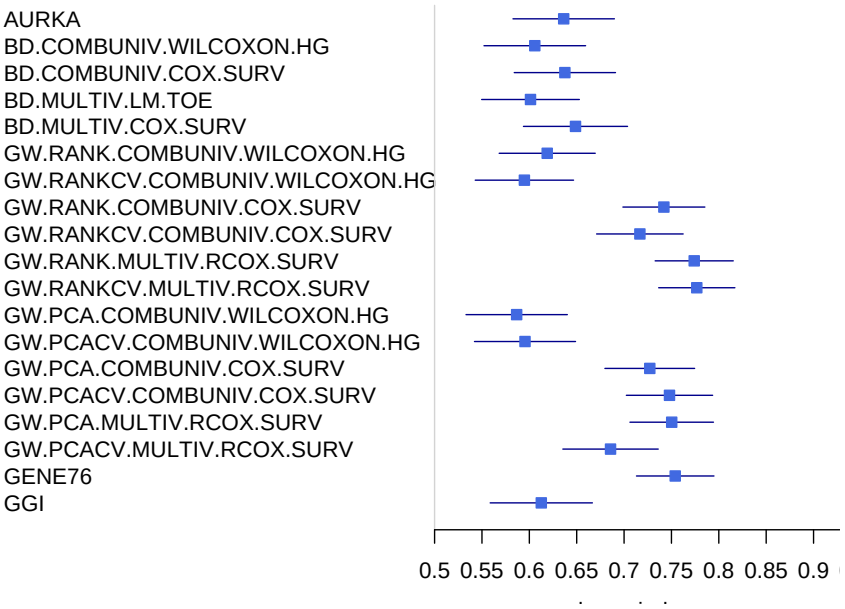



| | 3.years | 5.years | 10.years |
|---|---|---|---|
| Low | 0.97 [0.94,1.00] | 0.95 [0.91,0.99] | 0.87 [0.79,0.95] |
| High | 0.86 [0.81,0.91] | 0.78 [0.73,0.84] | 0.67 [0.60,0.74] |

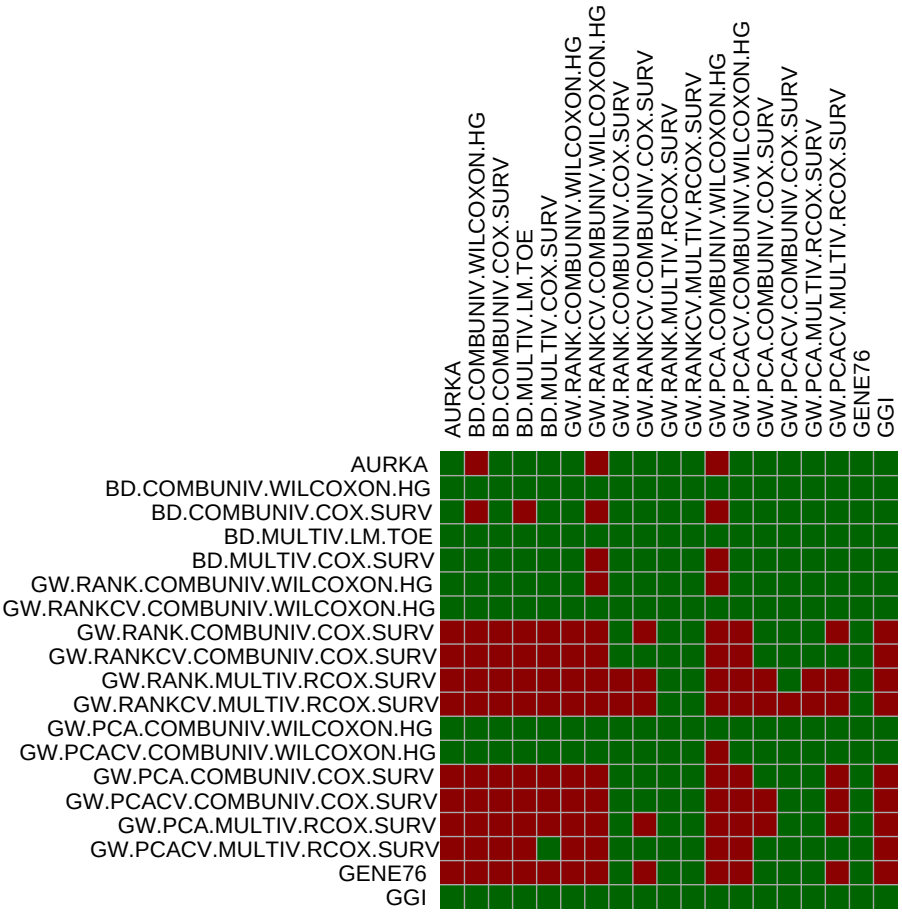
## 1.21 Model Comparison

### 1.21.1 Training Set

**Concordance Index for Risk Score** The following foresplot shows the concordance indices and their confidence interval for each method :



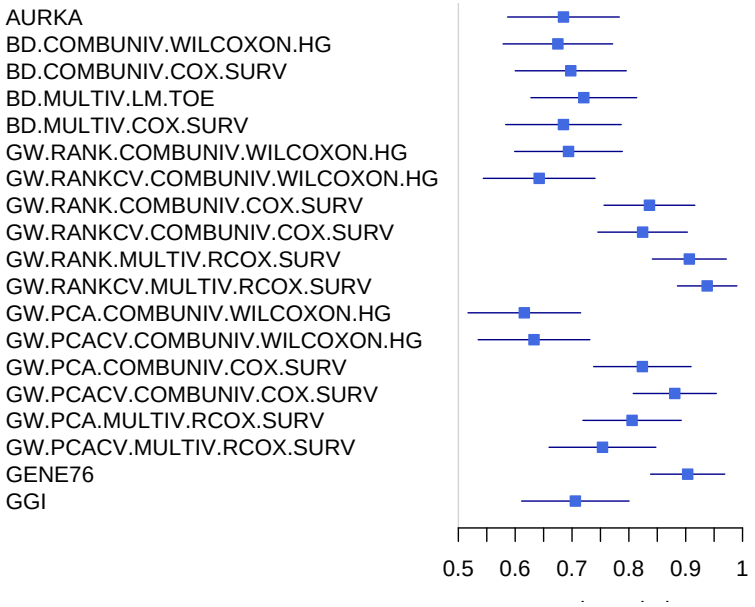

The following figure shows the pairwise comparison between the different methods. A red box represents a significant superiority (Student t test p-value $< 0.05$) of the method in row against the method in column. A green box represents no significant superiority (Student t test p-value $>= 0.05$) of the method in row against the method in column.



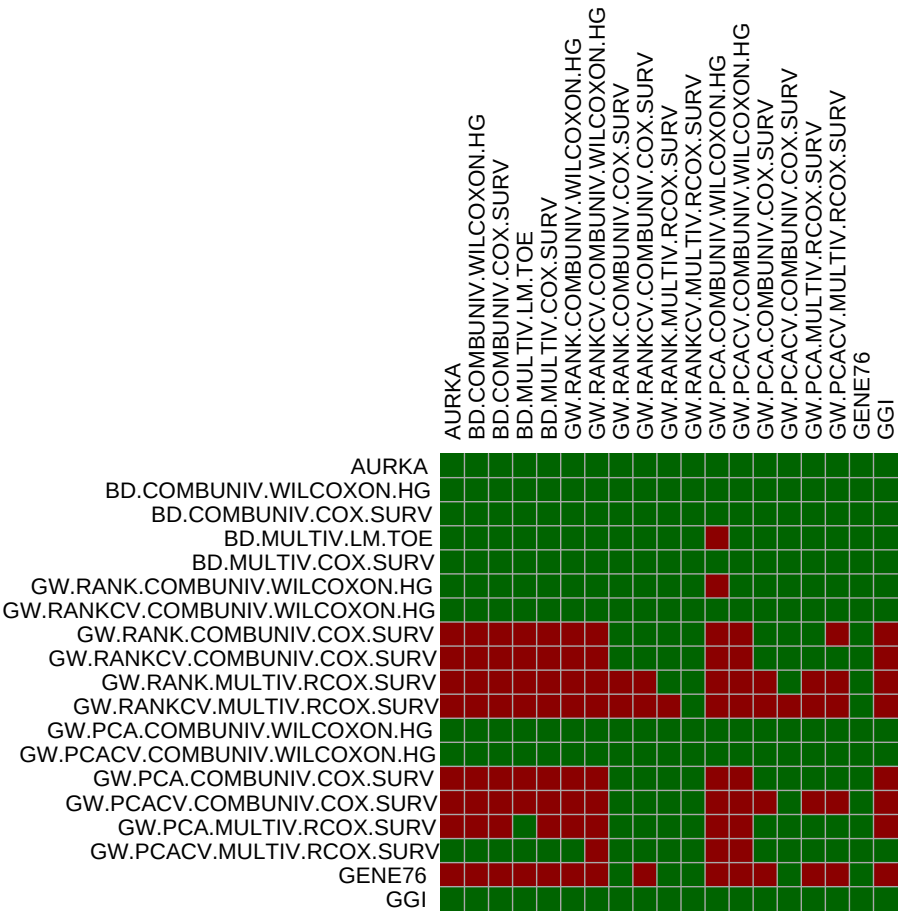

The following table shows the comparison between the different methods and AURKA. A text in bold represents a significant superiority (Student t test p-value $< 0.05$) of the method in row against AURKA.

| | concordance.index |
|---|---|
| AURKA | 0.636 |
| BD.COMBUNIV.WILCOXON.HG | 0.606 |
| BD.COMBUNIV.COX.SURV | 0.638 |
| BD.MULTIV.LM.TOE | 0.601 |
| BD.MULTIV.COX.SURV | 0.649 |
| GW.RANK.COMBUNIV.WILCOXON.HG | 0.619 |
| GW.RANKCV.COMBUNIV.WILCOXON.HG | 0.595 |
| GW.RANK.COMBUNIV.COX.SURV | **0.742** |
| GW.RANKCV.COMBUNIV.COX.SURV | **0.717** |
| GW.RANK.MULTIV.RCOX.SURV | **0.774** |
| GW.RANKCV.MULTIV.RCOX.SURV | **0.777** |
| GW.PCA.COMBUNIV.WILCOXON.HG | 0.587 |
| GW.PCACV.COMBUNIV.WILCOXON.HG | 0.595 |
| GW.PCA.COMBUNIV.COX.SURV | **0.727** |
| GW.PCACV.COMBUNIV.COX.SURV | **0.748** |
| GW.PCA.MULTIV.RCOX.SURV | **0.75** |
| GW.PCACV.MULTIV.RCOX.SURV | **0.686** |
| GENE76 | **0.754** |
| GGI | 0.613 |

**Concordance Index for Risk Group** The following foresplot shows the concordance indices and their confidence interval for each method :



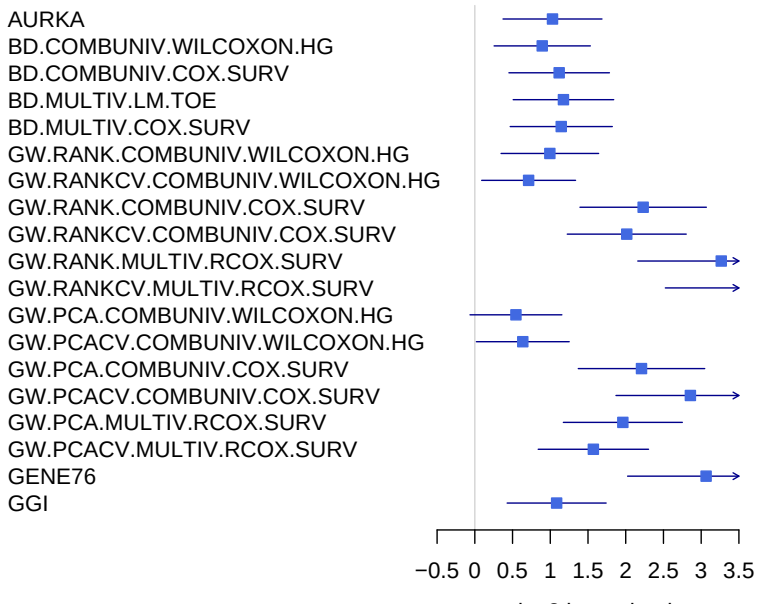

The following figure shows the pairwise comparison between the different methods. A red box represents a significant superiority (Student t test p-value $< 0.05$) of the method in row against the method in column. A green box represents no significant superiority (Student t test p-value $>= 0.05$) of the method in row against the method in column.



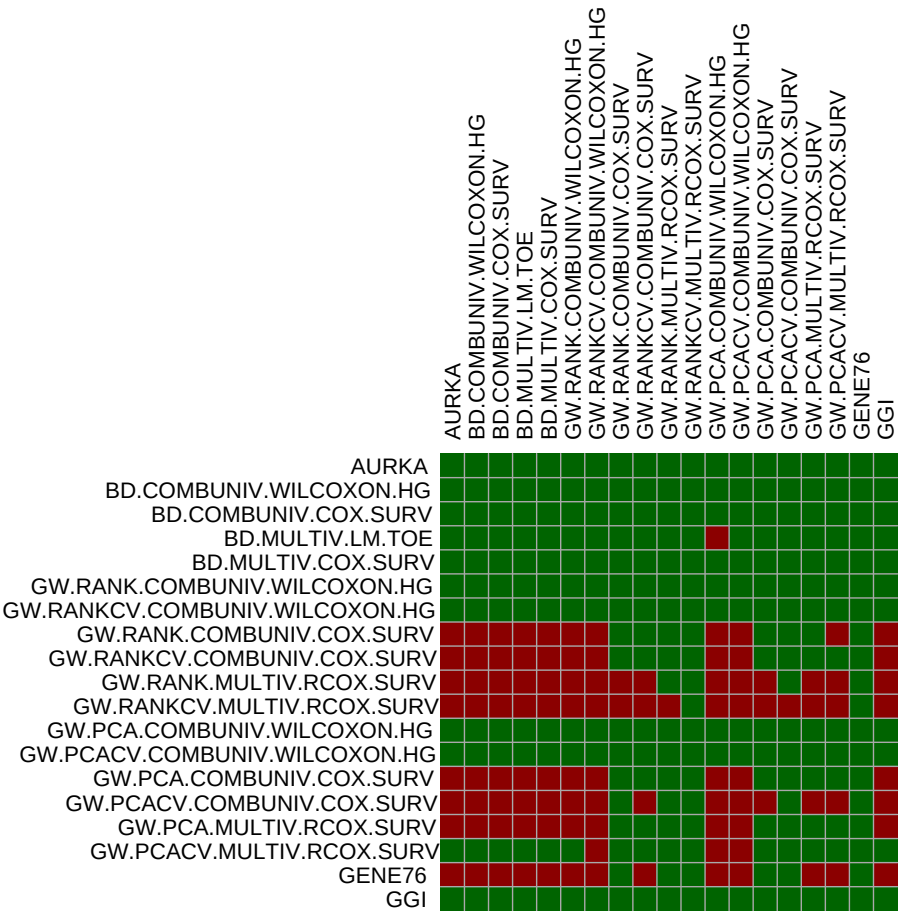

The following table shows the comparison between the different methods and AURKA. A text in bold represents a significant superiority (Student t test p-value $< 0.05$) of the method in row against AURKA.

| | concordance.index |
|---|---|
| AURKA | 0.685 |
| BD.COMBUNIV.WILCOXON.HG | 0.675 |
| BD.COMBUNIV.COX.SURV | 0.698 |
| BD.MULTIV.LM.TOE | 0.721 |
| BD.MULTIV.COX.SURV | 0.685 |
| GW.RANK.COMBUNIV.WILCOXON.HG | 0.694 |
| GW.RANKCV.COMBUNIV.WILCOXON.HG | 0.642 |
| GW.RANK.COMBUNIV.COX.SURV | **0.836** |
| GW.RANKCV.COMBUNIV.COX.SURV | **0.824** |
| GW.RANK.MULTIV.RCOX.SURV | **0.906** |
| GW.RANKCV.MULTIV.RCOX.SURV | **0.938** |
| GW.PCA.COMBUNIV.WILCOXON.HG | 0.616 |
| GW.PCACV.COMBUNIV.WILCOXON.HG | 0.633 |
| GW.PCA.COMBUNIV.COX.SURV | **0.824** |
| GW.PCACV.COMBUNIV.COX.SURV | **0.881** |
| GW.PCA.MULTIV.RCOX.SURV | **0.806** |
| GW.PCACV.MULTIV.RCOX.SURV | 0.754 |
| GENE76 | **0.903** |
| GGI | 0.706 |

**Hazard Ratio for Risk Group** The following foresplot shows the hazard ratios and their confidence interval for each method :



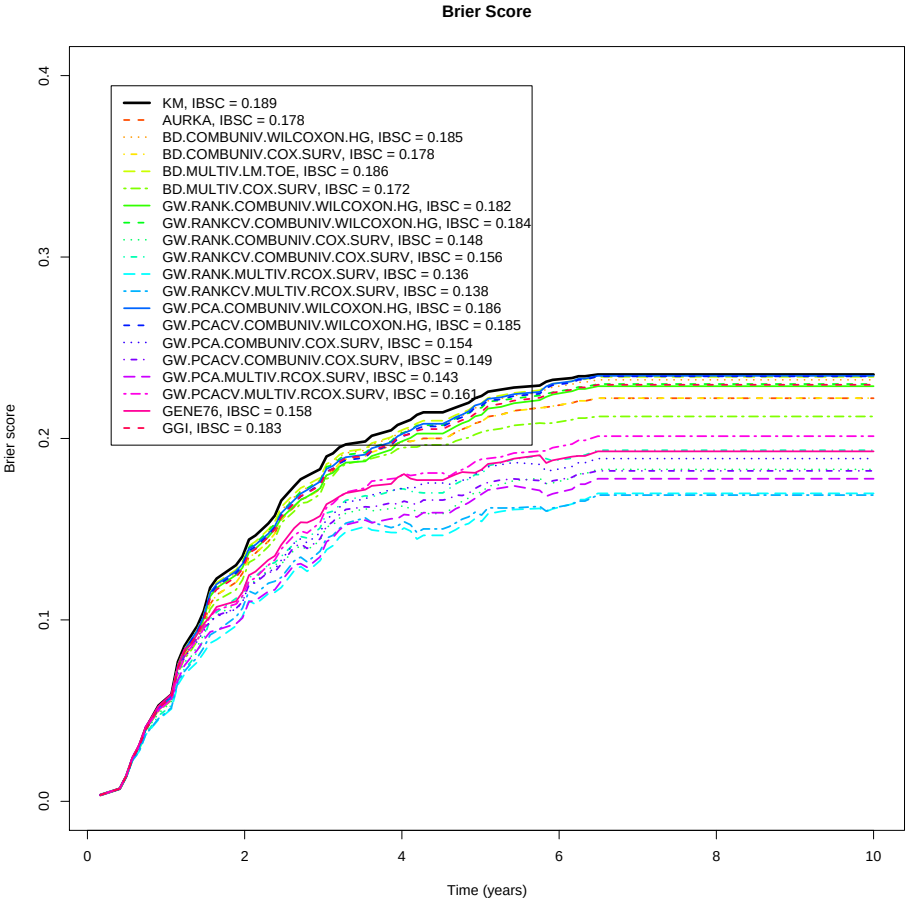

The following figure shows the pairwise comparison between the different methods. A red box represents a significant superiority (Student t test p-value $< 0.05$) of the method in row against the method in column. A green box represents no significant superiority (Student t test p-value $>= 0.05$) of the method in row against the method in column.



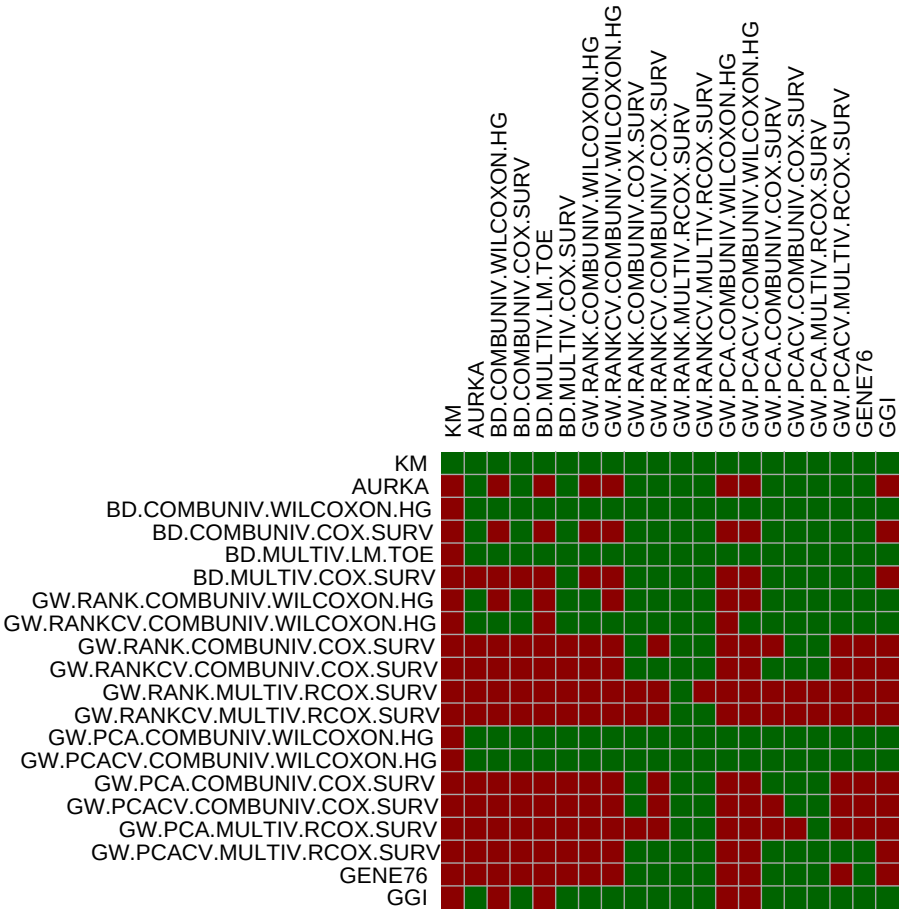

The following table shows the comparison between the different methods and AURKA. A text in bold represents a significant superiority (Student t test p-value $< 0.05$) of the method in row against AURKA.

| | HR |
|---|---|
| AURKA | 2.04 |
| BD.COMBUNIV.WILCOXON.HG | 1.86 |
| BD.COMBUNIV.COX.SURV | 2.17 |
| BD.MULTIV.LM.TOE | 2.26 |
| BD.MULTIV.COX.SURV | 2.21 |
| GW.RANK.COMBUNIV.WILCOXON.HG | 1.99 |
| GW.RANKCV.COMBUNIV.WILCOXON.HG | 1.64 |
| GW.RANK.COMBUNIV.COX.SURV | **4.69** |
| GW.RANKCV.COMBUNIV.COX.SURV | **4.04** |
| GW.RANK.MULTIV.RCOX.SURV | **9.62** |
| GW.RANKCV.MULTIV.RCOX.SURV | **14.1** |
| GW.PCA.COMBUNIV.WILCOXON.HG | 1.46 |
| GW.PCACV.COMBUNIV.WILCOXON.HG | 1.55 |
| GW.PCA.COMBUNIV.COX.SURV | **4.62** |
| GW.PCACV.COMBUNIV.COX.SURV | **7.25** |
| GW.PCA.MULTIV.RCOX.SURV | **3.9** |
| GW.PCACV.MULTIV.RCOX.SURV | 2.97 |
| GENE76 | **8.37** |
| GGI | 2.12 |

**Brier Score for Risk Score** The following figure shows the Brier score w.r.t. the time for each method :



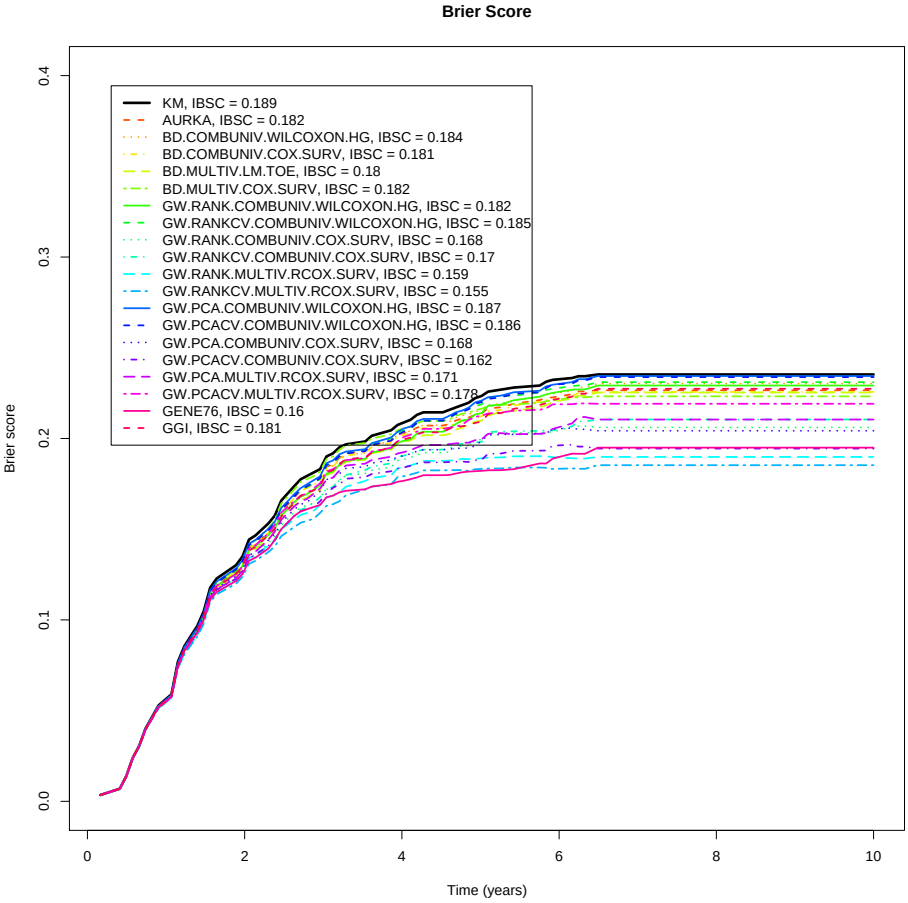

The following figure shows the pairwise comparison between the different methods. A red box represents a significant superiority (paired Wilcoxon rank sum test p-value $< 0.05$) of the method in row against the method in column. A green box represents no significant superiority (paired Wilcoxon rank sum test p-value $>= 0.05$) of the method in row against the method in column.



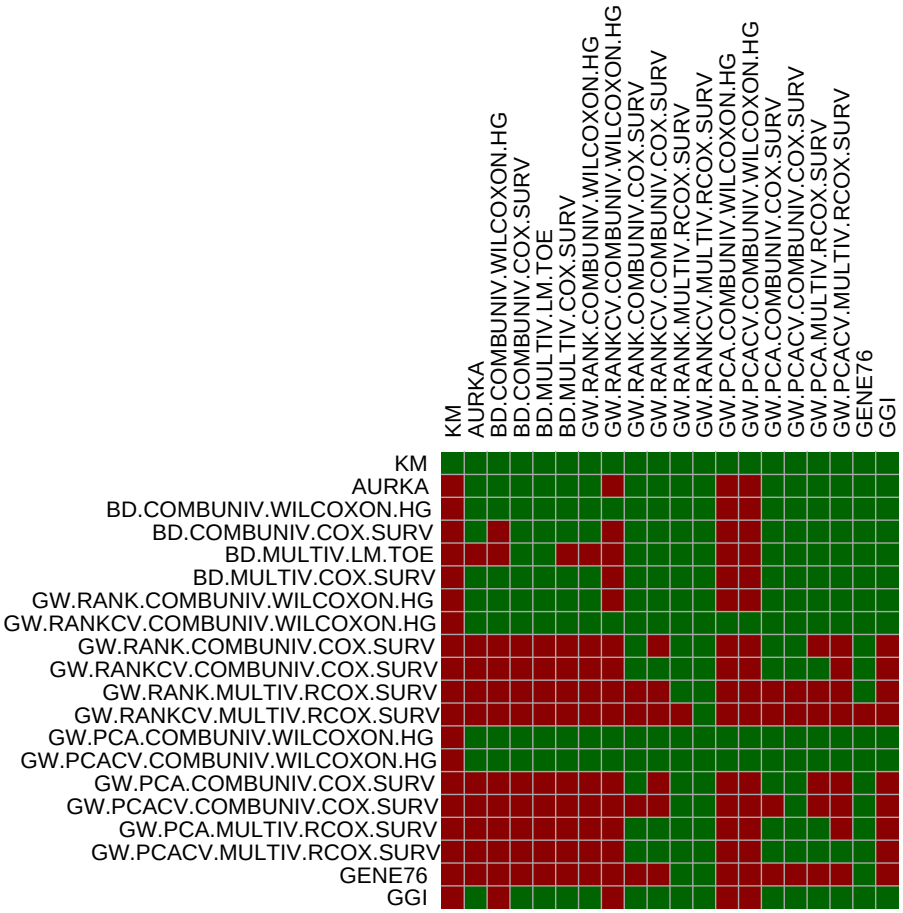

The following table shows the comparison between the different methods and AURKA. A text in bold represents a significant superiority (Student t test p-value $< 0.05$) of the method in row against AURKA.

| | IBSC |
|---|---|
| KM | 0.189 |
| AURKA | **0.178** |
| BD.COMBUNIV.WILCOXON.HG | 0.185 |
| BD.COMBUNIV.COX.SURV | 0.178 |
| BD.MULTIV.LM.TOE | 0.186 |
| BD.MULTIV.COX.SURV | **0.172** |
| GW.RANK.COMBUNIV.WILCOXON.HG | 0.182 |
| GW.RANKCV.COMBUNIV.WILCOXON.HG | 0.184 |
| GW.RANK.COMBUNIV.COX.SURV | **0.148** |
| GW.RANKCV.COMBUNIV.COX.SURV | **0.156** |
| GW.RANK.MULTIV.RCOX.SURV | **0.136** |
| GW.RANKCV.MULTIV.RCOX.SURV | **0.138** |
| GW.PCA.COMBUNIV.WILCOXON.HG | 0.186 |
| GW.PCACV.COMBUNIV.WILCOXON.HG | 0.185 |
| GW.PCA.COMBUNIV.COX.SURV | **0.154** |
| GW.PCACV.COMBUNIV.COX.SURV | **0.149** |
| GW.PCA.MULTIV.RCOX.SURV | **0.143** |
| GW.PCACV.MULTIV.RCOX.SURV | **0.161** |
| GENE76 | **0.158** |
| GGI | 0.183 |

**Brier Score for Risk Group** The following figure shows the Brier score w.r.t. the time for each method :



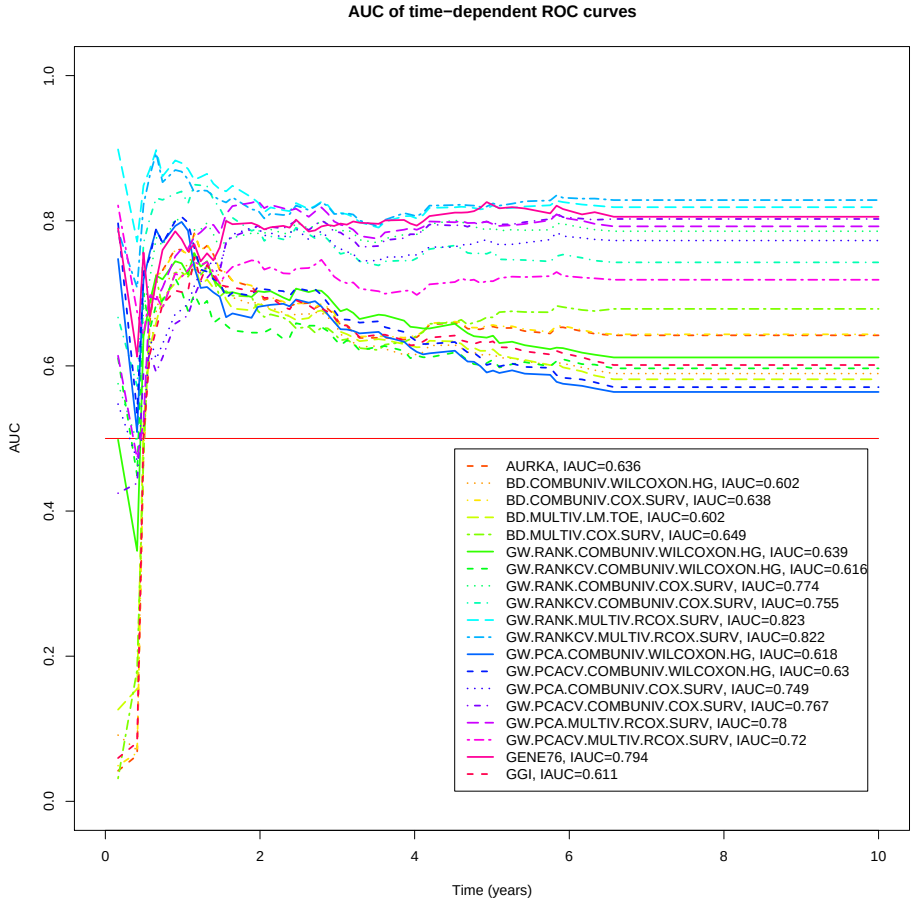

The following figure shows the pairwise comparison between the different methods. A red box represents a significant superiority (paired Wilcoxon rank sum test p-value $< 0.05$) of the method in row against the method in column. A green box represents no significant superiority (paired Wilcoxon rank sum test p-value $>= 0.05$) of the method in row against the method in column.



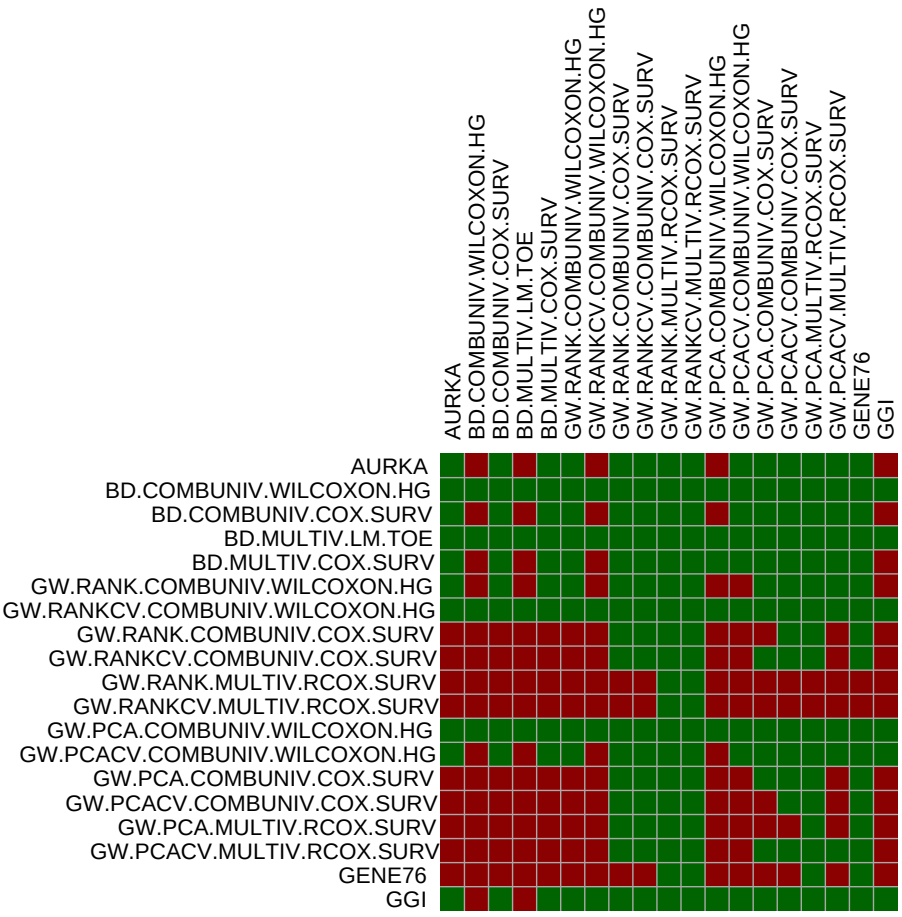

The following table shows the comparison between the different methods and AURKA. A text in bold represents a significant superiority (Student t test p-value $< 0.05$) of the method in row against AURKA.

| | IBSC |
|---|---|
| KM | 0.189 |
| AURKA | **0.182** |
| BD.COMBUNIV.WILCOXON.HG | 0.184 |
| BD.COMBUNIV.COX.SURV | 0.181 |
| BD.MULTIV.LM.TOE | **0.18** |
| BD.MULTIV.COX.SURV | 0.182 |
| GW.RANK.COMBUNIV.WILCOXON.HG | 0.182 |
| GW.RANKCV.COMBUNIV.WILCOXON.HG | 0.185 |
| GW.RANK.COMBUNIV.COX.SURV | **0.168** |
| GW.RANKCV.COMBUNIV.COX.SURV | **0.17** |
| GW.RANK.MULTIV.RCOX.SURV | **0.159** |
| GW.RANKCV.MULTIV.RCOX.SURV | **0.155** |
| GW.PCA.COMBUNIV.WILCOXON.HG | 0.187 |
| GW.PCACV.COMBUNIV.WILCOXON.HG | 0.186 |
| GW.PCA.COMBUNIV.COX.SURV | **0.168** |
| GW.PCACV.COMBUNIV.COX.SURV | **0.162** |
| GW.PCA.MULTIV.RCOX.SURV | **0.171** |
| GW.PCACV.MULTIV.RCOX.SURV | **0.178** |
| GENE76 | **0.16** |
| GGI | 0.181 |

**Time-Dependent ROC Curves for Risk Score** The following figure shows the evolution of the AUC of the time-dependent ROC curves with respect to the time for each method :



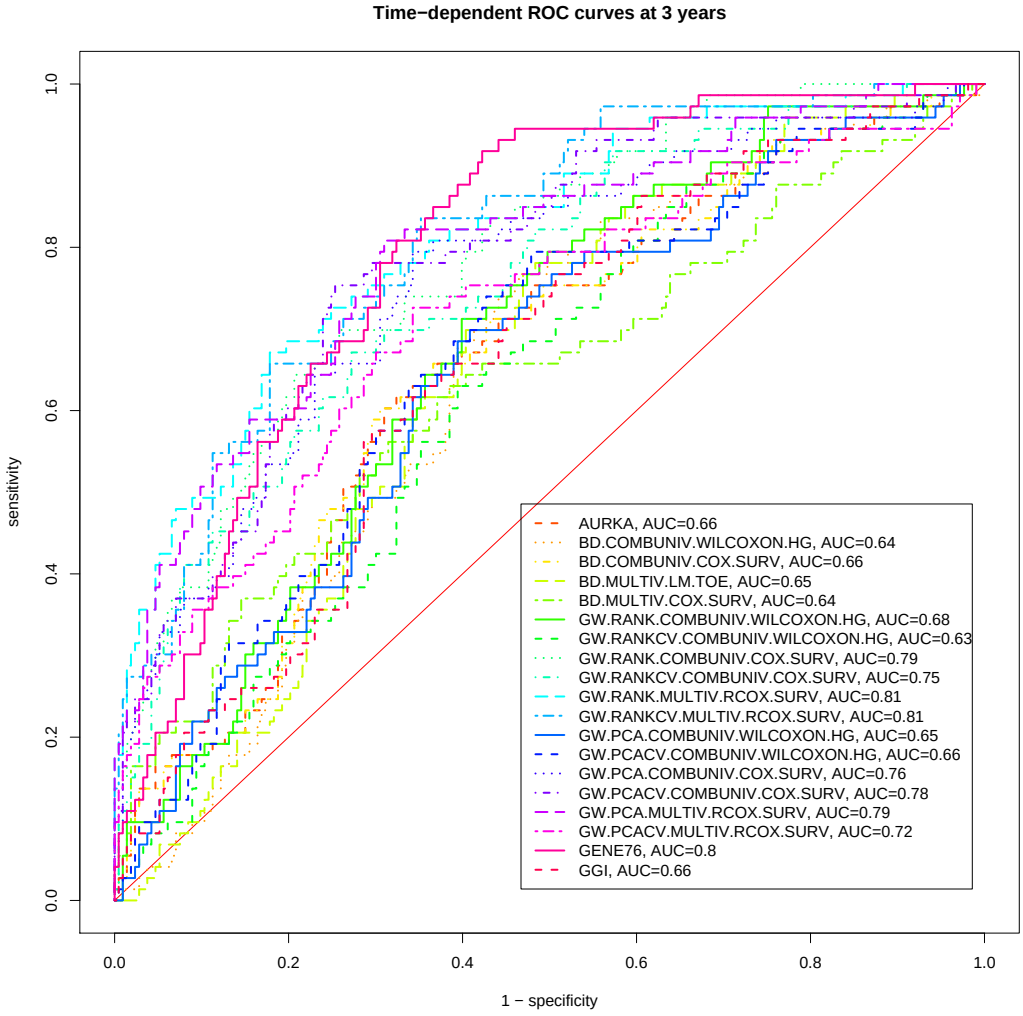

The following figure shows the pairwise comparison between the different methods. A red box represents a significant superiority (paired Wilcoxon rank sum test p-value $< 0.05$) of the method in row against the method in column. A green box represents no significant superiority (paired Wilcoxon rank sum test p-value $>= 0.05$) of the method in row against the method in column.

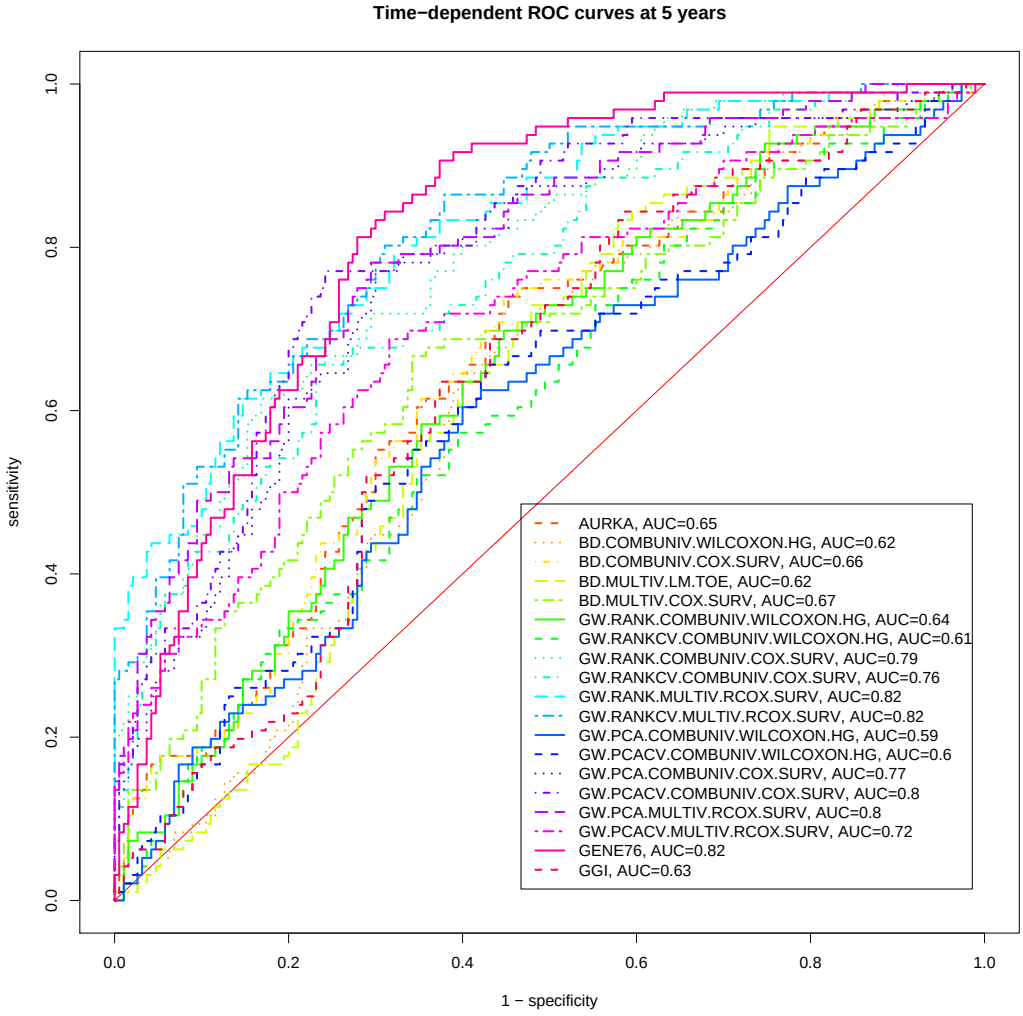


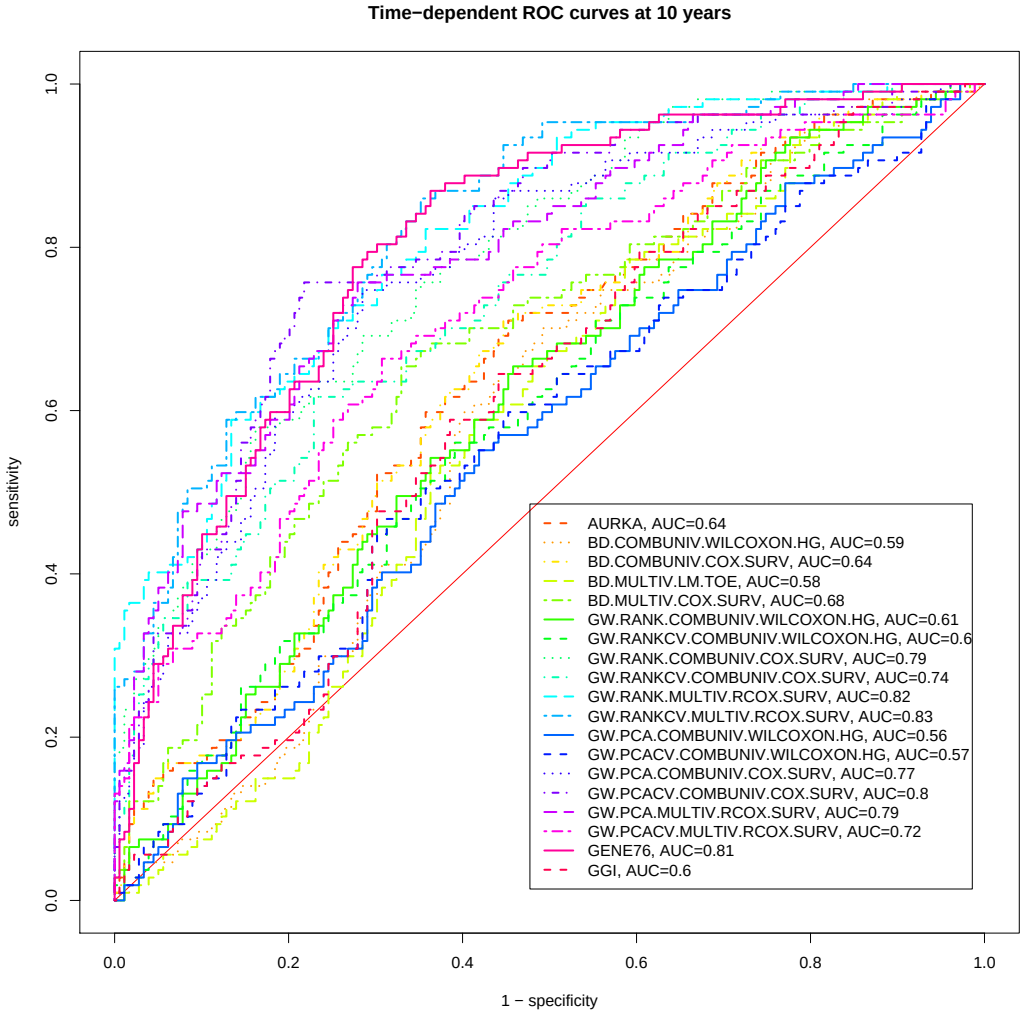

The following table shows the comparison between the different methods and AURKA. A text in bold represents a significant superiority (Student t test p-value $< 0.05$) of the method in row against AURKA.

| | IAUC |
|---|---|
| AURKA | 0.636 |
| BD.COMBUNIV.WILCOXON.HG | 0.602 |
| BD.COMBUNIV.COX.SURV | 0.638 |
| BD.MULTIV.LM.TOE | 0.602 |
| BD.MULTIV.COX.SURV | 0.649 |
| GW.RANK.COMBUNIV.WILCOXON.HG | 0.639 |
| GW.RANKCV.COMBUNIV.WILCOXON.HG | 0.616 |
| GW.RANK.COMBUNIV.COX.SURV | **0.774** |
| GW.RANKCV.COMBUNIV.COX.SURV | **0.755** |
| GW.RANK.MULTIV.RCOX.SURV | **0.823** |
| GW.RANKCV.MULTIV.RCOX.SURV | **0.822** |
| GW.PCA.COMBUNIV.WILCOXON.HG | 0.618 |
| GW.PCACV.COMBUNIV.WILCOXON.HG | 0.63 |
| GW.PCA.COMBUNIV.COX.SURV | **0.749** |
| GW.PCACV.COMBUNIV.COX.SURV | **0.767** |
| GW.PCA.MULTIV.RCOX.SURV | **0.78** |
| GW.PCACV.MULTIV.RCOX.SURV | **0.72** |
| GENE76 | **0.794** |
| GGI | 0.611 |

The following figure shows the time-dependent ROC curves and their corresponding AUC for each method :


Time−dependent ROC curves at 3 years
sensitivity
1 − specificity
0.0
0.2
0.4
0.6
0.8
1.0
AURKA, AUC=0.66
BD.COMBUNIV.WILCOXON.HG, AUC=0.64
BD.COMBUNIV.COX.SURV, AUC=0.66
BD.MULTIV.LM.TOE, AUC=0.65
BD.MULTIV.COX.SURV, AUC=0.64
GW.RANK.COMBUNIV.WILCOXON.HG, AUC=0.68
GW.RANKCV.COMBUNIV.WILCOXON.HG, AUC=0.63
GW.RANK.COMBUNIV.COX.SURV, AUC=0.79
GW.RANKCV.COMBUNIV.COX.SURV, AUC=0.75
GW.RANK.MULTIV.RCOX.SURV, AUC=0.81
GW.RANKCV.MULTIV.RCOX.SURV, AUC=0.81
GW.PCA.COMBUNIV.WILCOXON.HG, AUC=0.65
GW.PCACV.COMBUNIV.WILCOXON.HG, AUC=0.66
GW.PCA.COMBUNIV.COX.SURV, AUC=0.76
GW.PCACV.COMBUNIV.COX.SURV, AUC=0.78
GW.PCA.MULTIV.RCOX.SURV, AUC=0.79
GW.PCACV.MULTIV.RCOX.SURV, AUC=0.72
GENE76, AUC=0.8
GGI, AUC=0.66


Time−dependent ROC curves at 5 years
sensitivity
1 − specificity
0.0
0.2
0.4
0.6
0.8
1.0
AURKA, AUC=0.65
BD.COMBUNIV.WILCOXON.HG, AUC=0.62
BD.COMBUNIV.COX.SURV, AUC=0.66
BD.MULTIV.LM.TOE, AUC=0.62
BD.MULTIV.COX.SURV, AUC=0.67
GW.RANK.COMBUNIV.WILCOXON.HG, AUC=0.64
GW.RANKCV.COMBUNIV.WILCOXON.HG, AUC=0.61
GW.RANK.COMBUNIV.COX.SURV, AUC=0.79
GW.RANKCV.COMBUNIV.COX.SURV, AUC=0.76
GW.RANK.MULTIV.RCOX.SURV, AUC=0.82
GW.RANKCV.MULTIV.RCOX.SURV, AUC=0.82
GW.PCA.COMBUNIV.WILCOXON.HG, AUC=0.59
GW.PCACV.COMBUNIV.WILCOXON.HG, AUC=0.6
GW.PCA.COMBUNIV.COX.SURV, AUC=0.77
GW.PCACV.COMBUNIV.COX.SURV, AUC=0.8
GW.PCA.MULTIV.RCOX.SURV, AUC=0.8
GW.PCACV.MULTIV.RCOX.SURV, AUC=0.72
GENE76, AUC=0.82
GGI, AUC=0.63




The following table shows the specificity for a given sensitivity of 90% at some points in time:

| | years.3 | years.5 | years.10 |
|---|---|---|---|
| AURKA | 0.249 | 0.253 | 0.263 |
| BD.COMBUNIV.WILCOXON.HG | 0.263 | 0.247 | 0.246 |
| BD.COMBUNIV.COX.SURV | 0.272 | 0.268 | 0.279 |
| BD.MULTIV.LM.TOE | 0.263 | 0.268 | 0.223 |
| BD.MULTIV.COX.SURV | 0.174 | 0.205 | 0.257 |
| GW.RANK.COMBUNIV.WILCOXON.HG | 0.315 | 0.258 | 0.251 |
| GW.RANKCV.COMBUNIV.WILCOXON.HG | 0.258 | 0.242 | 0.184 |
| GW.RANK.COMBUNIV.COX.SURV | 0.432 | 0.400 | 0.419 |
| GW.RANKCV.COMBUNIV.COX.SURV | 0.437 | 0.353 | 0.369 |
| GW.RANK.MULTIV.RCOX.SURV | 0.432 | 0.468 | 0.497 |
| GW.RANKCV.MULTIV.RCOX.SURV | 0.488 | 0.526 | 0.553 |
| GW.PCA.COMBUNIV.WILCOXON.HG | 0.258 | 0.147 | 0.151 |
| GW.PCACV.COMBUNIV.WILCOXON.HG | 0.249 | 0.147 | 0.134 |
| GW.PCA.COMBUNIV.COX.SURV | 0.385 | 0.421 | 0.436 |
| GW.PCACV.COMBUNIV.COX.SURV | 0.469 | 0.511 | 0.503 |
| GW.PCA.MULTIV.RCOX.SURV | 0.380 | 0.442 | 0.402 |
| GW.PCACV.MULTIV.RCOX.SURV | 0.296 | 0.300 | 0.313 |
| GENE76 | 0.582 | 0.626 | 0.536 |
| GGI | 0.277 | 0.258 | 0.201 |

The following table shows the sensitivity of the methods in leaving 33% of the patients in the low-risk group at some points in time:

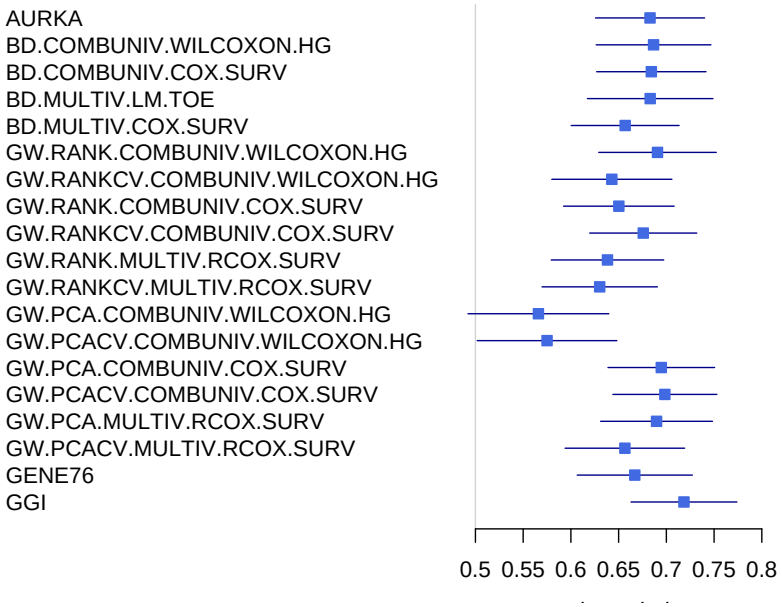
| | years.3 | years.5 | years.10 |
|---|---|---|---|
| AURKA | 0.808 | 0.802 | 0.785 |
| BD.COMBUNIV.WILCOXON.HG | 0.836 | 0.792 | 0.757 |
| BD.COMBUNIV.COX.SURV | 0.822 | 0.813 | 0.785 |
| BD.MULTIV.LM.TOE | 0.863 | 0.833 | 0.785 |
| BD.MULTIV.COX.SURV | 0.767 | 0.792 | 0.804 |
| GW.RANK.COMBUNIV.WILCOXON.HG | 0.863 | 0.813 | 0.776 |
| GW.RANKCV.COMBUNIV.WILCOXON.HG | 0.808 | 0.760 | 0.738 |
| GW.RANK.COMBUNIV.COX.SURV | 0.918 | 0.885 | 0.879 |
| GW.RANKCV.COMBUNIV.COX.SURV | 0.918 | 0.885 | 0.860 |
| GW.RANK.MULTIV.RCOX.SURV | 0.945 | 0.927 | 0.935 |
| GW.RANKCV.MULTIV.RCOX.SURV | 0.973 | 0.948 | 0.953 |
| GW.PCA.COMBUNIV.WILCOXON.HG | 0.795 | 0.740 | 0.720 |
| GW.PCACV.COMBUNIV.WILCOXON.HG | 0.808 | 0.760 | 0.738 |
| GW.PCA.COMBUNIV.COX.SURV | 0.877 | 0.885 | 0.879 |
| GW.PCACV.COMBUNIV.COX.SURV | 0.932 | 0.927 | 0.916 |
| GW.PCA.MULTIV.RCOX.SURV | 0.890 | 0.896 | 0.869 |
| GW.PCACV.MULTIV.RCOX.SURV | 0.836 | 0.823 | 0.832 |
| GENE76 | 0.945 | 0.958 | 0.925 |
| GGI | 0.863 | 0.844 | 0.785 |

The following table shows the specificity of the methods in leaving 33% of the patients in the low-risk group at some points in time:

| | years.3 | years.5 | years.10 |
|---|---|---|---|
| AURKA | 0.371 | 0.389 | 0.391 |
| BD.COMBUNIV.WILCOXON.HG | 0.385 | 0.389 | 0.380 |
| BD.COMBUNIV.COX.SURV | 0.380 | 0.400 | 0.397 |
| BD.MULTIV.LM.TOE | 0.394 | 0.411 | 0.397 |
| BD.MULTIV.COX.SURV | 0.362 | 0.389 | 0.408 |
| GW.RANK.COMBUNIV.WILCOXON.HG | 0.394 | 0.400 | 0.391 |
| GW.RANKCV.COMBUNIV.WILCOXON.HG | 0.376 | 0.374 | 0.369 |
| GW.RANK.COMBUNIV.COX.SURV | 0.413 | 0.437 | 0.453 |
| GW.RANKCV.COMBUNIV.COX.SURV | 0.413 | 0.437 | 0.441 |
| GW.RANK.MULTIV.RCOX.SURV | 0.423 | 0.458 | 0.486 |
| GW.RANKCV.MULTIV.RCOX.SURV | 0.432 | 0.468 | 0.497 |
| GW.PCA.COMBUNIV.WILCOXON.HG | 0.371 | 0.363 | 0.358 |
| GW.PCACV.COMBUNIV.WILCOXON.HG | 0.376 | 0.374 | 0.369 |
| GW.PCA.COMBUNIV.COX.SURV | 0.399 | 0.437 | 0.453 |
| GW.PCACV.COMBUNIV.COX.SURV | 0.418 | 0.458 | 0.475 |
| GW.PCA.MULTIV.RCOX.SURV | 0.404 | 0.442 | 0.447 |
| GW.PCACV.MULTIV.RCOX.SURV | 0.385 | 0.405 | 0.425 |
| GENE76 | 0.423 | 0.474 | 0.480 |
| GGI | 0.394 | 0.416 | 0.397 |

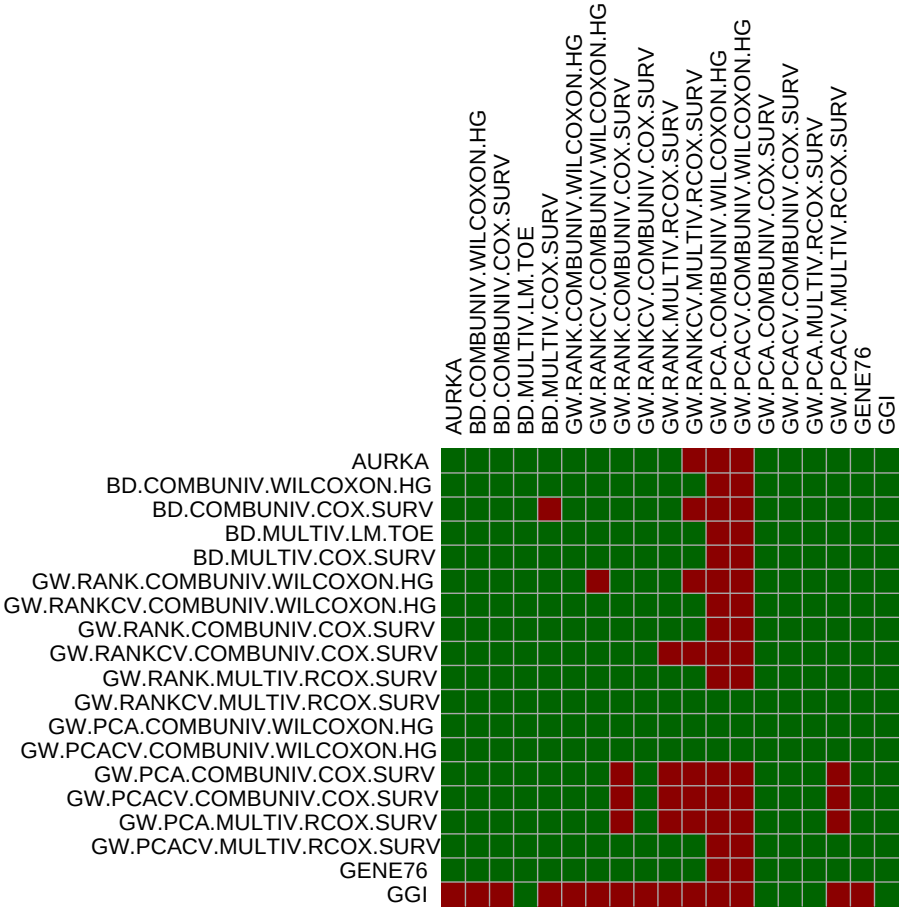
### 1.21.2 Test Set

**Concordance Index for Risk Score** The following foresplot shows the concordance indices and their confidence interval for each method :




The following figure shows the pairwise comparison between the different methods. A red box represents a significant superiority (Student t test p-value $< 0.05$) of the method in row

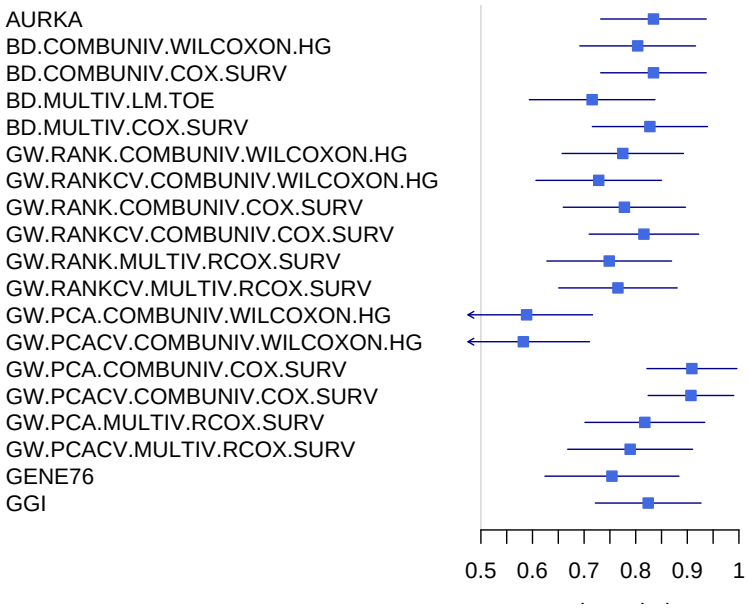
against the method in column. A green box represents no significant superiority (Student t test p-value >= 0.05) of the method in row against the method in column.



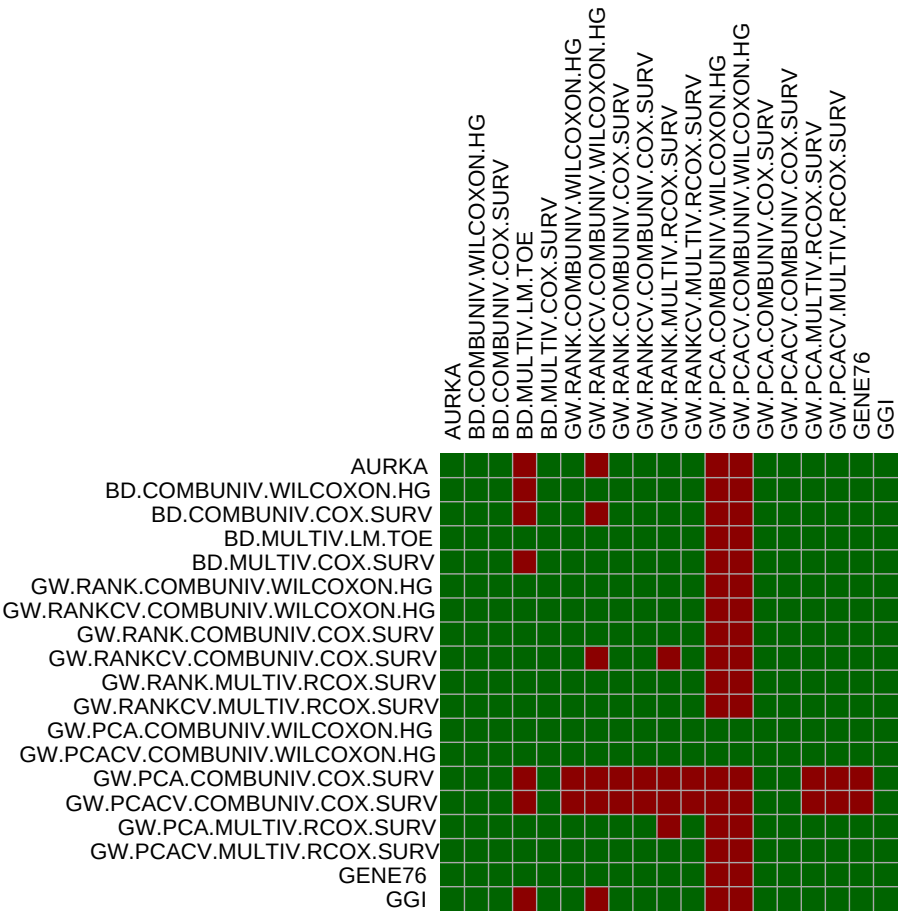

The following table shows the comparison between the different methods and AURKA. A text in bold represents a significant superiority (Student t test p-value $< 0.05$) of the method in row against AURKA.

| | concordance.index |
|---|---|
| AURKA | 0.683 |
| BD.COMBUNIV.WILCOXON.HG | 0.687 |
| BD.COMBUNIV.COX.SURV | 0.684 |
| BD.MULTIV.LM.TOE | 0.683 |
| BD.MULTIV.COX.SURV | 0.657 |
| GW.RANK.COMBUNIV.WILCOXON.HG | 0.691 |
| GW.RANKCV.COMBUNIV.WILCOXON.HG | 0.643 |
| GW.RANK.COMBUNIV.COX.SURV | 0.65 |
| GW.RANKCV.COMBUNIV.COX.SURV | 0.676 |
| GW.RANK.MULTIV.RCOX.SURV | 0.638 |
| GW.RANKCV.MULTIV.RCOX.SURV | 0.63 |
| GW.PCA.COMBUNIV.WILCOXON.HG | 0.566 |
| GW.PCACV.COMBUNIV.WILCOXON.HG | 0.575 |
| GW.PCA.COMBUNIV.COX.SURV | 0.695 |
| GW.PCACV.COMBUNIV.COX.SURV | 0.698 |
| GW.PCA.MULTIV.RCOX.SURV | 0.69 |
| GW.PCACV.MULTIV.RCOX.SURV | 0.657 |
| GENE76 | 0.667 |
| GGI | **0.718** |

**Concordance Index for Risk Group** The following foresplot shows the concordance indices and their confidence interval for each method :



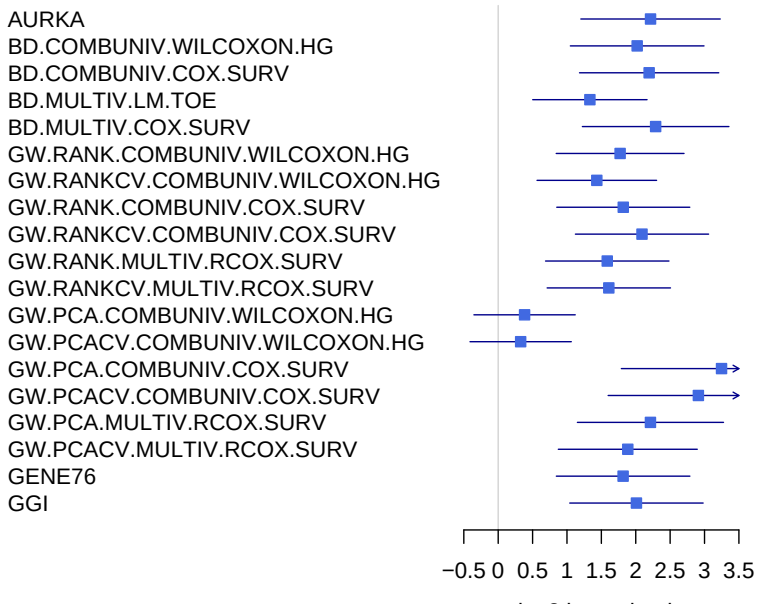

The following figure shows the pairwise comparison between the different methods. A red box represents a significant superiority (Student t test p-value $< 0.05$) of the method in row against the method in column. A green box represents no significant superiority (Student t test p-value $>= 0.05$) of the method in row against the method in column.



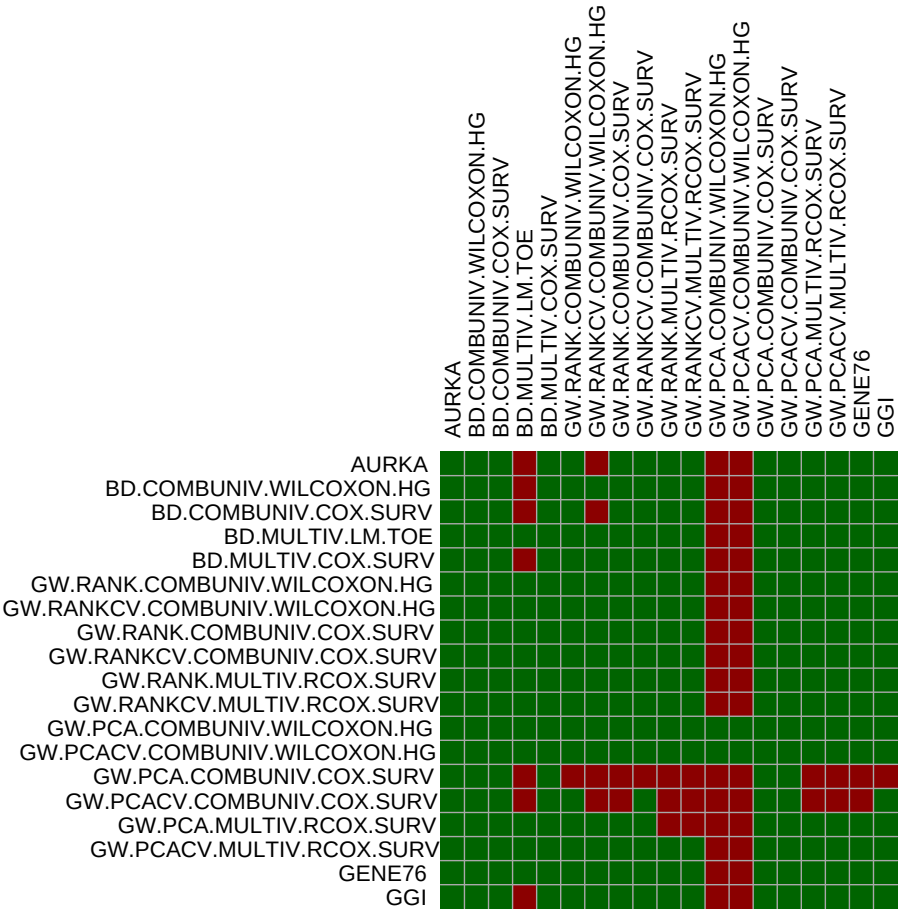

The following table shows the comparison between the different methods and AURKA. A text in bold represents a significant superiority (Student t test p-value $< 0.05$) of the method in row against AURKA.

| | concordance.index |
|---|---|
| AURKA | 0.834 |
| BD.COMBUNIV.WILCOXON.HG | 0.804 |
| BD.COMBUNIV.COX.SURV | 0.834 |
| BD.MULTIV.LM.TOE | 0.716 |
| BD.MULTIV.COX.SURV | 0.828 |
| GW.RANK.COMBUNIV.WILCOXON.HG | 0.775 |
| GW.RANKCV.COMBUNIV.WILCOXON.HG | 0.728 |
| GW.RANK.COMBUNIV.COX.SURV | 0.778 |
| GW.RANKCV.COMBUNIV.COX.SURV | 0.816 |
| GW.RANK.MULTIV.RCOX.SURV | 0.749 |
| GW.RANKCV.MULTIV.RCOX.SURV | 0.766 |
| GW.PCA.COMBUNIV.WILCOXON.HG | 0.589 |
| GW.PCACV.COMBUNIV.WILCOXON.HG | 0.582 |
| GW.PCA.COMBUNIV.COX.SURV | 0.909 |
| GW.PCACV.COMBUNIV.COX.SURV | 0.907 |
| GW.PCA.MULTIV.RCOX.SURV | 0.818 |
| GW.PCACV.MULTIV.RCOX.SURV | 0.789 |
| GENE76 | 0.754 |
| GGI | 0.824 |

**Hazard Ratio for Risk Group** The following foresplot shows the hazard ratios and their confidence interval for each method :



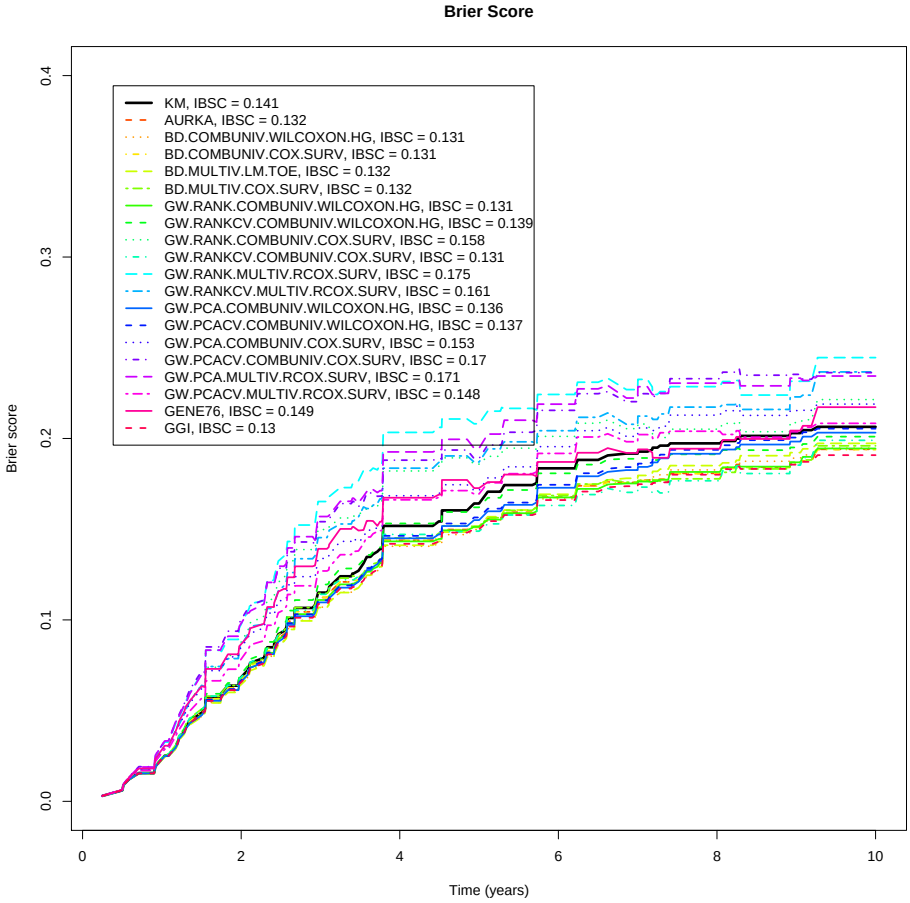

The following figure shows the pairwise comparison between the different methods. A red box represents a significant superiority (Student t test p-value $< 0.05$) of the method in row against the method in column. A green box represents no significant superiority (Student t test p-value $>= 0.05$) of the method in row against the method in column.



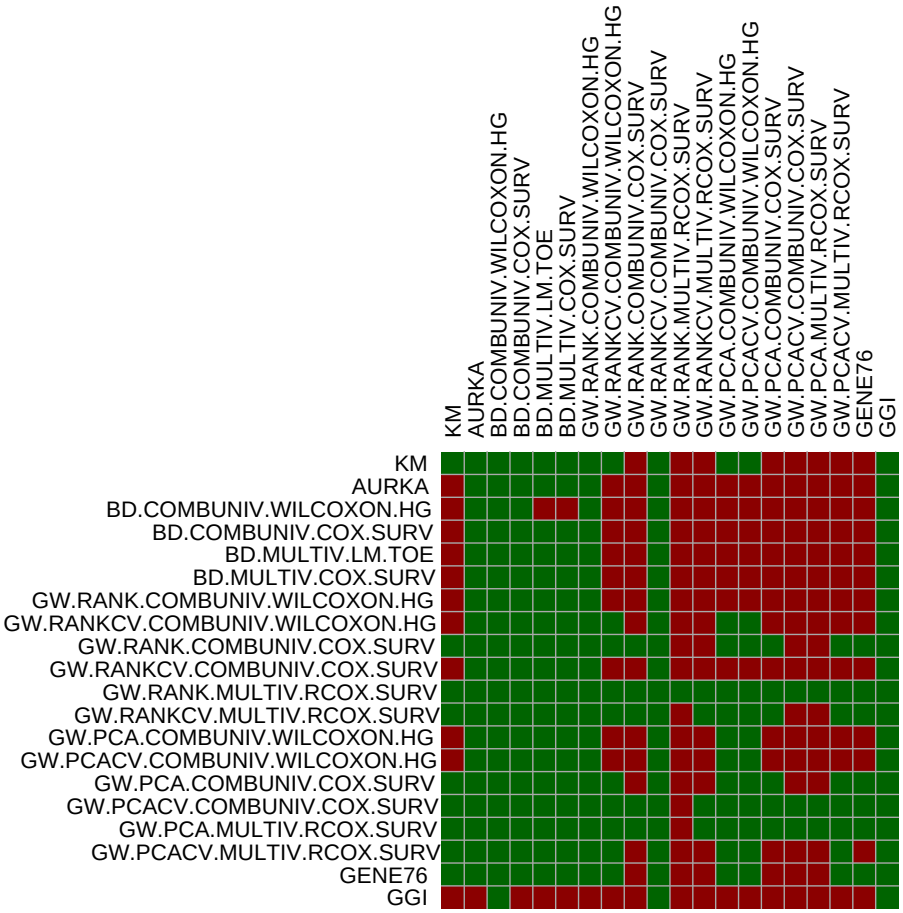

The following table shows the comparison between the different methods and AURKA. A text in bold represents a significant superiority (Student t test p-value $< 0.05$) of the method in row against AURKA.

| | HR |
|---|---|
| AURKA | 4.64 |
| BD.COMBUNIV.WILCOXON.HG | 4.06 |
| BD.COMBUNIV.COX.SURV | 4.58 |
| BD.MULTIV.LM.TOE | 2.52 |
| BD.MULTIV.COX.SURV | 4.89 |
| GW.RANK.COMBUNIV.WILCOXON.HG | 3.42 |
| GW.RANKCV.COMBUNIV.WILCOXON.HG | 2.7 |
| GW.RANK.COMBUNIV.COX.SURV | 3.53 |
| GW.RANKCV.COMBUNIV.COX.SURV | 4.26 |
| GW.RANK.MULTIV.RCOX.SURV | 3 |
| GW.RANKCV.MULTIV.RCOX.SURV | 3.05 |
| GW.PCA.COMBUNIV.WILCOXON.HG | 1.3 |
| GW.PCACV.COMBUNIV.WILCOXON.HG | 1.25 |
| GW.PCA.COMBUNIV.COX.SURV | 9.5 |
| GW.PCACV.COMBUNIV.COX.SURV | 7.52 |
| GW.PCA.MULTIV.RCOX.SURV | 4.64 |
| GW.PCACV.MULTIV.RCOX.SURV | 3.69 |
| GENE76 | 3.52 |
| GGI | 4.03 |

**Brier Score for Risk Score** The following figure shows the Brier score w.r.t. the time for each method :



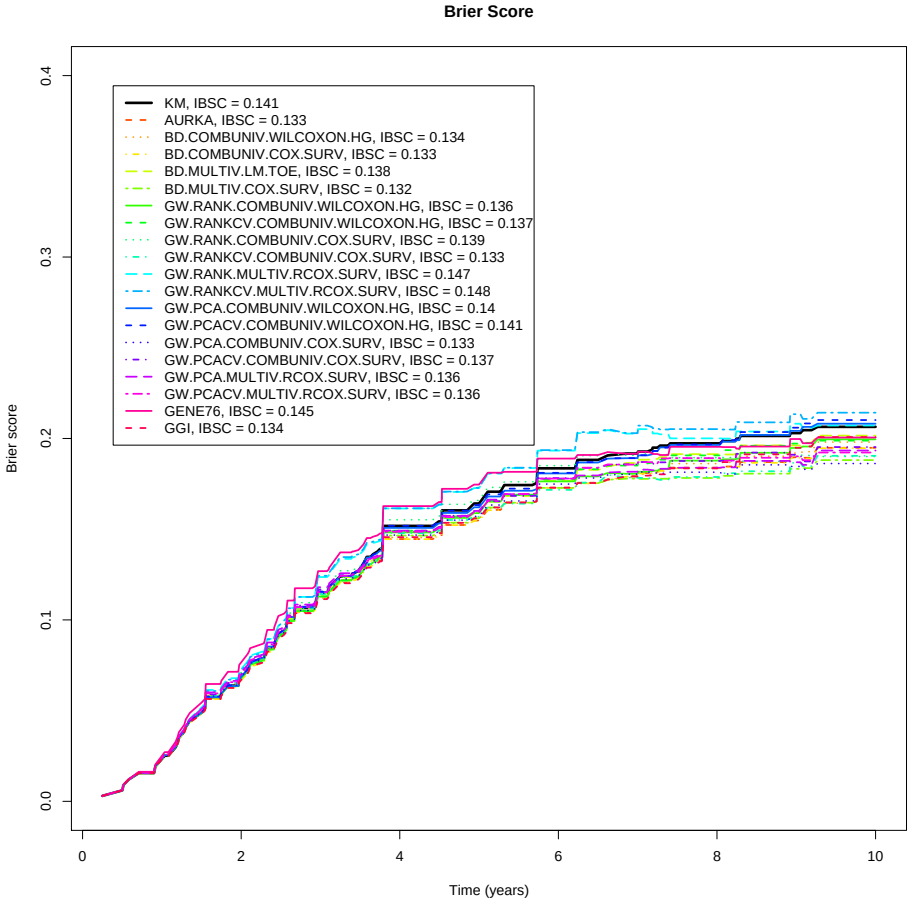

The following figure shows the pairwise comparison between the different methods. A red box represents a significant superiority (paired Wilcoxon rank sum test p-value $< 0.05$) of the method in row against the method in column. A green box represents no significant superiority (paired Wilcoxon rank sum test p-value $>= 0.05$) of the method in row against the method in column.



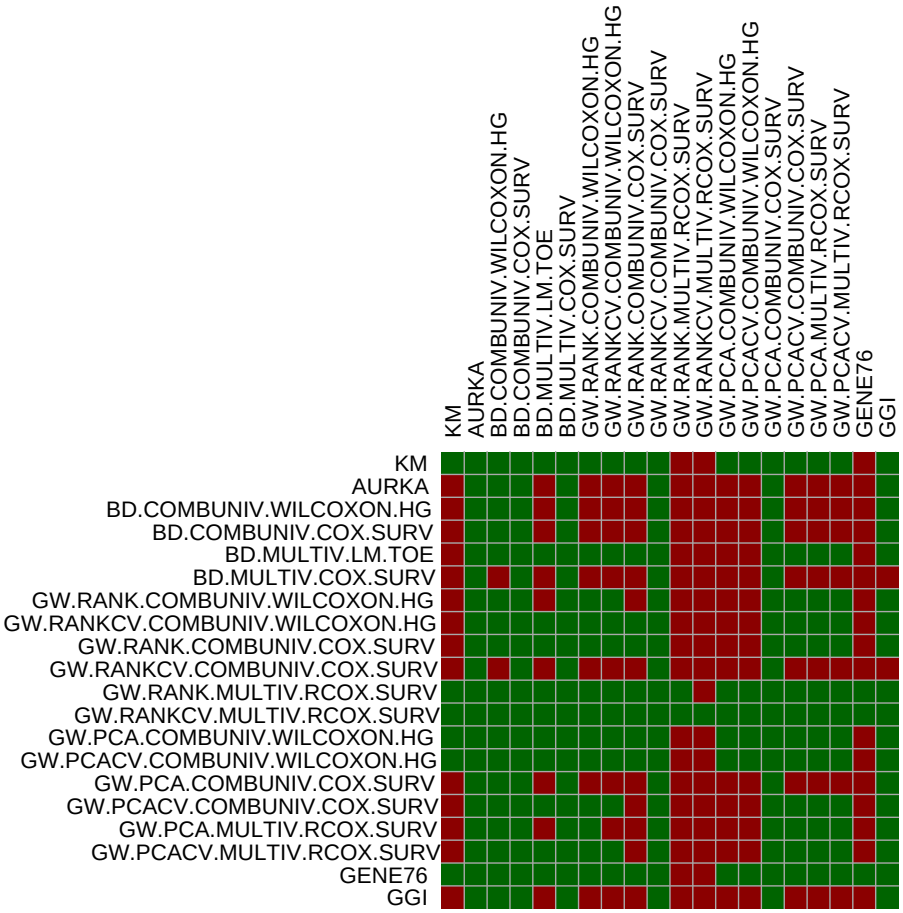

The following table shows the comparison between the different methods and AURKA. A text in bold represents a significant superiority (Student t test p-value $< 0.05$) of the method in row against AURKA.

| | IBSC |
|---|---|
| KM | 0.141 |
| AURKA | **0.132** |
| BD.COMBUNIV.WILCOXON.HG | 0.131 |
| BD.COMBUNIV.COX.SURV | 0.131 |
| BD.MULTIV.LM.TOE | 0.132 |
| BD.MULTIV.COX.SURV | 0.132 |
| GW.RANK.COMBUNIV.WILCOXON.HG | 0.131 |
| GW.RANKCV.COMBUNIV.WILCOXON.HG | 0.139 |
| GW.RANK.COMBUNIV.COX.SURV | 0.158 |
| GW.RANKCV.COMBUNIV.COX.SURV | 0.131 |
| GW.RANK.MULTIV.RCOX.SURV | 0.175 |
| GW.RANKCV.MULTIV.RCOX.SURV | 0.161 |
| GW.PCA.COMBUNIV.WILCOXON.HG | 0.136 |
| GW.PCACV.COMBUNIV.WILCOXON.HG | 0.137 |
| GW.PCA.COMBUNIV.COX.SURV | 0.153 |
| GW.PCACV.COMBUNIV.COX.SURV | 0.17 |
| GW.PCA.MULTIV.RCOX.SURV | 0.171 |
| GW.PCACV.MULTIV.RCOX.SURV | 0.148 |
| GENE76 | 0.149 |
| GGI | **0.13** |

**Brier Score for Risk Group** The following figure shows the Brier score w.r.t. the time for each method :



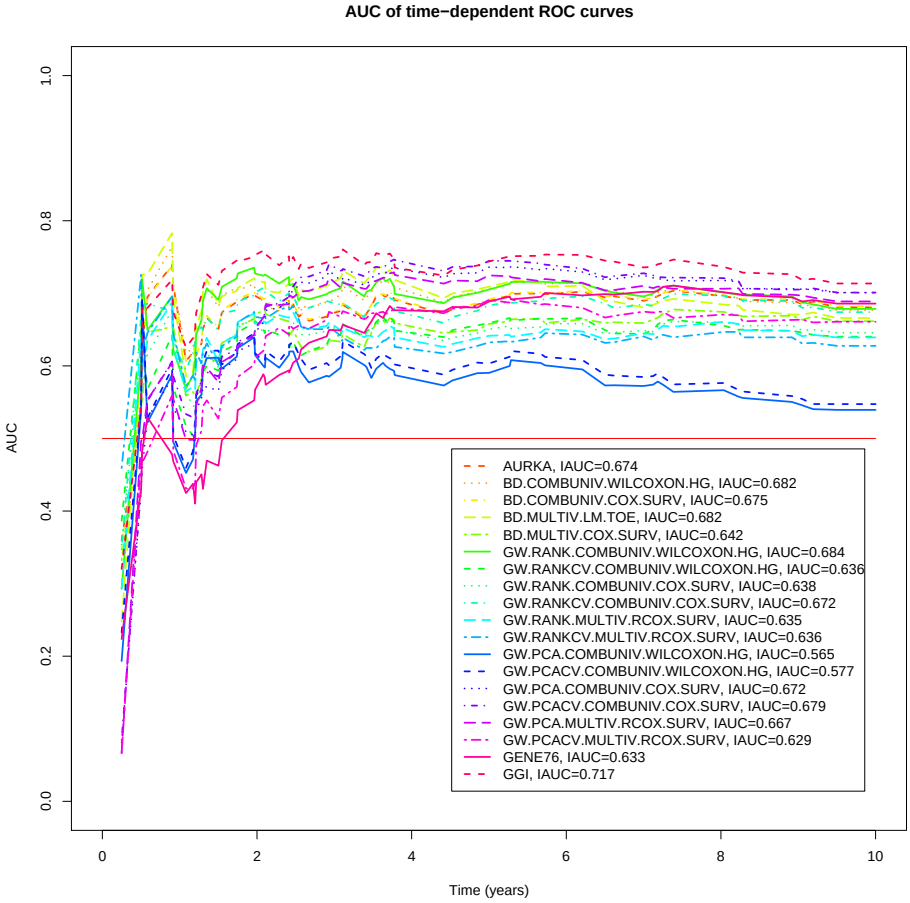

The following figure shows the pairwise comparison between the different methods. A red box represents a significant superiority (paired Wilcoxon rank sum test p-value $< 0.05$) of the method in row against the method in column. A green box represents no significant superiority (paired Wilcoxon rank sum test p-value $>= 0.05$) of the method in row against the method in column.



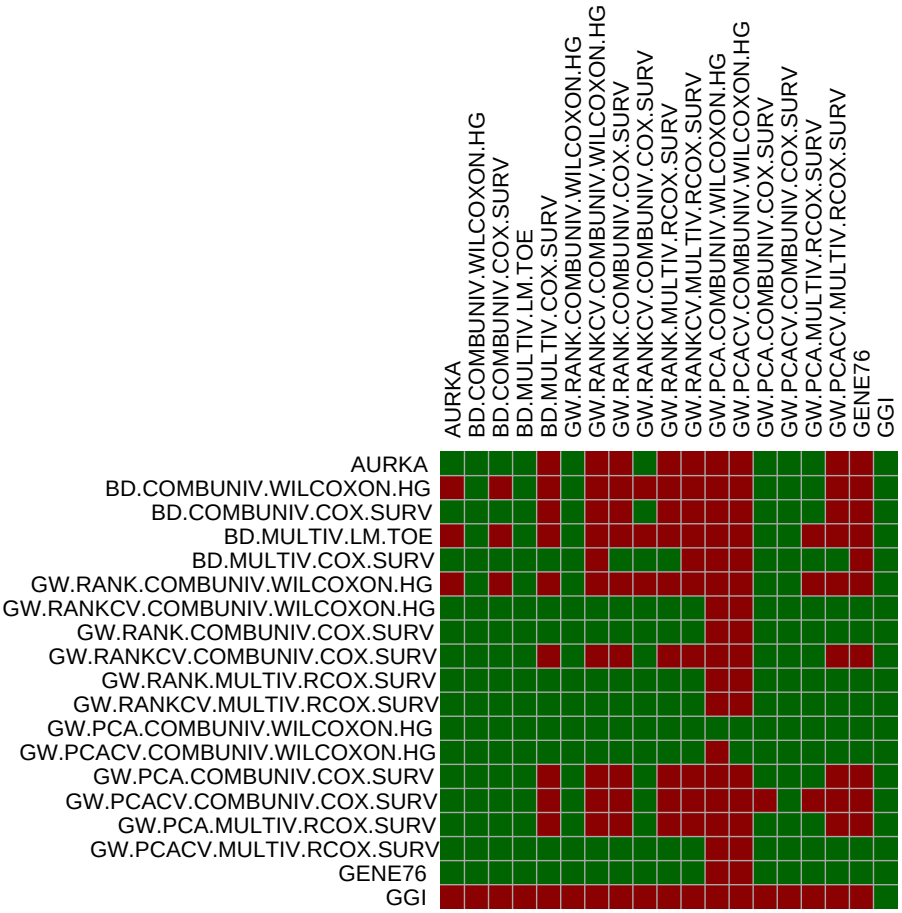

The following table shows the comparison between the different methods and AURKA. A text in bold represents a significant superiority (Student t test p-value $< 0.05$) of the method in row against AURKA.

| | IBSC |
|---|---|
| KM | 0.141 |
| AURKA | **0.133** |
| BD.COMBUNIV.WILCOXON.HG | 0.134 |
| BD.COMBUNIV.COX.SURV | 0.133 |
| BD.MULTIV.LM.TOE | 0.138 |
| BD.MULTIV.COX.SURV | 0.132 |
| GW.RANK.COMBUNIV.WILCOXON.HG | 0.136 |
| GW.RANKCV.COMBUNIV.WILCOXON.HG | 0.137 |
| GW.RANK.COMBUNIV.COX.SURV | 0.139 |
| GW.RANKCV.COMBUNIV.COX.SURV | 0.133 |
| GW.RANK.MULTIV.RCOX.SURV | 0.147 |
| GW.RANKCV.MULTIV.RCOX.SURV | 0.148 |
| GW.PCA.COMBUNIV.WILCOXON.HG | 0.14 |
| GW.PCACV.COMBUNIV.WILCOXON.HG | 0.141 |
| GW.PCA.COMBUNIV.COX.SURV | 0.133 |
| GW.PCACV.COMBUNIV.COX.SURV | 0.137 |
| GW.PCA.MULTIV.RCOX.SURV | 0.136 |
| GW.PCACV.MULTIV.RCOX.SURV | 0.136 |
| GENE76 | 0.145 |
| GGI | 0.134 |

**Time-Dependent ROC Curves for Risk Score** The following figure shows the evolution of the AUC of the time-dependent ROC curves with respect to the time for each method :



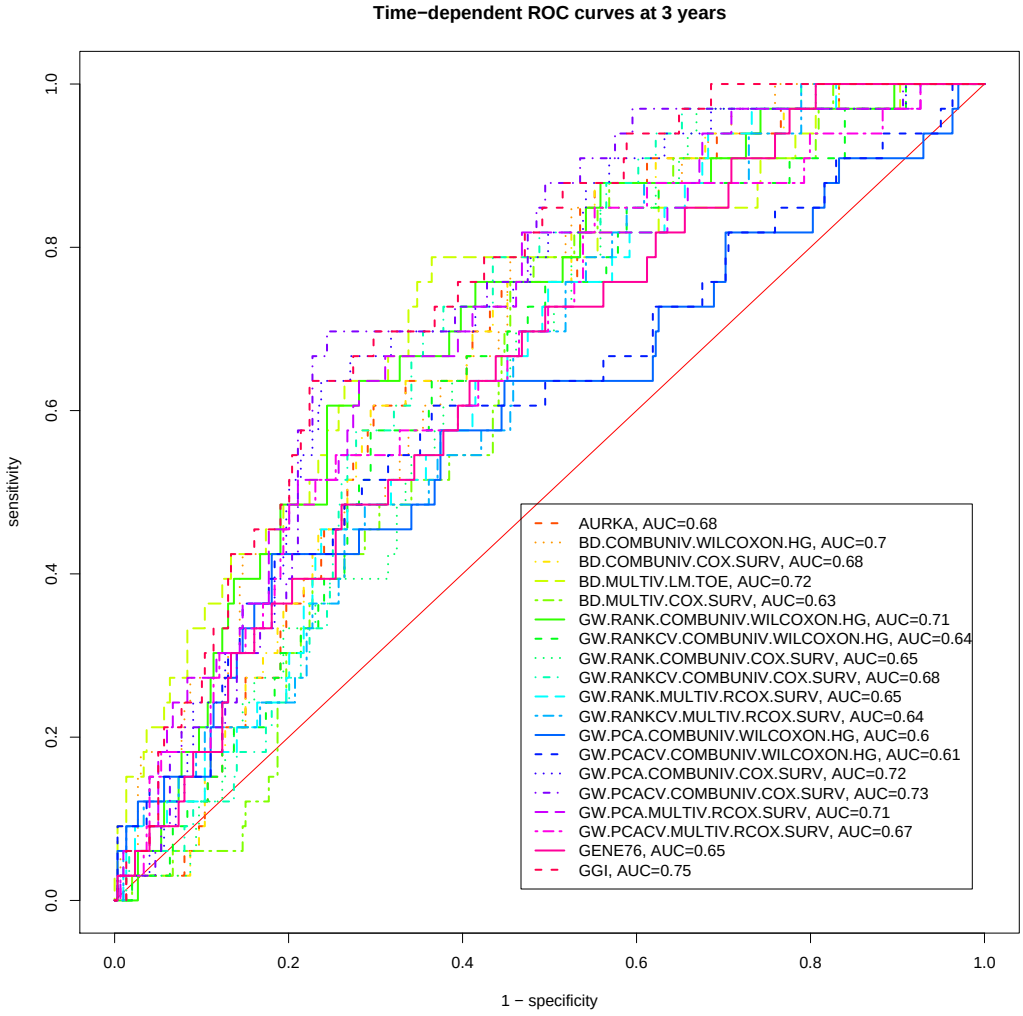

The following figure shows the pairwise comparison between the different methods. A red box represents a significant superiority (paired Wilcoxon rank sum test p-value $< 0.05$) of the method in row against the method in column. A green box represents no significant superiority (paired Wilcoxon rank sum test p-value $>= 0.05$) of the method in row against the method in column.

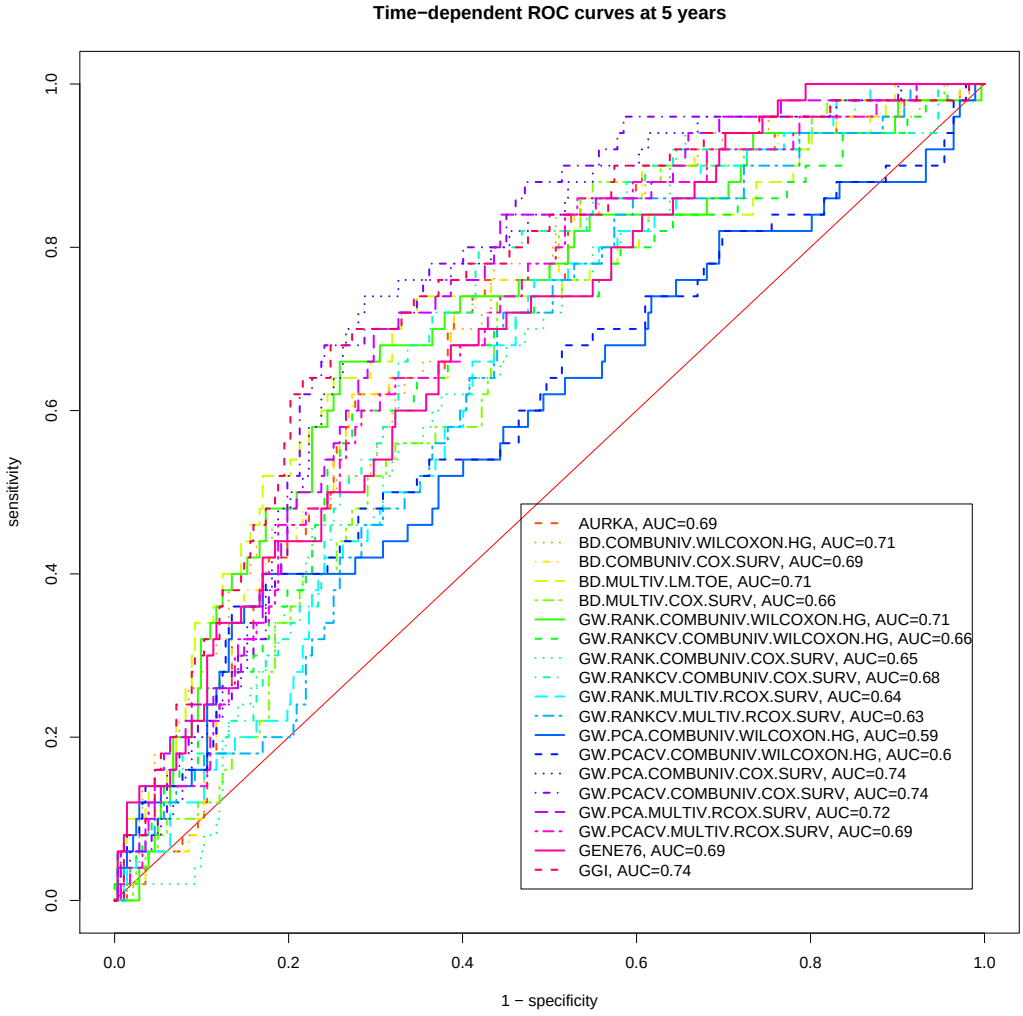


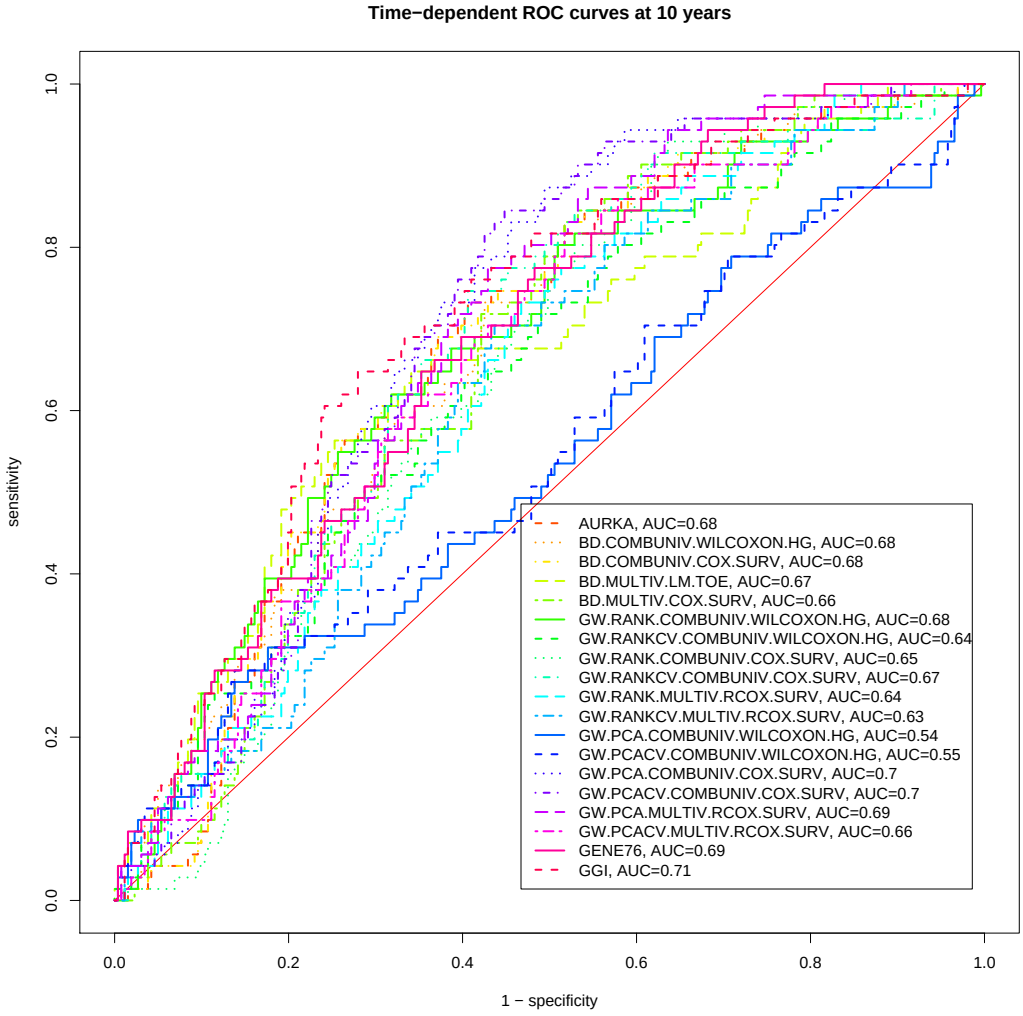

The following table shows the comparison between the different methods and AURKA. A text in bold represents a significant superiority (Student t test p-value $< 0.05$) of the method in row against AURKA.

| | IAUC |
|---|---|
| AURKA | 0.674 |
| BD.COMBUNIV.WILCOXON.HG | **0.682** |
| BD.COMBUNIV.COX.SURV | 0.675 |
| BD.MULTIV.LM.TOE | **0.682** |
| BD.MULTIV.COX.SURV | 0.642 |
| GW.RANK.COMBUNIV.WILCOXON.HG | **0.684** |
| GW.RANKCV.COMBUNIV.WILCOXON.HG | 0.636 |
| GW.RANK.COMBUNIV.COX.SURV | 0.638 |
| GW.RANKCV.COMBUNIV.COX.SURV | 0.672 |
| GW.RANK.MULTIV.RCOX.SURV | 0.635 |
| GW.RANKCV.MULTIV.RCOX.SURV | 0.636 |
| GW.PCA.COMBUNIV.WILCOXON.HG | 0.565 |
| GW.PCACV.COMBUNIV.WILCOXON.HG | 0.577 |
| GW.PCA.COMBUNIV.COX.SURV | 0.672 |
| GW.PCACV.COMBUNIV.COX.SURV | 0.679 |
| GW.PCA.MULTIV.RCOX.SURV | 0.667 |
| GW.PCACV.MULTIV.RCOX.SURV | 0.629 |
| GENE76 | 0.633 |
| GGI | **0.717** |

The following figure shows the time-dependent ROC curves and their corresponding AUC for each method :


Time−dependent ROC curves at 3 years
sensitivity
1 − specificity
0.0
0.2
0.4
0.6
0.8
1.0
AURKA, AUC=0.68
BD.COMBUNIV.WILCOXON.HG, AUC=0.7
BD.COMBUNIV.COX.SURV, AUC=0.68
BD.MULTIV.LM.TOE, AUC=0.72
BD.MULTIV.COX.SURV, AUC=0.63
GW.RANK.COMBUNIV.WILCOXON.HG, AUC=0.71
GW.RANKCV.COMBUNIV.WILCOXON.HG, AUC=0.64
GW.RANK.COMBUNIV.COX.SURV, AUC=0.65
GW.RANKCV.COMBUNIV.COX.SURV, AUC=0.68
GW.RANK.MULTIV.RCOX.SURV, AUC=0.65
GW.RANKCV.MULTIV.RCOX.SURV, AUC=0.64
GW.PCA.COMBUNIV.WILCOXON.HG, AUC=0.6
GW.PCACV.COMBUNIV.WILCOXON.HG, AUC=0.61
GW.PCA.COMBUNIV.COX.SURV, AUC=0.72
GW.PCACV.COMBUNIV.COX.SURV, AUC=0.73
GW.PCA.MULTIV.RCOX.SURV, AUC=0.71
GW.PCACV.MULTIV.RCOX.SURV, AUC=0.67
GENE76, AUC=0.65
GGI, AUC=0.75


Time−dependent ROC curves at 5 years
sensitivity
1 − specificity
0.0
0.2
0.4
0.6
0.8
1.0
AURKA, AUC=0.69
BD.COMBUNIV.WILCOXON.HG, AUC=0.71
BD.COMBUNIV.COX.SURV, AUC=0.69
BD.MULTIV.LM.TOE, AUC=0.71
BD.MULTIV.COX.SURV, AUC=0.66
GW.RANK.COMBUNIV.WILCOXON.HG, AUC=0.71
GW.RANKCV.COMBUNIV.WILCOXON.HG, AUC=0.66
GW.RANK.COMBUNIV.COX.SURV, AUC=0.65
GW.RANKCV.COMBUNIV.COX.SURV, AUC=0.68
GW.RANK.MULTIV.RCOX.SURV, AUC=0.64
GW.RANKCV.MULTIV.RCOX.SURV, AUC=0.63
GW.PCA.COMBUNIV.WILCOXON.HG, AUC=0.59
GW.PCACV.COMBUNIV.WILCOXON.HG, AUC=0.6
GW.PCA.COMBUNIV.COX.SURV, AUC=0.74
GW.PCACV.COMBUNIV.COX.SURV, AUC=0.74
GW.PCA.MULTIV.RCOX.SURV, AUC=0.72
GW.PCACV.MULTIV.RCOX.SURV, AUC=0.69
GENE76, AUC=0.69
GGI, AUC=0.74




The following table shows the specificity for a given sensitivity of 90% at some points in time:

| | years.3 | years.5 | years.10 |
|---|---|---|---|
| AURKA | 0.388 | 0.394 | 0.341 |
| BD.COMBUNIV.WILCOXON.HG | 0.348 | 0.362 | 0.345 |
| BD.COMBUNIV.COX.SURV | 0.378 | 0.394 | 0.364 |
| BD.MULTIV.LM.TOE | 0.258 | 0.220 | 0.238 |
| BD.MULTIV.COX.SURV | 0.358 | 0.372 | 0.395 |
| GW.RANK.COMBUNIV.WILCOXON.HG | 0.314 | 0.277 | 0.295 |
| GW.RANKCV.COMBUNIV.WILCOXON.HG | 0.224 | 0.206 | 0.234 |
| GW.RANK.COMBUNIV.COX.SURV | 0.351 | 0.362 | 0.383 |
| GW.RANKCV.COMBUNIV.COX.SURV | 0.398 | 0.390 | 0.387 |
| GW.RANK.MULTIV.RCOX.SURV | 0.324 | 0.326 | 0.284 |
| GW.RANKCV.MULTIV.RCOX.SURV | 0.271 | 0.277 | 0.291 |
| GW.PCA.COMBUNIV.WILCOXON.HG | 0.167 | 0.067 | 0.061 |
| GW.PCACV.COMBUNIV.WILCOXON.HG | 0.171 | 0.110 | 0.107 |
| GW.PCA.COMBUNIV.COX.SURV | 0.431 | 0.450 | 0.452 |
| GW.PCACV.COMBUNIV.COX.SURV | 0.465 | 0.486 | 0.467 |
| GW.PCA.MULTIV.RCOX.SURV | 0.328 | 0.358 | 0.379 |
| GW.PCACV.MULTIV.RCOX.SURV | 0.207 | 0.319 | 0.333 |
| GENE76 | 0.291 | 0.309 | 0.356 |
| GGI | 0.415 | 0.422 | 0.333 |

The following table shows the sensitivity of the methods in leaving 33% of the patients in the low-risk group at some points in time:

| | years.3 | years.5 | years.10 |
|---|---|---|---|
| AURKA | 0.909 | 0.900 | 0.873 |
| BD.COMBUNIV.WILCOXON.HG | 0.879 | 0.880 | 0.873 |
| BD.COMBUNIV.COX.SURV | 0.909 | 0.900 | 0.887 |
| BD.MULTIV.LM.TOE | 0.848 | 0.840 | 0.789 |
| BD.MULTIV.COX.SURV | 0.909 | 0.900 | 0.901 |
| GW.RANK.COMBUNIV.WILCOXON.HG | 0.879 | 0.840 | 0.845 |
| GW.RANKCV.COMBUNIV.WILCOXON.HG | 0.879 | 0.840 | 0.831 |
| GW.RANK.COMBUNIV.COX.SURV | 0.879 | 0.860 | 0.873 |
| GW.RANKCV.COMBUNIV.COX.SURV | 0.939 | 0.900 | 0.887 |
| GW.RANK.MULTIV.RCOX.SURV | 0.848 | 0.840 | 0.845 |
| GW.RANKCV.MULTIV.RCOX.SURV | 0.879 | 0.860 | 0.845 |
| GW.PCA.COMBUNIV.WILCOXON.HG | 0.727 | 0.760 | 0.718 |
| GW.PCACV.COMBUNIV.WILCOXON.HG | 0.727 | 0.740 | 0.704 |
| GW.PCA.COMBUNIV.COX.SURV | 0.939 | 0.940 | 0.944 |
| GW.PCACV.COMBUNIV.COX.SURV | 0.970 | 0.960 | 0.930 |
| GW.PCA.MULTIV.RCOX.SURV | 0.848 | 0.880 | 0.887 |
| GW.PCACV.MULTIV.RCOX.SURV | 0.879 | 0.880 | 0.873 |
| GENE76 | 0.818 | 0.840 | 0.859 |
| GGI | 0.939 | 0.900 | 0.859 |

The following table shows the specificity of the methods in leaving 33% of the patients in the low-risk group at some points in time:

| | years.3 | years.5 | years.10 |
|---|---|---|---|
| AURKA | 0.351 | 0.365 | 0.379 |
| BD.COMBUNIV.WILCOXON.HG | 0.351 | 0.365 | 0.383 |
| BD.COMBUNIV.COX.SURV | 0.355 | 0.369 | 0.387 |
| BD.MULTIV.LM.TOE | 0.348 | 0.358 | 0.360 |
| BD.MULTIV.COX.SURV | 0.355 | 0.369 | 0.391 |
| GW.RANK.COMBUNIV.WILCOXON.HG | 0.351 | 0.358 | 0.375 |
| GW.RANKCV.COMBUNIV.WILCOXON.HG | 0.351 | 0.358 | 0.372 |
| GW.RANK.COMBUNIV.COX.SURV | 0.351 | 0.362 | 0.383 |
| GW.RANKCV.COMBUNIV.COX.SURV | 0.358 | 0.369 | 0.387 |
| GW.RANK.MULTIV.RCOX.SURV | 0.348 | 0.358 | 0.375 |
| GW.RANKCV.MULTIV.RCOX.SURV | 0.351 | 0.362 | 0.375 |
| GW.PCA.COMBUNIV.WILCOXON.HG | 0.334 | 0.344 | 0.341 |
| GW.PCACV.COMBUNIV.WILCOXON.HG | 0.334 | 0.340 | 0.337 |
| GW.PCA.COMBUNIV.COX.SURV | 0.358 | 0.376 | 0.402 |
| GW.PCACV.COMBUNIV.COX.SURV | 0.361 | 0.379 | 0.398 |
| GW.PCA.MULTIV.RCOX.SURV | 0.348 | 0.365 | 0.387 |
| GW.PCACV.MULTIV.RCOX.SURV | 0.351 | 0.365 | 0.383 |
| GENE76 | 0.344 | 0.358 | 0.379 |
| GGI | 0.358 | 0.369 | 0.379 |