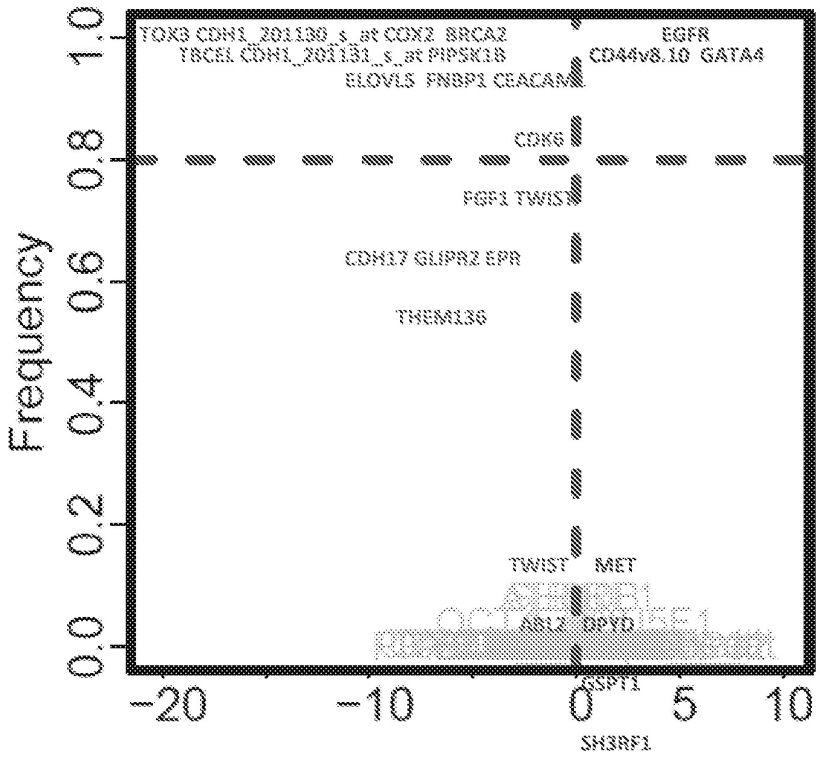

Frequency
0.0
0.2
0.4
0.6
0.8
1.0
−20
−10
0
5
10
TOX3 CDH1_201130_s_at COX2 BRCA2
TBCEL CDH1_201131_s_at PIPSK1B
ELOVL5 FNBP1 CEACAM
EGFR
CD44v8.10 GATA4
CDK6
FGF1 TWIST
CDH17 GLIPR2 EPR
THEM136
TWIST MET
ASL2 DPYD
SH3RF1

Fig. 1

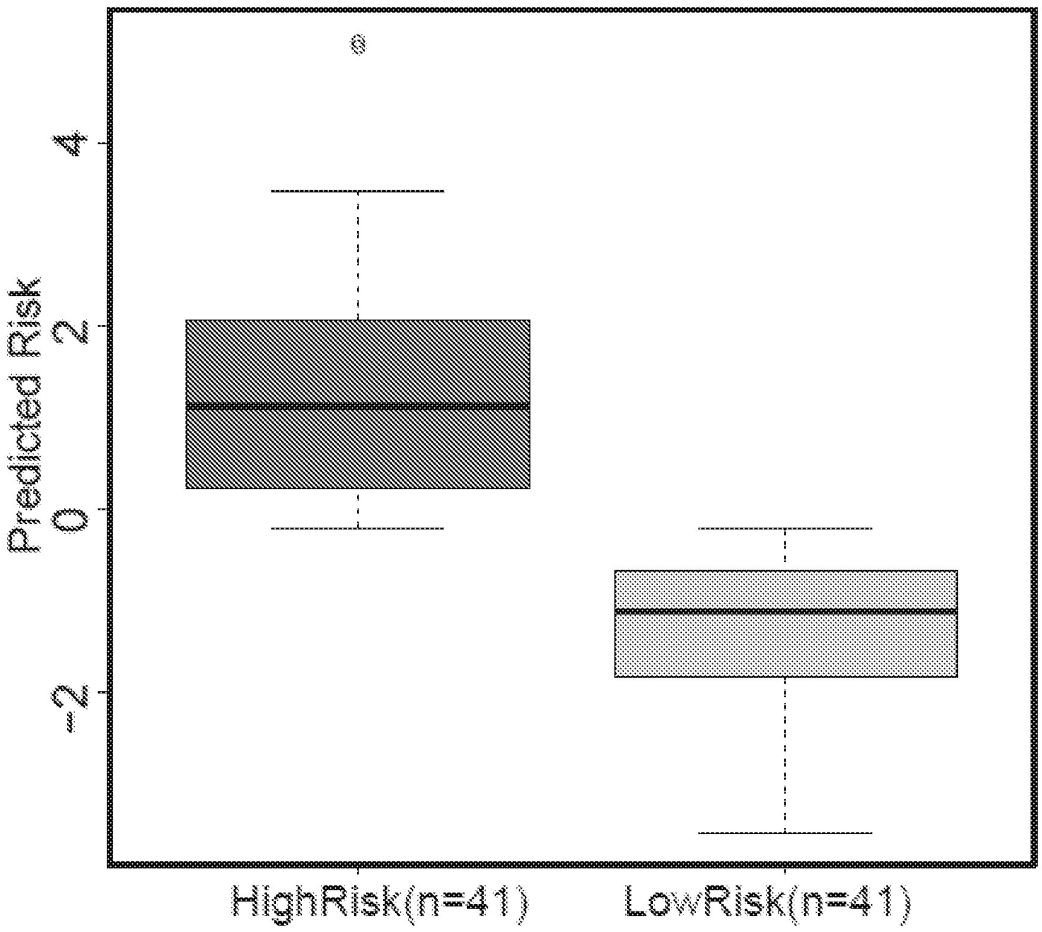


Fig. 2

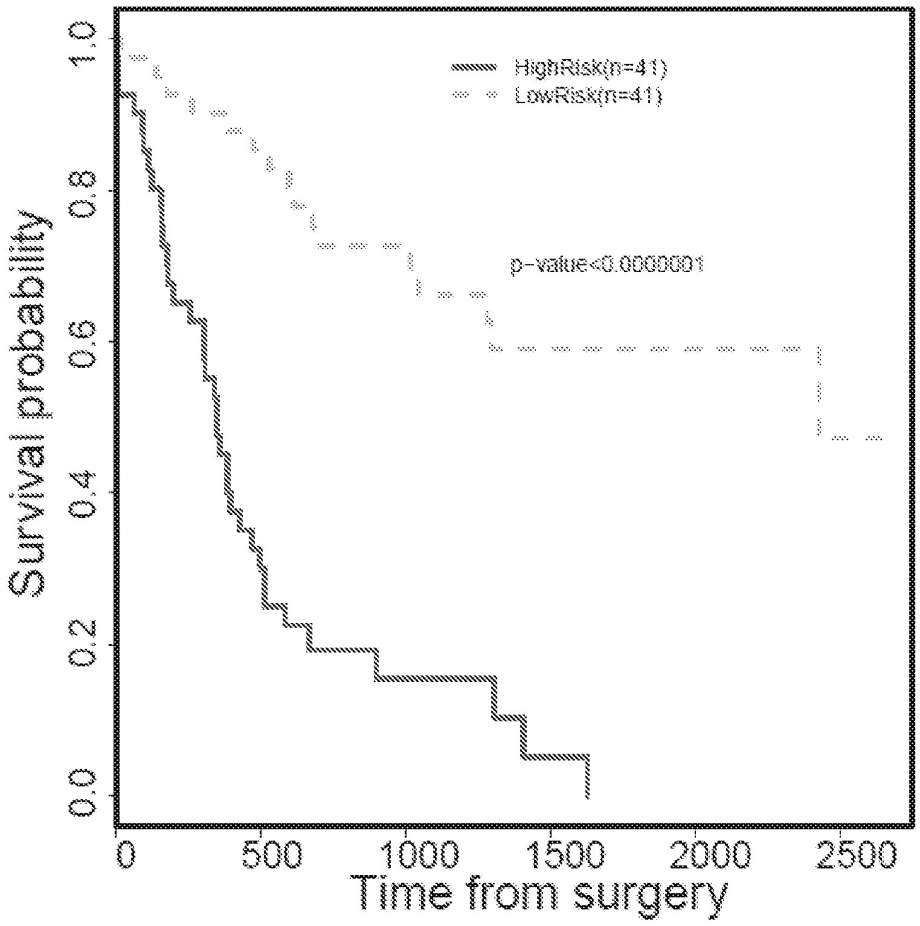

HighRisk(n=41)
LowRisk(n=41)
p-value<0.0000001
Survival probability
0.0
0.2
0.4
0.6
0.8
1.0
0
500
1000
1500
2000
2500
Time from surgery

Fig. 3

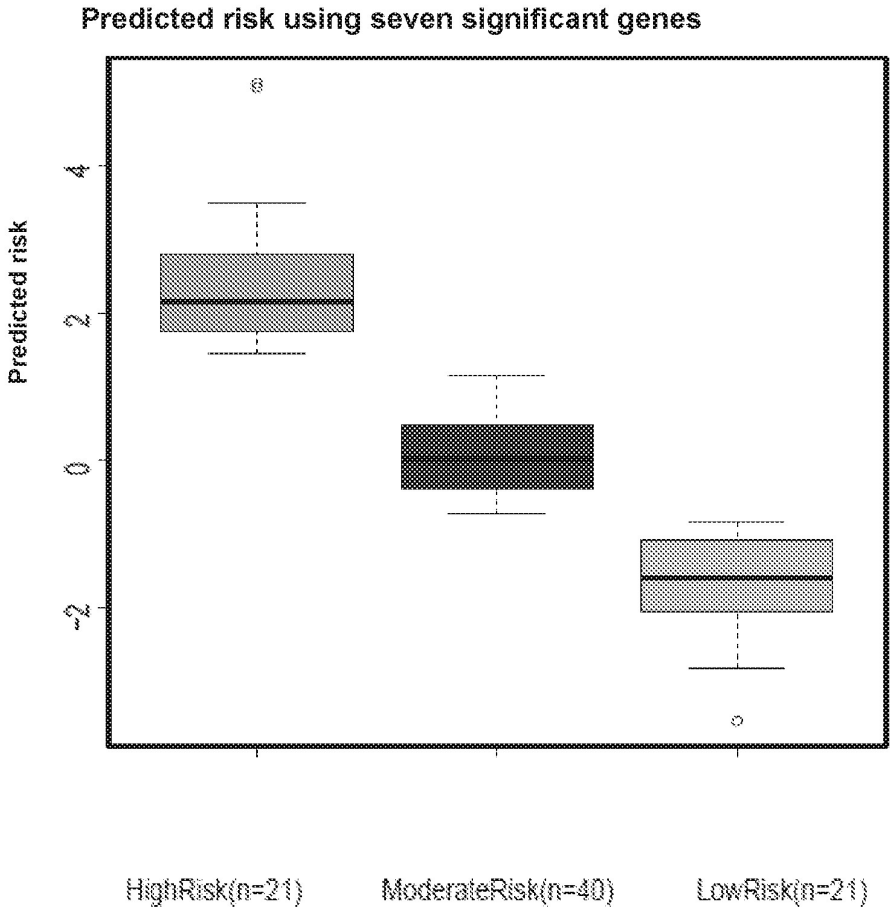

Predicted risk using seven significant genes
Predicted risk
4
2
0
-2
HighRisk(n=21)
ModerateRisk(n=40)
LowRisk(n=21)

Fig. 4

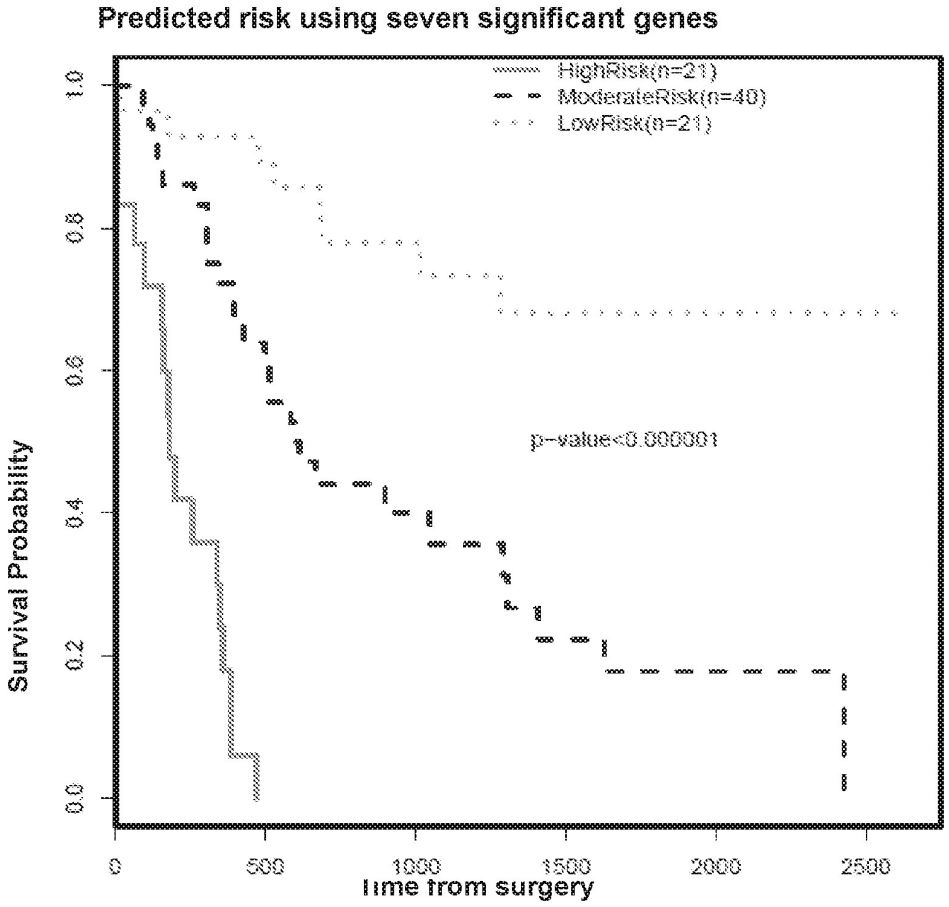

Predicted risk using seven significant genes
HighRisk(n=21)
ModerateRisk(n=40)
LowRisk(n=21)
p-value<0.000001
Survival Probability
0.0
0.2
0.4
0.6
0.8
1.0
0
500
1000
1500
2000
2500
Time from surgery

Fig. 5

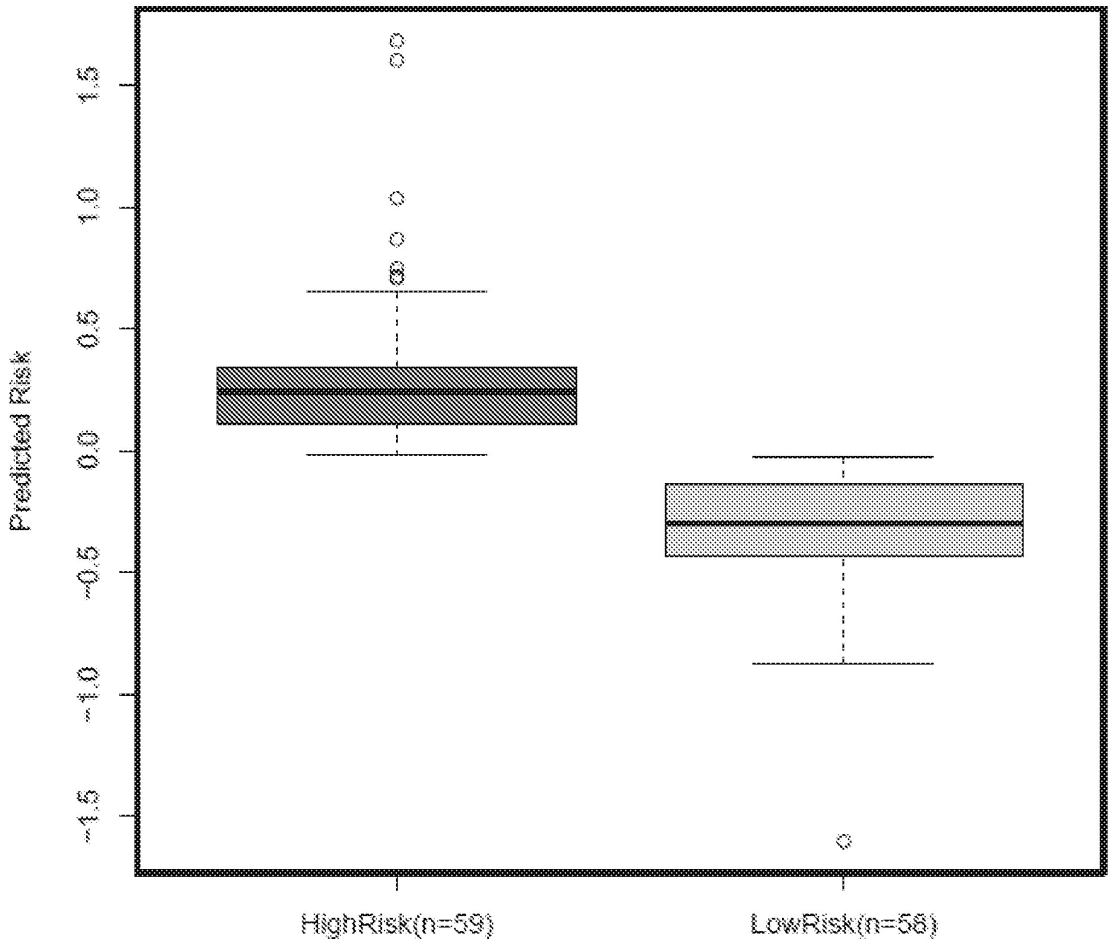


Fig. 6

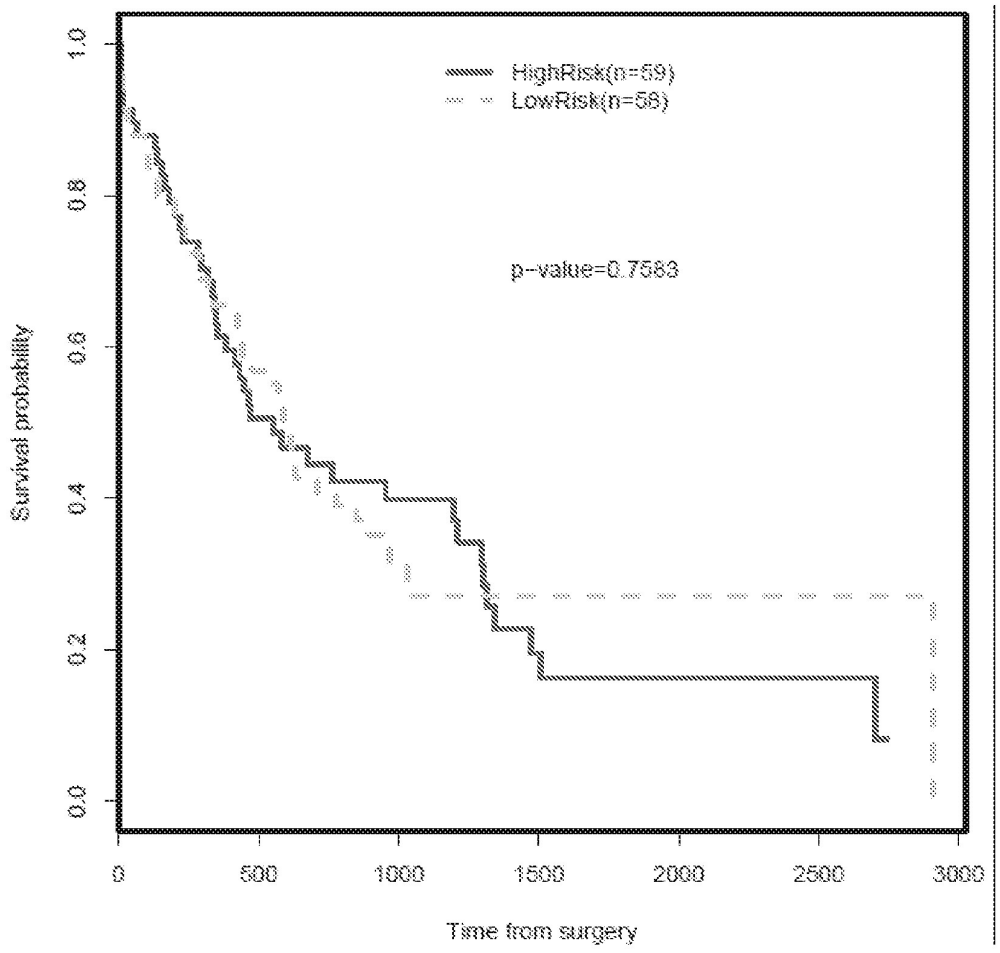

HighRisk(n=69)
LowRisk(n=58)
p-value=0.7583
Survival probability
0.0
0.2
0.4
0.6
0.8
1.0
0
500
1000
1500
2000
2500
3000
Time from surgery

Fig. 7

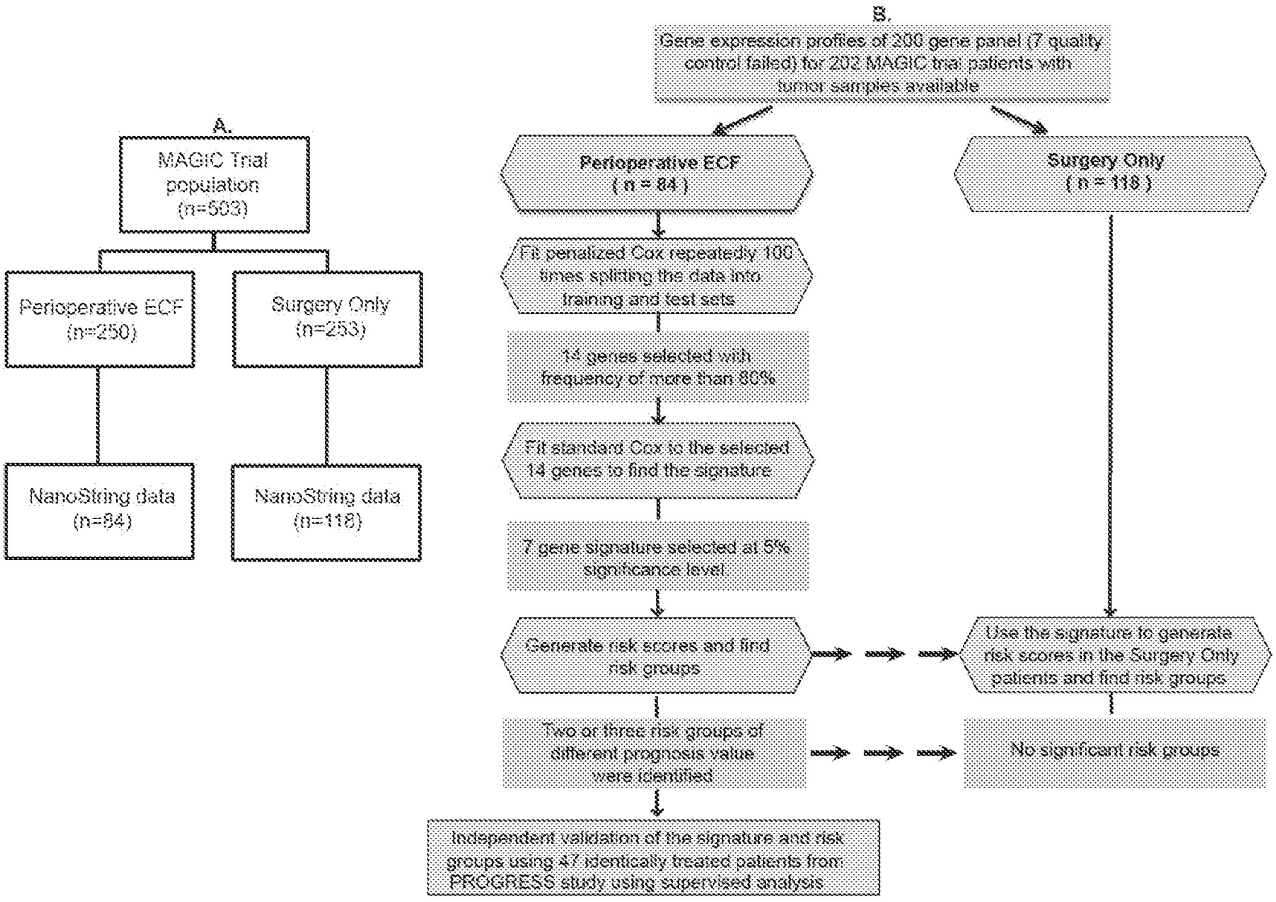

A.
MAGIC Trial population (n=503)
Perioperative ECF (n=250)
Surgery Only (n=253)
NanoString data (n=84)
NanoString data (n=118)
B.
Gene expression profiles of 200 gene panel (7 quality control failed) for 202 MAGIC trial patients with tumor samples available
Perioperative ECF (n = 84)
Surgery Only (n = 118)
Fit penalized Cox repeatedly 100 times splitting the data into training and test sets
14 genes selected with frequency of more than 80%
Fit standard Cox to the selected 14 genes to find the signature
7 gene signature selected at 5% significance level
Generate risk scores and find risk groups
Use the signature to generate risk scores in the Surgery Only patients and find risk groups
Two or three risk groups of different prognosis value were identified
No significant risk groups
Independent validation of the signature and risk groups using 47 identically treated patients from PROGRESS study using supervised analysis

Fig. 8

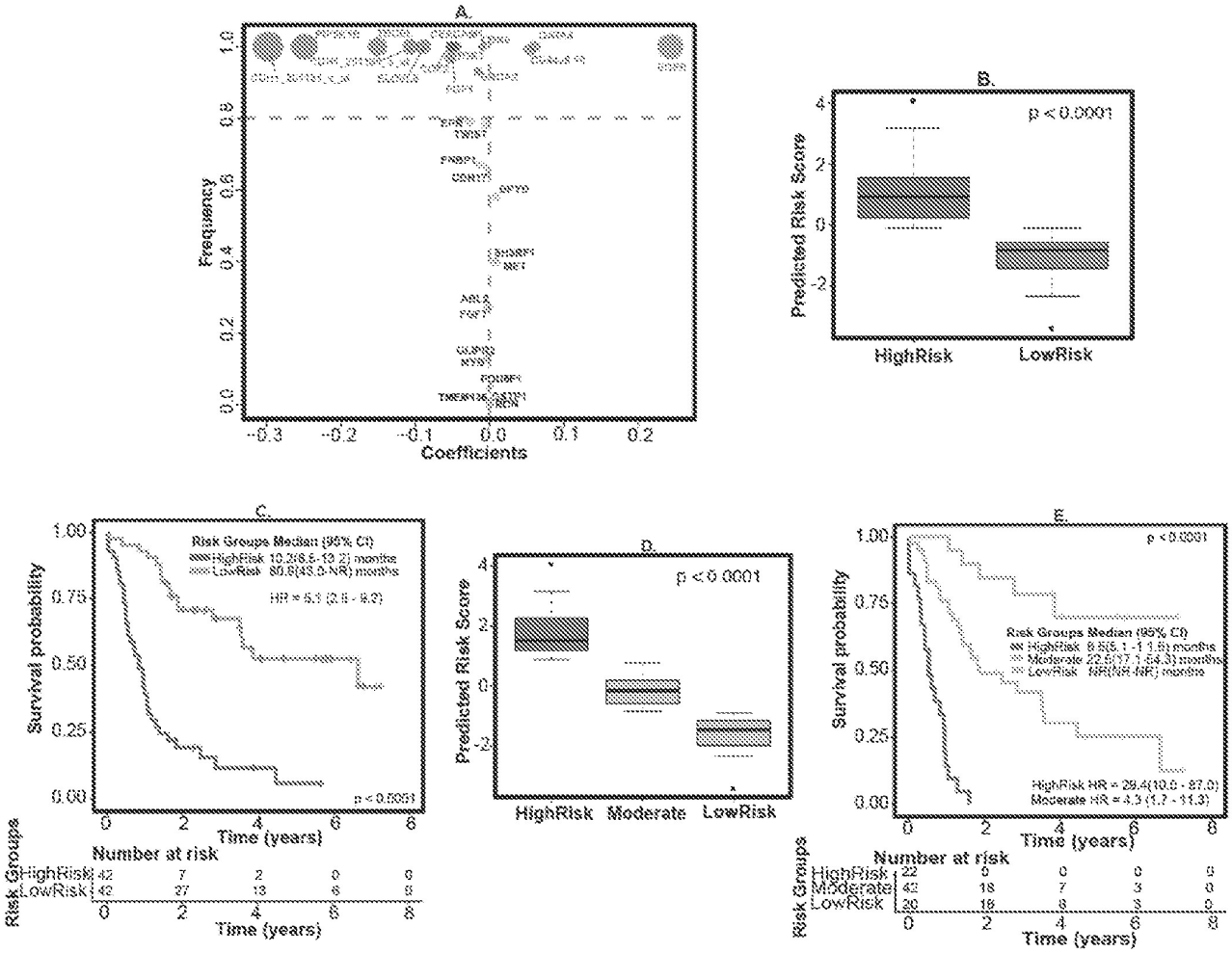

A.
Frequency
Coefficients
B.
p < 0.0001
Predicted Risk Score
HighRisk
LowRisk
C.
Survival probability
Risk Groups Median (95% CI)
HighRisk 10.2(8.5-13.2) months
LowRisk 80.8(43.0-NR) months
HR = 5.1 (2.8 - 9.2)
p < 0.0001
Time (years)
Number at risk
HighRisk 42 7 2 0 0
LowRisk 42 27 13 6 0
Risk Groups
D.
p < 0.0001
Predicted Risk Score
HighRisk Moderate LowRisk
E.
p < 0.0001
Survival probability
Risk Groups Median (95% CI)
HighRisk 8.5(5.1 - 11.5) months
Moderate 22.5(17.1-54.3) months
LowRisk NR(NR-NR) months
HighRisk HR = 29.4(10.0 - 87.0)
Moderate HR = 4.3 (1.7 - 11.3)
Time (years)
Number at risk
HighRisk 22 0 0 0 0
Moderate 42 18 7 3 0
LowRisk 20 18 6 3 0

Fig. 9

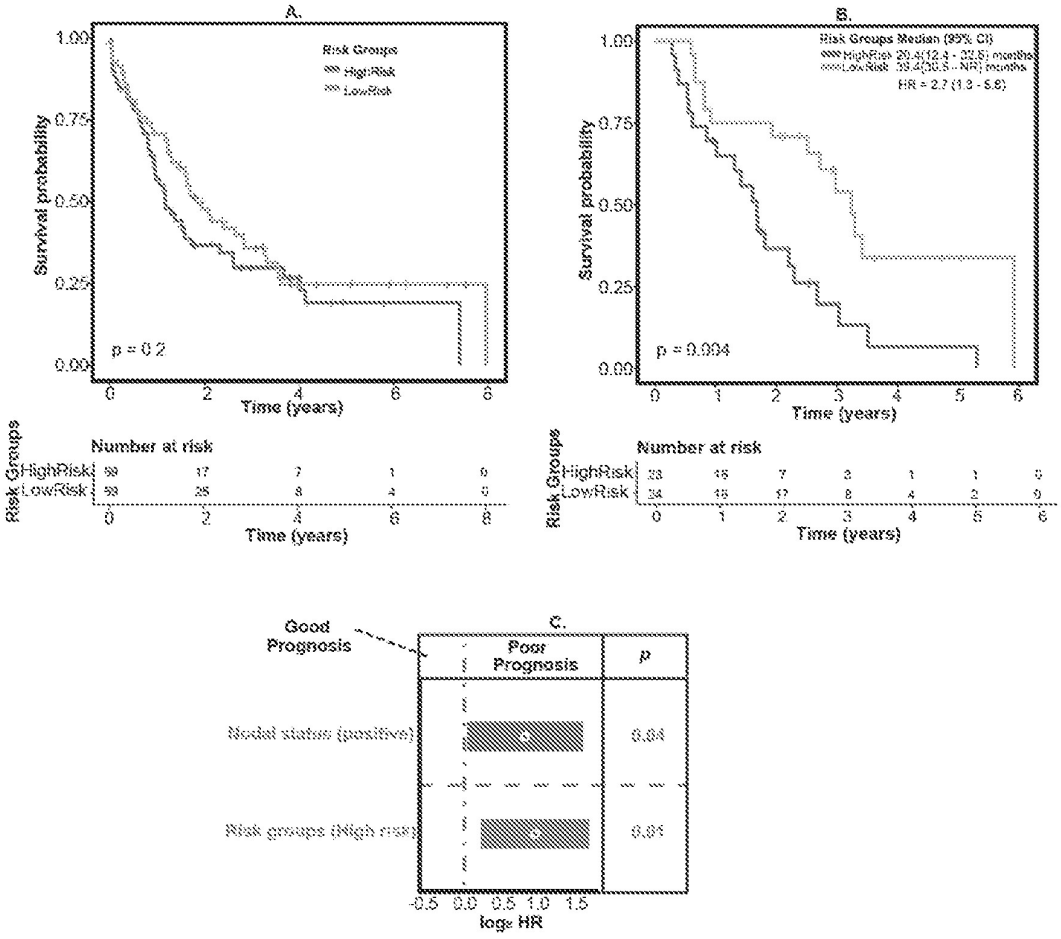

A.
Risk Groups
HighRisk
LowRisk
Survival probability
1.00
0.75
0.50
0.25
0.00
p = 0.2
0
2
4
6
8
Time (years)
Number at risk
Risk Groups
HighRisk 99 17 7 1 0
LowRisk 99 25 8 4 0
B.
Risk Groups Median (95% CI)
HighRisk 20.4 (12.4 - 32.6) months
LowRisk 39.4 (30.5 - NR) months
HR = 2.7 (1.8 - 5.8)
p = 0.004
Number at risk
HighRisk 33 15 7 3 1 1 0
LowRisk 34 15 12 8 4 2 0
C.
Good Prognosis
Poor Prognosis
p
Nodal status (positive)
0.04
Risk groups (High risk)
0.01
-0.5 0.0 0.5 1.0 1.5
log2 HR

Fig. 10

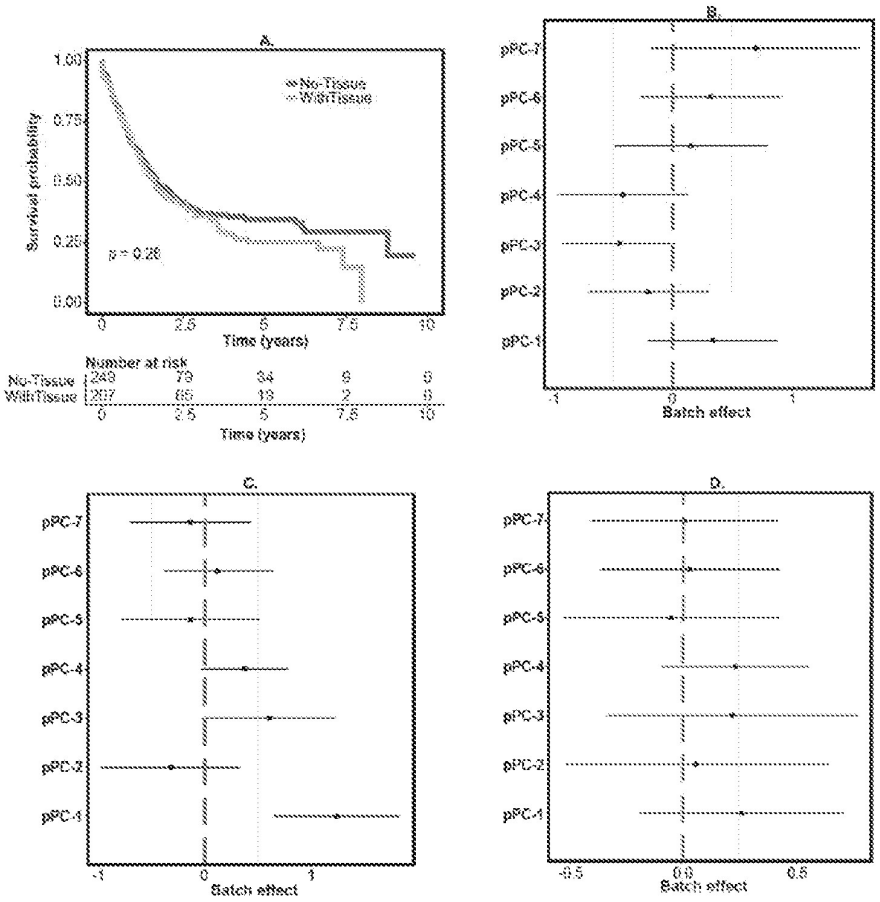

A.
No-Tissue
WithTissue
Survival probability
1.00
0.75
0.50
0.25
0.00
p = 0.28
0
2.5
5
7.5
10
Time (years)
Number at risk
No-Tissue 249 79 34 9 0
WithTissue 207 65 13 2 0
Time (years)
B.
pPC-7
pPC-6
pPC-5
pPC-4
pPC-3
pPC-2
pPC-1
-1
0
1
Batch effect
C.
pPC-7
pPC-6
pPC-5
pPC-4
pPC-3
pPC-2
pPC-1
-1
0
1
Batch effect
D.
pPC-7
pPC-6
pPC-5
pPC-4
pPC-3
pPC-2
pPC-1
-0.5
0.0
0.5
Batch effect

Fig. 11

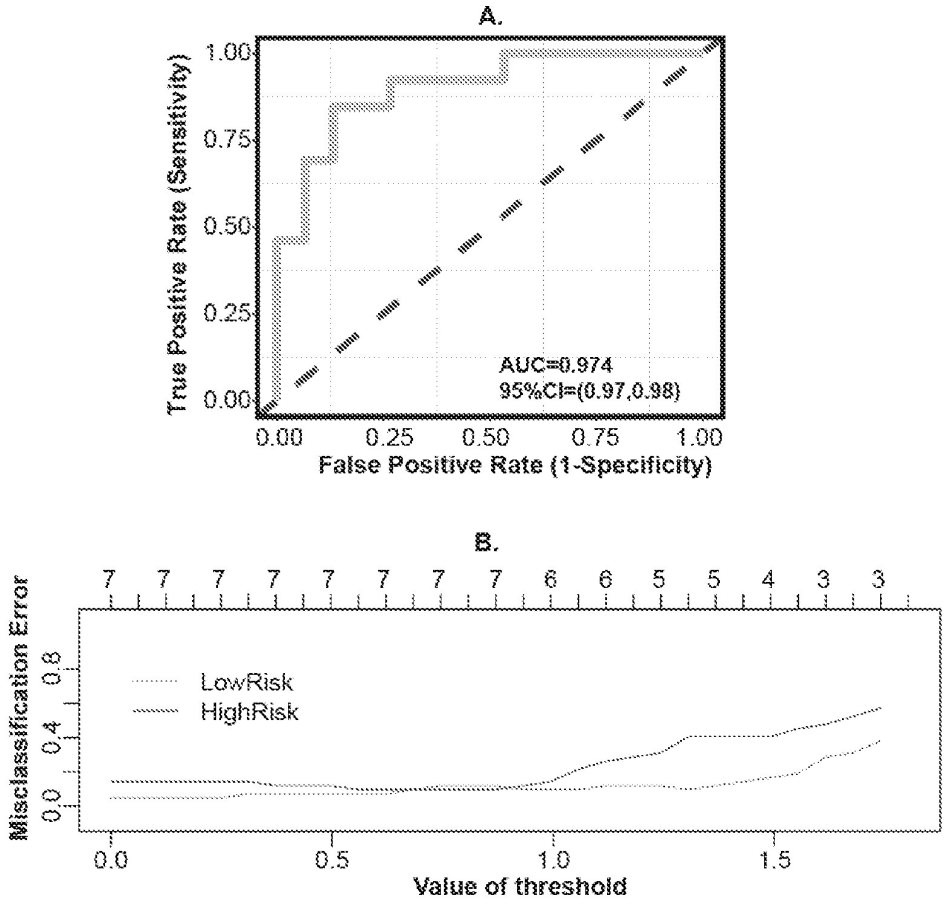

A.
True Positive Rate (Sensitivity)
1.00
0.75
0.50
0.25
0.00
AUC=0.974
95%CI=(0.97,0.98)
0.00
0.25
0.50
0.75
1.00
False Positive Rate (1-Specificity)
B.
Misclassification Error
0.0
0.4
0.8
7 7 7 7 7 7 7 7 6 6 5 5 4 3 3
LowRisk
HighRisk
0.0
0.5
1.0
1.5
Value of threshold

Fig. 12