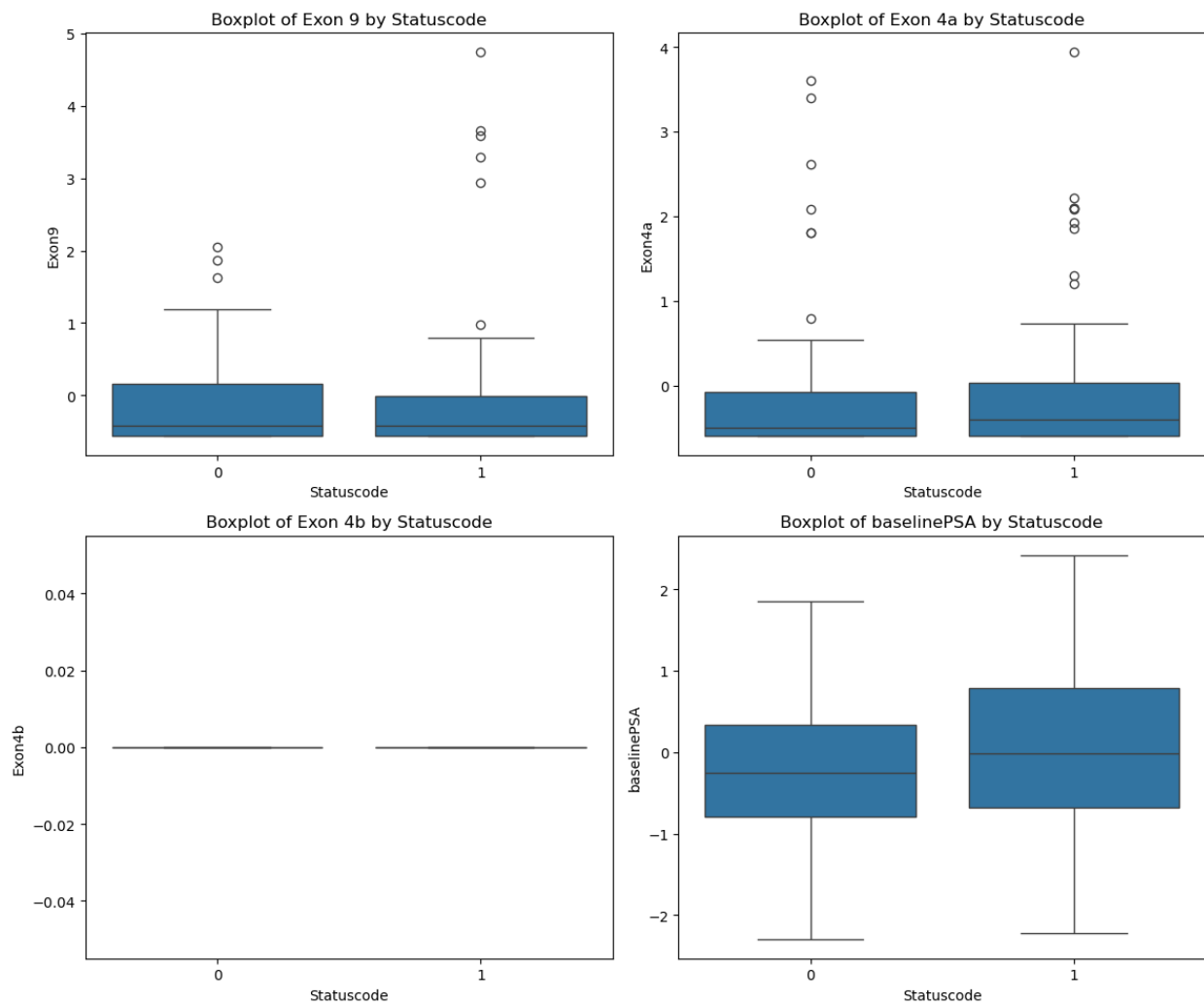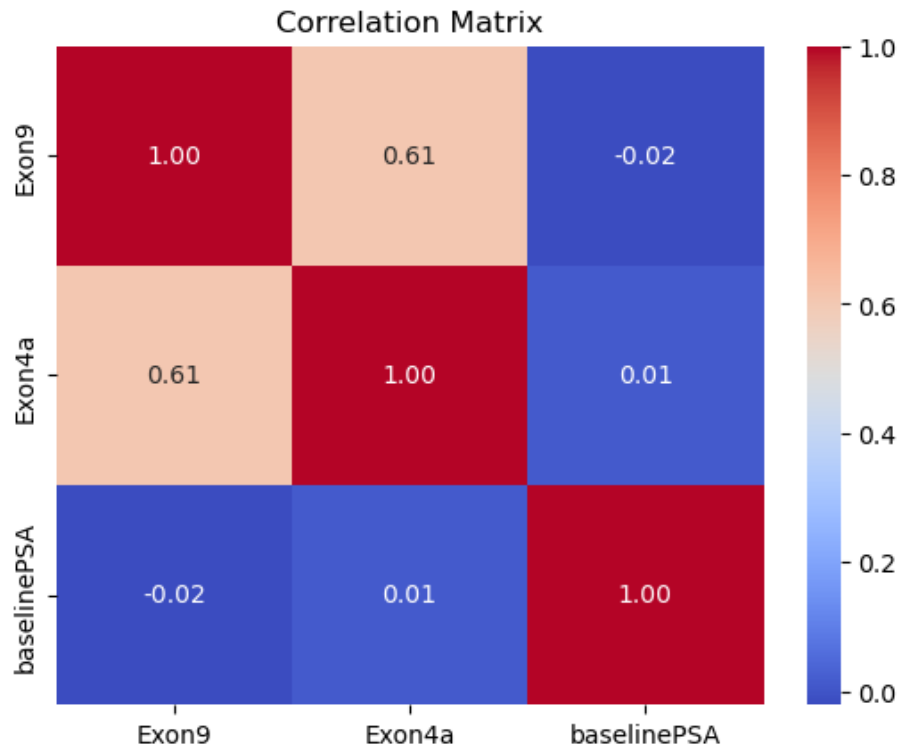**Pca Analysis Results**

Section 1: Status as the Response Variable

Data Boxplot Visualization

Status code: 1 – Cancer patients   Status code: 0 – Healthy patients

Correlation Matrix

Single Logistic Regression Results

Predictor Variable	P-value
PSA	0.044
Exon9	0.448
Exon4a	0.648

Multiple Logistic Regression Results

Model	$\beta 1$	$\beta 2$	$\beta 3$
$Pca \sim \beta 0 + \beta 1 PSA$	0.044		
$Pca \sim \beta 0 + \beta 1 PSA + \beta 2 Exon9$	0.042	0.421	
$Pca \sim \beta 0 + \beta 1 PSA + \beta 2 Exon4a$	0.045	0.664	
$Pca \sim \beta 0 + \beta 1 PSA + \beta 2 Exon4a + \beta 3 Exon9$	0.042	0.938	0.492

Model Comparison and Evaluation (10-fold cross validation)

Model	AUC	Accuracy	Precision	Recall	Sensitivity	Specificity
$Pca \sim \beta0 + \beta1PSA$	0.60	0.546	0.559	0.589	0.589	0.500
$Pca \sim \beta0 + \beta1PSA + \beta2Exon9$	0.62	0.574	0.589	0.589	0.589	0.558
$Pca \sim \beta0 + \beta1PSA + \beta2Exon4a$	0.61	0.556	0.571	0.571	0.571	0.538
$Pca \sim \beta0 + \beta1PSA + \beta2Exon4a + \beta3Exon9$	0.61	0.528	0.545	0.536	0.536	0.519

Likelihood Test Ratio Results

LRT for Model PSA + Exon 9 + Exon 4a vs PSA + Exon 9:

LR Stat=-0.363, p-value=1.000

LRT for Model PSA + Exon 4a vs PSA Only:

LR Stat=0.587, p-value=0.444
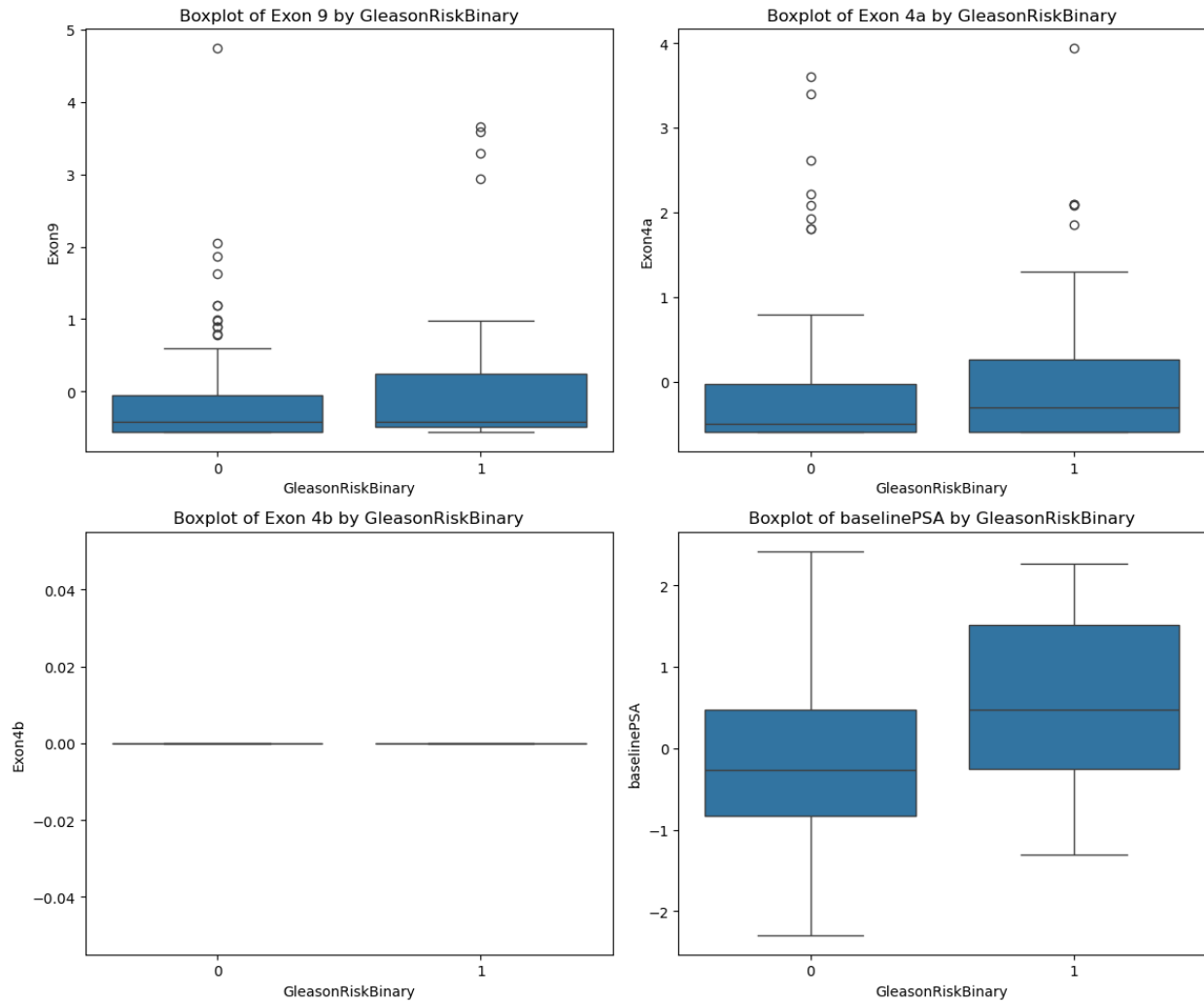
LRT for Model PSA + Exon 9 + Exon 4a vs PSA + Exon 4a:

LR Stat=-0.372, p-value=1.000

LRT for Model PSA + Exon 9 vs PSA Only:

LR Stat=0.578, p-value=0.447

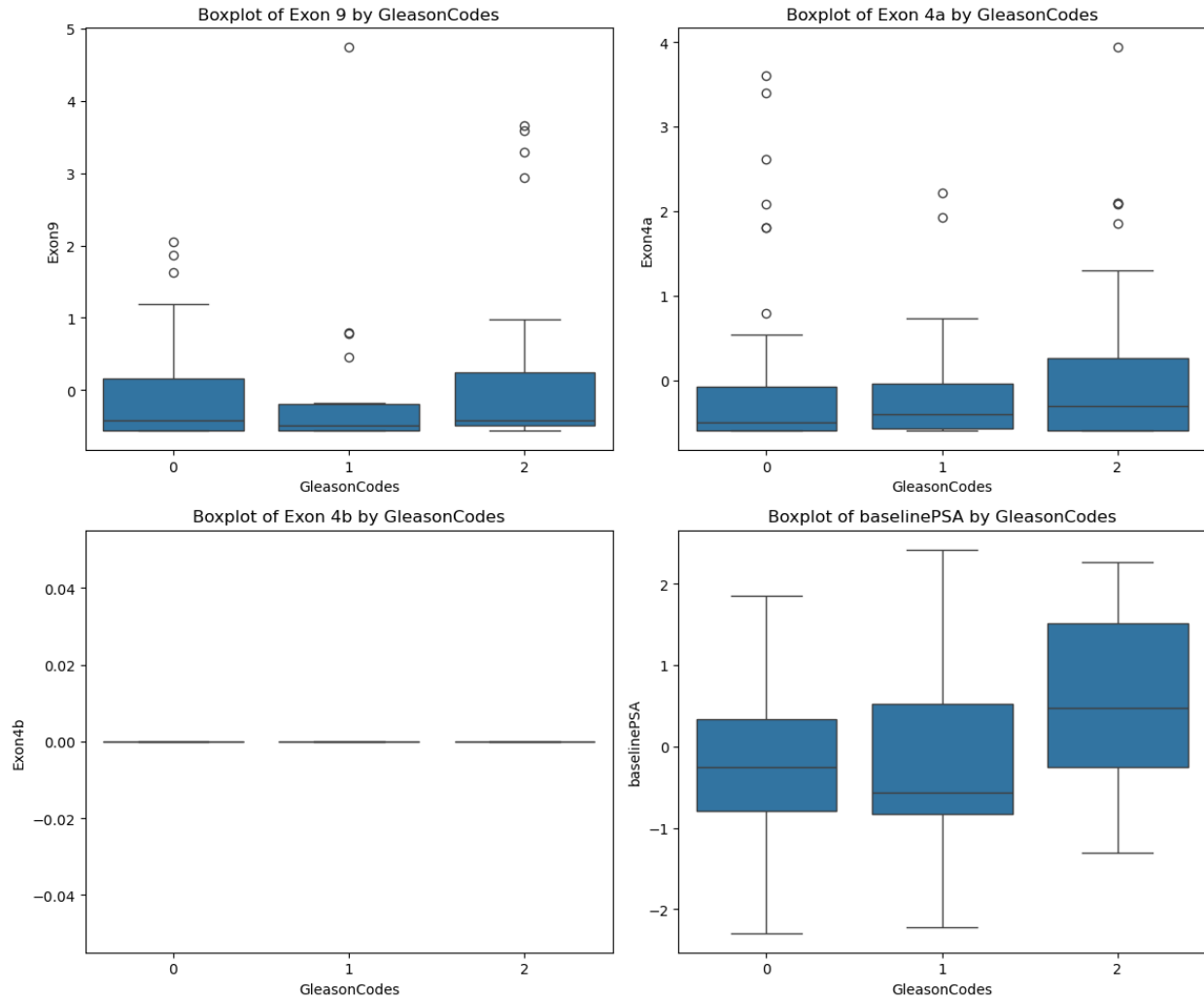# Section 2: Risk as the Response Variable using Gleason Score

Data Boxplot Visualization

Where

0 – Gleason score <=6

1 - Gleason score >=7

Boxplot of Exon 9 by GleasonCodes

Boxplot of Exon 4a by GleasonCodes

Boxplot of Exon 4b by GleasonCodes

Boxplot of baselinePSA by GleasonCodes

Where

0 – Gleason score 0

1 - Gleason score 6

2 - Gleason score >= 7

Single Logistic Regression Results

Predictor Variable	P-value
PSA	0.005
Exon9	0.182
Exon4a	0.271

Multiple Logistic Regression Results

Model	$\beta1$	$\beta2$	$\beta3$
$Pca \sim \beta0 + \ \beta1PSA$	0.005		
$Pca \sim \beta0 + \ \beta1PSA + \ \beta2Exon9$	0.004	0.157	
$Pca \sim \beta0 + \ \beta1PSA + \ \beta2Exon4a$	0.005	0.274	
$Pca \sim \beta0 + \ \beta1PSA + \ \beta2Exon4a + \ \beta3Exon9$	0.004	0.786	0.326

Model Comparison and Evaluation (10-fold cross validation)

Model	AUC	Accuracy	Precision	Recall	Sensitivity	Specificity
$Pca \sim \beta0 + \ \beta1PSA$	0.69	.750	0.667	0.200	0.200	0.962
$Pca \sim \beta0 + \ \beta1PSA + \ \beta2Exon9$	0.72	0.704	0.400	0.133	0.133	0.923
$Pca \sim \beta0 + \ \beta1PSA + \ \beta2Exon4a$	0.72	0.704	0.375	0.100	0.100	0.936
$Pca \sim \beta0 + \ \beta1PSA + \ \beta2Exon4a + \ \beta3Exon9$	0.73	0.685	0.333	0.133	0.133	0.897

<u>Likelihood Test Ratio Results</u>

LRT for Model PSA + Exon 9 + Exon 4a vs PSA + Exon 9:

LR Stat=-1.904, p-value=1.000

LRT for Model PSA + Exon 4a vs PSA Only:

LR Stat=7.676, p-value=0.006

LRT for Model PSA + Exon 9 + Exon 4a vs PSA + Exon 4a:

LR Stat=-2.897, p-value=1.000

LRT for Model PSA + Exon 9 vs PSA Only:

LR Stat=6.683, p-value=0.010