Results Tables

Table 1: Dataset Characteristics After Preprocessing

Characteristic	Value
Total observations after removing missing values	391 rows
Number of predictors	21
Training set size	274 (70%)
Testing set size	117 (30%)
Number of deaths (training)	166
Number of deaths (testing)	-

Table 2: Random Survival Forest Variable Importance

Rank	Variable	Importance
1	FIGO.stage	0.1163
2	Age.at.diagnosis	0.0210
3	Coitarche	0.0151
4	Reccurence	0.0149
5	Occupation	0.0100
6	Menarche	0.0085
7	Age.at.first.birth	0.0070
8	Marital.status	0.0063
9	Tumor.grade	0.0054
10	Tribe	0.0025

Table 3: Training Data Performance Metrics

Model	C-index	Brier Score	IAE	ISE
RSF	0.681	0.182	30.727	1157.993
SVM	0.527	0.312	21.081	532.622
KNN	0.000	0.343	0.000	0.000
NN	0.549	0.288	9.675	307.398

NB	0.349	0.362	8.336	173.317

Table 4: Testing Data Performance Metrics

Model	C-index	Brier Score	IAE	ISE
RSF	0.552	0.281	30.667	1229.449
SVM	0.441	0.339	22.055	791.394
KNN	0.433	0.331	14.199	714.862
NN	0.464	0.301	15.821	960.437
NB	0.453	0.377	15.344	774.979

Table 5: Overfitting Analysis

Model	C-index Train	C-index Test	C-index Diff	IAE Diff	ISE Diff	Overfitting Risk
RSF	0.681	0.552	0.129	-0.060	71.456	High
SVM	0.527	0.441	0.086	0.974	258.772	High
KNN	0.000	0.433	-0.433	14.199	714.862	High
NN	0.549	0.464	0.085	6.145	653.038	High
NB	0.349	0.453	-0.105	7.008	601.661	High

Positive C_index_Diff indicates lower test performance; Positive IAE_Diff and ISE_Diff indicate higher test errors

Table 6: Final Comprehensive Summary (Combined Training and Testing Performance)

Model	C-index Train	IAE Train	ISE Train	C-index Test	IAE Test	ISE Test
RSF	0.681	30.727	1157.993	0.552	30.667	1229.449
SVM	0.527	21.081	532.622	0.441	22.055	791.394
KNN	0.000	0.000	0.000	0.433	14.199	714.862
NN	0.549	9.675	307.398	0.464	15.821	960.437
NB	0.349	8.336	173.317	0.453	15.344	774.979

Table 7: Best Model Performance Summary

Metric	Model	Value
Best Training Performance (C-index)	RSF	0.681
Best Testing Performance (C-index)	RSF	0.552
Most Generalizable Model (Smallest C-index Gap)	NN	0.085

Table 8: Risk Category Distribution in Training Set

Risk Category	Frequency
Low	97
Moderate	96
High	81

Table 9: Risk Classification Model Performance

Model	Training Accuracy	Testing Accuracy
Random Forest (RF)	1.000	0.316
SVM	0.314	0.342
Gradient Boosting (GBM)	0.679	0.359

Table 10: Top 10 Risk Factors (Average Importance Across Models)

Rank	Variable	Importance Score
1	Parity	24.458
2	FIGO.stage	23.218
3	Occupation	10.683
4	Age.at.diagnosis	8.927
5	Age	8.842
6	Age.at.last.birth	7.849
7	Treatment.taken	7.327
8	Menopause	7.315

9	Coitarche	6.831
10	Age.at.first.birth	6.592

Table 11: Composite Risk Score Distribution in Test Set

Statistic	Value
Minimum	0.1598
1st Quartile	0.1903
Median	0.2105
Mean	0.2173
3rd Quartile	0.2405
Maximum	0.3018

Table 12: Composite Risk Category Distribution in Test Set

Risk Category	Frequency	Percentage
Low Risk	39	33.3%
Moderate Risk	38	32.5%
High Risk	40	34.2%

Table 13: Naive Bayes Risk Group Survival Times

Risk Group	Mean Survival Time (months)
High	1.424
Medium-High	3.907
Medium-Low	9.433
Low	29.603

Table 14: Sample Patient Demographics for Risk Assessment Demo

Variable	Value
Age	55
Marital Status	Divorced
Religion	Christianity
Parity	1
Menopause	46
FIGO Stage	Stage I A
Age at Diagnosis	54
Occupation	Business
Age at First Birth	18
Age at Last Birth	38
Recurrence	No
Family History	No
Smoking Status	No
Alcohol Consumption	No
Menarche	14
Coitarche	17
Tribe	Hausa
Histology	Adenocarcinoma
Tumor Grade	MD
Treatment Taken	Chemo/Radiotherapy
Comorbidity	No

Table 15: Sample Patient Risk Assessment Results

Model	High Risk Probability
Random Forest	0.252
SVM	0.293
Composite Risk Score	**0.272**
Risk Level	**LOW RISK**

Table 16: Model Recommendations Summary

Priority	Recommendation	Model/Approach
1	Highest Predictive Accuracy	Use model with highest test C-index (RSF)
2	Clinical Deployment	Choose most generalizable model (NN - smallest train-test gap)
3	Interpretability	Random Survival Forest (provides variable importance)
4	Robust Predictions	Ensemble methods combining top 2-3 models