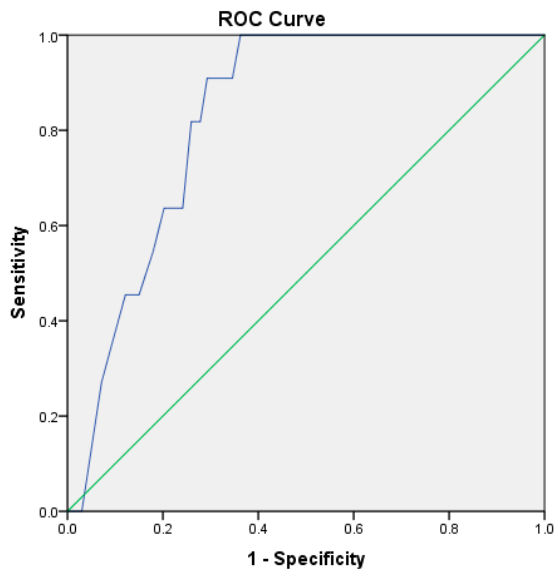


Supplementary Analysis Results

1. Logistic Regression Results for D30 Using Text-Based Risk Score as a Single Predictor

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B) with 95% CI
Text-based risk	6.22	2.29	7.36	1	0.01	502.32 [5.61–44966.12]
Constant	–9.40	1.94	23.52	1	0.00	



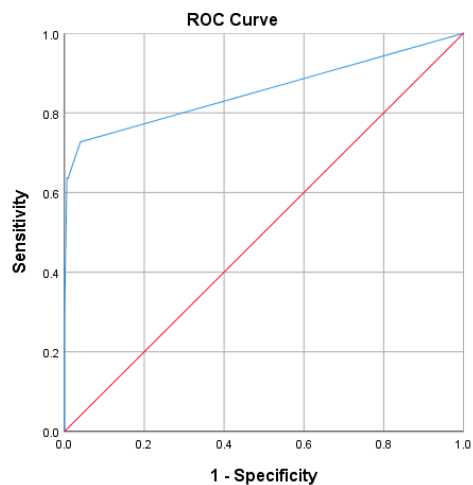
Area Under the Curve

Area	Std. Error	Asymptotic Sig.	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
0.83	0.03	0.00	0.77	0.89

2. Logistic Regression Results for D30 Using Structured Data Based Risk Factors

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B) with 95% CI
Ventilator	2.85	.98	8.51	1	0.00	17.21 [2.55–116.41]
Bleeding Disorder	4.18	1.27	10.93	1	0.00	65.50 [55.00–781.39]
Inotropic Support	2.66	.99	7.34	1	0.01	14.29 [2.09–97.87]
Constant	–6.50	.61	114.26	1	0.00	0.00



Diagonal segments are produced by ties.

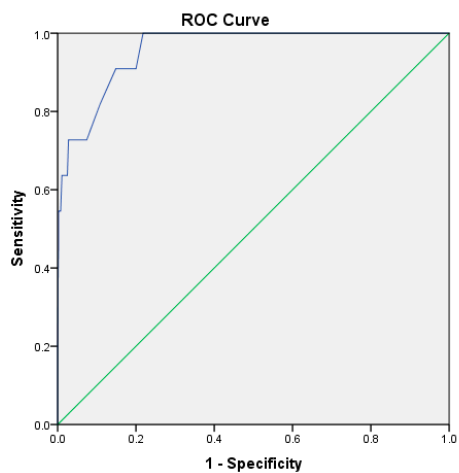
Area Under the Curve

Area	Std. Error ^a	Asymptotic Sig. ^b	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
0.86	0.08	0.00	0.70	1.00

3. Logistic Regression Results for D30 using both unstructured and structured data

Variables in the Equation

	B with 95% CI	Sig.	Exp(B) with 95% CI
Text-based risk	6.28 (0.42, 12.98)	0.066	533.40 [1.94–146823.50]
Ventilator	2.07 (0.15, 3.99)	0.035	7.96 [1.58–40.00]
Bleeding Disorder	3.56 (1.35, 5.81)	0.002	35.30 [5.33–233.67]
Inotropic Support	2.29 (0.19, 4.39)	0.032	9.84 [1.70–57.11]
Emergent Case	1.75 (0.03, 3.48)	0.046	5.76 [1.34–24.47]
Constant	–11.12 (–16.98, –5.26)	.0000	



Diagonal segments are produced by ties.

Area Under the Curve

Area	Std. Error ^a	Asymptotic Sig.	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
0.96	0.02	0.00	0.92	1.00

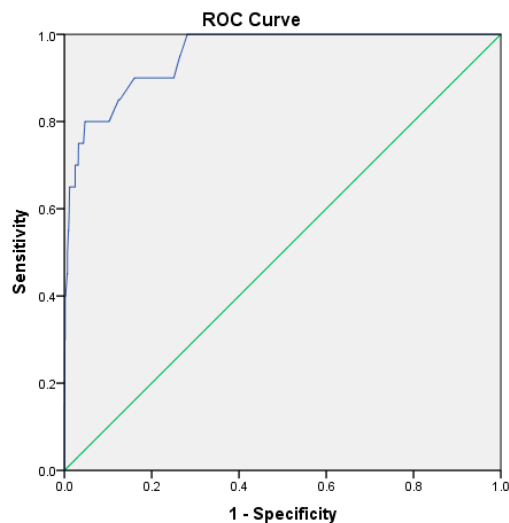
4. Five-Fold Cross Validation Using the Five Predictors Selected in Model C

	AUC:Train-Test	Regression Coefficients					
		Constant	Text-Based Risk	Ventilator	Bleeding Disorder	Inotropic Support	Emergent Case
Run 1	0.98 – 0.84	–15.028	9.320	3.684	4.015	2.599	2.395
Run 2	0.95 – 1.00	–10.307	5.772	1.626	2.281	2.398	1.697
Run 3	0.95 – 0.99	–10.486	5.657	1.105	4.079	3.325	2.044
Run 4	0.95 – 1.00	–11.294	7.013	1.687	3.623	1.897	1.359
Run 5	0.96 – 0.87	–11.576	6.504	2.505	3.854	1.678	1.500

5. Logistic Regression Results for D90 Using Both Unstructured and Structured Data

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B) with 95% CI
Ventilator	2.98	0.61	23.75	1	.00	19.64 [5.93–65.04]
Bleeding Disorder	2.91	0.98	8.80	1	.00	18.27 [2.68–124.52]
Neonate	1.41	0.58	5.85	1	.02	4.09 [1.31–12.80]
Emergent Case	1.41	0.67	4.48	1	.03	4.11 [1.11–15.25]
Text-based risk	4.42	1.96	5.09	1	.02	83.31 [1.79–3886.07]
Constant	–9.23	1.68	30.25	1	.00	.00



Diagonal segments are produced by ties.

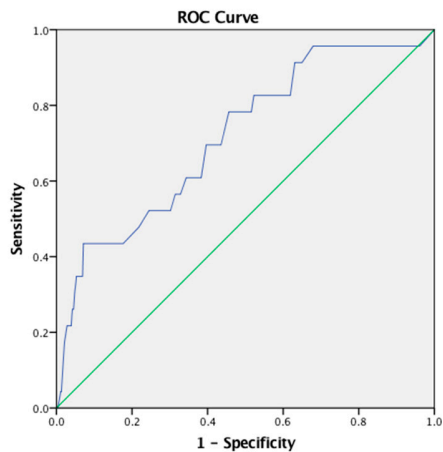
Area Under the Curve

Area	Std. Error	Asymptotic Sig.	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
0.95	0.02	0.00	0.92	0.99

6. Logistic Regression Results for Postoperative Superficial Incisional Surgical Site Infection Using Both Unstructured and Structured Data

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B) with 95% CI
Text-based risk	1.42	0.74	3.66	1	0.05	4.13 [0.97–17.66]
Neonate	1.65	0.44	14.22	1	0.00	5.19 [2.21–12.21]
Constant	–5.49	0.51	115.29	1	0.00	



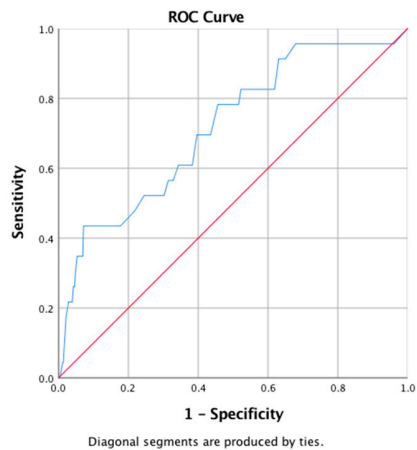
Area Under the Curve

Area	Std. Error	Asymptotic Sig.	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
0.72	0.06	0.00	0.61	0.83

7. Logistic Regression Results for Postoperative Superficial Incisional Surgical Site Infection using Both Unstructured and Structured Data

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B) with 95% CI
Text-based risk	1.42	0.74	3.66	1	0.05	4.13 [0.97–17.66]
Neonate	1.65	0.44	14.22	1	0.00	5.19 [2.21–12.21]
Constant	–5.49	0.51	115.29	1	0.00	0.00



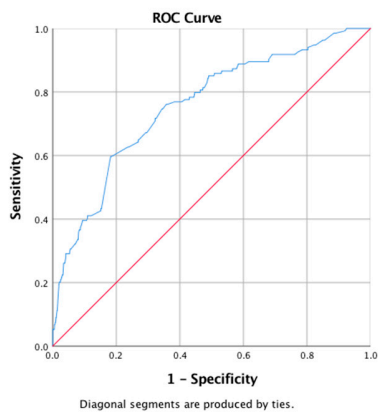
Area Under the Curve

Area	Std. Error	Asymptotic Sig.	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
0.72	0.06	0.00	0.61	0.83

8. Logistic Regression Results for Intra or Postoperative Blood Transfusion within 72 hours of Surgery Start Time Using Both Unstructured and Structured Data

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B) with 95% CI
Oxygen Support	1.06	0.40	7.092	1	0.01	2.87 [1.32–6.25]
Neuromuscular Disorder	1.31	0.24	30.42	1	0.00	3.72 [2.33–5.94]
Hematologic Disorders	1.20	0.33	12.96	1	0.00	3.32 [1.73–6.38]
Inotropic Support	2.53	0.62	16.76	1	0.00	12.52 [3.73–41.98]
Childhood Malignancy	0.52	0.25	4.116	1	0.04	1.67 [1.02–2.75]
Urgent Case	−1.95	0.74	7.014	1	0.01	.14 [0.03–0.60]
D30risk	−2.24	0.35	40.35	1	0.00	0.106
Constant	−2.11	0.15	198.08	1	0.00	0.12



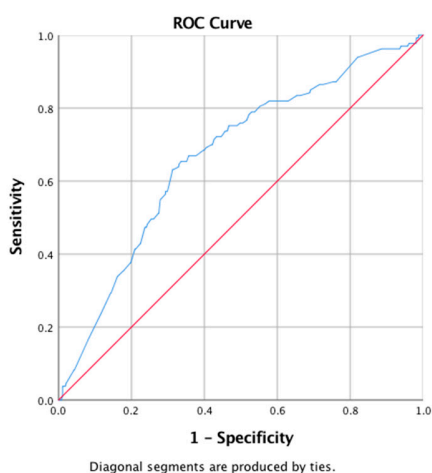
Area Under the Curve

Area	Std. Error	Asymptotic Sig.	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
0.76	0.02	0.00	0.71	0.80

9. Logistic Regression Results for Unplanned Readmission within 30 days of Surgery Using Both Unstructured and Structured Data

Variables in the Equation

	B	S.E.	Wald	df	Sig.	Exp(B) with 95% CI
SIRS	0.89	0.43	4.32	1	0.04	2.43 [1.05–5.63]
Sepsis	0.82	0.33	6.31	1	0.01	2.27 [1.20–4.49]
Neonate	−1.20	0.43	7.82	1	0.01	0.30 [0.13–0.70]
Text-based risk	1.54	0.29	28.14	1	0.00	4.67 [2.64–8.25]
Constant	−3.28	0.20	272.16	1	0.00	0.04



Area Under the Curve

Area	Std. Error	Asymptotic Sig.	Asymptotic 95% Confidence Interval	
			Lower Bound	Upper Bound
0.67	0.02	0.00	0.62	0.72