Predictors of Major Complication After Central Cord Syndrome Operation

Objective. The aim of this study is to determine the risk factors that would related to the major complication after central cord syndrome operations.

Summary of Background Data. Data were taken from the 2014 HCUP National Inpatient Sample (NIS) database. Risk factors in patient demographics as well as operative features were studied through t-test and chi-square test to determine their association with race and major complications. A step-wise multivariate logistic regression model was used to determine multivariate relationships with major complication.

Result. Elder patients discharged to a location that was not home, with a greater length of stay, and greater number of chronic conditions are more likely to have major complication after central cord syndrome operations.

1. Materials and Methods

Data were collected from 2014 NSQIP database. There are 946 observations and 178 variables in the data set. Ten variables, including sex, age (both continuous and in age ranges), length of stay (LOS), number of chronic conditions (NCHRONIC), Insurance Status (Primarypayer), Urban or Rural Status for patient location (PL_NCHS), median household income for patient location zip code (ZIPINC_QRTL), geographic region in the US (HOSP_DIVISION), Discharge destination (home or not home), and RACE (white and non-white) are of interest to determine if they would be the predictors of the occurrence of major complication. Table1 shows several important variables which are self-created according to other available information in the original data set. Appendix 1 shows relevant CPT code results by race.

As a preliminary study of effect of variable RACE (non-white, white), odds ratios were calculated for operative features and patient characteristics (Table 2). Continuous variables were compared between white and non-white races using two-sample t-tests. There are significant associations between white and non-white race with respect to Sex, AGE (both as continuous and age groups), length of stay (LOS), number of chronic conditions (NCHRONIC), Insurance Status

(Primarypayer), Urban or Rural Status for patient location (PL_NCHS), median household income for patient location zip code (ZIPINC_QRTL), and geographic region in the US (HOSP_DIVISION). Take insurance status for example. The odds of non-whites having Medicaid as their primary payer is 2.1921 times the odds of whites. Odds denoted by "1" in Table 2 denote the category used for comparison or the "baseline" group. Continuous variables were compared between those with major complication and those without using a two-sample t-test.

Table 1 Variables recoding notes

Variables	Notes
RACE	coded by collapsing original RACE's 5 non-white categories to be one
Discharge destination	coded by dividing DISPUNIFORM into 2 groups "home" and "not home"
primary payer	coded by collapsing Pay1's "Self-pay" and "No charge" to be "uninsured"
complication	coded based on other 30 diagnoses variables describe specific symptom of complication
reoperation	coded based on other 30 diagnoses variables describe specific reoperations
Major complication	created based on complication, reoperation and DIED

2. Univariate Analysis

Table 3 shows the odds ratios of a different risk factors and their 95% confidence intervals. Take discharge destination as an example. The odds of major complication after cervical fusion operations is 3.1148 times greater for those patients discharged to somewhere other than home compared to those discharged home. Odds denoted by "1" in Table 3 denote the category used for comparison or the "baseline" group. Continuous variables were compared between those with major complication and those without using a two-sample t-test. Appendix 2 shows relevant diagnostic codes by major complication.

3. Multivariate Analysis

A stepwise multivariate logistic regression was conducted on the data set with 945 non-missing observations and Table 4 shows the model results. A final model was chosen based on AIC value. Based on 0.05 significance level, AGE (continuous), length of stay (LOS), number of chronic conditions (NCHRONIC), and Discharge destination (home, not home) are significant predictors of major complications. Figure 1 shows the ROC of the final model and the area under the curve (AUC) is 0.7968 with a 95% confidence interval [0.7674, 0.8263].

The selection of cut-offs does not influence the modeling procedure (coefficients estimations, odds ratios and other statistics, Table 4) and resulting predicted probabilities. However, the misclassification rate and the power of the prediction are highly related to the cut-off setting used since these are used to predict an observation to be "1" or "0" (using the predicted probability). According to confusion table (Table 5), if the commonly used 0.5 is used as the prediction cut-off, the type I error will be extremely small but the power is small as well. But if the Youden's Index

Figure 1 ROC curve of the final model

value is used as the cut-off, this maximizes the sum of sensitivity and specificity, the power increases to 0.7027 while the type I error also increases dramatically. There is always a tradeoff between sensitivity and specificity. Youden's index is recommended here since it's a cut-off which equally weights sensitivity and specificity.
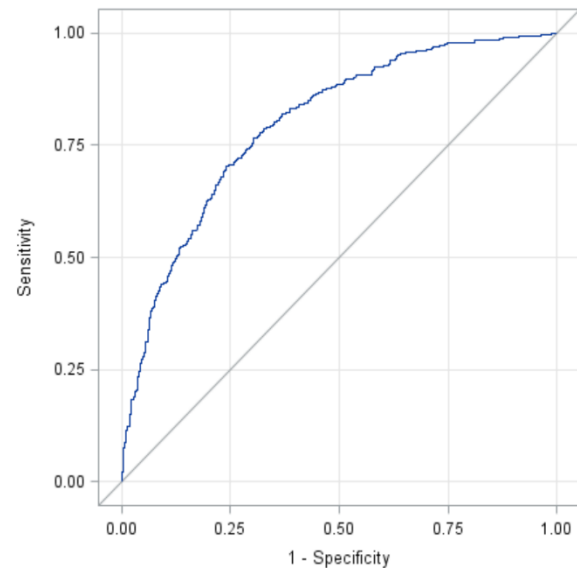
Table 2 Confusion matrix at different cutoffs of regular logistic regression

cutoff=0.5 (commen used)	predict	response 1	0	Sensitivity	Specificity	Type I error	Type II error
	1	163	75	0.489	0.877	0.123	0.511
	0	170	537				
cutoff= 0.382 (Youden's Index)	predict	response 1	0	Sensitivity	Specificity	Type I error	Type II error
	1	234	147	0.7027	0.7598	0.2402	0.2973
	0	99	465				
cutoff=0.031 (maxmize detection of majorcomplication)	predict	response 1	0	Sensitivity	Specificity	Type I error	Type II error
	1	333	604	1	0.0131	0.9869	0
	0	0	8				
cutoff=0.999 (minimize type I error)	predict	response 1	0	Sensitivity	Specificity	Type I error	Type II error
	1	1	0	0.0030	1	0	0.9970
	0	332	612				

4. Results and Discussion

Table 6 shows the comparison between univariate analysis and multivariate analysis for major complication. It seems that the univariate results identified more variables as risk factors for major complication than with multivariate relationships through the stepwise logistic regression. But they give consistent overall conclusions, [i.e., elder patients, discharged to a place other than home, with greater length of stay, and more chronic conditions, are more likely to have major complications after central cord syndrome operations.]

Table 6 Comparison between univariate analysis and multivariate analysis

Variable	Risk factor	significance	
		Univariate	**Multivariate**
SEX	male	non-significant	non-significant
AGE		significant	significant
Length of stay (LOS)		significant	significant
# of chronic conditions (NCHRONIC)		significant	significant
Primary Payer	medicaid	significant	non-significant
	medicare	significant	non-significant
	uninsured	non-significant	non-significant
Rural vs Urban (PL_NCHS)	Fringe	non-significant	non-significant
	>250,000 POP	non-significant	non-significant
	<250,000 POP	non-significant	non-significant
	Micropolitan	non-significant	non-significant
	Not metropolitan	significant	non-significant
Median Household income (ZIPINC_QRTL)	$39,000 - $47,999	non-significant	non-significant
	$48,000 - 62,999	non-significant	non-significant
	$63,000 or more	non-significant	non-significant
Region (HOSP_DIVISION)	MA	non-significant	non-significant
	ENC	non-significant	non-significant
	WNC	non-significant	non-significant
	SA	non-significant	non-significant
	ESC	non-significant	non-significant
	WSC	non-significant	non-significant
	Mountain	non-significant	non-significant
	Pacific	non-significant	non-significant
Discharge Destination	not home	significant	significant
race	Non-White	non-significant	non-significant

Table 3 Summary of Patients and Operative features by different race group

Variable	Category	All patients		White (N=593)		Non-white (N=292)		OR	95% CI	P-value
		N	%	N	%	N	%			
SEX	female	251	28.36%	181	30.52%	70	23.97%	1		
	male	634	71.64%	412	69.48%	222	76.03%	1.3933	[1.0112,1.9197]	0.0473
Age group	18-30	24	2.72%	13	2.20%	11	3.77%	1		
	31-40	51	5.78%	32	5.42%	19	6.51%	0.7017	[0.2625,1.8759]	0.6141
	41-50	101	11.45%	52	8.81%	49	16.78%	1.1136	[0.4561,2.7193]	0.8249
	51-60	231	26.19%	139	23.56%	92	31.51%	0.7822	[0.336,1.8211]	0.6632
	61-70	212	24.04%	147	24.92%	65	22.26%	0.5226	[0.2224,1.228]	0.1662
	70+	263	29.82%	207	35.08%	56	19.18%	0.3197	[0.1359,0.7522]	0.0108
Primary Payer	private	263	32.19%	184	32.74%	79	30.98%	1		
	medicaid	99	12.12%	51	9.07%	48	18.82%	2.1921	[1.3643,3.5221]	0.0013
	medicare	398	48.71%	298	53.02%	100	39.22%	0.7816	[0.5521,1.1064]	0.18
	uninsured	57	6.98%	29	5.16%	28	10.98%	2.2488	[1.2561,4.026]	0.0081
Rural vs Urban (PL_NCHS)	Central	244	27.85%	122	20.71%	122	42.51%	1		
	Fringe	202	23.06%	146	24.79%	56	19.51%	0.3836	[0.2578,0.5707]	<0.0001
	>250,000 POP	215	24.54%	151	25.64%	64	22.30%	0.4238	[0.2883,0.6231]	<0.0001
	<250,000 POP	68	7.76%	52	8.83%	16	5.57%	0.3077	[0.1665,0.5685]	0.0001
	Micropolitan	79	9.02%	62	10.53%	17	5.92%	0.2742	[0.1516,0.4958]	<0.0001
	Not metropolitan	68	7.76%	56	9.51%	12	4.18%	0.2143	[0.1094,0.4196]	<0.0001
Median Household income (ZIPINC_QRTL)	$1 - $38,999	263	30.58%	140	24.14%	123	43.93%	1		
	$39,000 - $47,999	228	26.51%	166	28.62%	62	22.14%	0.4251	[0.291,0.6211]	<0.0001
	$48,000 - 62,999	203	23.60%	155	26.72%	48	17.14%	0.3525	[0.2353,0.5281]	<0.0001
	$63,000 or more	166	19.30%	119	20.52%	47	16.79%	0.4495	[0.2967,0.6812]	0.0002
Region (HOSP_DIVISION)	NE	47	5.31%	44	7.42%	3	1.03%	1		
	MA	98	11.07%	60	10.12%	38	13.01%	9.2889	[2.6931,32.0388]	<0.0001
	ENC	134	15.14%	95	16.02%	39	13.36%	6.0211	[1.7643,20.5482]	0.0011
	WNC	50	5.65%	41	6.91%	9	3.08%	3.2195	[0.8147,12.723]	0.1229
	SA	190	21.47%	111	18.72%	79	27.05%	10.4384	[3.1295,34.8169]	<0.0001
	ESC	69	7.80%	57	9.61%	12	4.11%	3.0877	[0.8208,11.6155]	0.0977
	WSC	94	10.62%	54	9.11%	40	13.70%	10.8642	[3.1472,37.5035]	<0.0001

Variable	Category	All patients		White (N=593)		Non-white (N=292)		OR	95% CI	P-value
		N	%	N	%	N	%			
	Mountain	62	7.01%	40	6.75%	22	7.53%	8.0667	[2.2427,29.014]	0.0004
	Pacific	141	15.93%	91	15.35%	50	17.12%	8.0586	[2.3806,27.2791]	0.0001
Discharge Destination	home	377	42.65%	250	42.16%	127	43.64%	1		
	not home	507	57.35%	343	57.84%	164	56.36%	0.9412	[0.709,1.2494]	0.7176
complication	No	575	64.97%	377	63.58%	198	67.81%	1		
	Yes	310	35.03%	216	36.42%	94	32.19%	0.8286	[0.6157,1.1152]	0.2308
reoperation	No	876	98.98%	586	98.82%	290	99.32%	1		
	Yes	9	1.02%	7	1.18%	2	0.68%	0.5773	[0.1192,2.7968]	0.7255

Variable	All patients		White (N=593)		Non-white (N=292)		P-value t-test
	Mean Dev	Std	Mean Dev	Std	Mean Dev	Std	
AGE*	61.57	14.88	63.35	14.8	58.32	14.34	<0.0001
Length of stay (LOS)*	9.83	11.21	9.31	10.99	11.18	11.89	0.0208
# of chronic conditions* (NCHRONIC)	6.39	3.32	6.54	3.35	6.07	3.22	0.0472

* for numerical variables like AGE the means and the standard deviations are given for different race groups, and the p-values are based on two sample t-tests.

Table 4 Univariate Analysis of risk factors for Major Complication

Variable	Risk Factor	No Major Complications (N=612)		Major Complications (N=334)		OR	95% CI of OR	P-value
		N	%	N	%			
SEX	female	161	26.31%	104	31.14%	1		
	male	451	73.69%	230	68.86%	0.7895	[0.5888,1.0586]	0.1297
Age group	18-30	22	3.61%	4	1.20%	1		
	31-40	50	8.21%	4	1.20%	0.44	[0.1008,1.9211]	0.4268
	41-50	88	14.45%	20	6.01%	1.25	[0.3877,4.0307]	1
	51-60	179	29.39%	71	21.32%	2.1816	[0.726,6.5555]	0.2451
	61-70	143	23.48%	81	24.32%	3.1154	[1.0373,9.3562]	0.0472
	70+	127	20.85%	153	45.95%	6.626	[2.2254,19.7281]	0.0001
Primary Payer	private	219	39.46%	71	22.33%	1		
	medicaid	66	11.89%	37	11.64%	1.7292	[1.0664,2.804]	0.0292
	medicare	222	40%	199	62.58%	2.7649	[1.9897,3.8422]	<0.0001
	uninsured	48	8.65%	11	3.46%	0.7069	[0.3483,1.4345]	0.4013
Rural vs Urban (PL_NCHS)	Central	173	28.55%	85	25.68%	1		
	Fringe	134	22.11%	78	23.56%	1.1847	[0.8092,1.7345]	0.436
	>250,000 POP	145	23.93%	76	22.96%	1.0668	[0.7295,1.56]	0.7714
	<250,000 POP	52	8.58%	29	8.76%	1.1351	[0.6727,1.9154]	0.6863
	Micropolitan	60	9.90%	25	7.55%	0.848	[0.4972,1.4464]	0.5935
	Not metropolitan	42	6.93%	38	11.48%	1.8415	[1.1059,3.0661]	0.0235
Median Household income (ZIPINC_QRTL)	$1 - $38,999	173	28.93%	99	30.65%	1		
	$39,000 - $47,999	155	25.92%	96	29.72%	1.0823	[0.7591,1.5431]	0.7174
	$48,000 - 62,999	145	24.25%	74	22.91%	0.8918	[0.6139,1.2955]	0.5695
	$63,000 or more	125	20.90%	54	16.72%	0.7549	[0.5042,1.1302]	0.187
Region (HOSP_DIVISION)	NE	33	5.39%	14	4.19%	1		
	MA	68	11.11%	33	9.88%	1.1439	[0.5399,2.4236]	0.85
	ENC	91	14.87%	50	14.97%	1.2951	[0.6342,2.645]	0.5943
	WNC	51	8.33%	25	7.49%	1.1555	[0.5258,2.539]	0.8424
	SA	122	19.93%	77	23.05%	1.4877	[0.7483,2.9577]	0.3142
	ESC	43	7.03%	26	7.78%	1.4252	[0.6453,3.148]	0.43
	WSC	62	10.13%	32	9.58%	1.2166	[0.5708,2.5932]	0.7044
	Mountain	48	7.84%	28	8.38%	1.375	[0.6305,2.9988]	0.4421
	Pacific	94	15.36%	49	14.67%	1.2287	[0.6016,2.5097]	0.5981
Discharge Destination	home	316	51.63%	85	25.53%	1		
	not home	296	48.37%	248	74.47%	3.1148	[2.3238,4.175]	<0.0001
race	White	373	65.32%	220	70.06%	1		
	Non-White	198	34.68%	94	29.94%	0.8049	[0.5983,1.0829]	0.1565

Variable	Risk Factor	No Major (N=612)		Major (N=334)				P-value t-test
		Mean	Std Dev	Mean	Std Dev			
AGE		58.38	14.73	67.43	13.29			<0.0001

Variable	Risk Factor	No Major Complications (N=612)		Major Complications (N=334)		OR	95% CI of OR	P-value
		N	%	N	%			
LOS		7.17	7.64	14.69	14.59			<0.0001
# of chronic (NCHRONIC)		5.51	2.89	7.99	3.45			<0.0001

* for numerical variables like AGE the means and the standard deviations are given for different major complication groups, and the p-values are based on two sample t-tests.

Table 5 Coefficients of Logistic Regression

Variable	Risk Factor	Odds	Estimate	Std.Error	z value	p-value	Association type
Intercept			-4.4815	0.4226	112.436	<.0001	
Age (cont.)		1.034	0.0332	0.0062	28.7323	<.0001	positive
Length of stay (cont.)		1.07	0.0673	0.00972	47.8714	<.0001	positive
# chronic conditions NCHRONIC(cont.)		1.161	0.1497	0.0259	33.4827	<.0001	positive
Discharge destination	not home	2.026	0.353	0.0853	17.1346	<.0001	positive

NHIS

Background on Data[1]

The National Health Interview Survey (NHIS) has monitored the health of the nation since 1957. NHIS data on a broad range of health topics are collected through personal household interviews. For over 50 years, the U.S. Census Bureau has been the data collection agent for the National Health Interview Survey. Survey results have been instrumental in providing data to track health status, health care access, and progress toward achieving national health objectives. Some more information can be found here.

Typical Analyses Run

Analysis of NHIS data are similar with NIS data. The example report shown on the following section indicates the analyzing process. The purpose is to determine the risk factors that would

---

[1] This part was cited from here.

related to operative features (e.g., readmission) after some specific operations (e.g., cervical spine fusion surgery). Two sample t-test and chi-square test were used to analysis continuous and discrete variables to determine their association with operative features of interest. Then a step-wise multivariate logistic regression model was used to determine multivariate relationship with operative features of interest. As same with NIS data analysis, ROC curve and confusion matrix were used to decide the cut-off points. Consequently, the risk factors were determined by comparing result of univariate and multivariate analysis.

Example Report

The example report reveals the risk factors of readmission after taking cervical spine fusion surgery.

Retrospective MU Cervical Spine Fusion Study

Objective. The aim of this study is to determine the risk factors that would be related to different outcomes after cervical spine fusion surgery.

Summary of Background Data. Risk factors in patient demographics as well as operative features are studied through chi-square tests, t-tests, and multivariate logistic regression to determine their association with outcomes of interest.

Result. The risk factors for readmission that both univariate analyses and multivariate analyses identified is variables implant loosing, pseduoarthrosis, and adjacent segment disease.

1. Materials and Methods

Data were retrospectively collected from EMRs. There are 256 observations.

As a preliminary study of effect of variable RACE, a summary of all demographic, socioeconomic, and clinical variables were compared using chi-square tests for association and the associated odds ratios were calculated, as Table 1 shows. There are significant associations between white and non-white race with respect to wound complication - infection. Take it for example, the odds ratio of "Non-white" to "White" is 13.1111, and the p-value is less than 0.05,

which means the non-white patients have higher odds of Wound complication - infection compared to white patients.

Table 1. Demographic information for patients by different race group

Variable	Category	All patients （N=256）		White (N=237)		Non-white (N=19)		OR	95% CI	Pvalue
SEX	male	99	38.67%	95	40.08%	4	21.05%	1		
	female	157	61.33%	142	59.92%	15	78.95%	2.5088	[0.8079, 7.7906]	0.1019
Insurance	private	151	58.98%	141	59.49%	10	52.63%	1		
	Medicare/Medicare & Medicaid	23	8.98%	20	8.44%	3	15.79%	2.115	[0.5361, 8.3447]	0.2766
	Medicaid	78	30.47%	72	30.38%	6	31.58%	1.175	[0.4107, 3.3617]	0.7639
	Self-Pay	4	1.56%	4	1.69%	0	0.00%	0		
Diabetes	No	214	83.59%	200	84.39%	14	73.68%	1		
	Yes	42	16.41%	37	15.61%	5	26.32%	1.9305	[0.6558, 5.6829]	0.2263
Smoking	No	178	69.53%	167	70.46%	11	57.89%	1		
	Yes	78	30.47%	70	29.54%	8	42.11%	1.7351	[0.6693, 4.4977]	0.253
ASACLAS	1-2	138	53.91%	128	54%	10	52.63%	1		
	3-5	116	45.31%	107	45.15%	9	47.37%	1.0766	[0.4221, 2.7463]	0.8774
Age group	18-30	2	0.78%	2	0.84%	0	0%	0		
	31-40	35	13.67%	33	13.92%	2	10.53%	1		
	41-50	105	41.01%	98	41.35%	7	36.84%	1.1786	[0.2332, 5.9567]	0.8429
	51-60	74	28.91%	67	28.27%	7	36.84%	1.7239	[0.3392, 8.7617]	0.5091
	61-70	30	11.72%	28	11.81%	2	10.53%	1.1786	[0.1558, 8.9167]	0.8744
	70+	10	4.10%	9	3.80%	1	5.26%	1.8333	[0.1488, 22.5830]	0.6356
Hypertension	No	141	55.08%	134	56.54%	7	36.84%	1.0000		
	Yes	115	44.92%	103	43.46%	12	63.16%	2.2302	[0.8481, 5.8649]	0.0974
DIALYSIS	No	256	100%	237	100%	19	100%	1		
	Yes	0	0%	0	0%	0	0%	0		
Active Cancer	No	249	97.27%	230	97.05%	19	100%	1		
	Yes	7	2.73%	7	2.95%	0	0%	0		
	No	255	99.61%	236	99.58%	19	100%	1		

								OR	CI	p-value
wound complication – hematoma or seroma	Yes	1	0.39%	1	0.42%	0	0%	0		
Wound complication - infection	No	254	99.22%	236	99.58%	18	94.74%	1		
	Yes	2	0.78%	1	0.42%	1	5.26%	13.1111	[0.7870, 218.415]	0.02135
Implant loosening	No	227	88.67%	211	89.03%	16	84.21%	1		
	Yes	29	11.33%	26	10.97%	3	15.79%	1.5216	[0.4152, 5.5759]	0.5245
Implant breakage	No	251	98.05%	232	97.89%	19	100%	1		
	Yes	5	1.95%	5	2.11%	0	%	0		
any complication (wound, surgical, infection)	No	173	67.58%	162	68.35%	11	57.89%	1		
	Yes	83	32.42%	75	31.65%	8	42.11%	2.3582	[0.8803, 6.3167]	0.0809
Destination post-op	home	247	96.48%	228	96.20%	19	100%	1		
	Not home	9	3.52%	9	3.80%	0	0%	0		
comorbidities	no	84	32.81%	81	34.18%	3	15.79%	1		
	yes	172	67.19%	156	65.82%	16	84.21%	2.7692	[0.7839, 9.7822]	0.1012
pseudoarthrosis	no	230	89.84%	213	89.87%	17	89.47%	1		
	yes	26	10.16%	24	10.13%	2	10.53%	1.0441	[0.2273, 4.7966]	0.9558
Adjacent segment disease	no	234	91.41%	217	91.56%	17	89.47%	1		
	yes	22	8.59%	20	8.44%	2	10.53%	1.2765	[0.2750, 5.9250]	0.7552
Proximal junctional fracture	no	251	98.05%	233	98.31%	18	94.74%	1		
	yes	5	1.95%	4	1.69%	1	5.26%	3.2361	[0.3434, 30.4938]	0.2795
readmission	No	193	75.39%	181	76.37%	12	63.16%	1		
	Yes	63	24.61%	56	23.63%	7	36.84%	1.8854	[0.7082, 1.8854]	0.1991
Revision surgeries	No	203	79.30%	191	80.59%	12	63.16%	1		
	Yes	53	20.70%	46	19.41%	7	36.84%	2.4221	[0.9035, 6.4935]	0.0717
BMIgroup	<=35	202	78.91%	188	79.32%	14	73.68%			
	35+	54	21.09%	49	20.68%	5	26.32%	1.3703	[0.4708, 3.9885]	0.5628
COPD	No	229	89.45%	213	89.87%	16	84.21%	1		
	Yes	27	10.55%	24	10.13%	3	15.79%	1.6641	[0.4520,6.1263]	0.4403
BMI		30.65	6.77	30.54	6.76	32.05	6.69			0.3474
AGE		50.45	9.98	50.44	10.09	50.63	9			0.9357
Delta VAS		1.62	3.00	1.74	2.97	0.21	2.99			0.0323
Pre-op VAS		6.15	2.60	6.14	2.53	6.32	3.54			0.8342
Post-op VAS		4.46	3.13	4.34	3.09	6	3.32			0.0258
Median Income		45787.54	10458.01	45884.04	10505.58	44474.12	10364.01			0.5933

2. Univariate Analysis

Univariate analyses of readmission risk factors were performed through chi-square tests and the associated odds ratio was reported. The readmission variable has two measurements, "Yes" and "No". In addition, a stepwise multivariate logistic regression was performed to determine significant predictors.

Table 2 shows the odds ratios of a series potential risk factors and their 95% confidence intervals and p-values of the odds ratio tests. Take implant loosening for instance, the odds ratio of "Yes" to "No" is 11.5625, and the p-value is less than 0.05, which means the patients who have implant loosening will have higher odds of a readmission after cervical operations. And it is similar to the factors, complication, pseduoarthrosis, adjacent segment disease, and revision surgeries, which are other risk factors would result in readmission.

Table 2 Univariate Analysis of risk factors for any readmission

Variable	Category	Any readmission (N=63)		No readmission (N=193)		OR	95% CI	Pvalue
SEX	male	30	47.62%	69	35.75%	1		
	female	33	52.38%	124	64.25%	0.6121	[0.3443,1.0882]	0.10267
Insurance	private	39	61.90%	112	58.03%	1		
	Medicare/Medicare & Medicaid	3	4.76%	20	10.36%	0.4308	[0.1213,1.5294]	0.2940
	Medicaid	20	31.75%	58	30.05%	0.9903	[0.5298,1.8508]	1
	Self-Pay	1	1.59%	3	1.55%	0.9573	[0.0967,9.4749]	1
Diabetes	No	53	84.13%	161	83.42%	1		
	Yes	10	15.87%	32	16.58%	0.9493	[0.4374,2.0604]	1
Smoking	No	41	65.08%	137	70.98%	1		
	Yes	22	34.92%	56	29.02%	1.3127	[0.7175,2.4018]	0.4310
ASACLAS	1-2	30	47.62%	108	55.96%	1		
	3-5	33	52.38%	83	43.01%	1.4313	[0.8085,205341]	0.2446
Age group	18-30	0	0%	2	1.04%	0		
	31-40	8	12.70%	27	13.99%	1	[0.4021,2.4868]	1

	41-50	24	38.10%	81	41.97%	1		
	51-60	18	28.57%	56	29.02%	1.0848	[0.5389,2.1838]	0.8589
	61-70	9	14.29%	21	10.88%	1.4464	[0.5857,3.5719]	0.4722
	70+	4	6.35%	6	3.11%	2.25	[0.5864,8.6329]	0.2542
Hypertension	No	31	49.21%	110	56.99%	1		
	Yes	32	50.79%	83	43.01%	1.3681	[0.7735,2.4197]	0.3089
DIALYSIS	No	63	100%	193	100%	1		
	Yes	0	0%	0	0%	0		
Active Cancer	No	63	100%	186	96.37%	1		
	Yes	0	0%	7	3.63%	0		
wound complication – hemoatoma or seroma	No	63	100%	192	99.48%	1		
	Yes	0	0%	1	0.52%	0		
Wound complication - infection	No	61	96.83%	193	100%	1		
	Yes	2	3.17%	0	0%	0		
Implant loosening	No	42	66.67%	185	95.85%	1		
	Yes	21	33.33%	8	4.15%	11.5625	[4.7934,27.8909]	8.248e-9
Implant breakage	No	60	95.24%	191	98.96%	1		
	Yes	3	4.76%	2	1.04%	4.775	[0.7795,29.2520]	0.0974
comorbidities	no	17	26.98%	67	34.72%	1		
	yes	46	73.02%	126	65.28%	1.4388	[0.7661,2.7024]	0.2826
pseduoarthrosis	no	44	69.84%	186	96.37%	1		
	yes	19	30.16%	7	3.63%	11.4740	[4.5417,28.9875]	4.3138e-8
Adjacent segment disease	no	48	76.19%	186	96.37%	1		
	yes	15	23.81%	7	3.63%	8.3036	[3.2062,21.5052]	7.222e-6
Proximal junctional fracture	no	61	96.83%	190	98.44%	1		
	yes	2	3.17%	3	1.55%	2.0765	[0.3390,12.7175]	0.5993
Revision surgeries	No	12	19.05%	191	98.96%	1		
	Yes	51	80.95%	2	1.04%	405.875	[88.0191,1875.577]	1.6054e-39
BMIgroup	<=35	48	76.19%	154	79.79%	1		
	35+	15	23.81%	39	20.21%	1.23404	[0.6265,2.4306]	0.5943
RACE	White	56	88.89%	181	93.78%	1		
	Non-white	7	11.11%	12	6.21%	1.8854	[0.7082,5.0193]	0.2652

COPD	No	54	85.71%	175	90.67%	1		1
	Yes	9	14.29%	18	9.33%	1.6204	[0.68823.8152]	
Destination post-op	No	61	96.83%	186	96.37%	1		
	Yes	2	3.17%	7	3.63%	0.8712	[0.1763,4.3058]	0.8658
BMI		30.648	6.756	30.698	7.038	30.632	6.681	0.9462
AGE		50.453	9.985	51.413	10.932	50.140	9.667	0.3807
Delta VAS		1.624	2.996	1.651	2.891	1.615	3.037	0.9339
Pre-op VAS		6.153	2.606	6.349	2.528	6.089	2.634	0.4920
Post-op VAS		4.461	3.135	4.635	3.249	4.404	3.103	0.6128
Median Income		45787.54	10481.05	47020.10	8785.54	45398.66	10954.24	0.3011

3. Multivariate Analysis

A stepwise multivariate logistic regression was conducted on the data set with 243 observations and Table 3 shows the estimates of coefficients. Readmission was modeled as response, the variables of risk factors were included in the full model and dropped based on VIF value. Stepwise logistic regression was applied, and the final model was chosen by AIC. Based on a 0.05 significance level, implant loosing, adjacent segment disease, and pseduoarthrosis are significant predictors of readmission. Figure 1 shows the ROC of the final model and the AUC is 0.8242 with a 95% confidence interval being [0.7572, 0.8913].
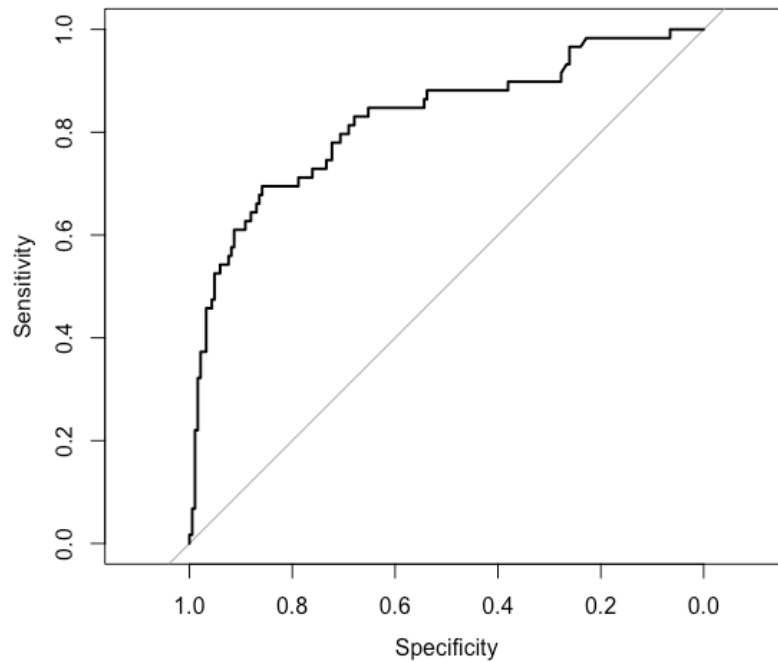
Figure 1. ROC curve of the final model

Table 3. Coefficient of Logistic Regression

coefficient	Odds	Estimate	Std.Error	z value	p-value	Association type
Intercept	0.01289	-4.3515	1.3231	-3.289	0.001	
Implant Loosing	12.7541	2.5459	0.5561	4.578	4.69e-6	Positive
Adjacent segment disease	9.2432	2.2239	0.5517	4.031	5.557e-5	Positive
pseduoarthrosis	7.5436	2.0207	0.5667	3.565	3.63e-4	Positive
Median Income	1	2.998e-5	1.746e-5	1.717	0.086	Positive
BMI	1.0422	0.0413	0.0268	1.541	0.1232	Positive
Sex	0.5754	-0.5527	0.3750	-1.474	0.1405	Negative

According to confusion table, the misclassification rate and the power the prediction are highly related to the cutoff setting. Because of the definition of variables, the cutoff point of common used and Youden's Index are 0.5 and 0.2091. To maximize detection of any readmission and minimize type I error, the cutoff point are set as 0.065 and 0.96. For these four cutoff points, the misclassification rates are 0.1564, 0.1811, 0.7160 and 0.2387 respectively.

Table 4. Confusion matrix at different cutoffs of regular logistic regression

	predict	response	Sensitivity	Specificity	Type I error	Type II error

cutoff=0.5 (common used)		1	0				
	1	31	10	0.5254237	0.9456522	0.05434783	0.4745763
	0	28	174				
cutoff=0.2091 (Youden's Index)	predict	response 1	0	Sensitivity	Specificity	Type I error	Type II error
	1	41	26	0.6949153	0.8586957	0.1413.43	0.3050847
	0	18	158				
cutoff=0.065 (maximize detection of any readmission)	predict	response 1	0	Sensitivity	Specificity	Type I error	Type II error
	1	59	174	1	0.05434783	0.9456522	0
	0	0	10				
cutoff=0.96 (minimize type I error)	predict	response 1	0	Sensitivity	Specificity	Type I error	Type II error
	1	1	0	0.01694915	1	0	0.9830508
	0	58	184				

4. Results and Discussion

Table 5 shows the comparison between univariate analysis and multivariate analysis for readmission. Both of univariate and multivariate analysis show significant for implant loosing, pseduoarthrosis, and adjacent segment disease.

Table 5 Comparison between univariate analysis and multivariate analysis

Variable	Risk factor	significance	
		Univariate	Multivariate
Sex	female	non-significant	non-significant
Insurance	Medi care or medi care + medic aid	non-significant	non-significant
	Medic aid	non-significant	non-significant
	Self-pay	non-significant	non-significant
Diabetes	Yes	non-significant	non-significant
Smoking	Yes	non-significant	non-significant
ASA class	more than severe disturb (3,4,5)	non-significant	non-significant
AGE (cont.)		non-significant	non-significant
BMI (cont.)		non-significant	non-significant
Preop VAS (cont.)		non-significant	non-significant
Postop VAS (cont.)		non-significant	non-significant
Delta VAS (cont.)		non-significant	non-significant
Median income (cont.)		non-significant	non-significant
Hypertension	Yes	non-significant	non-significant
Implant loosening	Yes	significant	significant
Implant brakeage	Yes	non-significant	non-significant

Comorbidities	Yes		non-significant	non-significant
pseduoarthrosis	Yes		significant	significant
Adjacent segment disease	Yes		significant	significant
Proximal junctional fracture	Yes		non-significant	non-significant
Destination post-op	Yes		non-significant	non-significant
Revision surgeries	Yes		significant	non-significant
RACE	non-white		non-significant	non-significant
COPD	Yes		non-significant	non-significant