Università degli studi di Milano-Bicocca

Bachelor's degree in Statistical and Economic Sciences

Research internship

Cervical Brace use in Post-Operative Cervical Spine Surgery

February 14 - May 06 2022

Department of Economics, Quantitative Methods and Business Strategies

Internship Report

Rambelli Filippo 857180

Supervised by:
Prof. Borrotti Matteo
Prof. Liberati Caterina

In collaboration with:
M.D. Fanti Andrea

Contents

Abstract

Cervical collars are often prescribed after a wide range of cervical surgeries. However, reasons that lead surgeons to prescribe both soft and rigid cervical collars don't always seem to have a counterpart in medical literature. The purpose of this study is to deeply analyse the behaviour of a large sample of Italian surgeons that have filled out a medical survey, with the aim of assessing the prescription patterns and understanding the rationale for post-operative bracing in Italy.

1 Introduction

This report will summarise the research internship that took place at the Department of Economics, Quantitative Methods and Business Strategies (DEMS) from 14 February to 6 May 2022. The whole project has been supervised by Prof. Borrotti Matteo and Prof. Liberati Caterina.

The data used during the internship were provided by the Italian neurosurgeon M.D. Fanti Andrea, who also suggested the main targets of the analysis.

First and foremost, we focused on understanding and studying the medical problem and on establishing the main research targets. The first goal was to provide a broad overview of the current status of postoperative bracing in Italy by using the information collected through the questionnaire. The second major aim was to identify main factors that lead surgeon to prescribe rigid, soft and no type of cervical collar after cervical surgeries. Once the main reasons behind prescriptions have been identified, it is possible for an expert in cervical spine surgery to conduct a deep supervision of the results and to check the consistency through a comparison with medical literature.

Concretely, using the available data from the survey, a deep exploratory analysis was carried out, in order to explore, learn and gain maximum insight into the data set. During the exploratory analysis, the steps

necessary for the data preparation were identified as well. Finally, the modelling phase took place with the creation of regression models, in order to identify the impact of each predictor into the prescription of cervical braces. In particular, two comparisons were performed: soft against rigid collar and soft collar against no collar. As we will see, not all the significant regressors for the first model were declared significant for the second one and vice versa. This is a positive outcome, as it means that factors that lead surgeons to prescribe collars are different for each type of brace.
Regarding the foundations of interest, it has come to light that surgeons that rely on literature have a positive association with prescription of soft collars in both the considered models.

2 Materials and Methods

2.1 Materials

2.1.1 Survey

A 45-items online survey was distributed to a sample of Italian surgeons with experience in cervical spine surgery. Overall the survey was filled out by 61 specialists; 58 out of those are neurosurgeons and the remaining 3 are orthopaedists.
The aim of the first part of the questionnaire is to collect information about the surgeons and the hospitals in which they operate; this information is really useful to perform stratified analysis. Surgeons were asked about their age, how many years of experience they have, the amount of surgeries they perform per year, the pathology they deal the most with and the type, location and catchment area of their hospitals.
The second section of the survey is focused on collecting information about surgeries themselves. In particular, surgeons were asked about 10 types of different surgeries, and they had to give information about the brace they prescribe for each surgery, the duration of the prescription and the clinical features that lead to their decision.
The last section contains few question about the foundations surgeons rely on for choosing whether or not to prescribe a specific type of cervical brace.

2.1.2 Datasets

Making use of the data collected through the survey, a dataset was created. The dataset consists of 61 rows, one for each surgeon, and 104 columns, containing demographic and professional information of the surgeons and the details of 10 specific surgeries they perform. This version of the dataset was used to explore the characteristics of the population, while it was necessary to create a second dataset in order to carry out statistical inference. Each row of the new dataset represents 1 of the 10 surgeries done by 1 of the 61 surgeons. Overall there were 30 surgeries not performed by some surgeons, so the second dataset has $61x10-30=580$ rows. The transformation of the dataset involved qualitative features, too. Each class of a categorical variable was converted into a $\{0,1\}$ dummy variable. For instance, the variable hospital type={university, polyclinic, other} generated 3 variables: university=$\{0,1\}$, polyclinic=$\{0,1\}$ and other=$\{0,1\}$, where 0 and 1 respectively indicate the absence or presence of the considered attribute. After making these adjustments, the total amount of columns in the second dataset was 42.

2.2 Methods

2.2.1 Preprocessing

The dataset required some preprocessing operations to make it suitable for the analysis.
Firstly, some answers to open questions needed to be fixed due to compilation errors. Secondly, coding adjustments to some occurrences were necessary in order to guarantee the comparability among themselves (e.g. Emilia Romagna and Emilia-Romagna). Lastly, few missing values were treated using an active strategy. In particular, 5 missing values for the variable 'number of annual cervical surgeries' were imputed using mean imputation conditioned to the catchment area of the hospital. Furthermore, other missing values related to hospitals were replaced through donations using values from other records referring to the same hospitals.

2.2.2 Exploratory analysis

The exploratory analysis was carried out using traditional techniques of data visualisation, tables of frequencies and statistical indexes. Missing data were discovered and evaluated through graphical visualisations, using especially of R's Naniar package. Box plots were the main tools to spot outliers, while histograms and non-parametric density plots were used to analyse distributions of quantitative variables. Furthermore, we used bar plots and contingency tables in order to evaluate the distribution into classes of categorical variables. Lastly, statistical indexes, conditioned box plots and bar plots were used to perform stratified and conditioned analysis.

2.2.3 Logistic regression

With reference to the inferential phase, two generalized linear models (glm) with binary outcomes and link logit were performed. In logistic regression models with multiple predictors $X = (X_1, X_2, \dots, X_p)$, the relationship of dependency between the posterior probability and the predictors is described by the logistic distribution as follows:

$$P(Y=1|x) = \pi(x) = \frac{e^{\beta_0 + \sum_{j=1}^{p} \beta_j x_j}}{1 + e^{\beta_0 + \sum_{j=1}^{p} \beta_j x_j}}$$

where $(\beta_0, \beta_1, \dots, \beta_p)$ are the coefficients obtained through maximum likelihood estimation (MLE).
Equivalently, we can consider the logit function and rewrite it as:

$$logit(\pi(x)) = \ln\left(\frac{\pi(x)}{1-\pi(x)}\right) = \beta_0 + \sum_{j=1}^{p} \beta_j x_j$$

Each regression coefficient β_j corresponds to the variation in the logit of the average outcome after an unitary increase in the j-th covariate, with everything else being equal. The interpretation of coefficients is carried out through odds ratio, as each β coefficient represents a log(odds ratio). So, considering the exponential transformation of the coefficients, we are evaluating p odds ratio.
An estimated coefficient grater than 0, or equivalently, an exp(coefficient) greater than 1, indicates positive association between the outcome and the class associated to the coefficient compared to the baseline level; while a coefficient smaller than 0 express positive association between the outcome and the baseline level. The exponential of the intercept expresses the odds of the outcome associated to an observation whose

qualitative variables take the baseline level and whose quantitative variables are equal to zero.

2.2.4 Stepwise bidirectional selection

In both comparison performed in the inferential part of the analysis, the final logistic regression model model was obtained performing a stepwise bidirectional selection based on AIC (Akaike Criterion Information). AIC is a statistical measure useful to assess models fitted on the same data, as it measures the goodness of fit of a model penalized for its complexity, intended as the number of predictors. AIC is computed as follows:

$$AIC = -2 \max_{\beta} logL(\beta; Y) + 2p$$

where $L(\beta; Y)$ is the likelihood of the model and p is the number of independent variables.
According to this criterion, when we compare two or more models fitted with the dependent variable and on the same data, we choose the one with the lowest AIC. The stepwise bidirectional selection procedure consists in fitting a new regression model for each step on a different set of explanatory variables obtained by adding or removing one covariate at a time until AIC can not be minimized anymore.

3 Results

3.1 Exploratory analysis

In the first steps of exploratory analysis, many characteristics and relationships of the population were inspected. In the following tables the main features regarding surgeons and hospitals are summarised.

Specialisation	
Orthopaedic	4.9%
Neurosurgery	95.1%
Age (years)	
< 30	0%
30-40	9.8%
40-50	31.2%
50-60	18.0%
> 60	41.0%
Experience (years)	
< 5	4.9%
5-10	8.2%
10-20	27.9%
> 20	59.0%
Cervical surgeries per year	
< 20	6.6%
20-50	29.5%
50-80	34.4%
> 80	29.5%

Table 1: Population's characteristics

Hospital	
Public	73.8%
Affiliated	26.2%
Main pathology	
Elective	85.2%
Traumatic	14.8%
Hospital type	
University	26.2%
Polyclinic	27.9%
Other	45.9%
Area of Italy	
North	50.8%
Center	18.0%
South and islands	31.2%
Hospital catchment area	
< 100.000	6.6%
100.000-500.000	41.0%
500.000-1.000.000	31.1%
$> 1.000.000$	21.3%

Table 2: Population's characteristics

As it is displayed, the majority of the respondents are neurosurgeons that mainly deal with elective pathology. The youngest surgeon is 35 years old, the oldest is 80 and the average age is 54.9 years (s.d. 11.6 years). The specialist with the least experience has been working for 2 years, while the most expert has 46 years of experience; the average experience is 22.0 years (s.d. 11.0 years). With regard to the amount of surgeries carried out by the respondents per year, it ranges from 10 to 250 with an average of 64.9 (s.d. 45.5). More than one half of the interviewed surgeons come from Northern Italy, followed respectively by Southern and Center Italy. With reference to the hospitals, almost one third of them are public and over two-fifths of them are neither polyclinic nor university. Lastly, over half of the considered hospitals has a catchment area greater than half a million people.
Considering surgeries, frequency and duration of prescriptions for each type of brace were evaluated. In particular, percentages of prescriptions were computed conditionally to 10 different types of surgeries. Moreover, duration of prescriptions were calculated conditionally to the two type of brace prescribed (soft and rigid). The following charts summarise the just mentioned information.

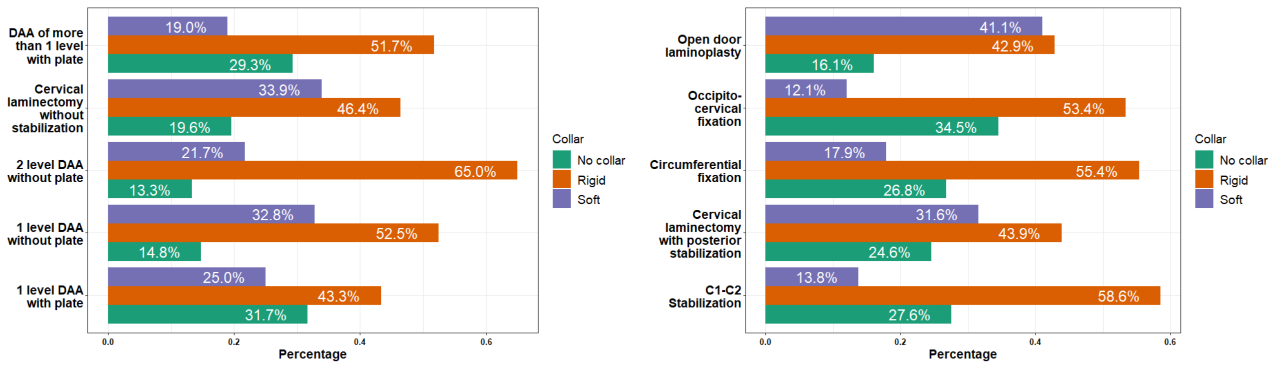

Figure 1: Collars' prescriptions frequencies

The surgery associated to the lowest brace prescription in post-operative is occipito-cervical fixation, while the one that lead to the highest brace prescription is 2 level discectomy and anterior arthrodesis without plate.

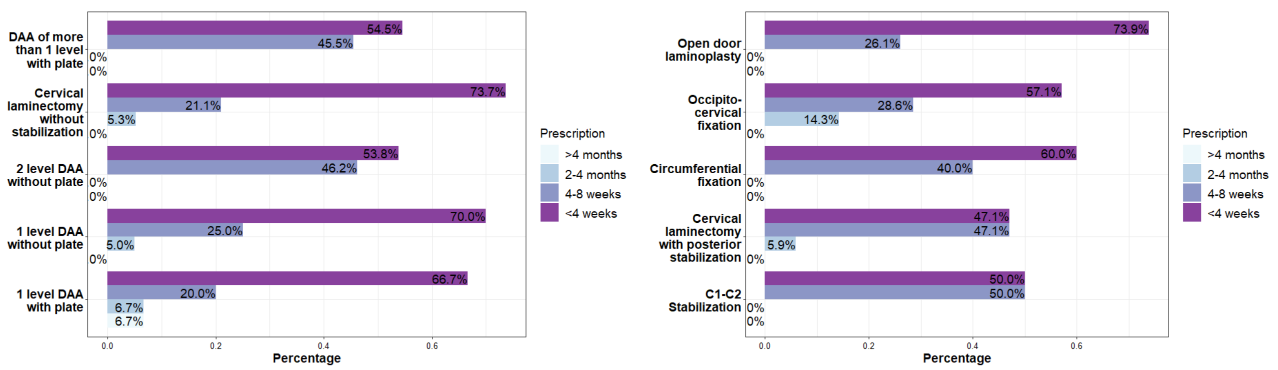

Figure 2: Soft collars' prescription duration

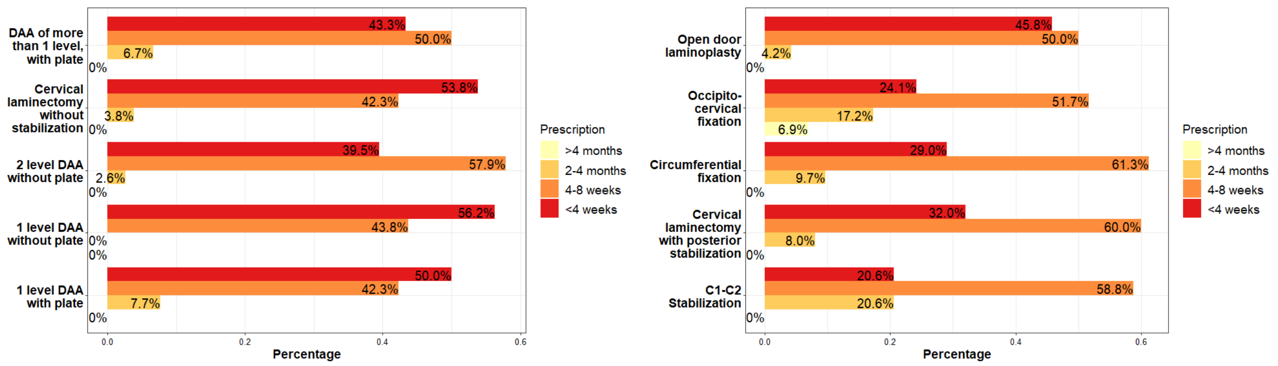

Figure 3: Rigid collars' prescriptions duration

With regard to the duration, soft collars usually have a short length of prescriptions, as the modal class is 4 weeks. On the other side, rigid braces are generally prescribed for a longer time, as the most common duration is 4-8 weeks and it is quite common for prescriptions to last up to 4 months. Lastly, clinical features and foundations behind prescriptions were analysed.

Clinical feature	Relevance
Advanced patient's age	62.2%
Previous operations in anamnesis	18.9%
Poor bone quality	18.9%

Table 3: Soft collar - clinical features' relevance

Clinical feature	Relevance
Advanced patient's age	23.5%
Previous operations in anamnesis	25.5%
Poor bone quality	51.0%

Table 4: Rigid collar - clinical feature's relevance

Foundation	Relevance
Literature	14.8%
Personal experience	32.7%
Colleagues' teaching	24.1%
Local customs	16.1%
Legal and medical protection	12.3%

Table 5: Foundations' relevance for prescription

Personal experience is the foundation surgeons rely the most on for their prescriptions, followed by literature and colleagues' teaching. Specifically for each brace, the most important clinical features that lead surgeon to prescribe respectively soft and rigid collars are old age of the patient and poor bone quality.

3.2 Inference

Before performing logistic regression, some transformations on the quantitative variables were applied. The variable 'year of birth' was excluded from the models in order to avoid multicollinearity, due to its strong correlation (-0.91) with another predictor 'years of experience in cervical surgery'. The logarithmic transformation was applied to the variable representing the number of cervical surgeries per year, aiming to adjust its positive asymmetry and to remove outliers. Finally, all the continuous predictors were standardized, in order to avoid scale effects. Standardized predictors are characterized by a zero mean and unit variance.

Standardization was applied specifically for each model using only the observation useful for each logistic regression, in order to avoid creating a bias.

3.2.1 Rigid against soft collar

The first logistic regression model was created to perform a comparison between rigid and soft cervical braces. Obviously, the observation included in this model are only the ones having the dependent variable that assumes value either rigid or soft. The target variable 'Collar' consists of 298 observations (67.4%) from the class 'rigid' and the remaining 144 from the class 'soft' (32.6%).
At the beginning we estimated the full model with all the predictors obtaining an AIC equal to 456.2. Then the AIC stepwise bidirectional selection was performed, leading to the optimal model according to this criterion. The final model has 21 predictors excluding the intercept, and has an AIC equal to 439.2. The model was deeply evaluated through different criteria before going any further.
Pearson chi-squared goodness-of-fit test was performed and led to the acceptance of the null hypothesis of goodness of the model, with a p-value of 0.7872. Also the Hosmer-Lemeshow test, considering $k+2$ partitions, led to the acceptance of the null hypothesis, as the p-value is equal to 0.6317. Diagnostic of single observation did not show any relevant trouble.
The most influential observation corresponds to a 56 years old orthopaedist from North Italy that prescribes a soft collar after open-door laminoplastic surgeries. A Bonferroni outlier test was performed, with the result that the suspected observation was not actually an outlier, so the record was not excluded from the model. The high influence is due to the fact that this surgeon is the only orthopaedist above the respondents that prescribes a soft collar. Even the residual analysis did not raise any doubts about the goodness of the model. As we can see, dispersion around x-axis is pretty homogeneous and no systematic trend exists.

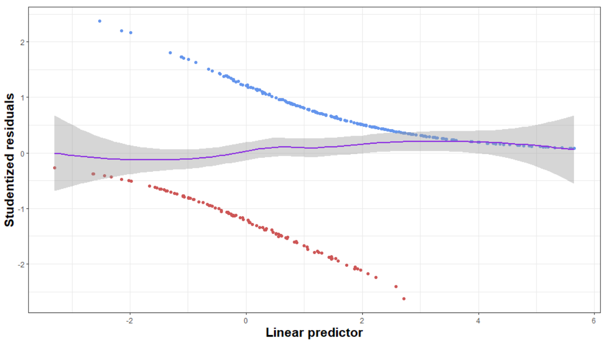

Figure 4: Residuals plot, soft against rigid collar

Once model evaluation is finished, we can move on the interpretation phase.
The target variable was recoded as follows:

- 0 if brace is soft
- 1 if brace is rigid

The following table shows the IRLS output obtained by estimating the best model according to AIC.

	Estimate	Odds ratio	Std.Error	Signif.
(Intercept)	−2.13	0.12	0.58	***
Experience in cervical surgery	0.43	1.54	0.17	***
Log(surgeries per year)	−0.33	0.72	0.15	**
Arthrodesis ratio increase	1.85	6.35	0.31	***
Patient 'expects' a brace	1.63	5.11	0.38	***
Literature	−0.71	0.49	0.32	**
Personal experience	1.58	4.84	0.50	***
Colleagues' teaching	−0.83	0.44	0.41	**
Local customs	−1.99	0.14	0.37	***
Legal and medical protection	0.67	1.96	0.34	**
Orthopaedic	3.06	21.34	1.13	***
Polyclinic hospital	1.61	4.99	0.42	***
University hospital	1.73	5.61	0.44	***
Affiliated hospital	0.59	1.81	0.38	
Southern Italy and islands	0.97	2.64	0.37	***
Catchment area <100.000	−1.78	0.17	0.72	**
Traumatic pathology	1.03	2.80	0.41	**
2 level discectomy and anterior arthrodesis without plate	0.86	2.37	0.42	**
Circumferential fixation	0.77	2.16	0.47	
Occipito-cervical fixation	1.31	3.69	0.53	**
Cervical laminectomy without stabilization	0.66	1.94	0.41	
C1-C2 stabilization	1.15	3.15	0.50	**
AIC	439.18			
Log Likelihood	−197.59			
Deviance	395.18			
Num. obs.	442			

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 6: IRLS output - soft against rigid collar

In the considered model, the exponential of the intercept, which is equal to 0.12, indicates the odds of a surgeon that is specialized in neurosurgery, works in a public hospital in the Northern Italy etc, and has performed a 1 level discectomy and anterior arthrodesis with plate surgery.
Surgeons that rely on literature, have an odds of rigid brace which is 51% smaller than surgeons who do not, with all the other regressors being the same. On the other hand, surgeons that prescribe a collar because patient 'expects' it, have an odds of the outcome which is over five times greater than surgeon who do not consider this factor when it comes to prescribe a brace.
Considering the other foundation, it turned out that personal experience and legal-medical protection have a positive association with the prescription of rigid braces, while colleagues' teaching and local customs have a positive association with soft braces' prescription. With reference to surgeries, a positive associations with the rigid collar exists switching from the baseline to all the 5 significant surgeries.

3.2.2 Soft against no collar

The second model was performed to compare soft brace's prescription with the prescription of no collar. The dependent variable consists of 144 observations from class 'soft' and 138 from class 'no collar'; data are nearly perfectly balanced, as proportions are 51.1 : 48.9 .
In this situation, dummy regressors representing clinical features that lead surgeon to prescribe collars were excluded from the model in order to avoid multicollinearity and perfect separation.
After an initial estimation of the full model (AIC 316.2), the stepwise selection based on AIC led to a model with 15 regressors and AIC equal to 303.4. Both Pearson's chi-squared and Hosmer Lemeshow (g=17) tests rejected the null hypothesis of model's goodness, respectively with p-values equals to 6.62 x 10^{-9} and 0.0043. Going into detail, there are 3 observations having both high influence on the model and large residuals. 2

out of 3 suspected statistical units refer to the same neurosurgeon, who prescribes soft collars in both the considered situations. The other influential observations refers to a different neurosurgeon that does not prescribe any brace after cervical laminectomy with posterior stabilization surgeries.
Then a new model was computed excluding these observations and it turned out that all the factors useful to evaluate the model have significantly improved.
The final model has 17 regressors and AIC equal to 270.9; this time both Pearson's chi-squared and Hosmer Lemeshow (g=19) tests failed to reject the null hypothesis, due to p-values respectively equal to 0.9776 and 0.1800. Neither diagnostic raised any doubt about the goodness of the final model, as there are no observation with large residuals and Bonferroni test denies the existence of outliers.

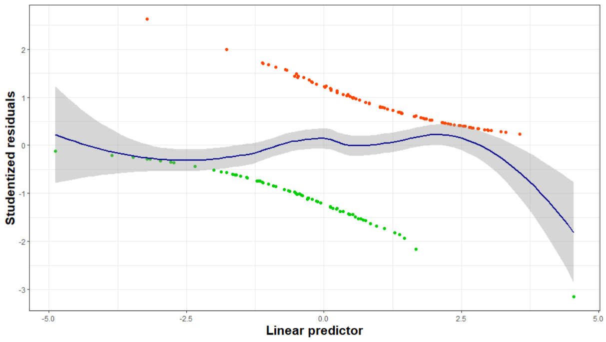

Figure 5: Residuals plot, soft against no collar before removing outliers

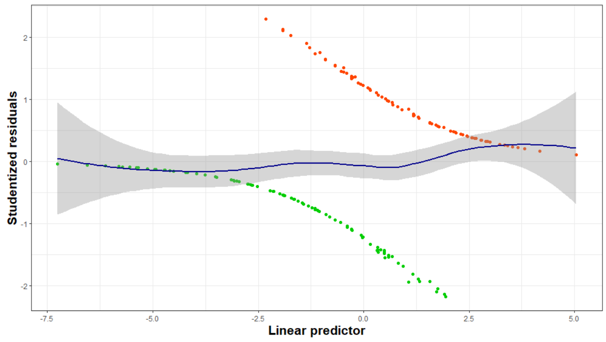

Figure 6: Residuals plot, soft against no collar after removing outliers

The target variable has been recoded as follows:

- 0 if brace is soft
- 1 if brace is missing

The final model's IRLS output is displayed in the following table.

	Estimate	Odds ratio	Std.Error	Signif.
(Intercept)	−0.90	0.41	0.80	
Literature	−1.20	0.30	0.52	**
Personal experience	2.59	13.29	0.76	***
Colleagues' teaching	−4.70	0.01	0.77	***
Local customs	−1.83	0.16	0.55	***
Orthopaedic	2.17	8.76	1.21	*
Polyclinic hospital	4.44	84.95	0.82	***
University hospital	1.86	6.41	0.74	**
Central Italy	−1.08	0.34	0.63	*
Southern Italy and islands	3.68	39.76	0.76	***
Catchment area 500.000-1.000.000	1.08	2.95	0.47	**
Catchment area >1.000.000	1.31	3.70	0.50	***
Traumatic pathology	−1.24	0.29	0.53	**
2 level discectomy and anterior arthrodesis without plate	−1.60	0.20	0.63	**
Occipito-cervical fixation	0.88	2.41	0.57	
Cervical laminectomy with posterior stabilization	−0.76	0.47	0.53	
Cervical laminectomy without stabilization	−1.17	0.31	0.59	**
Open door laminoplasty	−1.59	0.20	0.56	***
AIC	270.85			
Log Likelihood	−117.43			
Deviance	234.85			
Num. obs.	279			

$^{***}p < 0.01$; $^{**}p < 0.05$; $^{*}p < 0.1$

Table 7: IRLS output - soft collar against no collar

The intercept has the same interpretation it had in the previous model, as the baseline levels are still the same. This time there is not any significant continuous regressor.
Surgeons that rely on personal experience have on odds of not prescribing any collar that is approximately 13 times bigger than surgeons that do not rely on this foundation, with everything else being equal.
However, all the other significant foundations - literature, local's customs and colleagues' teaching - have a negative association with prescription of no collar. Considering the catchment area between 100.000 and 500.000 as baseline, the odds of no collar increase if we shift to hospitals with larger catchment areas.
With regard to specialisations, there is a positive association the lack of collars' prescriptions moving from neurosurgeons to orthopaedists.
Lastly, with reference to surgeries, only occipito-cervical fixation has an odds of no collar bigger than the baseline level, while the remaining 4 significant surgeries have a negative association with the outcome.

4 Conclusions

The performed analysis managed to provide an accurate overview of the current Italian post-operative bracing status. Through inference, factors that determinate the most prescriptions of each collar were identified. Despite baseline levels being all the same, significant regressors for each model are quite different, highlighting that factors behind prescription of different collars are actually different from each other.
Despite some estimated coefficient with large absolute values, the obtained models are certainly satisfactory. Extreme values in estimation are due to limited size of some classes of regressors.
Significance and effect on the outcome of foundations that underlie prescriptions are the ones that deserve the most an interpretation commitment. In particular, literature, if correctly interpreted, may provide answers regarding the consistency of prescriptions with medical literature. Moreover, the interpretation of local customs and legal-medical protection's coefficient may clarify whether or not there is any systematic trend in prescribing braces without any clinical evidence. However, further interpretation should be carried out together with an expert in cervical spine surgery.
An additional step of the analysis could be the creation of parametric and non-parametric classifiers with the aim of predicting the prescribed brace. Classification may be based on the specific surgery and on the characteristics of the surgeon and of the hospital that hosts the surgery.
Some of the possible classifiers are logistic regression itself, decision tree, random forest and neural networks. Considering the high presence of categorical variables, classification methods such as KNN and discriminant analysis can not be applied to considered data.
Possibly data could be enriched by adding clinical features specific for each patient, such as age, bone quality index ans number of previous cervical surgeries. In this case classification methods that work with numerical variables may be applied too.