Brain and Spine 4 (2024) 103918

Contents lists available at ScienceDirect

Brain and Spine

journal homepage: www.journals.elsevier.com/brain-and-spine

Machine learning predictive model for lumbar disc reherniation following microsurgical discectomy☆

Angel G. Mehandzhiyski [a,*], Nikola A. Yurukov [b], Petar L. Ilkov [a], Dilyana P. Mikova [c], Nikolay S. Gabrovsky [a]

[a] *Department of Neurosurgery, UMHATEM "N.I.Pirogov", Sofia, Bulgaria*
[b] *Staff Software Engineer, Canva Pty Ltd, Sydney, NSW, Australia*
[c] *Nuclear Medicine Clinic, UMHAT "St. Ivan Rilski", Sofia, Bulgaria*

ARTICLE INFO

Handling Editor: Prof F Kandziora

Keywords:
Machine learning
Lumbar disc herniation
Microsurgical discectomy
Recurrence
Same level
Reoperation

ABSTRACT

Introduction: The integration of machine learning (ML) algorithms into the field of neurosurgery has the potential to facilitate the decision-making process for the surgeons, improve the surgical outcomes and the overall patient satisfaction rates. Reoperations for same level lumbar disc reherniation are associated with poorer outcomes and greater rate of complications.
Research question: Proper preoperative patient evaluation could reveal the individuals at higher risk of reherniation. A novel machine learning algorithm was used for the creation of a predictive scoring system for lumbar disc reherniation for patients requiring microdiscectomy without fusion.
Material and methods: Retrospective chart review was completed of all adult patients that underwent microdiscectomy without fusion for symptomatic single level LDH, in a single center, over the last 3 years. 230 patients met the inclusion criteria. 19 of them required a second surgical intervention due to same level reherniation.
Results: Utilizing the Risk-SLIM model, the Lumbar Reherniation Score (LRS) was created. The score's accuracy was tested against other model architectures, and a standard five-fold cross-validation was performed. The LRS has AUC of 0.87, confusion matrix accuracy of 0.74, Matthews correlation coefficient of 0.36 and informedness of 0.62. The LRS individual reherniation risk probability ranges from 0% to 88.1%.
Discussion and conclusion: The LRS is a novel, easy-to-use, patient-specific tool for preoperative prediction of the individual patient-specific risk of same level symptomatic reherniation following microdiscectomy. Further validation and testing of the model is needed before it can be used in real-life patient treatment.

1. Introduction

Lumbar disc herniation (LDH) affects about 2–3% of the general population. Males are slightly more affected with approximately 4.8% of those over the age of 35 suffering from LDH compared to 2.5% of women in the same age group. Computed tomography (CT) and Magnetic resonance imaging (MRI) are the two main modalities used during the diagnostic process (Vialle et al., 2010). The clinical presentation of LDH varies. Progressive neurological deficits (motor or sensitive) in the lower extremities, partial loss of bladder or bowel function or cauda equina syndrome are clear indications for surgery (Yoon and Koch, 2021). The most performed spinal operation in the world for LDH is microsurgical discectomy (MSD) (Blamoutier, 2013). The most common complication following MSD is the recurrence (same level reherniation) of the disc herniation.

Different studies have identified lumbar disc reherniation (LDR) recurrence rates following microdiscectomy of 5–15% (Swartz and Trost, 2003; Zileli et al., 2024). Same level reherniation surgery is associated with poorer postoperative outcomes and may require fusion of the affected level due to complete intervertebral disc degeneration (Kim et al., 2021). Our literature review identified 5 separate factors that contribute to the lumbar disc herniation and reherniation - Disc extrusion or sequestration, Disc height reduction of 20% or more, Diabetes type I or II, Same level intervertebral joint degeneration and Smoking. (Kim et al., 2021; Mariscal et al., 2022; Huang et al., 2016a). The objective of this study was to assess the feasibility and viability of using a ML algorithm for the creation of a predictive score for lumbar

☆ Previous Presentations: The findings described in this paper will be presented as an oral poster at the 2024 EANS congress in Sofia, Bulgaria.

* Corresponding author.
E-mail address: mehandzhiyski.angel421@gmail.com (A.G. Mehandzhiyski).

https://doi.org/10.1016/j.bas.2024.103918
Received 15 July 2024; Received in revised form 10 September 2024; Accepted 9 October 2024
Available online 10 October 2024

List of used abbreviations and acronyms:

LDH -	lumbar disc herniation
LDR -	lumbar disc reherniation
MSD -	microsurgical discectomy
CT -	computed tomography
MRI -	magnetic resonance imaging
FJOA -	facet joint osteoarthritis
DM -	diabetes mellitus
LRS -	lumbar reherniation score
ROC –	receiver operating characteristic
AUC –	area under the curve
ML –	machine learning

disc reherniation following microdiscectomy without fusion. The Lumbar Reherniation Score (LRS) was created by utilizing the novel Risk-SLIM ML algorithm.

Abbreviations: LDH (lumbar disc herniation), LDR (lumbar disc reherniation), MSD (microsurgical discectomy), CT (computed tomography), MRI (magnetic resonance imaging), FJOA (facet joint osteoarthritis), DM (diabetes mellitus), LRS (lumbar reherniation score)

2. Materials and methods

2.1. Study population and data collection

A retrospective chart review was completed to identify patients, over the age of 18 years, that underwent microsurgical discectomy without fusion for symptomatic lumbar disc herniation, in a single institution from 2019 to 2022. Follow-up ranged from 1 to 3 years. Basic medical data was obtained – age, sex, comorbidities, previous spinal surgeries, genetic disorders, smoking and diabetes status. MRI and CT imaging were acquired. Exclusion criteria were: multilevel spinal pathology, previous thoraco-lumbar spine surgery, anatomical variations regarding the composition and/or the number of the lumbar vertebrates, known genetic anomalies associated with connective tissue disorders. In the end, 230 patients were included in the study. All of them had intracanalar lumbar disc hernias and were treated with microsurgical discectomy via an interlaminar approach. The surgeries were performed by 4 certified neurosurgeons working in the same department (2 professors with 15+ years of practice each, 1 associated professor with 10+ years of practice and 1 senior neurosurgeon with 5+ years of practice). 211 of them had no reherniation during the follow-up period.

19 individuals presented to the clinic during the follow-up period with same level symptomatic reherniation that required a second surgical intervention (MSD with or without fusion). LDRs were diagnosed by performing lumbar MRIs in patients with recurring neurological symptoms (severe pain or muscle weakness) during the follow-up period.Our literature review identified 5 separate patient-specific factors that contribute to lumbar disc degeneration and reherniation. (Kim et al., 2021; Mariscal et al., 2022; Huang et al., 2016a) (Table 1)

The radiological information was individually assessed by three separate authors (A.G.M, P.L.I, D.P.M).

2.2. Data analysis

Python's Numpy package (Harris et al., 2020) was used to generate some initial statistics for the data, including the data type, frequency for each possible value, the mean and the standard deviation (assuming that for binary variables "0" means "No" and "1" means "Yes") (Table 2)

The means of all independent variables split across the two samples were compared using an independent samples T-test, with the null hypothesis being that the 2 independent samples have identical means. The resulting p-values were tabulated (Table 1). Any independent variables with a p-value less than 0.05 are interpreted as statistically more significant for the reherniation process. The T-test only measures direct mean correlations, and does not account for non-linear interactions between the variables - which is why this analysis is simply used to identify biases in the data initially - they do not represent the relative importance of variables in the final model - the right proxy measure for this is the weight of each variable in the final model.

Please note we have used Welch's *t*-test (Welch, 1947) which does not assume equal population variances (West, 2021). This is important as we have no guarantee of equal variances in this case. Also please note that this is a preliminary analysis, where the p-value significance is a sufficient, but not necessary condition for variable inclusion in the model.

2.3. Creation of the model

The Risk-SLIM method was chosen as it has been previously used in multiple scoring systems with satisfactory results on external validation (Ustun and Rudin, 2019). The Risk-SLIM model was trained with all the available data. For each of the input variables, a weight was assigned

Table 2
Benchmark values for 80% of normal lumbar disc height at every lumbar level were calculated for males and females.

Intervertebral level/Gender	Male	Female
L1-L2	4.48 mm	3.92 mm
L2-L3	5.36 mm	4.64 mm
L3-L4	5.76 mm	5.12 mm
L4-L5	6.08 mm	5.52 mm
L5-S1	5.76 mm	5.44 mm

Table 1
Descriptive Statistics and Welch's *t*-test.

Risk factor	Data Type	Frequency	Mean	Standard Deviation	p-value	Supporting studies
Reherniation	Yes/No	Yes – 19 No – 211	0.08	0.275	–	–
Smoking	Yes/No	Yes – 69 No – 161	0.30	0.458	0.0558	Mariscal et al., 2022.; Huang, Weimin et al., 2016
Diabetes	Yes/No	Yes – 21 No – 209	0.09	0.288	0.1998	Huang, Weimin et al., 2016
Disc Extrusion	Yes/No	Yes – 197 No – 33	0.86	0.350	$2.38*10^{-9}$	Huang, Weimin et al., 2016
Disc Height Decrease 20%+	Yes/No	Yes – 85 No – 145	0.37	0.483	$1.67*10^{-4}$	Yong Guk Kim et al., 2021
Facet Joint Degeneration	Category 1, 2, 3 or 4	1–65 2–117 3–38 4–10	1.97	0.788	$7.55*10^{-5}$	Yong Guk Kim et al., 2021

that will be summed up to produce a total score for the patient. This risk score can then be interpreted as a probability of reherniation occurring for a particular patient (Table 3).

This scoring system has a minimum of 0 points and maximum is 13 points. Please note that all variables are binary, except Facet Joint Degeneration which has an integer value between 1 and 4 and can contribute up to 4 points to the risk score. The following mathematical formula is used to calculate a probability table:

Pr (Reherniation = 1) = 1 / (1+e^(11-score))

2.4. Model performance

In order to evaluate the accuracy of the machine learning model, a standard tool in binary classification analysis was used – the receiver operating characteristic curve (ROC curve) (Fawcett, 2006; Hanczar et al., 2010; Flach et al., 2011). It helped us visualize the effectiveness of our model, as well as the trade-offs between false positives and true positives as we tried to set a threshold for our model. Some caveats of this are explored in the literature (Hernández-et al., 2012; Hand, 2009; Saito and Rehmsmeier, 2015). The ROC curve is considered to be the gold standard for medical diagnostic tests (Hajian-Tilaki, 2013) (Fig. 1) (Nahm, 2022).

The AUC (area under the curve) of our model is 0.87 (Fig. 1). An AUC of 0.8–0.9 generally characterizes a good model, and an AUC of 0.9+ characterizes an excellent model. A model with AUC ≥ 0.8 is typically considered acceptable (Welch, 1947).Table. 4

The diagonal line in Fig. 1 represents a fully random model – one that has no information and is indistinguishable from random chance. The additional area covered by the curve characterizes the predictive power of the model. We can also see the thresholds at the various points on the curve, suggesting a cutoff probability of around 10% will work best for our model.

This means any patient who scores above 10% in our model (equivalent to having 9 or more score points) will be predicted to reherniate, and patients with less than that will be predicted not to have a re-occurrence.

Choosing the optimal threshold of 10% we can produce a confusion matrix (Fig. 2).

This matrix is a standard tool in binary classifiers and shows the number of true and false positives and negatives (Hicks et al., 2022). 153 patients were classified as true negatives (predicted to not re-herniate and indeed didn't re-herniate), 58 were false negatives (predicted to re-herniate but they didn't), 17 were true positives (predicted reherniation that actually occurred) and 2 were false positives (reherniation happened even though it was not predicted). The above model and confusion matrix indicate an accuracy of 0.74, a Matthews correlation coefficient (MCC) of 0.36 and an informedness of 0.62. These results indicate that the model is slightly biased towards over-predicting the probability of reherniation.

2.5. Cross-validation

To test the model for overfitting, a standard five-fold cross-validation was performed (Little et al., 2017a, 2017b; Tougui et al., 2021; Krstajic et al., 2014). This produced 5 variations of the model, based on 80% of the data. There were no significant changes to our risk score model through all five cross-validation training runs, which indicates the model is stable.

Table 3
RiskSLIM LDH Risk Score - Points are assigned to every risk factor according to its significance.

Independent Variable	Points
Disc Extrusion	5
Disc Height Decrease 20%+	2
Smoking	1
Diabetes	1
Facet Joint Degeneration	1–4

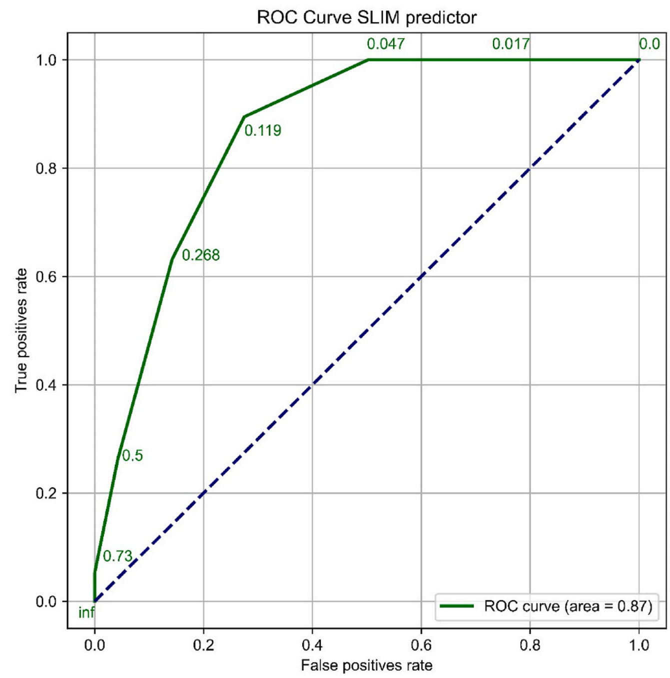

Fig. 1. ROC curve of the Lumbar Disc Reherniation model based on the RISK-SLIM machine learning algorithm.

Table 4
RiskSLIM Risk Score to Probability Mapping – The following table represents the individual probability of reherniation, calculated by our model, based on the sum total of all independent variables.

Total Patient Score	Probability of reherniation
0	0.0%
1	0.0%
2	0.0%
3	0.0%
4	0.0%
5	0.2%
6	0.7%
7	1.8%
8	4.7%
9	11.9%
10	26.9%
11	50.0%
12	73.1%
13	88.1%

2.6. Exploration of alternatives

Firstly, we explored the alternative of treating the facet joint degeneration (FJD) class as a binary variable, where class 1 and 2 would be set to "0", and class 3 and 4 would be set to "1". This resulted in a very slightly lower AUC from 0.8656 to 0.8591, a MCC from 0.364 to 0.35 and Informedness down from 0.6199 to 0.5425. The conclusion is it is better to treat FJD as an integer variable, as is presented in the final model.

Secondly, to illustrate the comparative performance of our model, we compared it to some standard binary classifiers. We used the RandomForestClassifier in Python's SKLearn (Pedregosa et al., 2011). After

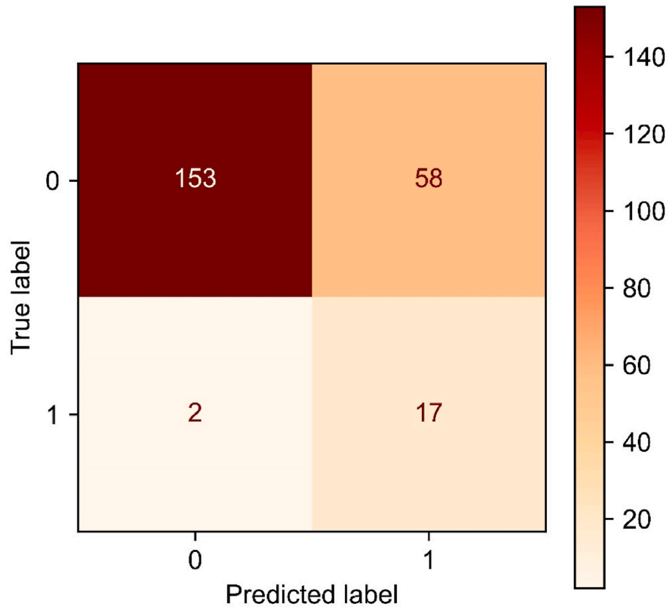

Fig. 2. Confusion matrix of the Lumbar Disc Reherniation model.

using various hyper-parameter settings, a maximum AUC of 0.88 was reached. This is similar to the accuracy we get with Risk-SLIM. Random Forest models wouldn't be useable in medical practice because they are not interpretable and don't give medical practitioners a reason why a certain risk score is provided or how the calculation was performed (Breiman, 2001).

Finally, we wanted to compare the Risk-SLIM model with a deep learning model (DLM) implemented in TensorFlow's Keras (Chollet F and others, 2015). After trying several neuron configurations, with varying network widths and depths, we produced an AUC of 0.75. Better results could likely be achieved through model adjustments; however deep learning models are not interpretable, thus they can not be used in this setting, and are purely used to contextualize the Risk-SLIM results (Alzubaidi et al., 2021).

3. Results

A total of 230 adult patients that underwent microdiscectomy without fusion for symptomatic lumbar disc herniation were retrospectively studied. Of them 19 (8.6%) experienced same level symptomatic reherniation during the follow-up period and required reoperation. Five factors were identified as statistically significant for symptomatic reherniation: disc extrusion or sequestration, smoking, diabetes, preoperative disc height reduction with 20% or more and facet joint. No significant difference in age was observed between the two groups ($p > 0.05$) (Table 5).

Table 5
Preoperative baseline characteristics. No significant difference was observed between the sexes in both reherniation and no reherniation groups.

Category	Males	Females	p-value
Total population	122	108	0.3559
Reherniation population	11	8	0.4913
No reherniation population	111	100	0.4489

Level of herniation	Total cases	Reherniation cases
L1-L2	6	0
L2-L3	12	0
L3-L4	17	3
L4-L5	91	6
L5-S1	104	10

The acquired data was compiled and the Risk-SLIM model was used to create a reherniation prediction score model. Following the compilation of the model, each of the 5 factors was assigned points (1–5) relative to their statistical significance for reherniation. Disc extrusion = 5; disc height decrease of 20% or more = 2; smoking = 1; diabetes = 1, facet joint degeneration (depending on severity) = 1 to 4. A patient specific score is calculated based on the presence or absence of the aforementioned factors. The probability of same level reherniation according to the model varies – from 0% for patients with none of the above-mentioned risk factors to 88.1% for patients with all of the above-mentioned risk factors. These results differ from the generalized populational reherniation risk following microdiscectomy, estimated to be between 5 and 15%, as they offer a personalized, rather than a population mean probability.

The model yielded an AUC of 0.87 which indicates good predictive ability, however some restrictions to the model's accuracy exist: the relatively small patient pool (230 individuals) and that the data is sourced from a single medical center. Nevertheless, the model does accord with the medical literature and further research with a wider patient pool and the enrolment of multiple centers could help broaden its applicability and increase precision.

4. Discussion and Conclusion

Symptomatic lumbar disc herniation affects millions of people worldwide. It's the most common reason for lumbar spine surgery globally. The economic burden of this disease is associated with the reduced productivity of the affected individuals due to sick-listing episodes and permanent disability benefits. The quality of life of the affected people usually improves in the first two years following surgery. Most patients return to their usual activities 6 months following surgery (Kang et al., 2020; Roiha et al., 2023).

Symptomatic reherniation and numerous lumbar surgeries are associated with increased likelihood of permanent neurological deficit, chronic pain syndrome, overall patient dissatisfaction and long-term health-related poor quality of life (Roiha et al., 2023). In different studies, the reherniation rates have been demonstrated to vary between 5% and 15% (Zileli et al., 2024). Several attempts have been conducted aiming to define the main risk factors associated with reherniation. Patient unrelated factors such as the surgical technique used and the learning curve of the surgeon have a direct impact over the reherniation rate (Huang et al., 2016a). Successive surgical interventions are associated with increased risk of intra and postoperative complications (Bombieri et al., 2022).

Machine learning algorithms are capable of creating predictive models based on certain parameters in large pools of heterogeneous data (Fawcett, 2006; Hanczar et al., 2010). Such algorithms can be used to improve the current state of medical care globally by analyzing multiple factors affecting a certain disease and creating predictive outcome scales. For example, Khor et al. produced a freely available web application that utilizes a ML model that predicts the post operative improvement in patients undergoing lumbar spinal fusion (Khor et al., 2018). ML models are currently involved in improving the diagnostics process by interpreting diagnostic imaging, assisting in intraoperative decision making, tumor labeling etc. ML models, and other AI tools, are slowly but surely increasing the effectiveness of the healthcare system (Schonfeld et al., 2024).

The LRS is a novel machine-learning based tool focused on the individual risk of same level reherniation. The LRS was created using the Risk-SLIM machine learning algorithm. It is based on 5 independent factors chosen after a literature review (Swartz and Trost, 2003; Zileli et al., 2024; Kim et al., 2021).

1. Diabetes mellitus (DM) – Patients suffering from DM (type I or II) have a statistically higher chance of developing lumbar disc reherniation (LDR). The current hypothesis is that the imbalance in the

anabolic and the catabolic metabolisms, caused by DM, leads to accelerated degeneration of the intervertebral disc tissue (Kim et al., 2021; Kakadiya et al., 2022; Park et al., 2021; Park et al., 2021, 2021; Mobbs et al., 2001). For the purpose of this study, the patients were separated into two groups – healthy individuals and diabetics (with type I or II DM). A binary coding system was used – diabetics were labeled as “1” and healthy individuals as “0”.
2. Smoking – Multiple cohort and case-control studies have demonstrated the positive association between smoking and LDH and LDR. Nicotine disrupts the production of the extracellular matrix and cell division. The chronic hypoxemia and the increased levels of carbon monoxide and dioxide in the bloodstream of smokers lead to local ischemic changes in the intervertebral discs (Kim et al., 2021; Huang et al., 2016b). Patients were separated again in two groups – nonsmokers and smokers (including users of all types of smokable tobacco products). A binary coding system was used – smokers were labeled as “1” and nonsmokers were labeled as “0”.
3. Same level intervertebral joint degeneration – Facet joint osteoarthritis (FJOA) is an independent factor associated with both LDH and LDR (Lumbar Disc Reherniation). The severity of FJOA correlates with the increase of the risk of LDH and LDR (Kim et al., 2021; Zhu et al., 2020; Fujiwara et al., 1999; Jentzsch et al., 2013). We used the FJOA CT-based grading system proposed by Pathria (Pathria et al., 1987). During the data evaluation, our team decided to change the initial values in the scoring system from 0-1-2-3 to 1-2-3-4, accordingly. Integer values were used for the incorporation of this parameter in the model.
4. Disc height decreased by 20% or more – Intervertebral disc height reduction is a clear sign of disc degeneration and annular ring degradation. These changes lead to improper distribution of the mechanical stress over the disc. We used data from the morphometric analysis published by K. Bach et al. to set the normal values for height of the lumbar intervertebral discs. We then calculated new “benchmark” values, for both males and females, that were 80% of the original averaged disc height for every level (Kim et al., 2021; Pathria et al., 1987; Shepard and Cho, 2019). The values we used are shown in Table 2. Disc heights were measured using a sagittal plane, T2 sequence MRI. All measurements were taken in midline view. The disc height was recorded as the average of three measurements – one in each portion of the disc – ventral, dorsal and middle. A binary coding system was used and patients with normal disc height or reduction of less than 20% were put in the “0” category while those with 20% reduction or more were labeled as “1”.
5. Disc extrusion or sequestration – Previous studies have described a strong connection between reherniation rates and disc extrusion. Severe annulus fibrosus degeneration is associated with nucleus pulposus herniation, extrusion and sequestration. To define these terms, we used The Lumbar disc nomenclature V. 2.0 (Huang et al., 2016a; Bach et al., 2018; Shin, 2014). Axial and sagittal window MRIs were used for the assessment of the level of lumbar disc degeneration. A binary coding system was used - patients without extrusion or sequestration were labeled as “0”, while those with extrusion or sequestration were labeled as “1” (Fardon et al., 2014).

The statistical significance of each contributing factor was identified by the algorithm. Integer score points were assigned accordingly: disc extrusion = 5; disc height reduction of 20% or more = 2; smoking = 1; diabetes = 1, facet joint degeneration (depending on severity) = 1 to 4. Summation of all points assigned to a patient gives us the Lumbar Reherniation Score. A patient scoring 0 points has 0% predicted risk for reherniation. Given the 10% threshold of the model, every score over 9 points is associated with elevated risk for reherniation. A patient scoring the maximum 13 points has predicted risk for same level reherniation of 88.1 %. The statistical risk associated with every LRS value in the range 0–13 is listed in Table 4.

The LRS was created as a tool for preoperative identification of patients with higher than the average risk for same level reherniation following lumbar disc surgery. Performing single step microdiscectomy and fusion in these individuals would reduce the reoperation rates thus decreasing the rates of intra and postoperative complications associated with lumbar disc herniation (Bombieri et al., 2022).

5. Conclusions

The Lumbar Reherniation Score is a proposed novel tool for preoperative patient assessment. It incorporates machine learning algorithms into the field of neurosurgery, aiding the clinical decision-making process. The LRS is the product of a well-structured and robust ML-based analysis performed in a single surgical center. Our predictive score aims to provide the surgeons with the ability to reliably inform the LDH patients about the individual reherniation risk following microdiscectomy without fusion. We hypothesize that a single stage discectomy and fusion could be used for patients with high LRS, thus reducing the reoperation rates and improving the patient's quality of life. As this is a pilot project, prospective validation of the model is needed, before it can be recognized as a global recurrence risk tool.

Relationships

There are no additional relationships to disclose.

Patents and intellectual property

There are no patents to disclose.

Other activities

There are no additional activities to disclose.

Research support

This research received no external financial or non-financial support.

References

Alzubaidi, L., Zhang, J., Humaidi, A.J., et al., 2021. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions. J Big Data 8 (1), 53. https://doi.org/10.1186/s40537-021-00444-8.

Bach, K., Ford, J., Foley, R., Januszewski, J., Murtagh, R., Decker, S., Uribe, J.S., 2018. Morphometric analysis of lumbar intervertebral disc height: an imaging study. World Neurosurg 20 (18), S1878–S8750. https://doi.org/10.1016/j.wneu.2018.12.014, 32836-5.Epub ahead of print. PMID: 30579030.

Blamoutier, A., 2013. Surgical discectomy for lumbar disc herniation: surgical techniques. J. Orthop. Traumatol.: Surgery & Research 99 (1), S187–S196. https://doi.org/10.1016/j.otsr.2012.11.005.

Bombieri, F.F., Shafafy, R., Elsayed, S., 2022. Complications associated with lumbar discectomy surgical techniques: a systematic review. J Spine Surg 8 (3), 377–389. https://doi.org/10.21037/jss-21-59. PMID: 36285095; PMCID: PMC9547702.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Chollet F and others. Keras, 2015. https://keras.io.

Fardon, D.F., Williams, A.L., Dohring, E.J., Murtagh, F.R., Gabriel Rothman, S.L., Sze, G. K., 2014. Lumbar disc nomenclature: version 2.0: recommendations of the combined task forces of the north American spine society, the American society of spine radiology and the American society of neuroradiology. Spine J. 14 (11), 2525–2545. https://doi.org/10.1016/j.spinee.2014.04.022. Epub 2014 Apr 24. PMID: 24768732.

Fawcett, Tom, 2006. An introduction to ROC analysis. Pattern Recogn. Lett. 27 (8), 861–874. https://doi.org/10.1016/j.patrec.2005.10.010. ISSN 0167-8655.

Flach, Peter, Hernandez-Orallo, Jose, Ferrari, Cesar, 2011. A coherent interpretation of AUC as a measure of aggregated classification performance. AppearinginProceedingsofthe28thInternationalConferenceon MachineLearning, Bellevue,WA,USA. Copyright2011by theauthor(s)/owner(s).

Fujiwara, A., Tamai, K., Yamato, M., An, H.S., Yoshida, H., Saotome, K., Kurihashi, A., 1999. The relationship between facet joint osteoarthritis and disc degeneration of the lumbar spine: an MRI study. Eur. Spine J. 8 (5), 396–401. https://doi.org/10.1007/s005860050193. PMID: 10552323; PMCID: PMC3611192.

Hajian-Tilaki, K., 2013. Receiver operating characteristic (ROC) curve analysis for medical diagnostic test evaluation. Caspian J Intern Med. 4 (2), 627–635. http s://www.ncbi.nlm.nih.gov/pmc/articles/PMC3755824/. (Accessed 22 September 2023).

Hanczar, Blaise, Hua, Jianping, Sima, Chao, Weinstein, John, Bittner, Michael, Dougherty, Edward R., 2010. Small-sample precision of ROC-related estimates. Bioinformatics 26 (6), 822–830. https://doi.org/10.1093/bioinformatics/btq037.

Hand, D.J., 2009. Measuring classifier performance: a coherent alternative to the area under the ROC curve. Mach. Learn. 77, 103–123. https://doi.org/10.1007/s10994-009-5119-5.

Harris, C.R., Millman, K.J., van der Walt, S.J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M.H., Brett, M., Haldane, A., Del Río, J.F., Wiebe, M., Peterson, P., Gérard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., Oliphant, T.E., 2020. Array programming with NumPy. Nature 585 (7825), 357–362. https://doi.org/10.1038/s41586-020-2649-2. Epub 2020 Sep 16. PMID: 32939066; PMCID: PMC7759461.

Hernández-Orallo, J., Flach, P., Ferri, C., 2012. A unified view of performance metrics: translating threshold choice into expected classification loss. J. Mach. Learn. Res. 13 (91), 2813–2869. http://jmlr.org/papers/v13/hernandez-orallo12a.html. (Accessed 22 September 2023).

Hicks, S.A., Strümke, I., Thambawita, V., et al., 2022. On evaluation metrics for medical applications of artificial intelligence. Sci. Rep. 12 (1), 5979. https://doi.org/10.1038/s41598-022-09954-8. Published 2022 Apr 8.

Huang, W., Han, Z., Liu, J., Yu, L., Yu, X., 2016a. Risk factors for recurrent lumbar disc herniation: a systematic review and meta-analysis. Medicine (Baltim.) 95, e2378.

Huang, W., Qian, Y., Zheng, K., Yu, L., Yu, X., 2016b. Is smoking a risk factor for lumbar disc herniation? Eur. Spine J. 25 (1), 168–176. https://doi.org/10.1007/s00586-015-4103-y. Epub 2015 Jul 10. PMID: 26160690.

Jentzsch, T., Geiger, J., Zimmermann, S.M., Slankamenac, K., Nguyen-Kim, T.D.L., Werner, C.M.L., 2013. Lumbar facet joint arthritis is associated with more coronal orientation of the facet joints at the upper lumbar spine. Radiology Research and Practice 2013, 1–9. https://doi.org/10.1155/2013/693971.

Kakadiya, G., Gandbhir, V., Soni, Y., Gohil, K., Shakya, A., 2022. Diabetes mellitus-A risk factor for the development of lumbar disc degeneration: a retrospective study of an Indian population. Global Spine J. 12 (2), 215–220. https://doi.org/10.1177/2192568220948035. Epub 2020 Sep 23. PMID: 32964735; PMCID: PMC8907643.

Kang, S.H., Yang, J.S., Cho, S.S., Cho, Y.J., Jeon, J.P., Choi, H.J., 2020. A prospective observational study of return to work after single level lumbar discectomy. J Korean Neurosurg Soc 63 (6), 806–813. https://doi.org/10.3340/jkns.2020.0227. Epub 2020 Nov 1. PMID: 33181867; PMCID: PMC7671783.

Khor, S., Lavallee, D., Cizik, A.M., et al., 2018. Development and validation of a prediction model for pain and functional outcomes after lumbar spine surgery. JAMA Surg 153 (7), 634–642. https://doi.org/10.1001/jamasurg.2018.0072.

Kim, Yong Guk, Yang, Joo Chul, Wan Kim, Tae, 2021. Risk factors for recurrence of disc herniation after single- level lumbar discectomy. Asian Journal of Pain 7, 1. https://doi.org/10.35353/ajp.2021.00003.

Krstajic, D., Buturovic, L.J., Leahy, D.E., Thomas, S., 2014. Cross-validation pitfalls when selecting and assessing regression and classification models. J Cheminform 6 (1), 10. https://doi.org/10.1186/1758-2946-6-10. Published 2014 Mar 29.

Little, M.A., Varoquaux, G., Saeb, S., et al., 2017a. Using and understanding cross-validation strategies. Perspectives on Saeb et al. GigaScience 6 (5), 1–6. https://doi.org/10.1093/gigascience/gix020.

Little, M.A., Varoquaux, G., Saeb, S., Lonini, L., Jayaraman, A., Mohr, D.C., Kording, K.P., 2017b. Using and understanding cross-validation strategies. Perspectives on Saeb et al. GigaScience 6 (5), 1–6. https://doi.org/10.1093/gigascience/gix020. PMID: 28327989; PMCID: PMC5441396.

Mariscal, G., Torres, E., Barrios, C., 2022. Incidence of recurrent lumbar disc herniation: a narrative review. J Craniovertebr Junction Spine 13 (2), 110–113. https://doi.org/10.4103/jcvjs.jcvjs_38_22. Epub 2022 Jun 13. PMID: 35837428; PMCID: PMC9274669.

Mobbs, R.J., Fau, N.R., Chandran, K.N., 2001. Lumbar discectomy and the diabetic patient: incidence and outcome. J. Clin. Neurosci. 8, 10–13.

Nahm, F.S., 2022. Receiver operating characteristic curve: overview and practical use for clinicians. Korean J Anesthesiol 75 (1), 25–36. https://doi.org/10.4097/kja.21209.

Park, C.H., Min, K.B., Min, J.Y., et al., 2021. Strong association of type 2 diabetes with degenerative lumbar spine disorders. Sci. Rep. 11, 16472. https://doi.org/10.1038/s41598-021-95626-y.

Pathria, M., Sartoris, D.J., Resnick, D., 1987. Osteoarthritis of the facet joints: accuracy of oblique radiographic assessment. Radiology 164 (1), 227–230.

Pedregosa, F., Varoquaux, G., Gramfort, A., et al., 2011. Scikit-learn: machine learning in python. J. Mach. Learn. Res. 12, 2825–2830. https://www.jmlr.org/papers/volume12/pedregosa11a/pedregosa11a.pdf.

Roiha, M., Marjamaa, J., Siironen, J., et al., 2023. Favorable long-term health-related quality of life after surgery for lumbar disc herniation in young adult patients. Acta Neurochir. 165, 797–805. https://doi.org/10.1007/s00701-023-05522-9.

Saito, T., Rehmsmeier, M., 2015. The precision-recall plot is more informative than the ROC plot when evaluating binary classifiers on imbalanced datasets. PLoS One 10 (3), e0118432. https://doi.org/10.1371/journal.pone.0118432. Published 2015 Mar 4.

Schonfeld, E., Mordekai, N., Berg, A., Johnstone, T., Shah, A., Shah, V., Haider, G., Marianayagam, N.J., Veeravagu, A., 2024. Machine learning in neurosurgery: toward complex inputs, actionable predictions, and generalizable translations. Cureus 16 (1), e51963. https://doi.org/10.7759/cureus.51963. PMID: 38333513; PMCID: PMC10851045.

Shepard, N., Cho, W., 2019. Recurrent lumbar disc herniation: a review. Global Spine J. 9 (2), 202–209. https://doi.org/10.1177/2192568217745063. Epub 2017 Dec 18. PMID: 30984501; PMCID: PMC6448208.

Shin, B.J., 2014. Risk factors for recurrent lumbar disc herniations. Asian Spine J 8 (2), 211–215. https://doi.org/10.4184/asj.2014.8.2.211. Epub 2014 Apr 8. PMID: 24761206; PMCID: PMC3996348.

Swartz, K.R., Trost, G.R., 2003. Recurrent lumbar disc herniation. Neurosurg. Focus 15 (3), 1–4. https://doi.org/10.3171/foc.2003.15.3.10.

Tougui, I., Jilbab, A., Mhamdi, J.E., 2021. Impact of the choice of cross-validation techniques on the results of machine learning-based diagnostic applications. Healthc Inform Res 27 (3), 189–199. https://doi.org/10.4258/hir.2021.27.3.189.

Ustun, B., Rudin, C., 2019. Learning optimized risk scores. J. Mach. Learn. Res. 20, 1–75.

Vialle, L.R., Vialle, E.N., Suárez Henao, J.E., Giraldo, G., 2010. Lumbar disc herniation. Revista Brasileira de Ortopedia (English Edition) 45 (1), 17–22. https://doi.org/10.1016/s2255-4971(15)30211-1.

Welch, B.L., 1947. The generalization of 'Student's' problem when several different population variances are involved. Biometrika 34 (1/2), 28–35. https://doi.org/10.2307/2332510.

West, R.M., 2021. Best practice in statistics: use the Welch t-test when testing the difference between two groups. Ann. Clin. Biochem. 58 (4), 267–269. https://doi.org/10.1177/0004563221992088.

Yoon, W.W., Koch, J., 2021. Herniated discs: when is surgery necessary? EFORT Open Rev 6 (6), 526–530. https://doi.org/10.1302/2058-5241.6.210020. PMID: 34267943; PMCID: PMC8246101.

Zhu, K., Su, Q., Chen, T., Zhang, J., Yang, M., Pan, J., Wan, W., Zhang, A., Tan, J., 2020. Association between lumbar disc herniation and facet joint osteoarthritis. BMC Musculoskelet Disord 21 (1), 56. https://doi.org/10.1186/s12891-020-3070-6. PMID: 31996194; PMCID: PMC6990568.

Zileli, Mehmet, Oertel, Joachim, Sharif, Salman, Zygourakis, Corinna, 2024. Lumbar disc herniation: prevention and treatment of recurrence: WFNS spine committee recommendations. World Neurosurgery: X 22, 100275. https://doi.org/10.1016/j.wnsx.2024.100275. ISSN 2590-1397.