Journal of Health Informatics

J Health Inform. 2026;18:e1856
DOI: https://doi.org/10.59681/2175-4411.v18.iEspecial.2026.1856

ARTIGO COMPLETO
CBIS 2026

A geometrical exploration of vertebral column conditions

Uma exploração geométrica das condições da coluna vertebral

Felipe Oliveira Santos[1], Solange da Silva[1], Juliana Paula Felix[2] ✉

ORCID: 0009-0005-1708-9587 · 0009-0006-7646-4196 · 0000-0003-4095-1639

[1] Pontifícia Universidade Católica de Goiás, Goiânia (GO), Brasil.

[2] Universidade Federal de Goiás, Goiânia (GO), Brasil.

ARTICLE INFORMATION

TYPE
Artigo completo · CBIS 2026

IDENTIFIER
e1856

DOI
10.59681/2175-4411.v18.iEspecial.2026.1856

RECEIVED
17 May 2026

ACCEPTED
20 July 2026

REVIEW PROCESS
Double Blind Review

HANDLING EDITOR
Juliano de Souza Gaspar

ACCESS
Open, CC BY 4.0

CORRESPONDING AUTHOR
Juliana Paula Felix
julianafelix@ufg.br

LICENSE

This is an open access article published under the Creative Commons CC BY 4.0 license, which permits use, distribution, and reproduction in any medium, provided the original work is properly cited.

CC BY 4.0

ABSTRACT

Objective: To propose a geometric exploration of three vertebral column conditions: normality, disc herniation, and spondylolisthesis, using interpretable machine learning to understand how they organize in the attribute space and to complement the performance-focused literature. **Method:** We utilized a public dataset with 310 patients, unevenly distributed in the classes with six biomechanical features. Three models were employed alongside exploratory principal component analysis (PCA), with performance metrics, odds ratio and feature importance as interpretative tools. **Results:** The analysis reveals a strong biomechanical axis associated with spinopelvic alignment that separates spondylolisthesis into a distinct region, resembling a peninsula, confirmed by both logistic regression and random forest models. Moreover, the overlap between the remaining classes persisted across all three models, with logistic regression outperforming the others, suggesting an inherent data property. **Conclusion:** The interpretable model reveals a meaningful clinical structure, supporting clinical decision-making in digital health where explicability is essential for model reliability.

Keywords: Machine Learning; Intervertebral Disc Displacement; Spondylolisthesis.

RESUMO

Objetivo: Propor a análise geométrica de três condições da coluna vertebral: normalidade, hérnia de disco e espondilolistese, utilizando aprendizado de máquina interpretável para compreender sua organização no espaço de atributos e complementar a literatura focada em performance. **Método:** Utilizamos um dataset público com 310 pacientes, distribuídos desigualmente entre as classes com seis medidas biomecânicas. Três modelos foram empregados com análise de componentes principais (PCA) exploratória, extraindo métricas de performance, odds ratio e importância de atributos como ferramentas interpretativas. **Resultados:** A análise revelou um eixo biomecânico associado ao alinhamento espinopélvico que separa a espondilolistese em região distinta, como uma península, confirmado pela regressão logística e floresta aleatória. A sobreposição entre as demais classes persistiu nos três modelos, com regressão logística superando os demais, indicando uma propriedade inerente dos dados. **Conclusão:** O modelo interpretável revela uma estrutura clínica relevante, corroborando o processo decisório clínico em Saúde Digital, onde a explicabilidade é essencial para a confiabilidade no modelo.

Descritores: Aprendizado de Máquina; Deslocamento do Disco Intervertebral; Espondilolistese.

CBIS 2026 XXI CONGRESSO BRASILEIRO DE INFORMÁTICA EM SAÚDE

Paper presented at the 21st Brazilian Congress of Health Informatics (CBIS 2026), held from September 23 to 25, 2026, in Brasília (DF), Brazil, and published in the CBIS 2026 section of the Journal of Health Informatics.

HOW TO CITE THIS ARTICLE: Santos FO, da Silva S, Felix JP. A geometrical exploration of vertebral column conditions. J Health Inform 2026;18:e1856. doi: https://doi.org/10.59681/2175-4411.v18.iEspecial.2026.1856

1 INTRODUCTION

Lumbar pain, including its chronic form, has been a public health issue for decades. It is estimated that between 70% and 85% of all people experience lumbar pain at some point in their lives [1]. In Brazil, lumbar pain generates significant demand on healthcare services, yet no methodologically robust estimate of national prevalence is currently available, as existing studies present a high risk of bias [2].

Frequently, lumbar pain is associated with two conditions of the vertebral column: disc herniation, which affects 1% to 3% of the population annually [3], and spondylolisthesis, which affects 5.4% of adults and 2.6% of children under six years of age [4]. Disc herniation can be broadly defined as "a localized or focal displacement of disc material beyond the limits of the intervertebral disc space" [3]. Spondylolisthesis, in turn, is understood "as a translation of one vertebra over another in the anterior or posterior direction" [4].

The differential diagnosis of vertebral column conditions is usually challenging, since, even though they can present distinct physiopathological mechanisms, the symptoms may overlap, and disc herniation has been overdiagnosed in patients with spondylolisthesis [5]. Tied to that, there is the issue that diagnostic confirmation is dependent on imaging exams that are expensive, such as magnetic resonance imaging [4], making it often inaccessible, thus making the use of machine learning techniques in diagnostics very attractive, as it can potentially reduce costs while maintaining high accuracy.

However, many studies concentrate on developing algorithms that focus on disease classification, without emphasizing, therefore, a geometric approach to the issue. Usually, they do not investigate how these conditions are organized in the attribute space, which could contribute to understanding the nature of the condition and also an analysis of the separability of the classes, i.e., if the groups are well-defined or if there is an overlap between them. For instance, one could question whether disc herniation and spondylolisthesis present completely distinct morphological patterns or whether they share nearby regions in the attribute space, and what would be considered a "normal" state.

This geometric approach requires an interpretable model, which is consistent with the growing demand for ethical and responsible artificial intelligence (AI) in digital health, particularly in supporting clinical decision-making. Therefore, we explore these vertebral column conditions through machine learning techniques, such as data visualization, dimensionality reduction through Principal Component Analysis (PCA), and three machine learning models: logistic regression (LR), random forest (RF) and support vector machine (SVM). We aim to understand whether the classes form distinct groups in the attribute space and what this reveals about the conditions. By employing these three models, and focusing on logistic regression, results suggest that the spinopelvic alignment variables play a central role in structuring the dataset and separating spondylolisthesis from other conditions while the distinction between herniation and normal individuals remains less clear due to how disc herniation affects soft tissue rather than pelvic alignment.

2 METHODS

This section presents the definitions necessary to understand the clinical conditions studied in this paper. Moreover, it details the dataset composition, the preprocessing steps taken, the classification models employed and their justification, and the validation and metrics employed.

2.1 Clinical Definitions of the Studied Conditions

Among the classifications used by the dataset, it is important to understand, in light of the specific literature, what their definitions and characteristics are, since the author did not specify in the dataset or a related article.

Here, the "normal" class comprises patients without disc herniation or spondylolisthesis. Furthermore, following the definition of Fardon *et al.* [6], these patients are understood as having morphologically normal intervertebral discs, that is, without structural displacement or pathological deformation.

Disc herniation is considered here as a condition in which "the nucleus pulposus becomes displaced through the annulus fibrosus, usually posteriorly, causing narrowing, or stenosis, of the central canal or neuroforamina, leading to mechanical compression of the thecal sac or associated nerve roots" [3]. It is important to highlight that this condition is often idiopathic, i.e., it is not commonly associated with any other disease or known cause. The only exception known is that it can occur due to trauma or physiological disc degeneration from the natural aging process.

Lastly, according to Tebet [4], spondylolisthesis is considered to be a condition difficult to understand for vertebral column specialists as it can present in a myriad of anatomical and clinical forms. Given that it is a condition involving the translation of a portion of the column over an adjacent one, the displaced part shifts the torso, potentially leading to severe clinical consequences. Much like disc herniation, its etiology is not perfectly clear, which means its real causes, pathogenesis and development are not well established, only that it is a multifactorial condition.

Table 1. Biomechanical features of the dataset.

Feature	Description	Unit	Reference
Pelvic Incidence (PI)	Angle between sacral plate and femoral axis	Degrees	[8]
Pelvic Tilt (PT)	Angle between vertical and femoral head axis	Degrees	[8]
Lumbar Lordosis Angle (LL)	Cobb angle between L1 and L5	Degrees	[9]
Sacral Slope (SS)	Angle between sacral plate and horizontal	Degrees	[8]
Pelvic Radius (PR)	Distance between femoral head and S1	mm	[10]
Degree of Spondylolisthesis (DS)	Slippage of one vertebra over another	-	[4]

Source: Prepared by the authors (2026).

2.2 The Dataset

This study uses exclusively a public, anonymized dataset, openly available at the University of California, Irvine Machine Learning Repository [7]. Because the study relies on secondary, publicly available data with no direct interaction with human participants or collection of sensitive personal information, submission to a Research Ethics Committee (CEP) was not required, in accordance with Resolution No. 510, of April 7, 2016.

The data was collected by Dr. Henrique da Mota, and the dataset was constructed and made available by Guilherme Barreto and Ajalmar Neto. Each of the 310 patients "is represented by six biomechanical attributes derived from the shape and orientation of the pelvis and lumbar spine" [7], which are presented in Table 1. The patients are distributed as follows in each class: 100 patients categorized as Normal (NO), 60 as Disc Herniation (DH) and 150 as Spondylolisthesis (SL), making it an unbalanced dataset.

It is worth noting that the measurement of degree of spondylolisthesis has no associated unit in the dataset. Therefore, we opted to treat it as dimensionless. The normalization process also makes all the units irrelevant to the models utilized.

2.3 Data Preprocessing

All experimentation was performed using Python. To ensure comparability between features, as they use different units of measurement and have different value ranges, all variables were standardized using StandardScaler from the scikit-learn library. For the exploratory visual analysis, the standardization and the PCA were applied to the entire dataset as it allows us to visualize the geometric structure of the data and interpret the principal components in a simple two-dimensional space. Therefore, the PCA is only a tool used for visual exploration and was not employed in the pipeline, especially due to the influence it would have on the interpretability of the classification model employed.

2.4 Classification Models

The logistic regression model was chosen due to its simplicity and, mostly, due to its interpretability, as it allows us to analyze the contribution of each feature through its coefficients, unlike "black-box models". Since this problem is a classification one involving three classes, a multinomial model with a softmax function was used. Additionally, to provide a more robust analysis and to understand whether the attributes had a non-linear relationship between each other, the Support Vector Machine (SVM) and the Random Forest (RF) models were also chosen and used to compare in terms of performance metrics. The Random Forest model was chosen over a Decision Tree model since it handles unbalanced datasets more effectively, reducing the high variance and the risk of overfitting inherent to decision trees.

In order to address the imbalance in the dataset, we applied the parameter "class_weight='balanced'" to all three models, which assigns higher weights to minority classes, such as the DH class. In tandem, we utilized balanced accuracy, which is the average of recall obtained on each class, as an optimization metric in the internal GridSearch.

2.5 Validation and Performance Metrics

To guarantee the statistical integrity of the results, a classifier pipeline was implemented to apply StandardScaler and call all three models using a nested stratified cross-validation with 5 folds. This ensures that the parameters are computed solely from the training data and then applied to the test fold, preventing any information from the test set from influencing the training process. Moreover, the nested cross-validation allows the optimization of hyperparameters used in each model, through GridSearch, for the problem without contaminating the test dataset and having an overfitting issue.

The following metrics were extracted for the analysis: accuracy, precision, recall, and F1-score. The overall performance, F1-score and correct classifications of all models are reported in the next section. Finally, from the RF model we extracted the mean of attribute importance, and employed odds ratio to the LR model so we could understand which attributes influenced the most in the classification of each condition, allowing for a better clinical understanding.

Table 2. Comparative performance metrics and correct classifications per class

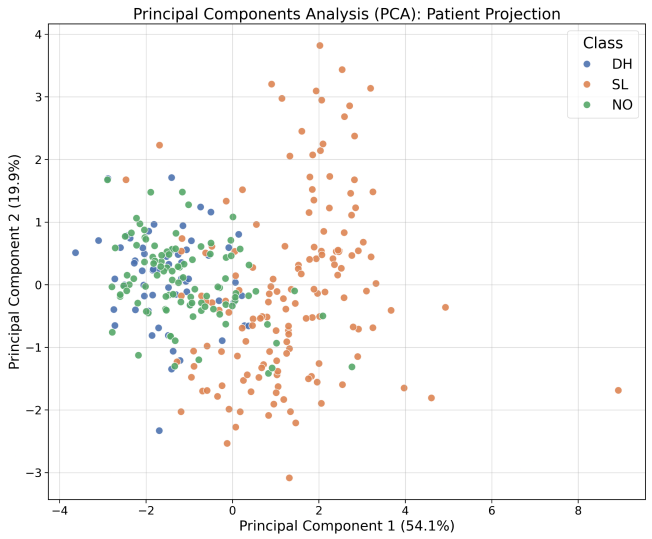
Model	Overall Performance		F1-Score per Class			Correct Classifications		
	Accuracy	SD (±)	DH	NO	SL	DH	NO	SL
LR	86.77	0.0493	0.74	0.81	0.96	47 (78.3%)	80 (80%)	141 (94%)
SVM	83.55	0.0472	0.71	0.76	0.94	47 (78.3%)	72 (72%)	137 (91.3%)
RF	85.81	0.0624	0.69	0.79	0.97	40 (66.7%)	78 (78%)	147 (98%)

Source: Prepared by the authors (2026).

3 RESULTS

Figure 1 presents the projection of patients onto the first two principal components. The SL class (in orange) is structured seemingly as a peninsula, almost isolated but still close to the region of the other two classes, which overlap in the left region, with DH and NO in blue and green, respectively. These two components account for 74% of the total variance.

Figure 1. Principal Component Analysis Graph



Source: Prepared by the authors (2026), adapted from [7].

Table 2 presents the comparison between each machine learning (ML) model, focusing on general accuracy and standard deviation (SD) from all five folds, F1-score per class and the correct classification of each model, so it is possible to evaluate how it performs for each class. The logistic regression model performed marginally better than the other two, with an 86.77% accuracy.

Table 3 shows a feature analysis, focusing on contributing to the interpretability of the conditions studied by using odds ratio on the logistic regression model and feature importance from random forest. The last column, the most associated class (Most Assoc.), relates the attribute to the class it contributes the most in the classification process. It is worth noting that the absolute value for the DS should not be interpreted literally and the high SD confirms instability across each fold due to a near-perfect separation regarding this attribute in between the SL and the other classes.

4 DISCUSSION

4.1 The Biomechanical Axis

The principal component analysis explained 74% of the total variance in the dataset, with PC1 alone accounting for 54.1%. This component is mainly influenced by four biomechanical features: (a) pelvic incidence (PI), (b) lumbar lordosis angle (LL), (c) sacral slope (SS), and (d) degree of spondylolisthesis (DS). This finding resonates with the results of the odds ratio (Table 3) for SL, in which DS is the main attribute associated with the condition, followed by SS, PI, and LL. Together, they can be interpreted as representing a biomechanical axis related to spinopelvic alignment.

This representational axis is consistent with the clinical literature, which views the vertebral column as an interdependent structure where global sagittal balance is determined by the interaction between the pelvis and lumbar spine [4]. The relationship between

Table 3. Feature-level interpretability metrics: odds ratio and feature importance

Feature	Odds Ratio (Logistic Regression)				Feature Importance (Random Forest)		Most Assoc.
	DH	NO	SL	SD (±)	Importance	SD	Class
Degree of Spondylolisthesis	0.214	0.124	301.754	236.206	38.01%	2.81	SL
Sacral Slope	0.354	1.637	1.800	0.248	17.03%	1.17	SL/NO
Pelvic Incidence	0.654	1.087	1.425	0.152	11.70%	1.18	SL
Lumbar Lordosis Angle	0.687	1.248	1.222	0.292	11.70%	1.21	SL/NO
Pelvic Radius	0.478	2.298	0.936	0.044	12.83%	1.14	NO
Pelvic Tilt	2.021	0.603	0.839	0.067	8.73%	0.85	DH

Source: Prepared by the authors (2026).

these features is not arbitrary, but rather explained by how the pelvic incidence is defined as the sum of pelvic tilt and sacral slope (PI = PT + SS), and it strongly correlates with lumbar lordosis [9]. The odds ratio for SL (Table 3) corroborates these principles, validating the geometric interpretation of the data.

4.2 Geometric Organization of the Classes

The PCA projection (Figure 1) reveals a clear separation of the SL from NO and DH, which appear largely overlapped. This indicates that spondylolisthesis has a distinct biomechanical profile, while normal and herniated patients share similar morphological characteristics or ones that cannot be easily distinguished just by these six features.

This interpretation is strongly supported by the results in Table 3, as the odds ratio for degree of spondylolisthesis is the highest among all features and while the absolute value is inflated due to near-perfect class separation, the direction is unequivocally clear: any increase in DS heightens substantially the probability of an SL classification in the LR model. This is corroborated by the 38% importance in the RF model, which has a completely distinct mechanism but points to the same feature hierarchy. This is further reflected in the F1-scores (Table 2) for all three models, as they are 0.94 (SVM), 0.96 (LR) and 0.97 (RF), with correct classification being above 90% in all three.

However, some SL patients are located within the overlapping region, creating the image of a peninsula, a region that is still related to a larger region despite being mostly isolated. This suggests the existence of a frontier zone where biomechanical profiles approximate one another. This is possibly explained by how early stages of SL have yet to substantially affect the sagittal alignment to the point it is evident the morphological profile has altered in comparison to what is expected in the normal conditional range.

4.3 Overlap between NO and DH patients

The substantial overlap between NO and DH observed in the PCA (Figure 1) is reflected in the models' performances (Table 2). LR was stable, albeit low, in classifying NO and DH with a correct classification of 80% and 78.3%, respectively. SVM was able to correctly classify the same for DH but its NO classification dropped to 72%. RF was the worst at classifying DH patients, with only 66.7%, and did not handle NO individuals any better than LR, with 78% accuracy.

This difficulty could be due to a non-linear relationship between the attributes when it comes to the differentiation of the NO and DH patients. However, both RF and SVM, models better suited to handle such a scenario, underperform compared to LR. The F1-scores for SVM's NO (0.76) and RF's DH (0.69) are the worst among them. This worsening in scores, compared to LR, argues against the hypothesis of a hidden non-linear boundary, indicating that the overlap is an inherent property of the data structure.

Therefore, the results corroborate the clinical understanding that disc herniation is harder to classify based on these parameters alone because it affects soft tissue rather than the spinopelvic alignment in a fundamental way as spondylolisthesis does. The only structural differentiator was the pelvic tilt attribute, which most likely reflects the understanding that DH patients tend to adopt a functional compensatory posture, either by retroverting or tilting the pelvis, to alleviate mechanical compression on the nerve roots.

4.4 Comparison with Other ML Studies

As Table 2 shows, the LR model performed marginally better (86.77%) than RF (85.81%) and SVM (83.55%) models. The RF model appears highly competitive by optimizing the majority class (SL 98% correct), but it does so by sacrificing the minority class (DH only 66.7%) whereas LR has a more balanced performance (SL with 94% correct classification and DH with 78.3%), making it preferable as it provides clinical insight unavailable in the other two.

It is important, moreover, to contextualize this experiment within broader literature on machine learning applications to this same dataset. Our LR model is a reflection of our prioritization of geometric understanding and interpretability over predictive optimization. Compared to studies that use the original, unmodified dataset, our models perform remarkably well. Handayani [11], Oyedotun *et al.* [12], and Unal *et al.* [13] reported accuracies of 83% (K-NN), 88.06% (neural networks), and 74–83% (SVM/K-NN/RBF-NN), respectively, using the original dataset.

Studies that reported substantially higher accuracies achieved this through considerable trade-offs: Rojas-Lopez *et al.* [14] used Adaptive Synthetic Sampling (ADASYN) to create synthetic samples to address the imbalance in the dataset, achieving 95.11% with ensemble methods, while Unal *et al.* [13] transformed the feature space via attribute weighting (MSCBAW) to exceed 99% accuracy. In both cases, this near-perfect accuracy came at the cost of interpretability, as the original biomechanical meaning of the features is diminished or lost entirely.

A more recent study, Amilo *et al.* [15], applied a fractional-order differential equation framework coupled with machine learning to the same dataset, modeling disease progression through pseudo-time and viscoelastic memory effects, which is a substantially more complex approach than ours, both computationally and mathematically. Despite this higher complexity, their correlational findings (e.g., pelvic incidence and sacral slope correlation of 0.87) converge with the biomechanical axis identified through our geometric exploration (Section 4.1), reinforcing that

this relationship is robust across substantially different methodological approaches. Notably, their binary classification, as they merged DH and SL into a single "Abnormal" class, is explicitly noted by the authors as a limitation, one that our three-class approach directly addresses.

It is also worth mentioning the distinct uses of PCA between our work and Rojas-Lopez *et al.* [14]: their analysis used PCA to identify redundant attributes affecting model performance. In contrast, we used it as an exploratory and interpretive tool, revealing the "peninsular" structure of spondylolisthesis and the overlap between DH and NO classes.

Across studies, SL is the most distinguishable class, reflecting a property of the data itself, while DH-NO separation remains difficult. Handayani [11] reports NO as the lowest-accuracy class (78%) and DH intermediate (84%); Unal *et al.* [13] report DH-NO as the hardest pairwise group (around 75%); Rojas-Lopez *et al.* [14] note the same pattern. This consistent difficulty across studies with different models supports our geometric explanation for why this boundary persists.

Therefore, this approach demonstrates that interpretability is a requirement for models and support systems when it comes to reliable clinical decisions, especially in digital health where model explicability is as important as its accuracy.

5 CONCLUSION

This paper aimed to investigate geometrically three conditions of the lumbar spine: (a) normal, (b) disc herniation, and (c) spondylolisthesis, by using a dataset with 310 patients and six biomechanical attributes. Through visualization techniques, dimensionality reduction (PCA), and machine learning models such as logistic regression, support vector machine and random forest, we sought to complement the existing performance-driven literature by understanding how these conditions are organized in the attribute space and what this organization reveals about their nature through an interpretable model.

The geometric analysis revealed a strong biomechanical axis, captured by the first principal component, which clearly separates spondylolisthesis from other classes. This axis is primarily defined by pelvic incidence, lumbar lordosis angle, sacral slope, and degree of spondylolisthesis, which are all features whose relationships are well established in the clinical literature. This is corroborated by both the odds ratio and RF feature importance. Together, these convergent methods confirm the degree of spondylolisthesis as the dominant structural separator for SL classification.

However, the PCA projection also revealed a frontier zone where some SL points lie close to or within the NO-DH overlapping region, suggesting the presence of borderline cases. They may correspond to early-stage spondylolisthesis, where the degree of slippage is still low and sagittal alignment has not yet been substantially altered. The considerable overlap between NO and DH could suggest the hypothesis of non-linear relationships between attributes but that is unlikely given that both RF and SVM underperformed in comparison to LR and had difficulty separating the two classes, as did other models used in the literature. Thus, it is more likely that the overlap is a limitation of the dataset as the attributes focus on spinopelvic alignment rather than the soft tissue that disc herniation impacts.

The main contribution of this paper lies in demonstrating how a simple, interpretable model can provide clinically meaningful explicability to the geometric organization of spinal conditions. By analyzing the extracted odds ratio, patient projection in the PCA and its loadings, and feature importance mean from the RF model, we observe a convergence of independent methods pointing towards what is already clinically known. Consequently, the geometric exploration aids in explaining why high-complexity black-box models struggle with the DH-NO boundary, validating that transparent models are sufficient and clinically robust, especially when interpreted together. These findings directly support clinical decision-making in digital health, where such transparency is mainly lost in black-box approaches.

Regarding the limitations in this study, the main ones are related to the dataset itself. First, it is a relatively small dataset, with an unbalanced distribution of patients among classes. Furthermore, the features available pertain only to spinopelvic parameters, which do not capture soft tissue alterations that could be strong predictors for disc herniation. The dataset is also cross-sectional, preventing the analysis of disease progression over time, and it lacks risk factor information that could enrich clinical interpretation of the results.

Future work could potentially use non-linear models with explicability techniques to further investigate the structure of the DH-NO overlap. Moreover, a dataset with longitudinal data would provide more information to understand the conditions' progression and the frontier zones identified. Finally, the question of what the geometry of the data can reveal about a clinical condition can be extended to other conditions, especially with temporal progression.

Author contributions. Felipe Oliveira Santos contributed to conceptualization, methodology, software, validation, visualization, and writing – original draft. Solange da Silva contributed to formal analysis, supervision, and writing – review & editing. Juliana Paula Felix contributed to conceptualization, methodology, formal analysis, supervision, and writing – review & editing. All authors read and approved the final version of the manuscript.

Funding. This research received no external funding.

Data and materials availability. This study exclusively used a publicly available, anonymized dataset obtained from the UCI Machine Learning Repository. The dataset is accessible at https://archive.ics.uci.edu/dataset/212/vertebral+column.

Conflicts of interest. The authors declare that they have no conflicts of interest.

Use of artificial intelligence. Claude by Anthropic and Google's Gemini were used in aiding the coding process and in methodological review (pipeline, structure of nested CV, and extraction and formatting of odds ratios). They were also used in text revision for clarity and coherence. All research, interpretations of results, and conclusions are the authors' responsibility. The AI tools were used to support the intellectual work and not as a means to replace authorship.

REFERENCES

1. Andersson GB. Epidemiological features of chronic low-back pain. Lancet. 1999;354(9178):581-5. doi: https://doi.org/10.1016/S0140-6736(99)01312-4
2. Nascimento PR, Costa LO. Low back pain prevalence in Brazil: a systematic review. Cad Saude Publica. 2015;31(6):1141-56. doi: https://doi.org/10.1590/0102-311X00046114
3. Zhang AS, Xu A, Ansari K, Hardacker K, Anderson G, Alsoof D, et al. Lumbar Disc Herniation: Diagnosis and Management. Am J Med. 2023;136(7):645-51. doi: https://doi.org/10.1016/j.amjmed.2023.03.024
4. Tebet MA. Conceitos atuais sobre equilíbrio sagital e classificação da espondilólise e espondilolistese. Revista Brasileira de Ortopedia. 2014;49(1):3-12. doi: https://doi.org/10.1016/j.rbo.2013.04.011
5. Semaan H, Curnutte B, Cooper M, Obri J, Elsamaloty M, Obri T, et al. Overreporting of the disc herniation in lumbar spine MRI scans performed for patients with spondylolisthesis. Acta Radiol. 2021;62(3):388-93. doi: https://doi.org/10.1177/0284185120925483
6. Fardon DF, Williams AL, Dohring EJ, Murtagh FR, Gabriel Rothman SL, Sze GK. 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. 2014;14(11):2525-45. doi: https://doi.org/10.1016/j.spinee.2014.04.022
7. Barreto G, Neto A. Vertebral Column [dataset]. 2005. UCI Machine Learning Repository. Available from: https://doi.org/10.24432/C5K89B.
8. Le Huec JC, Aunoble S, Philippe L, Nicolas P. Pelvic parameters: origin and significance. Eur Spine J. 2011;20 Suppl 5(Suppl 5):564-71. doi: https://doi.org/10.1007/s00586-011-1940-1
9. Roussouly P, Pinheiro-Franco JL. Biomechanical analysis of the spino-pelvic organization and adaptation in pathology. Eur Spine J. 2011;20 Suppl 5(Suppl 5):609-18. doi: https://doi.org/10.1007/s00586-011-1928-x
10. Sergides IG, McCombe PF, White G, Mokhtar S, Sears WR. Lumbo-pelvic lordosis and the pelvic radius technique in the assessment of spinal sagittal balance: strengths and caveats. Eur Spine J. 2011;20 Suppl 5(Suppl 5):591-601. doi: https://doi.org/10.1007/s00586-011-1926-z
11. Handayani I. Application of K-Nearest Neighbor Algorithm on Classification of Disk Hernia and Spondylolisthesis in Vertebral Column. Indonesian J. of Inf. Syst. 2019;2(1):57-66. doi: https://doi.org/10.24002/ijis.v2i1.2352
12. Oyedotun OK, Olaniyi EO, Khashman A. Disk hernia and spondylolisthesis diagnosis using biomechanical features and neural network. Technol Health Care. 2016;24(2):267-79. doi: https://doi.org/10.3233/THC-151126
13. Unal Y, Polat K, Kocer HE. Classification of vertebral column disorders and lumbar discs disease using attribute weighting algorithm with mean shift clustering. Measurement. 2016;77:278-91. doi: https://doi.org/10.1016/j.measurement.2015.09.013
14. Rojas-López AG, Rodríguez-Molina A, Uriarte-Arcia AV, Villarreal-Cervantes MG. Vertebral Column Pathology Diagnosis Using Ensemble Strategies Based on Supervised Machine Learning Techniques. Healthcare (Basel). 2024;12(13):1324. doi: https://doi.org/10.3390/healthcare12131324
15. Amilo D, Sadri K, Hincal E, Izuchukwu C, Hafez M, Farman M, et al. A Fractional-Order Machine Learning Framework for Modeling Vertebral Column Pathology and Biomechanical Dynamics. CMES. 2026;147(3):1-10. doi: https://doi.org/10.32604/cmes.2026.077921

EDITORIAL NOTES

The Journal of Health Informatics is a scientific journal maintained by the Brazilian Society of Health Informatics (Sociedade Brasileira de Informática em Saúde), a non-profit scientific association dedicated to the development of Health Informatics and Digital Health in Brazil. The editorial process is conducted with scientific independence, free from institutional, commercial, or sponsor interference, and follows international best practices in publication ethics.

This article was double-blind peer reviewed by specialists in the field. The editorial team used artificial intelligence tools, under human supervision, in the initial screening, in checking the text, references and numbers and, when needed, in redrawing figures, charts and tables, which the authors approved in the proof. The tools and safeguards are described in the journal's Policy on the use of artificial intelligence. Editorial decisions were made by people.

SBIS Sociedade Brasileira de Informática em Saúde