

ORIGINAL ARTICLE

IJPHY

Prospective Study Protocol for Developing an Artificial Intelligence Model to Diagnose Lumbar Intervertebral Disc Prolapse

¹Sandeep Pattnaik²Rajneesh Kumar^{*3}Manu Goyal

ABSTRACT

Background: Lumbar prolapsed intervertebral disc (PIVD) is a common musculoskeletal disorder that is routinely diagnosed clinically and confirmed by magnetic resonance imaging (MRI). Although there have been various artificial intelligence (AI) models studied as diagnostic options, they are mostly nonspecific for lumbar PIVD and do not incorporate clinical and physical examination parameters, which may contribute to misdiagnosis. This planned study aims to develop an AI-driven diagnosis tool using clinical and physical parameters and to compare its findings with MRI-based diagnoses.

Methods: The research involves two concurrent phases. Phase one includes finding applicable clinical and physical diagnostic parameters from interviews with patients, literature, and expert questionnaires. These parameters are to be used to construct a dataset from chosen participants for developing the AI model. In phase two, there will be a prospective cross-sectional study comparing the diagnostic output of the AI tool with MRI results. The performance of the AI model will be measured using standard metrics and iteratively optimized according to the results.

Results: Phase one started in October 2023 to determine diagnostic parameters. AI development data collection started in February 2024 and took 12–14 months. The future cross-sectional validation phase will follow the development of the AI model. Final study findings are expected by mid-2025.

Conclusion: Upon the development of an efficient AI tool, the model can serve as a diagnostic substitute for MRI in diagnosing lumbar PIVD, thereby improving accessibility and reducing diagnostic delays, particularly in areas with limited imaging capacity.

Keywords: Artificial Intelligence, Low Back Pain, Intervertebral Disc, Magnetic Resonance Imaging, Intervertebral Disc Displacement.

Received 23rd May 2025, accepted 27th August 2025, published 09th September 2025



www.ijphy.com

10.15621/ijphy/2025/v12i3/1861

CORRESPONDING AUTHOR

^{*3}Manu Goyal

Department of Physiotherapy, Maharishi Markandeshwar Institute of Physiotherapy and Rehabilitation, Maharishi Markandeshwar (Deemed to be University), Mullana-Ambala, Haryana, India.

Email: manu.goyal@mmumullana.org

¹Department of Physiotherapy, Maharishi Markandeshwar Institute of Physiotherapy and Rehabilitation, Maharishi Markandeshwar (Deemed to be University), Mullana-Ambala, Haryana, India.

²Department of Computer Science Engineering, Maharishi Markandeshwar Engineering College, Maharishi Markandeshwar (Deemed to be University), Mullana-Ambala, Haryana, India.



INTRODUCTION

The healthcare system has always been challenging in a developing country like India. It faces a substantial shortage of healthcare professionals compared to the World Health Organization standard, especially in the government sector and rural India. The statistics for the physiotherapists are no different [1]. This leaves paramedics and medical assistants to provide healthcare facilities, such as conducting diagnoses and prescribing interventions accordingly [2]. However, the lack of adequate knowledge to conduct such tasks in the context of these professionals has led chiefly to misdiagnosis and improper treatment. The impact of COVID-19 further challenged the Indian healthcare system. So, following the pandemic, India has turned rapidly towards digitizing healthcare [3]. One step towards achieving this is using artificial intelligence (AI) in different healthcare sectors.

AI has been the state-of-the-art technology used in the healthcare sector, mainly for the last one or two decades [4]. It has evolved a lot in diagnosing different conditions. These conditions also include some musculoskeletal disorders like lateral epicondylitis [5], peri arthritis shoulder [6], and primarily low back pain (LBP) [7].

Of all the musculoskeletal conditions responsible for LBP, physiotherapists and physicians most commonly encounter the lumbar prolapsed intervertebral disc (PIVD), otherwise known as the herniated disc, during their clinical practice. PIVD can be defined as a posteriorly concentrated displacement of the intervertebral disc content, comprising the inner gel-like nucleus pulposus and the outer fibrous annulus fibrosus, beyond its natural limits, as a manifestation of degenerative disc disease [8].

It is predominantly diagnosed based on the patient's demographic details, clinical history, symptoms, and physical examination results. It is further confirmed with lumbar magnetic resonance imaging (MRI), computed tomography, and conventional radiography. MRI is considered the confirmatory and benchmark diagnostic technique, with an accuracy rate as high as 97% [9]. Even with such high accuracy, there has always been a disparity between the MRI findings and the actual symptoms. MRI findings have shown false positives in the case of asymptomatic individuals [10]. Moreover, analyzing such radiological indicators requires skills that could be lacking among general healthcare practitioners [11]. Lastly, it is one of the more expensive methods of investigation and is also considered a contraindication for a few individuals. So, if clinicians are equipped with a low-cost and more accurate diagnostic method, it will boost the diagnosis process and the treatment of patients with lumbar PIVD.

This is where AI could aid, thus filling the gaps that often lead to misdiagnosis. There have been a few AI-based diagnostic tools that help correctly diagnose lumbar PIVD and other musculoskeletal conditions associated with the lumbar spine. Their accuracy in diagnosing lumbar PIVD ranged between 81.1 and 95.38%. Most of these

dealt with AI interaction with MRI findings, bypassing clinical symptoms and findings [12]. These AI-based tools thus don't address the primary drawbacks associated with the radiological diagnostic tools available for diagnosing lumbar PIVD.

Aim of the Study

An AI-based diagnosis tool should correlate with clinical findings rather than radiological findings to arrive at a proper diagnosis. This proposed study aims to develop an AI-based clinical assessment system interacting with clinical findings and parameters for diagnosing patients with lumbar PIVD and to correlate it with the MRI findings, which is considered the gold standard.

METHODS

Study Design and Procedure

This prospective study is planned to be conducted, consisting of two simultaneous phases:

- In the first phase, we will develop an AI tool to diagnose patients with lumbar PIVD based on clinical parameters. This will be achieved by following the steps described in Textbox-01.

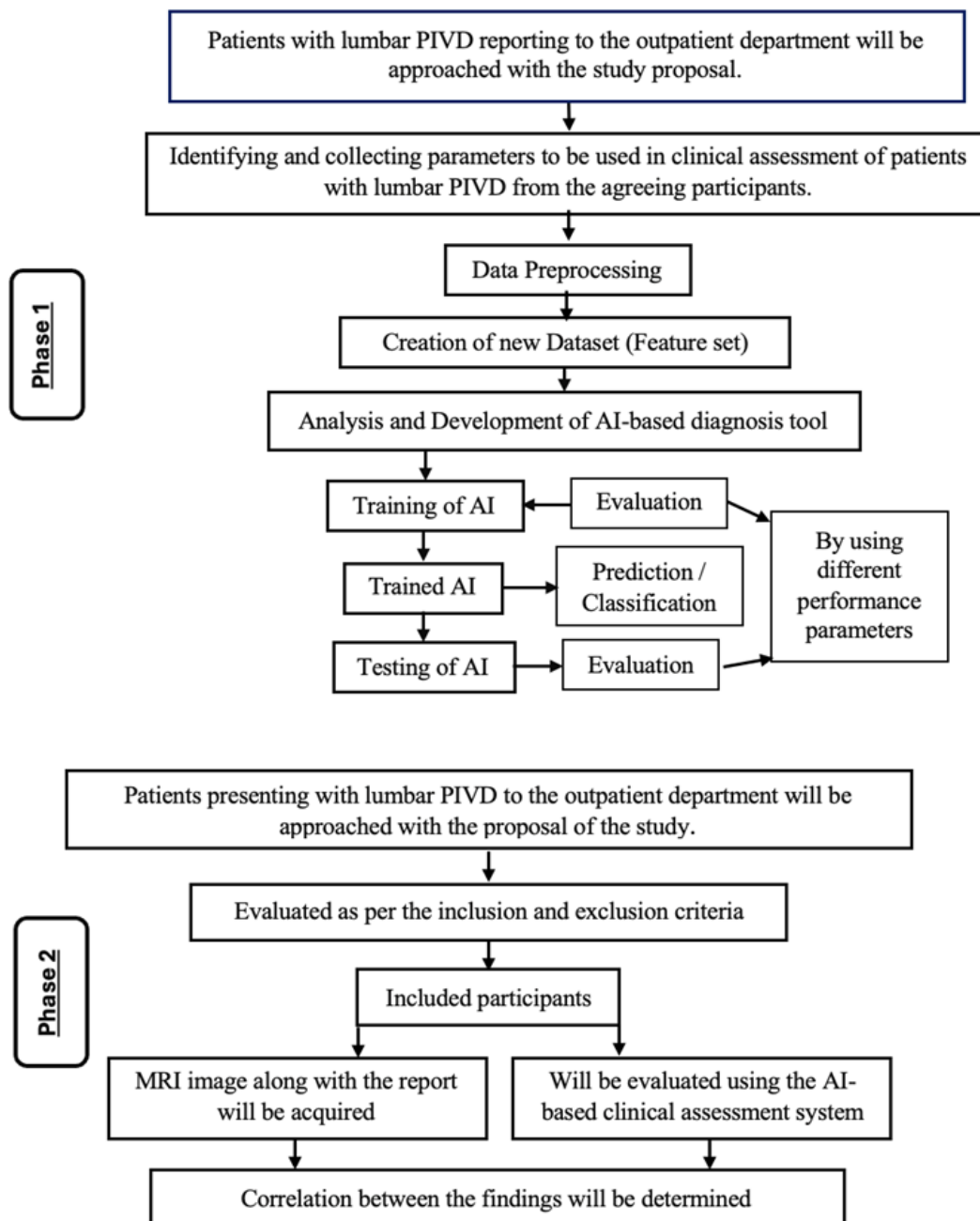
Textbox 1: Steps of the first phase.

Steps of the first phase.

1. Determination of clinical parameters responsible for diagnosing lumbar PIVD based on the patients' interview, extensive literature review, and online expert opinion survey.
2. The determined clinical parameters will be collected to form the dataset for the AI algorithm to diagnose patients with lumbar PIVD. This will include:
 - Data preprocessing
 - Data cleaning
 - Data transformation
 - Data reduction
 - Data augmentation (if needed)
3. Machine learning. The continual process has three parts: training, evaluation, and adjustments of the supervised classifier until satisfactory results regarding performance parameters are obtained.
 - Machine training with the acquired dataset and procurement of the results.
 - Evaluation and reevaluation of the results based on performance parameters.
 - They are making the necessary adjustments in the algorithm to improve it.
4. The installation of the AI tool in the workplace is to be used at once.

- The second phase will advance a prospective cross-sectional study, allowing correlation between the AI-based diagnosis and the MRI diagnosis of the lumbar PIVD. A detailed procedure is explained in Figure-01.

Figure 01: AI-based algorithm for the diagnosis of lumbar prolapsed intervertebral disc



Site, Period, and Participants

The study will be performed at the Physiotherapy Research Laboratory, Maharishi Markandeshwar Institute of Physiotherapy and Rehabilitation, and the Super Speciality Hospital, Maharishi Markandeshwar (Deemed to be University), Mullana-Ambala, Haryana, India, between 2023 and 2025.

The reference population for the prospective study will include the population reporting at the Physiotherapy Outpatient Department of the Super Specialty Hospital diagnosed with lumbar PIVD. The study will only include participants who meet the eligibility criteria, as explained in Textbox 2. They will be recruited using a consecutive sampling method.

Textbox 2: Eligibility criteria.

Inclusion Criteria

1. Both males and females
2. Age between 20 to 50 years.
3. Patient with low back pain.
4. Patient with or without radicular pain along a specific dermatome.
5. Patient with positive straight leg raise test.
6. Patient with or without neurological signs and symptoms.
7. Patient with M.R.I. of lumbar spine.

Exclusion Criteria

1. MRI showing degenerative disc disease.
2. Patient with red flags (spondylosis, spondylolisthesis, spondylodiscitis, Spinal canal stenosis, metastasis, bladder–bowel involvement, osteoporosis, trauma, spinal fractures and dislocations, etc.)

The required data for the development of the AI-based assessment tool has been acquired in 14 months, starting in February 2024. The inclusion of the participants will continue until a sturdy AI model has been created. For the prospective cross-sectional study, participant recruitment will commence once the AI tool is developed and will continue until the target sample size is achieved (approximately 2-3 months).

Ethical Consideration

Ethical approval was obtained from the Institutional Ethical Committee of Maharishi Markandeshwar Institute of Medical Sciences and Research, Haryana, India, which approved the study with project no. IEC-2659. Registration for this study has been completed with the 'Clinical Trial Registry-India' (C.T.R.I./2023/09/058094). The ethical principles of the Helsinki Declaration of 1964, reviewed by the World Health Organization in 2000, will be followed during the study [13]. All information collected will be treated confidentially. All participants will be requested to provide written consent after being informed about the study details.

Data Collection

The patients diagnosed with lumbar PIVD, reporting at the Physiotherapy Outpatient Department, and the study and its objectives will be explained. They will then be evaluated as per the eligibility criteria. Patients meeting the criteria will be given an information sheet, and informed consent will be obtained. Then, the clinical parameters will be documented in a data collection form to develop the AI tool. These collected data will then be utilized to form the new dataset, which will be further used by the researcher responsible for creating the AI-based diagnosis tool.

Similar procedures will be undertaken for the cross-sectional study. After obtaining informed consent, participants' clinical parameters and MRI findings will be obtained in the data collection form. The collected data will then be associated with a unique identifier to relate it to its corresponding MRI diagnosis and eliminate nonrandomized information. The lead researcher will then feed the collected data into the AI tool to generate a diagnostic result, which will later be correlated against MRI outcomes.

Sample Size Calculation

It will be convenient to have at least 300 samples to develop the AI-based diagnosis tool. The proposed sample has been calculated by taking into account the available literature [14, 15].

For the prospective cross-sectional study, we calculated that a minimum of 29 consecutive patients would be needed to establish the correlation. The calculation was made using G*Power 3.1.9.7, applying an alpha level of 0.05 and statistical power of 80% [16].

Data Analysis

During the development phase of the AI model, its output will be systematically compared with MRI-based

diagnoses to validate its predictive accuracy. With this validation dataset, the 'true positive (TP)', 'true negative (TN)', 'false positive (FP)', and 'false negative (FN)' data will be calculated. From them, the confusion matrix will be obtained. Additional performance metrics, such as 'sensitivity', 'specificity', 'false positive rate (FPR)', 'false negative rate (FNR)', 'accuracy', 'precision', 'recall', and the 'F1-score', will be derived from the confusion matrix. These metrics reflect the overall performance of the classification model.

The prospective cross-sectional study will be analyzed using IBM SPSS Statistics software, version 26 (IBM Corp., Armonk, NY) [17]. The Shapiro-Wilk test will be employed to evaluate the normality of the demographic and outcome variables. Descriptive statistics will be used to summarize participant demographics. According to the data type, the chi-square correlation test will be used to determine the correlation between the efficient AI-based diagnosis system for patients with lumbar PIVD and the MRI findings.

RESULTS

The three-staged procedure to determine the physical and clinical parameters responsible for the diagnosis of lumbar PIVD began in October 2023, and the compiled data was collected by January 2024. The procurement of data for the development of the AI-based tool started in February 2024. Now that the dataset is prepared, the development of the AI-based tool has started, which is estimated to last for 2 to 3 months. The inclusion of participants will begin within this period, and September 2025 is expected to be the final result of the study.

DISCUSSION

Summary

This study presents several benefits. This study, to our understanding, is the first to explore the development of an AI-based diagnosis tool for diagnosing lumbar PIVD based on physical and clinical parameters. Secondly, it will facilitate the development of a new dataset that can present better results in diagnosing lumbar PIVD compared to the previously developed AI tools. Thirdly, if it succeeds, it can serve as an effective and cost-effective alternative to conventional MRI-based diagnosis of lumbar PIVD, saving valuable time. Lastly, no similar dataset has been developed with this particular population in northern India, thus making it a unique study by itself.

Cicero M et al. (2017), Rajpurkar P et al. (2018), and Wu JT et al. (2020) have compared the findings of AI models with those of professionals based on investigative findings [18-20]. Some of them have proven the high potential of AI models as a diagnostic tool; however, they need to be studied for all pathologies and disease conditions.

Limitations

Some of the factors may hold limitations to the study's conduct. They include:

- It is a single-centered study. Due to the nature

of the study, it is confined to a geographic site. A multi-centered study would have increased the generalizability.

- The number of participants estimated was informed by findings reported in prior studies. Thus, fewer patients will be recruited to train the AI. However, a larger training sample would have been better for generalizability.

CONCLUSION

Should the study yield favourable results, then the AI model could potentially serve as a supportive diagnostic tool in future clinical workflows and be used as an alternative for the diagnosis of lumbar PIVD instead of the conventionally used investigatory findings that can reduce time and testing costs, increase diagnostic accuracy, and improve patient safety.

DATA AVAILABILITY

Due to the inclusion of confidential health data, only anonymized datasets will be shared. These can be provided to reviewers or researchers upon request, subject to ethical and institutional guidelines.

ABBREVIATION

PIVD: Prolapsed Intervertebral Disc

MRI: Magnetic Resonance Imaging

AI: Artificial Intelligence

LBP: Low Back Pain

CTRI: Clinical Trial Registry-India

TP: True Positive

TN: True Negative

FP: False Positive

FN: False Negative

FPR: False Positive Rate

FNR: False Negative Rate

CONFLICTS OF INTEREST

None declared.

DISCLAIMER

There is nothing to disclose.

SOURCE OF FUNDING

None

REFERENCES

- [1] Rao KD, Sundararaman T, Bhatnagar A, Gupta G, Kokho P, Jain K. Which doctor for primary health care? Quality of care and non-physician clinicians in India. *Soc Sci Med*. 2013 May 1;84:30–4.
- [2] Rao KD, Bhatnagar A, Berman P. So many, yet few: Human resources for health in India. *Hum Resour Health* [Internet]. 2012 Aug 13 [cited 2023 Oct 1];10(1):1–9. Available from: <https://link.springer.com/articles/10.1186/1478-4491-10-19>
- [3] Bhambere S. Rapid Digitization of Healthcare-A Review of COVID-19 Impact on our Health systems. 2021 [cited 2023 Oct 1]; Available from: <https://www.researchgate.net/publication/349581091>
- [4] Davenport T, Kalakota R. The potential for artificial intelligence in healthcare. *Future Healthc J* [Internet]. 2019 Jun [cited 2023 Oct 2];6(2):94. Available from: <https://pmc/articles/PMC6616181/>
- [5] Droppelmann G, Tello M, García N, Greene C, Jorquera C, Feijoo F. Lateral elbow tendinopathy and artificial intelligence: Binary and multilabel findings detection using machine learning algorithms. *Front Med (Lausanne)*. 2022 Sep 23;9:945698.
- [6] Yu L, Li Y, Wang XF, Zhang ZQ. Analysis of the value of artificial intelligence combined with musculoskeletal ultrasound in the differential diagnosis of pain rehabilitation of scapulohumeral periarthritis. *Medicine (United States)*. 2023 Apr 7;102(14).
- [7] D'antoni F, Russo F, Ambrosio L, Bacco L, Vollero L, Vadalà G, et al. Artificial Intelligence and Computer Aided Diagnosis in Chronic Low Back Pain: A Systematic Review. *Int J Environ Res Public Health* [Internet]. 2022 May 1 [cited 2023 Oct 2];19(10):5971. Available from: <https://www.mdpi.com/1660-4601/19/10/5971/htm>
- [8] Deyo RA, Mirza SK. Herniated Lumbar Intervertebral Disk. *New England Journal of Medicine* [Internet]. 2016 May 5 [cited 2023 Oct 2];374(18):1763–72. Available from: <https://www.nejm.org/doi/full/10.1056/NEJMcp1512658>
- [9] Kim JH, van Rijn RM, van Tulder MW, Koes BW, de Boer MR, Ginai AZ, et al. Diagnostic accuracy of diagnostic imaging for lumbar disc herniation in adults with low back pain or sciatica is unknown; a systematic review. *Chiropr Man Therap*. 2018 Dec 21;26(1):37.
- [10] Singla DrS, Sharma DrR, Sharma DrR, Singh DrA, Dhillon DrAS, Satti DrSK. Comparison between clinical finding and magnetic resonance imaging finding of lumbar prolapsed intervertebral disc. *International Journal of Orthopaedics Sciences*. 2020 Oct 1;6(4):670–4.
- [11] Mohan C. Quality program in radiology: Persue or perish. *Indian Journal of Radiology and Imaging*. 2017 Jan 27;27(01):1–3.
- [12] Pattnaik S, Goyal M. Artificial Intelligence as a Diagnostic Tool for Prolapsed Intervertebral Disc: A Literature Review. In: Shukla SK, Yadav DK, Singh R, editors. *Research in Mathematical Statistical and Computational Sciences*. Lucknow: MKSES Publication; 2023. p. 55–66.
- [13] Salako SE. The declaration of Helsinki 2000: ethical principles and the dignity of difference. *Med Law*. 2006 Jun;25(2):341–54.
- [14] MB BM, PhD DM, HENDRY D, WADDELL G. Artificial Intelligence in the Diagnosis of Low-Back Pain and Sciatica. *Spine (Phila Pa 1976)*. 1988 Feb;13(2):168–72.
- [15] Staartjes VE, Quddusi A, Klukowska AM, Schröder ML. Initial classification of low back and leg pain based on objective functional testing: a pilot study of machine learning applied to diagnostics. *European Spine Journal*. 2020 Jul 18;29(7):1702–8.
- [16] Faul F, Erdfelder E, Buchner A, Lang AG. Statistical

-
- power analyses using G*Power 3.1: Tests for correlation and regression analyses. *Behav Res Methods*. 2009 Nov;41(4):1149–60.
- [17] Hinton P, McMurray I, Brownlow C. *SPSS Explained*. Routledge; 2014.
- [18] Cicero M, Bilbily A, Colak E, Dowdell T, Gray B, Perampaladas K, et al. Training and Validating a Deep Convolutional Neural Network for Computer-Aided Detection and Classification of Abnormalities on Frontal Chest Radiographs. *Invest Radiol* [Internet]. 2017 May 1 [cited 2024 Jan 7];52(5):281–7. Available from: <https://pubmed.ncbi.nlm.nih.gov/27922974/>
- [19] Rajpurkar P, Irvin J, Ball RL, Zhu K, Yang B, Mehta H, et al. Deep learning for chest radiograph diagnosis: A retrospective comparison of the CheXNeXt algorithm to practicing radiologists. *PLoS Med* [Internet]. 2018 Nov 1 [cited 2024 Jan 7];15(11). Available from: <https://pubmed.ncbi.nlm.nih.gov/30457988/>
- [20] Wu JT, Wong KCL, Gur Y, Ansari N, Karargyris A, Sharma A, et al. Comparison of Chest Radiograph Interpretations by Artificial Intelligence Algorithm vs Radiology Residents. *JAMA Netw Open* [Internet]. 2020 Oct 1 [cited 2024 Jan 7];3(10). Available from: <https://pubmed.ncbi.nlm.nih.gov/33034642/>