

Article

Not peer-reviewed version

Validating the Accuracy of a Patient-Facing Clinical Decision Support System in Predicting Lumbar Disc Herniation: Diagnostic Accuracy Study

Fatima Badahman , [Mashaal Alsobhi](#) , Almaha Alzahrani , [Mohamed Faisal Chevidikunnan](#) , [Ziyad Neamatallah](#) , [Abdullah Alqarni](#) , [Umar Alabasi](#) , [Ahmed Abduljabbar](#) , [Reem Basuodan](#) , [Fayaz Khan](#) *

Posted Date: 17 July 2024

doi: 10.20944/preprints202407.1412.v1

Keywords: Artificial Intelligence; Machine Learning; Back Pain; Clinical Decision Support System; Lumbar; Disc Herniation.



Preprints.org is a free multidiscipline platform providing preprint service that is dedicated to making early versions of research outputs permanently available and citable. Preprints posted at Preprints.org appear in Web of Science, Crossref, Google Scholar, Scilit, Europe PMC.

Copyright: This is an open access article distributed under the Creative Commons Attribution License which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Article

Validating the Accuracy of a Patient-Facing Clinical Decision Support System in Predicting Lumbar Disc Herniation: Diagnostic Accuracy Study

Fatima Badahman ¹, Mashael Alsobhi ¹, Almaha Alzahrani ², Mohamed Faisal Chevidikunnan ¹, Ziyad Neamatallah ¹, Abdullah Alqarni ¹, Umar Alabasi ¹, Ahmed Abduljabbar ³, Reem Basuodan ⁴ and Fayaz Khan ^{1,*}

¹ Department of Physical Therapy, Faculty of Medical Rehabilitation Sciences, King Abdulaziz University, Jeddah, Saudi Arabia; fatimahassanb@gmail.com; mgalsobhi@kau.edu.sa; mfaisal@kau.edu.sa; zneamatallah@kau.edu.sa; amalqarni@kau.edu.sa; ualabasi@kau.edu.sa

² Department of Physical Therapy, King Faisal Hospital, Makkah, Saudi Arabia; aahmedalzahrani0035@stu.kau.edu.sa.

³ Department of Radiology, Faculty of Medicine, King Abdulaziz University Hospital, Jeddah, Saudi Arabia; aabduljabbar@kau.edu.sa.

⁴ Department of Rehabilitation Sciences, College of Health and Rehabilitation Sciences, Princess Nourah Bint Abdulrahman University, Riyadh, Saudi Arabia. rmbasoudan@pnu.edu.sa

* Correspondence: rkhan2@kau.edu.sa

Abstract: Background: Low back pain (LBP) is a major cause of disability globally, and the diagnosis of LBP is challenging for clinicians. Objective: This study aimed to assess the accuracy level of artificial intelligence as a Clinical Decision Support System (CDSS) called Therapha compared to MRI in predicting lumbar disc herniated patients. Methods: One hundred low back pain patients aged ≥ 18 years old were included in the study. The study was conducted in 3 stages, at first, a case series was conducted by matching MRI and Therapha diagnosis for 10 patients. Subsequently, Delphi methodology was employed to establish a clinical consensus. Lastly, to determine the accuracy of the newly developed software, a cross-sectional study was undertaken, involving 100 patients. Results: The software showed a significant diagnostic accuracy with the area under the curve in the ROC analysis determined as 0.84 with a sensitivity of 88% and a specificity of 80%. Conclusion: The study's findings revealed that CDSS using Therapha has a reasonable level of efficacy, and this can be utilized clinically to acquire faster and more accurate screening of patients with lumbar disc herniation.

Keywords: artificial intelligence; machine learning; back pain; clinical decision support system; lumbar; disc herniation

1. Introduction

Low back pain (LBP) is a highly prevalent condition globally [1], with a prevalence rate of approximately 53.5% in Saudi Arabia [2]. LBP and its related disabilities are influenced by a combination of psychological, social, and biological factors, making it a complex disorder [3]. LBP can appear in different clinical and pathognomonic manifestations such as lumbar disc herniation and radiculopathy which are the two primary factors responsible for causing LBP [4]. However, lumbar disc herniation can be the cause of different types of pain which may lead to the overlap of dermatome and myotome, resulting in misdiagnosis of nerve root involvement [5]. Therefore, numerous clinical examinations have been conducted to determine the most accurate diagnostic test for detecting the likelihood and presence of radiculopathy caused by lumbar disc herniation, and some of these examinations include straight leg raising, crossed straight leg raising, tendon reflexes, atrophy, and sensory deficits tests [6].

In addition, radiographic examination has traditionally been employed in primary healthcare settings and is broadly undertaken as a valid reference standard for identifying lumbar herniated discs [7]. Magnetic Resonance Imaging (MRI) and Computed Tomography (CT) have a considerable degree of accuracy in terms of specificity and sensitivity. The specificity of CT in detecting lumbar

herniated discs with nerve root involvement ranges from 0.7 to 0.87, while its sensitivity ranges from 0.62 to 0.9. On the other hand, MRI has a specificity ranging from 0.43 to 0.97 and a sensitivity ranging from 0.6 to 1.0 in the same diagnosis [7]. However, the diagnosis of LBP is challenging for clinicians and needs a high level of skill, experience, and understanding of the complex physiological and anatomical structure of the human body [8]. Therefore, this existing dilemma might be resolved by using CDSS to support the diagnosis of LBP and improve timely knowledge access[8].

Artificial intelligence (AI) will assist healthcare practitioners by providing reliable and expedient medical consultation services to improve the quality of treatment. It has the ability to supplement and enhance decision-making in various contexts of high importance [9–11]. CDSS can process patient information intelligently and present it to clinicians [12], or physiotherapists to support decision-making on the administration and assessment of various medical conditions such as assessing gait deviation in anterior cruciate ligament (ACL) injury [13], diagnosing and classifying peripheral neuropathy [14], predicting elderly quality of life [15], detecting coronary heart disease by using fuzzy rule-based system with multi-objective genetic algorithm [16] and helping clinician in cancer treatment [17].

In the current study, researchers conducted a diagnostic accuracy study to assess the accuracy of an AI-enabled platform and an algorithm such as CDSS. The study examined a new web-based software called Therapha, which incorporates CDSS technology. This software aims to assist clinicians in diagnosing LBP by providing them with real-time support and guidance. Therapha is a triaging software that would guide the clinician through a comprehensive patient history consisting of medical history, social history, family medical history, self-reported pain evaluation, functional deficits, psychological issues, sleep patterns, and posture. Upon concluding the history taking, it will generate the most likely differential diagnosis related to the LBP. This emerging technology has the potential to serve as a valuable instrument for medical healthcare professionals in the process of making clinical decisions [18,19]. Consequently, it has the potential to enhance both the accuracy of diagnoses and the promptness of accessing knowledge [8]. The objective of this study is to assess the diagnostic precision of AI as a CDSS in comparison to MRI for diagnosing patients with lumbar disc herniation.

2. Materials and Methods

2.1. Study Design

The study was designed in 3 stages; (i) a case series was conducted to generate hypotheses and examine the effectiveness of a novel patient-facing assessment. (ii) the Delphi method was used to validate the prediction accuracy of Therapha and evaluate the capability of the CDSS in comparison to experts' opinions. Then was followed by (iii) a diagnostic accuracy study to assess the prediction accuracy of Therapha web-software in contrast to MRI in diagnosing lumbar radiculopathy.

2.2. Participants

Male and female patients above the age of 18 years and complaining of LBP with radiating symptoms extending down to the leg. Patients who were suffering from any other LBP pathology, history of cancer, previous spine surgery and failed back syndrome were excluded. Before obtaining the patient's clinical and personal background, all participants provided their signature on a written informed consent document. The ethical approval was obtained from the research ethics committee at King Abdulaziz University Hospital (KAUH) (HA-02-J-008), Jeddah, Saudi Arabia.

2.3. Case Series

The current study started with a preliminary case series to determine Therapha web software's accuracy in diagnosing patients with herniation of lumbar disc. This study comprised 10 patients with LBP who met the inclusion criteria. All patients underwent an MRI, and then patients' history was collected by a blind assessor using Therapha web software on the same day before the MRI procedure. Therapha analyzed patient data and generated the most likely differential diagnosis of the LBP. The study assistant with 7 years' experience working in the field of radiology, matched the

software’s final report to the MRI report and tabulated the data (Therapha and MRI final diagnosis) for each patient. (Figure 1)

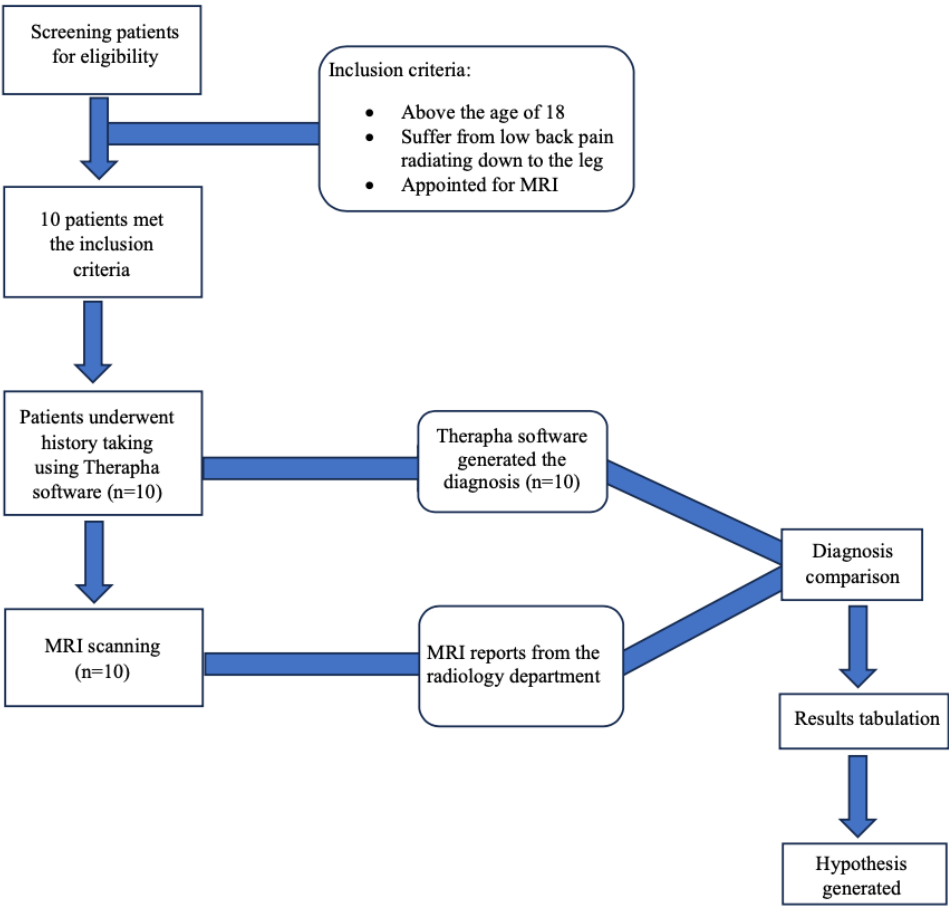


Figure 1. Design of case series to generate the study hypothesis.

2.4. Delphi Method

The Delphi method has been used in various medical and healthcare settings [20,21] to determine expert consensus in situations where there is insufficient evidence available [22]. The Delphi method is an effective process that uses multiple rounds of expert voting, in which a questionnaire is distributed to a group of experts who respond anonymously to the questions. The survey results are then arranged in a tabular form and reported back to the experts, and each panel is requested to complete the questionnaire once again [21]. (Figure 2)

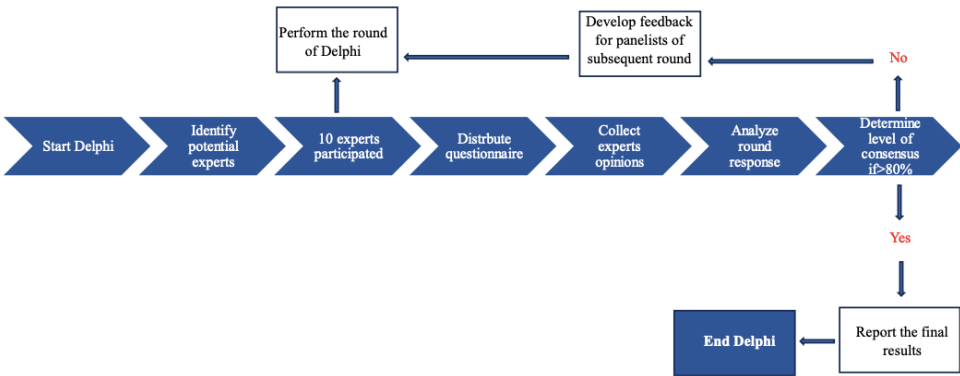


Figure 2. Delphi study steps for Therapha software validation on patients with low back pain.

2.4.1. Panel Selection: A total of ten experts have participated in this study to validate Therapha software. The panel was chosen based on their research and clinical experience in assessing and treating patients with LBP. Experts were from multiple countries, including Saudi Arabia, India, the United Arab Emirates, the United Kingdom, and the United States.

2.4.2. Round 1: In the first round, all experts were emailed a set of statements, instructions for rating, and an explanation of the study aims. Experts were asked to read patient clinical history summaries that were automated by the Therapha software from subjective reporting and were asked to give a final diagnosis for each patient. Then, a rating was made on the expert’s final diagnosis (agree/disagree) with Therapha diagnosis on a scale of 1 to 9, where a score of 1 to 3 score was classified as “inappropriate”, 4 to 6 as “uncertain”, and those in the range of 7 to 9 as “appropriate” [21]. Panelists were asked to justify their vote if they chose to vote for “disagreement”.

Therapha prediction was consolidated with an 80% agreement among the experts (8 out of 10 panellists) to accept or remove a statement during the construction of the final diagnosis. Based on Lynn’s et. al. findings, 80% was determined as an adequate cut-off [23]. Statements that did not receive 80% agreement were amended based on feedback from the expert panel and circulated to panellists in the second round.

2.5. Diagnostic Accuracy

2.5.1. Sample Size: The sample size was calculated by using G*power software (version 3.1.9.7, Dusseldorf, Germany), with an alpha of 0.05, a power of 0.8 and an effect size of 0.5. Based on the software calculation, the minimum required number was 90.

2.5.2. Assessment: Upon receiving the patient’s consent, the personal and clinical history were taken using Therapha software on the same day before the patient underwent an MRI. Each software-assisted patient evaluation session lasted around 10-15 minutes. The history encompassed initial demographic details and questions about the patient’s medical history, family history, social history, functional deficits, the severity of the back pain, nature of the pain, duration, onset, location, easing and aggravating factors, abnormal sensations, mechanism of injury and its relationship to the emergence of pain. Also, patients were asked about symptoms that may indicate systemic disease infection, fracture, or neurological impairments.

2.5.3. Procedure: Patients were referred to the Radiology department from orthopedic, neurology, and spine surgery clinics. The Therapha software is only available in English, hence requiring the use of a blind assessor. An experienced physical therapist, who was blinded to the patient's information, utilized Therapha software to gather the patient's personal and clinical history before the MRI. The final report received from the software was compared to the gold standard (MRI reports) by the study assistant and excluded any other pathology such as infection, fracture or tumor. (Figure 3)

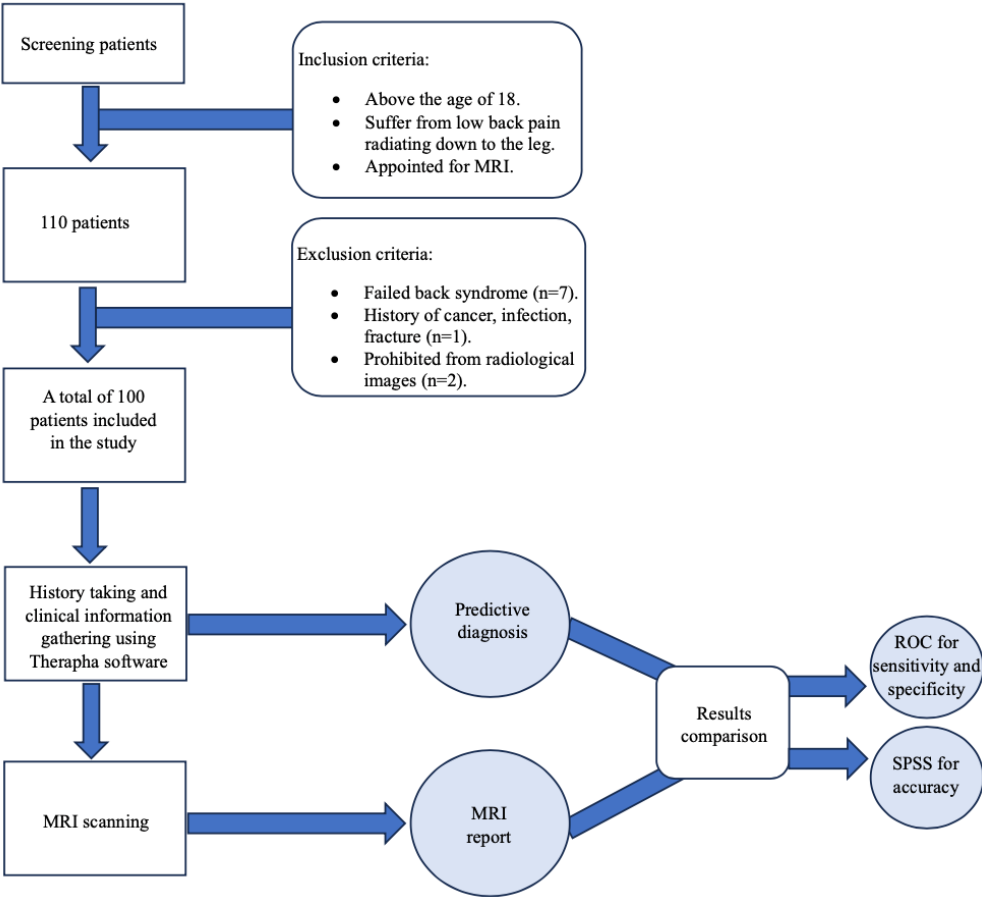


Figure 3. Flowchart demonstrating the diagnostic accuracy study process of the Therapha software.

2.6. Outcome Measures

2.6.1. Therapha Software: Therapha web software is a CDSS that is considered a comprehensive screening tool and triaging system for spine pathology which supports self-service by the patient or clinician. This system utilizes the power of information technology to identify medically established patterns by incorporating over 4000 independent variables, which is evidently superior to the clinical reasoning by a human clinician. This algorithmic approach enables the tool to predict the possible differential diagnoses in an order of probability percentile. The system divides spine disorders into subgroups besides conventional classification. The supervised AI model identifies patterns in the same manner as clinician conducts an evaluation and the Therapha system is capable of modifying their algorithm in the light of newer research findings. This improves the solution’s predictive accuracy over the period through an iterative process. CDSS uses data from the clinician or patient self-administered questionnaire to refine clinical reasoning and predictions.

The information gathered from the patient generates a 'digital twin,' which serves as the primary input to this clinical decision support solution. The other main component is the knowledge repository, which utilizes the data captured by the chatbot to generate the top three hypothetical diagnoses. The algorithm is the software that embodies AI; which is the system's most innovative component.

2.6.2. MRI: In this study, all patients were scanned with MRI, using 3 Tesla Verio and Skyra systems (Siemens, Erlangen, Germany). A slice thickness of 3 mm disc by disc was used with an interslice gap of 1-0.5 mm.

2.7. Statistical Analysis

Data were statistically analyzed by using SPSS for Windows (version 23.0; SPSS Inc, Chicago, IL) to evaluate the predictive accuracy of Therapha web software with MRI reports. Anthropometric data were analyzed using descriptive analysis. The accuracy of the software performance was measured

using the Receiver Operating Curve (ROC) to find the Sensitivity and Specificity. P-value < .05 was considered statistically significant.

3. Results

3.1. Case Series

Lumbar disc herniation was verified by MRI for all patients (n=10) except for cases # (4&6) who had a different spine diagnosis by AI software. Patient’s Therapha software inferences and MRI reports are illustrated in Table 1.

Table 1. MRI findings and Therapha software inferences.

Patient no.	Age	BMI	MRI Diagnosis	THERAPHA Prediction
1	55	38	Mild diffuse disc bulge	Lumbar radiculopathy/Lumbar disc herniation/ Lumbago with Sciatica
2	67	26	At the level of T12-L1 there is mild central disc bulge	Lumbar facet joint pain/lumbar zygapophyseal/non-specific low back pain
3	59	22	Mild diffuse disc bulge	Lumbar radiculopathy/lumbar disc herniation/ lumbago with sciatica
4	40	21	No significant disc disease	Lumbar radiculopathy/lumbar disc herniation/ lumbago with sciatica
5	33	25	Diffuse posterior bulge	Lumbar radiculopathy/lumbar disc herniation/ lumbago with sciatica
6	38	25	Extradural meningeal cyst	Lumbar radiculopathy/lumbar disc herniation/ lumbago with sciatica
7	65	32	Minimal diffuse posterior disc bulges	Lumbar radiculopathy/lumbar disc herniation/ lumbago with sciatica
8	61	34	Lumbar disc bulge L4-5, L5-S1	Lumbar radiculopathy/lumbar disc herniation/ lumbago with sciatica
9	40	29	Lumbar disc bulge L4-5, L5-S1	Lumbar radiculopathy/lumbar disc herniation/ lumbago with sciatica
10	69	25	Lumbar disc bulge L4-5, L5-S1	Lumbar radiculopathy/lumbar disc herniation/ lumbago with sciatica

MRI findings for case# 4 showed "no significant disc disease"; however, Therapha identified it as "Lumbar Radiculopathy/Lumbar Disc Herniation/ Lumbago with Sciatica". This can be explained by the nature of the pain. As the patient fed into the chatbot, they reported dull and radiating pain from the back spreading to the L3-L4 dermatomal regions, along with tingling and numbness covering the lateral aspect of the thigh and lower leg corresponding to the L4-S1 dermatomes. There was also tenderness in the midline corresponding to the L4 segment. Further, the patient mentioned trouble with functional activities such as squatting, climbing stairs, and walking. Also, reports decreased intolerance to activities like bending forward, standing, and sitting. As a result, patients complain of restriction in daily living activities and notice guarding and limited movement of the involved areas.

Regarding case# 6, the MRI report was "extradural meningeal cyst", while Therapha predicted the patient as "Lumbar Radiculopathy/Lumbar Disc Herniation/ Lumbago with Sciatica". Researchers can interpret this from the patient’s medical history since the patient reported his/her

pain as a dull ache radiating in nature and marked the corresponding area to T12- S1 dermatomal distribution. Besides, having abnormal sensations of tingling and numbness in the feet in L4-S1 dermatomal regions and the presence of tender spots. In addition, the patient’s pain worsened in the morning and got better as the day progressed. Also, the pain was exacerbated by deep breathing, laughing, sneezing or coughing. The patient reported difficulty in moving the affected area, as well as functional activities like squatting, climbing stairs, and walking. The complaints noted by the patient were aggravated with standing, walking, and sustained positions. Moreover, the patient reports decreased tolerance in activities such as coming up from a bent-over position and from sitting to standing positions, bending forward, standing, and sitting. As a consequence, the patient complained of restriction in daily living activities and noticed guarding or limiting movement of the involved area. Furthermore, the patient had a history of falling. The patient reported an easing of their symptoms with rest in the supine position and medications.

3.2. Delphi Method

Following the completion of the case series, the assessor collected and tabulated experts’ rating sheets. In the first round, all respondents agreed (n=10) that the patient’s medical history is congruent with Therapha's prediction, thereby achieving consensus. As all 10 cases had a consensus of above 80%, there was no need to re-circulate rating sheets for a second round.

Regarding case# (2,3,4,5,8,9,10), there was 100% consensus that the patient’s medical history was consistent with Therapha predictions. However, regarding cases #1, #6, and #7, there was a consensus of 90% among all 10 experts.

3.3. Diagnostic Predictive Accuracy

The demographic characteristics of patients who participated in this study are displayed in Table 2. The accuracy of predictions from the software was determined by using ROC. The curve was up toward the left with an area under the curve of .84 (p = .001, 95% CI; 0.6 to 1.0) (Figure 4). Therapha had a sensitivity of 88%, specificity of 80%, positive predictive value of 99%, negative predictive value of 27%, positive likelihood ratio of 4.4, and negative likelihood ratio of .15.

Table 2. Descriptive statistics (n=100).

			Mean ± SD	Median (min-max)
Age (Years)			48.44 ± 13.96	48.50 (18-84)
Gender	Male	21		
	Female	79		
Height (cm)			161.85 ± 7.83	160 (149-185)
Weight (kg)			76.83 ± 15.05	75 (50-130)
BMI			29.34 ± 5.68	28.5 (21-54)

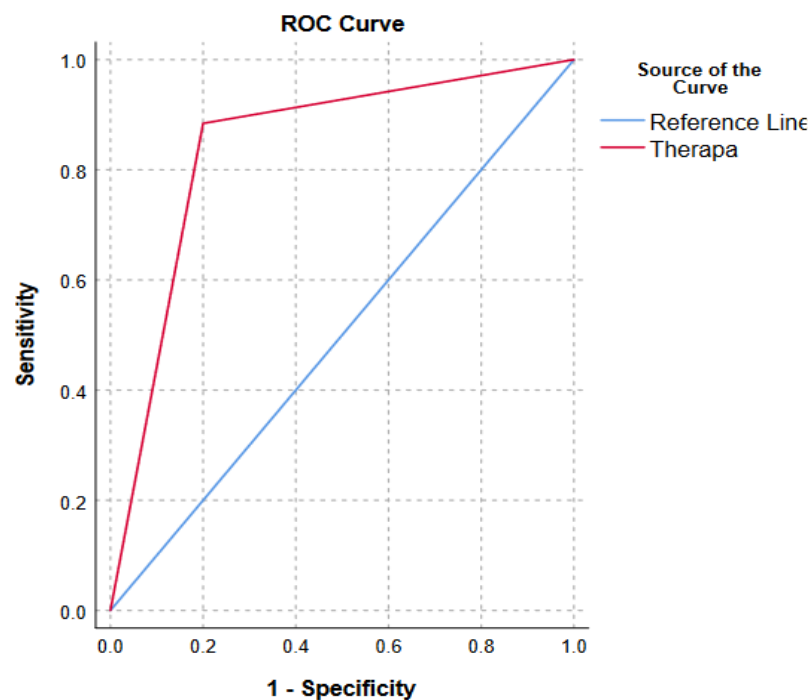


Figure 4. Receiver Operating Curve (ROC).

4. Discussion

In this study, a novel web software has been proposed for predicting lumbar disc herniation. The results of this study showed that Therapha software can be used for predicting and assisting with the diagnosis of lumbar disc herniation with a high level of accuracy. Findings revealed that Therapha technology has a satisfactory level of effectiveness, making it suitable for expedited and more effective screening of patients with lumbar disc herniation. To the best of our knowledge, this is the first study to examine the efficacy of an AI-enabled tool compared to the gold standard (MRI) for diagnosing patients with lumbar disc herniation.

4.1. Delphi Method

The Delphi method was used to reach an agreement in diagnosing patients with lumbar herniated discs. In the present study, experts agreed on statements representing the final diagnosis of patients that were generated by Therapha software which can be used as a guideline for clinical practice to help in decision-making process. This reinstates the significance of the tool's potential to be used as an adjunct to provide care that is backed by the power of technology.

Each patient's signs and symptoms were in match to their MRI findings. Since most patients have the same history of lumbar herniated disc, patient compliance with LBP accompanied by radiculopathy and abnormal sensations like numbness and tingling were reported. Most patients in this study have shown an increase in their symptoms with flexion activities and the same was eased with extension-based movements. The majority of patients complain of pain with prolonged activity such as standing, sitting, or even walking for a long time. Also, patients report that sneezing, deep breathing, or coughing exacerbate their symptoms. In addition, patients have the same type of pain "dull, aching, radiating to the lateral side down to the leg and foot". Furthermore, muscle spasms and weakness of either the back or thigh muscles have been reported.

Concerning cases # (4 and 6), the Therapha predictions aligned with the expert diagnostic findings. However, experts' agreements were not entirely consistent with the results obtained from the MRI scans. They had a consensus of 100% and 90% for cases # (4 and 6), respectively. Based on the findings, it could be stated that Therapha can identify patients with lumbar disc herniation efficiently which could help clinicians in their decision-making process by serving as a cognitive aid.

Therapha software is an innovative tool that can assist healthcare professionals in analyzing patient history (screening and diagnostic questions) and provide the most likely differential diagnosis without disturbing the clinical workflow. The solution accomplishes this through an innovative way: replacing the traditional patient intake forms with a conversational chatbot. The entire interaction is automated into an organized clinical summary and provides an actionable insightful report that includes probable differential diagnoses. Screening questions help to identify patients who require special or alternative care pathways [20], while diagnostic questions help to identify the proper pathology, allowing for suitable treatment selections [20].

The use of proper diagnostic and screening questions can improve patient care by successfully identifying problems and discovering correct diseases [20]. History taking alone has been proven to accurately identify clinical problems by 56% to 82.5% of the time [24,25]. Therapha has an accuracy of 84% in predicting lumbar disc herniation pathology only through history taking incorporated with AI-based technology.

However, this procedure does not diminish the significance of physical examination. To confirm a diagnosis, a physical examination should be done in conjunction with a history. Physical examinations can sometimes reveal unexpected diagnoses or contribute to previously obtained data [26]. Investigations can also help in LBP diagnosis, particularly medical imaging advances have offered healthcare professionals with new non-invasive approaches to improve patient care [27].

4.2. Diagnostic Predictive Accuracy

Various systems have been developed to assist in diagnosing different pathologies. However, there are limited systems in the literature that are specifically designed to help in the diagnosis and treatment of LBP [28–30]. Although AI-based solutions have been developed in a variety of medical disciplines, research incorporating AI in LBP is still in its embryonic stage [31].

The result of this study was in line with Kadhim et. al., where they have studied a fuzzy expert system for diagnosing LBP patients depending on medical observation symptoms through using Fuzzy rule and other factors that may influence back pain diagnoses such as body mass index, gender, age, and patient history. Thereafter, depending on the information provided, the system offers the most likely spine pathology and treatment. This system showed an accuracy of 90% when compared to other specialist diagnoses[32].

Furthermore, Lin et al., examined a web-based decision support system that used production for knowledge representation to evaluate patient's information and suggest a diagnosis for LBP. The system infers multi-part diagnoses rapidly using a mini-Bayesian approach. It includes two interfaces for knowledge update and convenient system access, an inference engine, a knowledge base and a case repository. The CDSS shows performance equal to human experts [8]. However, this web-based decision support system needs to expand on its knowledge replenishment support. A clinician's diagnostic knowledge is likely to accumulate over time, making the system's knowledge replenishment support important. On the other hand, Therapha web software is currently built on a hybrid algorithm that utilizes over 4000 independent variables and is periodically updated depending on feedback from clinicians on the predictions made. They verify and discuss the feedback from clinicians and then make the required changes to improve the power of algorithmic. The vision is to have a learning system as the system absorbs large quantities of clinical data over the course of time.

Sari et al., examined two expert systems (adaptive neuro-fuzzy inference system and artificial neural network) to objectively assess the level of pain intensity in low back patients. Researchers used two input variables including skin resistance, a visual analog scale, and one output variable which was LBP intensity [33]. They found higher accuracy in predicting LBP when the two systems were used in combination. Sari et al., proposed an objective and subjective method that can be used to scale a patient's level of pain [33]. However, this cannot be used as a diagnostic tool, rather it can be used as a screening tool. On the other hand, the current study presents a clinical decision support tool (Therapha) which, is part of its algorithm to determine LBP intensity subjectively by using a visual analog scale and detailed history related to pain.

In recent years, the area of AI has been growing rapidly, with multiple studies demonstrating high levels of accuracy in various AI applications. Indeed, a growing body of literature has reported the high level of accuracy of AI systems in various medical fields, especially in medical imaging. Lee et al., converted lumbar CT scans into axial T2-weighted MRI slices by using generative adversarial networks. Then, they compare it to each patient's actual lumbar MRI images [11]. The results show an 80.2% similarity to the actual lumbar MRI images. Moreover, Ramirez et al., used AI techniques such as decision trees, support vector machines, and a logistic regression classifier to identify scoliosis curve changes. They found that the support vector machine generated the best results on 141 radiological spine scans, with an accuracy of 86% and applied AI on MRI scans to estimate the surgical level of patients having disc decompression surgery [34,35]. Harada et al., recently developed an algorithm that identified patients at risk of re-herniation after microdiscectomy. After validation, the prediction method showed a recall of 80%, an area under the ROC of 72%, and an accuracy of 70% [36].

The above-mentioned studies showed high accuracy and performance as Therapha software has shown. However, Therapha is a triaging system assessing physical problems related to the musculoskeletal system, but it does not assess and read the MRI itself. Image processing and computer vision are gaining momentum as computational resources such as powerful graphics processing units become more readily accessible. Indeed, current spine research studies using AI methods are focused on medical imaging. However, AI research in other areas such as spine pathology diagnosis is lacking, and more studies and research are required. This study had several limitations, the software was only used for the lumbar area to identify patients with radiculopathy, excluding other conditions or lesions associated with back pain. Additionally, although the software was intended to detect various pathologies connected to lower back pain, it did not encompass all possible conditions. Furthermore, the study exhibited a generalization bias by exclusively recruiting individuals with symptoms related to the lower back and lower limbs, which accounts for the lower negative predictive value (NPV) of 27%. Also, the age group of the patient in the study is worth a mention. Additionally, it is noteworthy to state the age group of the participants involved in the study as a limitation to generalising the findings.

Also, the study had no control over the MRI devices, radiology technicians, and radiology consultants. Another limitation of the study was the software is only available in the English language. As mentioned before, Therapha software supports patient self-service, the majority of the patients in this study didn't know English, so translation to other languages is important to increase the access to patient self-service features in the software. However, in this study patient's history were taken and entered into the software by the clinician. Further, this study had several strengths, the software is a cost-effective and quick screening tool that may aid patients who are contraindicated to MRI such as claustrophobic individuals. For future research, it is recommended to investigate and study other spinal conditions such as cervical, thoracic, sacroiliac, and other conditions related to the lumbar region.

5. Conclusions

Given the increasing expenses associated with diagnosing and treating LBP, Therapha could be a highly cost-efficient tool to improve diagnostic accuracy and deliver effective care. Therapha's patient-centred approach utilizing AI to augment decision-making will lead to enhanced clinical outcomes in spine care. This study provides new insights for healthcare researchers, practitioners, and policymakers into the potential of implementing emerging technologies, especially AI in healthcare delivery and diagnosis. As evidenced in the study, CDSS can be an efficient cognitive support for clinicians to render optimal patient care. Future studies are required to investigate additional spinal illnesses such as sacroiliac, facet joint disorders, etc., not limited to the lumbar region, but the thoracic, and cervical segments as well.

Author Contributions: "Conceptualization, F.B, A.A, M.F.C, Z.N, U.A, A.Q, F.K; methodology, F.B, M.A, A.A, A.J, Z.N, M.F.C, F.K; validation, Z.N, U.A, A.Q, A.J, R.B, F.K; formal analysis, F.B, M.A, M.F.C, R.B, F.K; investigation, A.A, Z.N, U.A, A.Q, A.J, F.K; resources, M.A, Z.N, A.Q, A.J, F.K; data curation, F.B, M.A, A.A, R.B,

M.F.C; writing—original draft preparation, F.B, M.A, A.A, M.F.C, R.B; writing—review and editing, Z.N, A.Q, U.A, A.J, F.K; visualization, Z.N, M.F.C, A.J, F.K; supervision, M.A, U.A, A.Q, F.K; project administration, M.A, A.Q, R.B, F.K; funding acquisition, R.B. All authors have read and agreed to the published version of the manuscript.

Funding: This research was funded by the Princess Nourah Bint Abdulrahman University researchers supporting project number (PNURSP2024R267), Princess Nourah Bint Abdulrahman University, Riyadh, Saudi Arabia.

Institutional Review Board Statement: “The study was conducted in accordance with the Declaration of Helsinki, and approved by the Institutional Review Board of Research Ethics Committee at King Abdulaziz University Hospital (KAUH) (No. 125-22; HA-02-J-008), Jeddah, Saudi Arabia.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: The data presented in this study are available on request from the corresponding author. The data are not publicly available due to privacy restrictions.

Acknowledgments: In this section, you can acknowledge any support given which is not covered by the author contribution or funding sections. This may include administrative and technical support, or donations in kind (e.g., materials used for experiments).

Conflicts of Interest: The authors declare no conflicts of interest.

References

1. Walker, B.F. The Prevalence of Low Back Pain: A Systematic Review of the Literature from 1966 to 1998. *J Spinal Disord* 2000, 13, 205–217, doi:10.1097/00002517-200006000-00003.
2. Alotaibi, M.A.; Alfaifi, R.M.; Alhowimel, A.S.; Alodaibi, F.A.; Alzahrani, H.; Alenazi, A.M.; Alqahtani, B.A.; Elnaggar, R.K. The Key Determinants of Low Back Pain among Lifestyle Behaviors in Adolescents: A Cross-Sectional Study from Saudi Arabia. *Medicine (Baltimore)* 2024, 103, e37669, doi:10.1097/MD.00000000000037669.
3. Hartvigsen, J.; Hancock, M.J.; Kongsted, A.; Louw, Q.; Ferreira, M.L.; Genevay, S.; Hoy, D.; Karppinen, J.; Pransky, G.; Sieper, J.; et al. What Low Back Pain Is and Why We Need to Pay Attention. *Lancet* 2018, 391, 2356–2367, doi:10.1016/S0140-6736(18)30480-X.
4. Scaia, V.; Baxter, D.; Cook, C. The Pain Provocation-Based Straight Leg Raise Test for Diagnosis of Lumbar Disc Herniation, Lumbar Radiculopathy, and/or Sciatica: A Systematic Review of Clinical Utility. *J Back Musculoskelet Rehabil* 2012, 25, 215–223, doi:10.3233/BMR-2012-0339.
5. Al Nezari, N.H.; Schneiders, A.G.; Hendrick, P.A. Neurological Examination of the Peripheral Nervous System to Diagnose Lumbar Spinal Disc Herniation with Suspected Radiculopathy: A Systematic Review and Meta-Analysis. *Spine J* 2013, 13, 657–674, doi:10.1016/j.spinee.2013.02.007.
6. Rebain, R.; Baxter, G.D.; McDonough, S. A Systematic Review of the Passive Straight Leg Raising Test as a Diagnostic Aid for Low Back Pain (1989 to 2000). *Spine (Phila Pa 1976)* 2002, 27, E388-395, doi:10.1097/00007632-200209010-00025.
7. Jarvik, J.G.; Deyo, R.A. Diagnostic Evaluation of Low Back Pain with Emphasis on Imaging. *Ann Intern Med* 2002, 137, 586–597, doi:10.7326/0003-4819-137-7-200210010-00010.
8. Lin, L.; Hu, P.J.-H.; Liu Sheng, O.R. A Decision Support System for Lower Back Pain Diagnosis: Uncertainty Management and Clinical Evaluations. *Decision Support Systems* 2006, 42, 1152–1169, doi:10.1016/j.dss.2005.10.007.
9. Cai, C.J.; Reif, E.; Hegde, N.; Hipp, J.; Kim, B.; Smilkov, D.; Wattenberg, M.; Viegas, F.; Corrado, G.S.; Stumpe, M.C.; et al. Human-Centered Tools for Coping with Imperfect Algorithms During Medical Decision-Making. *Proceedings of the 2019 CHI Conference on Human Factors in Computing Systems* 2019, 1–14, doi:10.1145/3290605.3300234.
10. Cai, C.J.; Winter, S.; Steiner, D.; Wilcox, L.; Terry, M. “Hello AI”: Uncovering the Onboarding Needs of Medical Practitioners for Human-AI Collaborative Decision-Making. *Proc. ACM Hum.-Comput. Interact.* 2019, 3, 104:1-104:24, doi:10.1145/3359206.
11. Lee, M.H.; Siewiorek, D.; Smailagic, A.; Bernardino, A.; Badia, S. Opportunities of a Machine Learning-Based Decision Support System for Stroke Rehabilitation Assessment. *ArXiv* 2020.
12. Musen, M.A.; Middleton, B.; Greenes, R. Clinical Decision-Support Systems. In *Biomedical Informatics*; Springer London, 2014; pp. 643–674 ISBN 978-1-4471-4473-1.
13. Park, J.H.; Choi, M.-H.; Lee, J.; Han, H.-S.; Lee, M.C.; Ro, D.H. Gait Deviations of Patients with Ruptured Anterior Cruciate Ligament: A Cross-Sectional Gait Analysis Study on Male Patients. *Knee Surg Relat Res* 2021, 33, 45, doi:10.1186/s43019-021-00128-w.

14. Kunhimangalam, R.; Ovalath, S.; Joseph, P.K. A Clinical Decision Support System with an Integrated EMR for Diagnosis of Peripheral Neuropathy. *J Med Syst* 2014, 38, 38, doi:10.1007/s10916-014-0038-9.
15. Ahmadi, M.; Nopour, R. Clinical Decision Support System for Quality of Life among the Elderly: An Approach Using Artificial Neural Network. *BMC Med Inform Decis Mak* 2022, 22, 293, doi:10.1186/s12911-022-02044-9.
16. Lahsasna, A.; Ainon, R.N.; Zainuddin, R.; Bulgiba, A. Design of a Fuzzy-Based Decision Support System for Coronary Heart Disease Diagnosis. *J Med Syst* 2012, 36, 3293–3306, doi:10.1007/s10916-012-9821-7.
17. Emani, S.; Rui, A.; Rocha, H.A.L.; Rizvi, R.F.; Juaçaba, S.F.; Jackson, G.P.; Bates, D.W. Physicians' Perceptions of and Satisfaction With Artificial Intelligence in Cancer Treatment: A Clinical Decision Support System Experience and Implications for Low-Middle-Income Countries. *JMIR Cancer* 2022, 8, e31461, doi:10.2196/31461.
18. Azimi, P.; Yazdani, T.; Benzel, E.C.; Aghaei, H.N.; Azhari, S.; Sadeghi, S.; Montazeri, A. A Review on the Use of Artificial Intelligence in Spinal Diseases. *Asian Spine J* 2020, 14, 543–571, doi:10.31616/asj.2020.0147.
19. Azimi, P.; Mohammadi, H.R.; Benzel, E.C.; Shahzadi, S.; Azhari, S.; Montazeri, A. Artificial Neural Networks in Neurosurgery. *J Neurol Neurosurg Psychiatry* 2015, 86, 251–256, doi:10.1136/jnnp-2014-307807.
20. Eubank, B.H.; Mohtadi, N.G.; Lafave, M.R.; Wiley, J.P.; Bois, A.J.; Boorman, R.S.; Sheps, D.M. Using the Modified Delphi Method to Establish Clinical Consensus for the Diagnosis and Treatment of Patients with Rotator Cuff Pathology. *BMC Med Res Methodol* 2016, 16, 56, doi:10.1186/s12874-016-0165-8.
21. The Rand/UCLA Appropriateness Method User's Manual; Fitch, K., Ed.; MR / Rand DG-XII/RE; Rand: Santa Monica, 2001; ISBN 978-0-8330-2918-8.
22. Meshkat, B.; Cowman, S.; Gethin, G.; Ryan, K.; Wiley, M.; Brick, A.; Clarke, E.; Mulligan, E. Using an E-Delphi Technique in Achieving Consensus across Disciplines for Developing Best Practice in Day Surgery in Ireland. 2014, doi:10.5430/jha.v3n4p1].
23. Lynn, M.R. Determination and Quantification of Content Validity. *Nurs Res* 1986, 35, 382–385.
24. Peterson, M.C.; Holbrook, J.H.; Von Hales, D.; Smith, N.L.; Staker, L.V. Contributions of the History, Physical Examination, and Laboratory Investigation in Making Medical Diagnoses. *West J Med* 1992, 156, 163–165.
25. Hampton, J.R.; Harrison, M.J.; Mitchell, J.R.; Prichard, J.S.; Seymour, C. Relative Contributions of History-Taking, Physical Examination, and Laboratory Investigation to Diagnosis and Management of Medical Outpatients. *Br Med J* 1975, 2, 486–489.
26. Davis, J.L.; Murray, J.F. History and Physical Examination. Murray and Nadel's Textbook of Respiratory Medicine 2016, 263-277.e2, doi:10.1016/B978-1-4557-3383-5.00016-6.
27. Chou, R.; Qaseem, A.; Owens, D.K.; Shekelle, P.; Clinical Guidelines Committee of the American College of Physicians Diagnostic Imaging for Low Back Pain: Advice for High-Value Health Care from the American College of Physicians. *Ann Intern Med* 2011, 154, 181–189, doi:10.7326/0003-4819-154-3-201102010-00008.
28. Iqbal, K.; Yin, X.-C.; Hao, H.-W.; Ilyas, Q.M.; Ali, H. An Overview of Bayesian Network Applications in Uncertain Domains. *IJCTE* 2015, 7, 416–427, doi:10.7763/IJCTE.2015.V7.996.
29. Matar, H.E.; Navalkisoor, S.; Berovic, M.; Shetty, R.; Garlick, N.; Casey, A.T.H.; Quigley, A.-M. Is Hybrid Imaging (SPECT/CT) a Useful Adjunct in the Management of Suspected Facet Joints Arthropathy? *Int Orthop* 2013, 37, 865–870, doi:10.1007/s00264-013-1811-y.
30. Hurri, H.; Karppinen, J. Discogenic Pain. *Pain* 2004, 112, 225–228, doi:10.1016/j.pain.2004.08.016.
31. D'Antoni, F.; Russo, F.; Ambrosio, L.; Vollero, L.; Vadalà, G.; Merone, M.; Papalia, R.; Denaro, V. Artificial Intelligence and Computer Vision in Low Back Pain: A Systematic Review. *Int J Environ Res Public Health* 2021, 18, 10909, doi:10.3390/ijerph182010909.
32. Kadhim, M.A.; Afshar, M.; Alam, Kaur, H. Design and Implementation of Fuzzy Expert System for Back Pain Diagnosis.; 2011.
33. Sari, M.; Gulbandilar, E.; Cimbiz, A. Prediction of Low Back Pain with Two Expert Systems. *J Med Syst* 2012, 36, 1523–1527, doi:10.1007/s10916-010-9613-x.
34. Cabitza, F.; Locoro, A.; Banfi, G. Machine Learning in Orthopedics: A Literature Review. *Front Bioeng Biotechnol* 2018, 6, 75, doi:10.3389/fbioe.2018.00075.
35. Ramirez, L.; Durdle, N.G.; Raso, V.J.; Hill, D.L. A Support Vector Machines Classifier to Assess the Severity of Idiopathic Scoliosis from Surface Topography. *IEEE Trans Inf Technol Biomed* 2006, 10, 84–91, doi:10.1109/titb.2005.855526.
36. Harada, G.K.; Siyaji, Z.K.; Mallow, G.M.; Hornung, A.L.; Hassan, F.; Basques, B.A.; Mohammed, H.A.; Sayari, A.J.; Samartzis, D.; An, H.S. Artificial Intelligence Predicts Disk Re-Herniation Following Lumbar Microdiscectomy: Development of the "RAD" Risk Profile. *Eur Spine J* 2021, 30, 2167–2175, doi:10.1007/s00586-021-06866-5.

Disclaimer/Publisher's Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.