Diagnostic accuracy of contrast and non-contrast 1.5 tesla magnetic resonance imaging for lumbar herniated nucleus pulposus based on surgical findings: A systematic review and meta-analysis of diagnostic test accuracy studies

Endra Wibisono Harmawan
Neurosurgeon. Department of Neurosurgery. Faculty of Medicine, Airlangga University- Dr. Soetomo General Academic Hospital, Surabaya, Indonesia

Tedy Apriawan
Neurosurgeon Staff. Department of Neurosurgery. Faculty of Medicine, Airlangga University- Dr. Soetomo General Academic Hospital, Surabaya, Indonesia
Corresponding author email: drtedyapri@fk.unair.ac.id

Eko Agus Subagio
Spine Neurosurgeon Staff. Department of Neurosurgery. Faculty of Medicine, Airlangga University- Dr. Soetomo General Academic Hospital, Surabaya, Indonesia

Muhammad Faris
Spine Neurosurgeon Staff. Department of Neurosurgery. Faculty of Medicine, Airlangga University- Dr. Soetomo General Academic Hospital, Surabaya, Indonesia

Sri Andreani Utomo
Neuroradiologist Staff. Department of Radiology. Faculty of Medicine, Airlangga University- Dr. Soetomo General Academic Hospital, Surabaya, Indonesia

Budi Utomo
Statistics lecturer. Department of Public Health and Preventive Medicine. Faculty of Medicine, Airlangga University- Dr. Soetomo General Academic Hospital, Surabaya, Indonesia

Abdul Hafid Bajamal
Prof of Neurosurgeon. Department of Neurosurgery. Faculty of Medicine, Airlangga University- Dr. Soetomo General Academic Hospital, Surabaya, Indonesia

Abstract---This study aims to provide the latest information regarding the accuracy of contrast and non-contrast 1.5 Tesla MRI in the diagnosis of lumbar Herniated Nucleus Pulposus (HNP) compared with surgical findings. We included studies with subjects diagnosed with lumbar HNP due to degenerative processes, preoperative 1.5 Tesla contrast, and non-contrast MRI and described the number of findings of MRI diagnostic tests (index test) on the actual conditions found during surgery (reference standard). We searched literature from 5 databases: ProQuest, Pubmed, Cochrane Library, Biomed Central, and ScienceDirect. The sensitivity and specificity of contrast and non-contrast 1.5 Tesla MRI ranged between 64-95% and 55-100% (95% CI) with the area under the curve above the threshold on the ROC curve. Two studies compared the accuracy of MRI and Computed Tomography (CT) myelography with a wider ROC curve on CT myelography than on MRI. The ROC curve which has a wide area under the curve above the threshold depicts the relationship between sensitivity and specificity, shows that contrast non-contrast 1.5 Tesla MRI has a good accuracy in HNP diagnostic.

Keywords---diagnostic test accuracy, MRI, herniated nucleus pulposus, surgical findings.

Introduction

Significant radiology findings of disc degeneration, facet hypertrophy, and disc protrusion were associated with the cause of Low Back Pain (LBP) (Hutchins TA et al., 2021). Low back pain is the main cause of work related disorders as well as disability, especially in the productive age population (Wu A et al., 2020). According to The Global Burden (2020) of Disease, LBP ranks fourth in the 10 most common diseases between people aged 25-49 years. Increasingly sophisticated imaging technology is expected to improve diagnostic accuracy and help the decision making of effective therapy for patient recovery. Imaging is needed primarily as a planning tool before surgery in patients with LBP who have previously failed conservative therapy (Johnson SM and Shah LM., 2019; Rosyid AN, Yamin M, and Puspitasari AD., 2019; Yates M et al., 2020).

MRI has been chosen because it is non-invasive, provides good anatomical morphological details, and does not expose the patient to ionizing radiation. Its widespread use has a good intraobserver agreement on reproducibility test statistically, especially in the assessment of spondylolisthesis, disc degeneration, infection process, Modic changes, fracture, and facet arthropathy (Kusmiati T, Narendrani HP., 2019). Reports of a high prevalence of lumbar spine abnormalities seen on MRI in asymptomatic subjects are also available (Hebelka

H et al., 2019; Jiang X, Chen D et al., 2018; Kim JH et al., 2018). In addition, a false positive imaging on MRI can lead to inappropriate therapy, increase the cost of treatment and result in suboptimal treatment (Bajamal AH et al., 2021).

The importance of assessing the accuracy and use of MRI imaging which is the current modality of choice in the diagnosis of LBP is related to the findings from the clinical examination and during surgery. The findings during the surgery can assess the accuracy of preoperative MRI imaging as a diagnostic tool (Ekedahl H et al., 2018; Kim JH et al., 2018). The use of 1.5 Tesla MRI is superior in the diagnosis of pathological abnormalities of the spine. This is due to the less possibility of artifacts due to bowel movements and breath than 3 Tesla MRI which is very sensitive to movement (Wu A et al., 2020). Patients who experience pain during MRI are very susceptible to movements that cause artifacts (Vargas MI, Boto J and Meling TR., 2021; Varlotta CG et al., 2020).

This systematic review and meta-analysis aim to provide information about the diagnostic accuracy of MRI and identify the factors that influence it. The conclusion of the accuracy test value of MRI obtained in this meta-analysis is expected to be a reference for clinicians to use MRI imaging as a determinant of decision making more wisely and understand the limitations that imaging has while considering the clinical condition of the patient.

Method

Research Design, Protocol, and Registration

This review followed the Diagnostic Test Accuracy (DTA) Protocol of The Cochrane Collaboration protocol. The authors adopted the Cochrane DTA format as it helps readers find review results quickly and to assess the validity, applicability, and implications of these results. Journal selection was carried out using guidelines from PRISMA (Preferred Reporting Items for Systematic Review and Meta-Analysis Protocol) flow diagrams and PRISMA checklist assessments adapted to the purpose of this study. This research has been registered in the International Prospective Register of Systematic Reviews/PROSPERO with registration number CRD42021277779.

Population, Sample, and PICO Research Questions

The research population was all journal articles obtained from the literature search results according from the Population Intervention Comparison Outcome (PICO). The samples were obtained by filtering the search results of articles from the publication database based on the PRISMA flow according to the inclusion and exclusion criteria. The strategy used to find articles using the PICO framework can be seen in table 1.

Table 1. PICO Research Questions

Population (P)	Intervention (I)	Comparison (C)	Outcome (O)
Patients diagnosed with lumbar HNP	1.5 Tesla Contrast and Non-Contrast MRI Preoperative Imaging (index test)	Surgical Findings (reference standard)	Accuracy value : sensitivity and specificity

Research Inclusion Criteria
Literatures with subjects aged over 18 years old diagnosed with lumbar HNP due to degenerative processes, number of samples more than 10 patients, preoperative MRI examination 1.5 Tesla contrast and non-contrast with routine MRI protocol, literature that describes the number of findings of MRI diagnostic tests (index test) on the actual conditions found during surgery (reference standard) resulted in True Positive (TP), False Positive (FP), False Negative (FN) and True Negative (TN) numbers, cohort research design, full-text literature that has been published without limitation of publication year and in English.

Research Exclusion Criteria
Literatures with subjects diagnosed with HNP due to trauma, infection, malignancy or causes other than degenerative disorders, cervical and thoracic HNP, preoperative MRI examination 0.35 Tesla and 3 Tesla without a routine protocol, cross-sectional and RCT study design, literatures in the form of abstract, without full text and articles without peer review process.

Source of information and Journal search keywords
The literature search in this systematic review used 5 databases, namely ProQuest, Pubmed, Cochrane Library, Biomed Central, and ScienceDirect. We searched the articles or journals using keywords and boolean operators (AND, OR NOT or AND NOT) which were used both to broaden and narrow the search, making it easier to determine the articles or journals used. The keywords in this systematic review were adjusted to the Medical Subject Heading (MeSH) and can be seen in table 2.

Table 1 : Journal search keywords

Diagnos*		*MRI*		*Herniated Nucleus Pulposus*			
OR		*OR*				*Compar**	
	AND		*AND*		*AND*		*AND* *Surg**
Accura*		*Magnetic Resonance Imaging*		*OR* HNP		*OR*	
OR		*OR*		*OR*		*Correlat**	

Examinat*	MR Imaging	Lumbar Disc Herniation

Journal selection
All journal articles that met the inclusion criteria were checked for duplication, then screened in two stages for eligibility. In the first stage, they were checked based on the title and abstract. It was conducted by a team consisting of five reviewers (EWH, TA, BU, EAS, and MF) and if there were doubts, they discussed it with the senior reviewer (AH). Full texts of the selected journals were collected, and the relevant data were extracted in a characteristic table. The characteristics table included variables that can be measured. The literature selection was carried out based on the search flow/PRISMA 2020 flow chart based on the recommendation from the Cochrane Review.

Literature quality assessment and risk of bias
The quality of the literature included in this study was assessed separately by five reviewers (EWH, TA, BU, EAS, MF, and SAU) using the Newcastle-Ottawa Quality Assessment Form for Cohort Studies. Any disagreements were resolved by discussion until we reached a conclusion. The reviewers conducted a qualitative analysis of the risk of bias in the literature obtained independently using the QUADAS2 tools.

Data Analysis
Meta-analysis was done using Microsoft Office program and Review Manager (Revman) version 5.4 Cochrane. Statistical analysis of the effect size data was performed using the Review Manager (Revman) application version 5.4 Cochrane. The reviewer extracted data from each study and if it was not available, a recalculation was carried out based on numerical data in the research journal, in the form of sensitivity, specificity, as well as the values of TP, FP, TN, and FN. In this study, the effect size assessed the pooled proportion of sensitivity and specificity with a 95% confidence interval. The results of the sensitivity and specificity conclusions was included in the ROC curve.

Discussion

Literature Search Results
Literature search was done in online journal databases which included Pubmed (96 literature), Cochrane Library (6 literature), Biomed Central (119 literature), ProQuest (1175 literature,) and Science Direct (7 literature) with a total of 1403 literatures which were the first hit according to with the MeSH terms used. Filtering through titles and abstracts and eliminating duplications obtained 76 filtered literatures. Then, we carried out a more in-depth assessment by selecting journals with prospective research designs that looked for True Positive, False Positive, True Negative, and False Negative values from the interpretation of preoperative MRI imaging compared with findings during surgery. From these 76 literatures, the full text and exclusion criteria were filtered, and 8 literatures that met the criteria for the synthesis of qualitative and quantitative data were obtained. The flow chart of the literature search results can be seen in Figure 1. A

summary of data from the 8 literatures included in the meta-analysis can be seen in Tables 3 and 4.
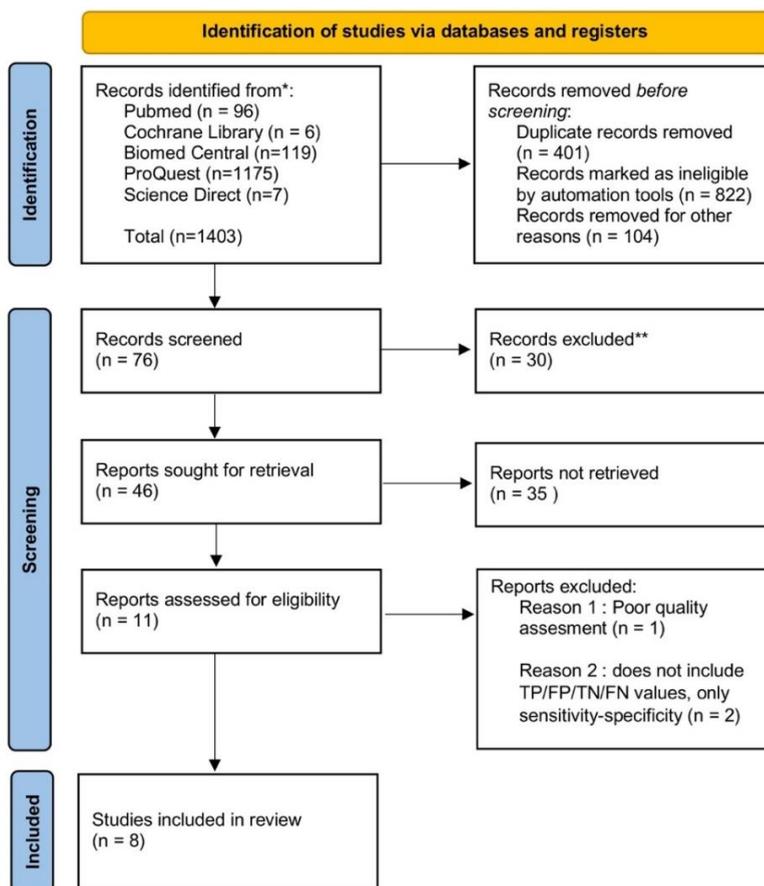
Figure 1: PRISMA 2020 flow diagram (Page MJ et al., 2020)

Table 4. Summary of literature included in the calculation of meta-analysis statistics

Description	Total
Total numbers of patients	551
- Conservative	163 (29,6%)
- Surgery	388 (70,4%)
Sex	
- Male	310 (56%)
- Female	241 (44%)
Mean Age (Range)	43 (18-83)
Total Number of Disc Levels Operated	586

Assessment of Study Quality and Risk of Bias

The summary results of the study quality assessment can be seen in table 5. In this meta-analysis, eight out of nine literatures were judged to have good quality (Good) and only one literature was deemed sufficient (Fair). The literature that has sufficient criteria is the research of Chawalparit et al. in (2006), where the assessment did not get an asterisk on the selection criteria because the number of discs studied was not described, and in the group selection criteria, there was no data showing the period between MRI and operating time, so the data could not be processed further to assess the relationship. Other assessments received an asterisk on all criteria and were representative enough to be included in the meta-analysis.

The possibility of bias in the reference plane is because some studies do not mention the type of surgery performed, making it difficult to equalize the type of surgery assessed in this meta-analysis. The possibility of bias in several studies in the field of Flow and Timing is influenced by the fact that not all studies mention the period between MRI and surgery time, so it was not possible to examine the relationship between the time of preoperative MRI and findings during surgery. A summary of the risk assessment of research bias can be seen in table 6 and graph 1.
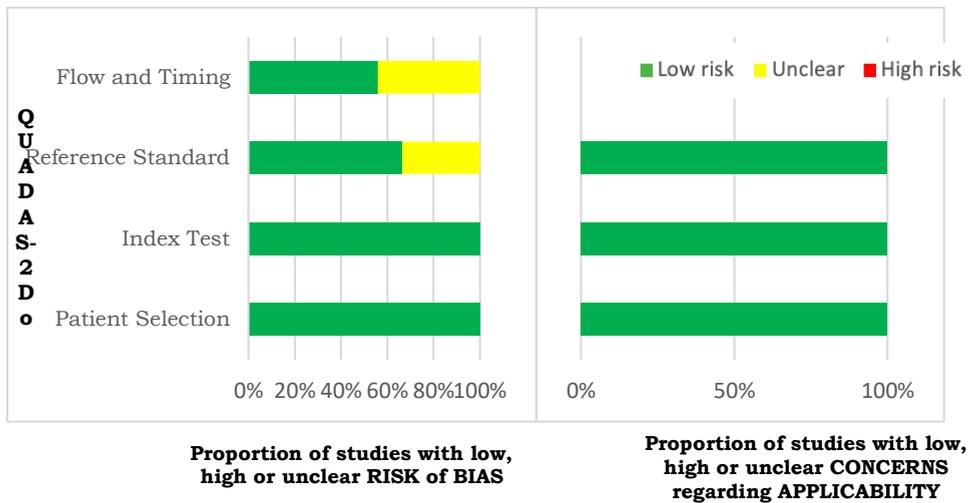
Graph 1. QUADAS-2 tools bias risk assessment chart

Table 6 : Assess the risk of bias with QUADAS-2 tools for research

Reference	Risk of Bias				Applicability Concerns		
	Patients Selection	Index Test	Reference Standard	Flow and Timing	Patients Selection	Index Test	Reference Standard
Bernard et al. 1994	+	+	+	+	+	+	+
Bischoff et al. 1993	+	+	+	?	+	+	+
Chawalparit et al. 2006	+	+	?	?	+	+	+
Forristal et al. 1988	+	+	?	?	+	+	+
Jackson et al. 1989	+	+	+	?	+	+	+
Thornbury et al. 1993	+	+	+	+	+	+	+
Pfirmann et al. 2004	+	+	+	?	+	+	+
Parmar et al. 2018	+	+	+	+	+	+	+

Description: + (*Low Risk*); ? (*Unclear*); − (*High Risk*)

MRI Accuracy of Surgical Findings

The MRI accuracy value of 1.5 Tesla contrast and non-contrast was assessed based on the concordance of the results of preoperative imaging interpretation (index test) which was confirmed with the current best diagnosis of HNP, namely the findings during surgery (reference standard). The data was then processed in a 2x2 table analysis to get the value of TP/FP/TN/FN which can produce sensitivity and specificity numbers in assessing the accuracy of a diagnostic test. The number entered was the number of disc levels examined for findings during surgery. Data is presented in a forest plot pooled prevalence. Eight studies with a total of 586 disc levels explored were included in this meta-analysis. The sensitivity and specificity ranged between 64-95% and 55-100% with wide confidence interval. The forest plot of the sensitivity and specificity of MRI accuracy can be seen in Figure 2.

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Bernard 1994	33	9	13	11	0.72 [0.57, 0.84]	0.55 [0.32, 0.77]
Bischoff 1993	25	10	10	27	0.71 [0.54, 0.85]	0.73 [0.56, 0.86]
Chawalparit 2006	19	3	4	7	0.83 [0.61, 0.95]	0.70 [0.35, 0.93]
Forristal 1988	22	0	2	7	0.92 [0.73, 0.99]	1.00 [0.59, 1.00]
Jackson 1989	38	8	21	53	0.64 [0.51, 0.76]	0.87 [0.76, 0.94]
Parmar 2018	30	2	15	25	0.67 [0.51, 0.80]	0.93 [0.76, 0.99]
Pfirrmann 2004	62	2	3	27	0.95 [0.87, 0.99]	0.93 [0.77, 0.99]
Thornbury 1993	7	4	4	7	0.64 [0.31, 0.89]	0.64 [0.31, 0.89]

Figure 2: Forest plot MRI diagnostic accuracy of 1.5 Tesla contrast and non-contrast. The range of sensitivity and specificity obtained ranged between 64-95% and 55-100% (95% confidence intervals).

The sensitivity and specificity values summarized from the literature can also represent the True Positive Rate / TPR (sensitivity) value, which is the proportion of samples with positive test results for the entire diseased population, and the False Positive Rate/FPR (1-specificity) value, i.e. samples with positive tests for the entire population without the disease. TPR and FPR were used to calculate the Receiver Operating Characteristic curve (ROC curve). ROC curve shows the meeting point (trade-off) between FPR (x-axis) and TPR (y-axis) with the law that the higher sensitivity value will be accompanied by a decrease in specificity. In this meta-analysis, the ROC curve can be seen in Figure 3, the curve appears close to the left and the top. From the ROC space, it shows that MRI has good accuracy (accurate).
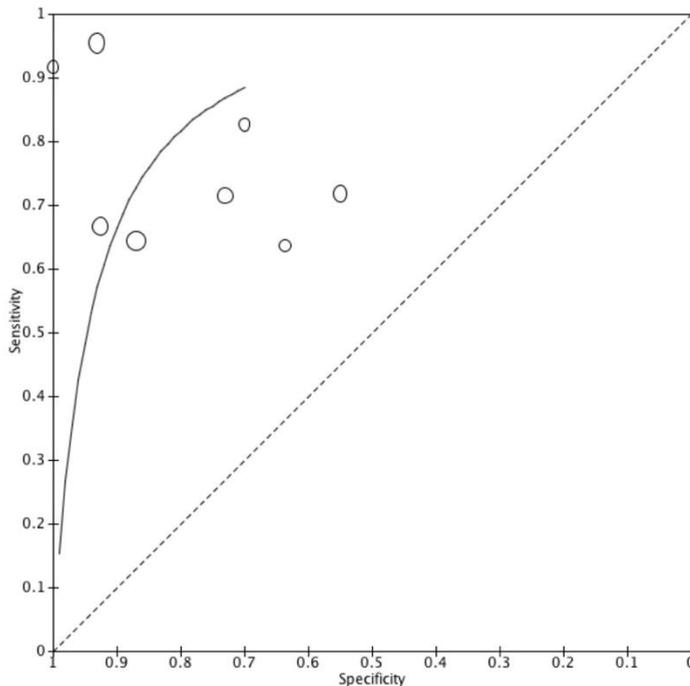
Figure 3: ROC curve MRI accuracy 1.5 Tesla contrast and non-contrast. The actual test curve indicates that MRI has good accuracy (accurate) because it is close to the top left corner point (perfect classifier).

Comparison of MRI and Other Imaging Accuracy
In this meta-analysis, there were literatures that examined the comparison of accuracy between CT myelography and MRI 1.5 Tesla contrast and non-contrast in the diagnosis of lumbar HNP, namely the study by Bischoff et al. 1993 and Thornbury et al. 1993. The range of sensitivity and specificity of CT myelography accuracy was between 73-78% and 73-75%, while for MRI it was 64-71% and 63-73%, respectively. The forest plot graph of a comparison of the sensitivity and specificity of the accuracy of CT myelography and MRI can be seen in Figure 4. The ROC curve of comparison of the accuracy between CT myelography and MRI 1.5 Tesla contrast and non-contrast can be seen in Figure 5. The CT myelography curve (red line) approaches the left and top sides of the ROC space more than the MRI curve (black line) which shows that CT myelography has a better accuracy

compared to MRI. An explanation of the things that affect the comparison of accuracy will be discussed further in the discussion chapter with a systematic review approach.

Akurasi MRI

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Bischoff 1993	25	10	10	27	0.71 [0.54, 0.85]	0.73 [0.56, 0.86]
Thornbury 1993	7	4	4	7	0.64 [0.31, 0.89]	0.64 [0.31, 0.89]

Akurasi CT Mielografi

Study	TP	FP	FN	TN	Sensitivity (95% CI)	Specificity (95% CI)
Bischoff 1993	28	9	8	27	0.78 [0.61, 0.90]	0.75 [0.58, 0.88]
Thornbury 1993	8	3	3	8	0.73 [0.39, 0.94]	0.73 [0.39, 0.94]

Figure 4 Forest plot comparison of the accuracy of CT myelography and MRI 1.5 Tesla contrast and non-contrast. The range of sensitivity and specificity of the accuracy of CT myelography was between 73-78% and 73-75%, while for MRI it was 64-71% and 63-73% (95% confidence intervals).
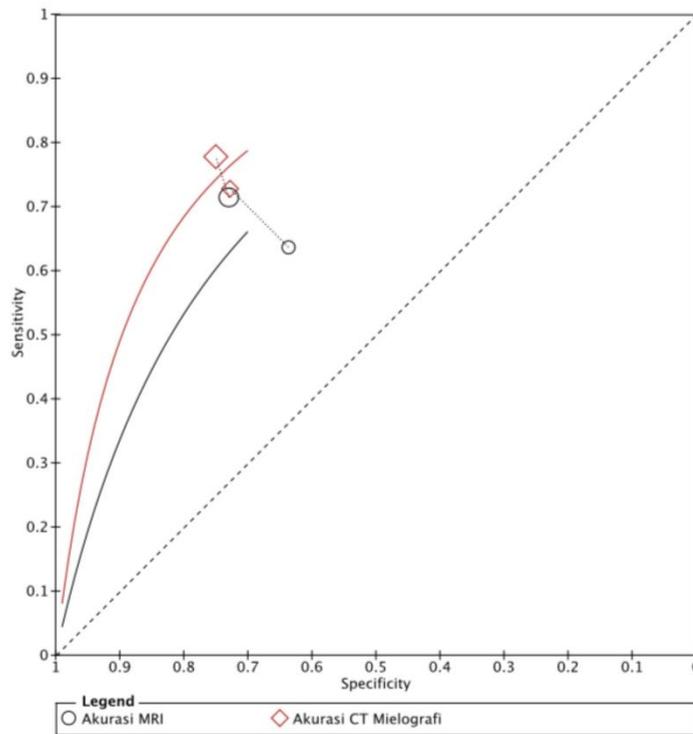
Figure 5 ROC curve comparison of accuracy between CT myelography and MRI 1.5 Tesla contrast and non-contrast. It can be seen that the CT myelography curve is closer to the upper left corner point (perfect classifier) than the MRI depicting better accuracy.

Discussion

Assessment of MRI accuracy is an interesting discussion because until now MRI is the imaging of choice for the most reliable diagnosis (reference standard) to date, namely findings during surgery (Vali Y et al., 2021). This study helps clinicians in the decision making of the operative action plans apart from the clinical examination of the patient because preoperative preparation is very influential to the outcome (Zileli M et al., 2020). A false positive (FP) value represents the surgery that is actually not needed by the patient and conversely, a high false-negative (FN) value reflects the preoperative examination is at risk of misdiagnosis so that the patient does not get the best treatment. The FN number has a more detrimental meaning for patients because it cannot screen groups of patients who have anatomic abnormalities and are indications for surgery (Dehmoobad Sharifabadi A et al., 2019; Lathyris D, Haidich AB et al., 2021).

Research Characteristics
The eight filtered journal literature were published between the year of 1988—2018, which shows that MRI imaging has always been a continuous evaluation material from time to time to study and evaluate its accuracy as a diagnostic tool. The journal literatures included in this meta-analysis have a global background of research locations. From a total of eight literatures, one study is from Georgia, one study is from Thailand, five studies are from the United States of America (USA), and one study is from Switzerland and one study is from India. The samples of these studies have been asked for consent to be included in the cohort study conducted. Some of the literature included in this meta-analysis lists the accuracy of MRI and other preoperative imaging (CT discography and CT myelography) whose data will be discussed separately in comparison with MRI examinations.

The eight literatures resulted in a total sample of 438 patients (68%) diagnosed with lumbar HNP who were operated between the total population of 641 patients, of which the remaining 203 patients (32%) of HNP were treated conservatively. Patients who were decided for surgery were patients who had received previous conservative therapy and did not have satisfactory results after therapy. There were 358 (56%) male and 283 (44%) female. The mean age was 43 years with the youngest age was 18 years and the oldest was 83 years. The total number of discs studied was 586 discs. Research by Chawalparit et al. and the study of Bernard et al. were the study with the largest number of samples involving 123 and 33 patients, respectively. This met the inclusion criteria which require the number of study samples to be more than 10 patients to gain more representative results. Male was the most common sex to suffer HNP in 7 of the 8 synthesized studies. This refers to several risk factors for HNP such as smoking history and a tendency to do a rough manual labor. The average age in all studies was the productive age of the adult population, which the risk increases with age (Bernard TN Jr., 1994; Vlaeyen JWS., 2018).

The MRI used in the eight synthesized studies was 1.5 Tesla contrast and non-contrast MRI. The magnetic strength of an MRI is directly proportional to the number of signals received from the human body during an MRI scan. Since signals from the body are then used to create the image, the higher the magnetic

strength, the more detailed the image, so the higher the magnetic strength, the better the detail in the image and the less the artifacts that appear Varlotta CG et al., 2020; Kasch R et al., 2022). The use of 1.5 Tesla MRI is superior in the diagnosis of pathological abnormalities in the spine due to the less possibility of artifacts on imaging duration due to bowel movements and breath than 3 Tesla MRI which is very sensitive to movements. The condition of the patient who is in pain during MRI is very susceptible to movements that causes artifacts.[33] Different MRI brands were also used in the eight studies so that in general there was no tendency for one brand to stand out more and the data obtained were not much different between several MRI brands. The MRI sequences used in the eight studies had uniformity, namely the T1WI and T2WI sequences in the axial sagittal section, so that the preoperative HNP interpretation had the same reference, especially in determining the location and size of the HNP in the disc (D'Aprile P et al., 2018).

The period between the MRI and the operating time affects the interpretation of the accuracy of the preoperative imaging with the findings during the operation because progressive HNP will produce different findings if the MRI is performed at a longer period before surgery (Benzakour T et al., 2019) In the eight studies assessed, there were only 4 studies that provided data regarding the period between the onset of preoperative MRI, 5 studies did not provide information on when MRI was performed. This is still one of the risks of bias in this meta-analysis. The longest period for MRI was reported to be 1 year and the shortest period was 1 month preoperative. Eight studies that were reviewed had a distribution of types of surgery, namely 7 open lumbar surgery with laminectomy discectomy and 1 study with endoscopic spine surgery discectomy. The operations were carried out with the same purpose. Discectomy was done to be able to see the HNP material clearly during the surgery so that a diagnosis could be made. The findings during surgery are the golden standard in establishing the diagnosis of HNP where the condition of the disc can be observed directly, so it was used as a reference standard in this meta-analysis of diagnostic accuracy.

Heterogeneity
Between eight studies that assessed the accuracy of 1.5 Tesla contrast and non-contrast MRI compared with the clinical findings of surgery, there is still a fairly high heterogeneity, which can be caused by different pathologies in each case, the brand, and specificity of the MRI, the different MRI techniques used by the radiologists, inter-individual interpretation, population and patient characteristics, and research methods used. This affects the accuracy of the results of each study which varies in this meta-analysis. Kim et al. In 2018, his systematic review that assessed the accuracy of CT scans and myelography on sciatica patients also obtained varying results because the characteristics of diagnostic tests as index tests have different brands, specificities, and techniques in each study studied (Kim JH et al., 2018).

1.5 Tesla Contrast and Non-Contrast MRI Imaging Accuracy
The use of MRI began to develop after 1980 until now. In this meta-analysis, there are 2 studies conducted more than 2 decades ago, namely research by Forristal 1988 and Jackson 1989, where at that time imaging and diagnostic technology were still not as advanced as today, so the findings of accuracy values need to be

considered and compared with the latest research. The findings obtained on imaging must be consistent with their usefulness in clinical practice. The benefit of MRI in identifying pathological abnormalities of the spine, especially the lumbar spine, depends on how big the role of MRI is in helping the decision-making for the next HNP management that affects the clinical outcome of the patient (Ekedahl H et al., 2018; Kasch R et al., 2022). This could be the role of MRI to exclude patients who have no anatomical abnormalities to avoid invasive surgical therapy (in this case avoiding false-positive findings) and vice versa to find and identify as many patients as possible with anatomic abnormalities where delaying surgical therapy could result in the poorer clinical outcome of patients (in this case avoiding false-negative numbers), so it is important to study and assess the accuracy of the selected preoperative imaging tool more deeply (Michelini G et al., 2018).

The results of the pooled prevalence of the study resulted in a sensitivity and specificity ranging from 64-95% and 55-100% with wide confidence intervals. The accuracy value illustrates that the MRI findings vary widely and are not absolute values if used as a guide to surgical management in patients. There are still false positive values that can occur so decision making should still consider the patient's history and clinical findings. The accuracy value can be assessed for quality from the ROC graph obtained from the True Positive Rate/TPR and False Positive Rate/FPR values, where the curve approaching the point in the upper left corner (perfect classifier) describes the good accuracy of MRI (Kim JH et al., 2018; Lathyris D, Haidich AB et al., 2021; Mander GTW and Munn Z., 2021).

1.5 Tesla Imaging Accuracy Compared to CT Myelography
This meta-analysis also analyzes the comparison of the accuracy of 1.5 Tesla contrast and non-contrast MRI with other imaging techniques, namely CT myelography. The sensitivity and specificity of CT myelography accuracy ranged between 73-78% and 73-75%, while for MRI it was 64-71% and 63-73%, respectively. In general, the comparison of accuracy seen on the ROC curve shows that CT myelography is more sensitive and specific in the diagnosis of HNP than MRI. Some of these things possibly caused by several factors, including: First, CT examination by inserting a contrast agent is more focused on anatomical pathology locations than MRI which relies on magnetic strength for detailed imaging; and Second, the MRI imaging equipment used at that time was probably not as advanced as it is today in terms of imaging technology and techniques so that in terms of accuracy it was lower than CT with contrast that was used and has been developed (Weisenthal BW et al., 2021). In line with the times and technology, MRI is becoming more and more often used because it is less invasive and comfortable for patients, as well as being a reliable imaging tool without many people knowing the accuracy range of previous studies. This makes MRI not re-examined for its accuracy of surgical findings because it has become a safer and more comfortable option for patients than CT imaging with contrast agents at the pathological location (Patel DM, Weinberg BD and Hoch MJ., 2020).

CT myelography is an important imaging modality that combines the advantages of myelography and the high resolution of CT. It provides a detailed description of pathological spinal conditions, especially those involving the dural sac and its contents (Weisenthal BW et al., 2021). However, the role of CT myelography has

decreased dramatically and precisely with the advent of MRI, which provides a non-invasive method for demonstrating pathological spinal conditions with high signal intensity in soft tissues. Currently, CT myelography is often performed on patients who require dural sac evaluation but have contraindications for undergoing MRI such as the presence of a pacemaker, metallic foreign body in the eye, deep brain stimulator device, foreign body in the body made of metal, and close to vital organs and cerebral clip (Kato S et al., 2019; Patel DM, Weinberg BD and Hoch MJ., 2020; Price DB and Ortiz AO., 2017).

Research Advantages

This study is the first systematic review and meta-analysis to assess the accuracy of 1.5 Tesla contrast and non-contrast MRI in the diagnosis of lumbar HNP. The journals were synthesized from 2 decades ago to the latest research to provide an objective picture of the conclusions about the accuracy of MRI from time to time. The selected studies came from various parts of the world with different population characteristics based on race and country so that the results of the analysis can describe the global conditions in the use of MRI in the world.

Research Limitations

The results of this study cannot be used as a conclusive conclusion for spinal pathological abnormalities because the research conducted is limited to herniation at the lumbar level, therefore, these results cannot be generalized to other pathological abnormalities at the spinal level. The research synthesized in this meta-analysis would be better and more complete if it had the latest research sources in the last 5 years, but from the results of a search for journals in the database, there was no recent primary study that re-examined the accuracy of MRI. This trend is possible because until now there is no preferred imaging examination over MRI. After all, some examinations such as CT discography/myelography tend to be invasive. This can cause clinicians to trust the results of MRI imaging without wanting to examine its accuracy more deeply.

Conclusion

1.5 Tesla contrast and non-contrast MRI has a good accuracy in the diagnosis of lumbar HNP concerning surgical findings, assessed from the Receiver Operating Characteristic curve with a large Area Under Curve above the threshold which represents the relationship between sensitivity and specificity. CT myelography examination is more sensitive and specific than 1.5 Tesla contrast and non-contrast MRI. The use of CT myelography is not routine compared to MRI because it is more invasive and uncomfortable for the patient. CT myelography can be used as an alternative in patients with MRI contraindications.

Suggestion; 1) Research and analysis of the accuracy of MRI in the diagnosis of spinal pathological abnormalities at the cervical and thoracal levels to complete the accuracy data. 2) The study of the accuracy of the 3 Tesla contrast and no contrast MRI diagnostic test in the diagnosis of HNP compared with the surgical findings can be used as an evaluation and comparison with 1.5 Tesla MRI. 3) Regular seminars and workshops on techniques and interpretation of MRI imaging of the spine to increase knowledge and discuss the latest issues regarding the development of MRI accuracy.

Acknowledgments

References

Bajamal AH, Permana KR, Faris M, Zileli M, Peev NA. Classification and Radiological Diagnosis of Thoracolumbar Spine Fractures: WFNS Spine Committee Recommendations. Neurospine. 2021 Dec;18(4):656-666. doi: 10.14245/ns.2142650.325.

Benzakour T, Igoumenou V, Mavrogenis AF, Benzakour A. Current concepts for lumbar disc herniation. Int Orthop. 2019 Apr;43(4):841-851. doi: 10.1007/s00264-018-4247-6.

Bernard TN Jr. Using computed tomography/discography and enhanced magnetic resonance imaging to distinguish between scar tissue and recurrent lumbar disc herniation. Spine (Phila Pa 1976). 1994 Dec 15;19(24):2826-32. doi: 10.1097/00007632-199412150-00017.

Bischoff RJ, Rodriguez RP, Gupta K, Righi A, Dalton JE, Whitecloud TS. A comparison of computed tomography-myelography, magnetic resonance imaging, and myelography in the diagnosis of herniated nucleus pulposus and spinal stenosis. J Spinal Disord. 1993 Aug;6(4):289-95. doi: 10.1097/00002517-199306040-00002.

Chawalparit O, Churojana A, Chiewvit P, Thanapipatsir S, Vamvanij V, Charnchaowanish P. The limited protocol MRI in diagnosis of lumbar disc herniation. J Med Assoc Thai. 2006 Feb;89(2):182-9.

D'Aprile P, Nasuto M, Tarantino A, Cornacchia S, Guglielmi G, Jinkins JR. Magnetic Resonance Imaging in degenerative disease of the lumbar spine: Fat Saturation technique and contrast medium. Acta Biomed. 2018 Jan 19;89(1-S):208-219. doi: 10.23750/abm.v89i1-S.7024.

Dehmoobad Sharifabadi A, Leeflang M, Treanor L, Kraaijpoel N, Salameh JP, Alabousi M, Asraoui N, Choo-Foo J, Takwoingi Y, Deeks JJ, McInnes MDF. Comparative reviews of diagnostic test accuracy in imaging research: evaluation of current practices. Eur Radiol. 2019 Oct;29(10):5386-5394. doi: 10.1007/s00330-019-06045-7.

Ekedahl H, Jönsson B, Annertz M, Frobell RB. Accuracy of Clinical Tests in Detecting Disk Herniation and Nerve Root Compression in Subjects With Lumbar Radicular Symptoms. Arch Phys Med Rehabil. 2018 Apr;99(4):726-735. doi: 10.1016/j.apmr.2017.11.006.

Forristall RM, Marsh HO, Pay NT. Magnetic resonance imaging and contrast CT of the lumbar spine. Comparison of diagnostic methods and correlation with surgical findings. Spine (Phila Pa 1976). 1988 Sep;13(9):1049-54. doi:

10.1097/00007632-198809000-00013.

Global burden of 369 diseases and injuries in 204 countries and territories, 1990-2019: a systematic analysis for the Global Burden of Disease Study 2019. Lancet. 2020 Oct 17;396(10258):1204-1222. doi: 10.1016/S0140-6736(20)30925-9. Erratum in: Lancet. 2020 Nov 14;396(10262):1562.

Hebelka H, Lagerstrand K, Brisby H, Owen PJ, Quittner MJ, Rantalainen T, Belavy DL. The importance of level stratification for quantitative MR studies of lumbar intervertebral discs: a cross-sectional analysis in 101 healthy adults. Eur Spine J. 2019 Sep;28(9):2153-2161. doi: 10.1007/s00586-019-06059-1.

Hutchins TA, Peckham M, Shah LM, Parsons MS, Agarwal V, Boulter DJ, Burns J, Cassidy RC, Davis MA, Holly LT, Hunt CH, Khan MA, Moritani T, Ortiz AO, O'Toole JE, Powers WJ, Promes SB, Reitman C, Shah VN, Singh S, Timpone VM, Corey AS. ACR Appropriateness Criteria® Low Back Pain: 2021 Update. J Am Coll Radiol. 2021 Nov;18(11S):S361-S379. doi: 10.1016/j.jacr.2021.08.002.

Jackson RP, Cain JE Jr, Jacobs RR, Cooper BR, McManus GE. The neuroradiographic diagnosis of lumbar herniated nucleus pulposus: II. A comparison of computed tomography (CT), myelography, CT-myelography, and magnetic resonance imaging. Spine (Phila Pa 1976). 1989 Dec;14(12):1362-7. doi: 10.1097/00007632-198912000-00013.

Jiang X, Chen D. Magnetic resonance imaging analysis of work-related chronic low back pain: comparisons of different lumbar disc patterns. J Pain Res. 2018 Oct 31;11:2687-2698. doi: 10.2147/JPR.S162988.

Johnson SM, Shah LM. Imaging of Acute Low Back Pain. Radiol Clin North Am. 2019 Mar;57(2):397-413. doi: 10.1016/j.rcl.2018.10.001.

Kasch R, Truthmann J, Hancock MJ, Maher CG, Otto M, Nell C, Reichwein N, Bülow R, Chenot JF, Hofer A, Wassilew G, Schmidt CO. Association of Lumbar MRI Findings with Current and Future Back Pain in a Population-based Cohort Study. Spine (Phila Pa 1976). 2022 Feb 1;47(3):201-211. doi: 10.1097/BRS.0000000000004198.

Kato S, Demura S, Matsubara H, Inaki A, Shinmura K, Yokogawa N, Murakami H, Kinuya S, Tsuchiya H. Utility of bone SPECT/CT to identify the primary cause of pain in elderly patients with degenerative lumbar spine disease. J Orthop Surg Res. 2019 Jun 20;14(1):185. doi: 10.1186/s13018-019-1236-4.

Kim JH, van Rijn RM, van Tulder MW, Koes BW, de Boer MR, Ginai AZ, Ostelo RWGJ, van der Windt DAMW, Verhagen AP. 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 Aug 21;26:37. doi: 10.1186/s12998-018-0207-x.

Kusmiati T, Narendrani HP. 'POTT'S disease'. *Jurnal Respirasi*. 2019. 2(3), pp. 99–109. doi: 10.20473/jr.v2-I.3.2016.99-109.

Lathyris D, Haidich AB. Meta-analysis of diagnostic test accuracy: Just another meta-analysis? Intensive Crit Care Nurs. 2021 Jun;64:103028. doi: 10.1016/j.iccn.2021.103028.

Mander GTW, Munn Z. Understanding diagnostic test accuracy studies and systematic reviews: A primer for medical radiation technologists. J Med Imaging Radiat Sci. 2021 Jun;52(2):286-294. doi: 10.1016/j.jmir.2021.02.005.

Michelini G, Corridore A, Torlone S, Bruno F, Marsecano C, Capasso R, Caranci F, Barile A, Masciocchi C, Splendiani A. Dynamic MRI in the evaluation of the spine: state of the art. Acta Biomed. 2018 Jan 19;89(1-S):89-101. doi:

10.23750/abm.v89i1-S.7012.

Page MJ, McKenzie JE, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, Shamseer L, Tetzlaff JM, Akl EA, Brennan SE, Chou R, Glanville J, Grimshaw JM, Hróbjartsson A, Lalu MM, Li T, Loder EW, Mayo-Wilson E, McDonald S, McGuinness LA, Stewart LA, Thomas J, Tricco AC, Welch VA, Whiting P, Moher D. The PRISMA 2020 statement: an updated guideline for reporting systematic reviews. BMJ. 2021 Mar 29;372:n71. doi: 10.1136/bmj.n71.

Parmar J, Gulati Y, Vora M, Patel B, Mohan C. Accuracy of the kissing sign on lumbar spine MRI in cases of axillary disc herniation and the surgical correlation: an Indian multi-center study. Rev Bras Ortop. 2018 Feb 6;53(6):681-686. doi: 10.1016/j.rboe.2017.10.011.

Patel DM, Weinberg BD, Hoch MJ. CT Myelography: Clinical Indications and Imaging Findings. Radiographics. 2020 Mar-Apr;40(2):470-484. doi: 10.1148/rg.2020190135.

Pfirrmann CW, Dora C, Schmid MR, Zanetti M, Hodler J, Boos N. MR image-based grading of lumbar nerve root compromise due to disk herniation: reliability study with surgical correlation. Radiology. 2004 Feb;230(2):583-8. doi: 10.1148/radiol.2302021289.

Phalke VV, Gujar S, Quint DJ. Comparison of 3.0 T versus 1.5 T MR: imaging of the spine. Neuroimaging Clin N Am. 2006 May;16(2):241-8, ix. doi: 10.1016/j.nic.2006.02.005.

Price DB, Ortiz AO. Myelography: From Lipid-Based to Gadolinium-Based Contrast Agents. Magn Reson Imaging Clin N Am. 2017 Nov;25(4):713-724. doi: 10.1016/j.mric.2017.06.005.

Rosyid AN, Yamin M, Puspitasari AD. The role of imaging in the diagnosis of pulmonary embolism. Biomolecular and Health Science Journal. 2019. 2(1), pp. 57–62. doi: 10.20473/bhsj.v2i1.13281.

Suryasa, I. W., Rodríguez-Gámez, M., & Koldoris, T. (2021). Get vaccinated when it is your turn and follow the local guidelines. *International Journal of Health Sciences, 5*(3), x-xv. https://doi.org/10.53730/ijhs.v5n3.2938

Sitch AJ, Dekkers OM, Scholefield BR, Takwoingi Y. Introduction to diagnostic test accuracy studies. Eur J Endocrinol. 2021 Feb;184(2):E5-E9. doi: 10.1530/EJE-20-1239.

Thornbury JR, Fryback DG, Turski PA, Javid MJ, McDonald JV, Beinlich BR, Gentry LR, Sackett JF, Dasbach EJ, Martin PA. Disk-caused nerve compression in patients with acute low-back pain: diagnosis with MR, CT myelography, and plain CT. Radiology. 1993 Mar;186(3):731-8. doi: 10.1148/radiology.186.3.8267688.

Vali Y, Yang B, Olsen M, Leeflang MMG, Bossuyt PMM. Reporting of test comparisons in diagnostic accuracy studies: A literature review. Res Synth Methods. 2021 May;12(3):357-367. doi: 10.1002/jrsm.1469.

Vargas MI, Boto J, Meling TR. Imaging of the spine and spinal cord: An overview of magnetic resonance imaging (MRI) techniques. Rev Neurol (Paris). 2021 May;177(5):451-458. doi: 10.1016/j.neurol.2020.07.005.

Varlotta CG, Ge DH, Stekas N, Frangella NJ, Manning JH, Steinmetz L, Vasquez-Montes D, Errico TJ, Bendo JA, Kim YH, Stieber JR, Varlotta G, Fischer CR, Protopsaltis TS, Passias PG, Buckland AJ. MRI Radiological Predictors of Requiring Microscopic Lumbar Discectomy After Lumbar Disc Herniation. Global Spine J. 2020 Feb;10(1):63-68. doi: 10.1177/2192568219856345.

Vlaeyen JWS, Maher CG, Wiech K, Van Zundert J, Meloto CB, Diatchenko L,

Battié MC, Goossens M, Koes B, Linton SJ. Low back pain. Nat Rev Dis Primers. 2018 Dec 13;4(1):52. doi: 10.1038/s41572-018-0052-1.

Weisenthal BW, Glassman SD, Mkorombindo T, Nelson L, Carreon LY. When does CT-Myelography add value beyond MRI for lumbar degenerative disease? Spine J. 2021 Nov 27:S1529-9430(21)01050-0. doi: 10.1016/j.spinee.2021.11.016.

Widana, I.K., Sumetri, N.W., Sutapa, I.K., Suryasa, W. (2021). Anthropometric measures for better cardiovascular and musculoskeletal health. *Computer Applications in Engineering Education*, *29*(3), 550–561. https://doi.org/10.1002/cae.22202

Wu A, March L, Zheng X, Huang J, Wang X, Zhao J, Blyth FM, Smith E, Buchbinder R, Hoy D. Global low back pain prevalence and years lived with disability from 1990 to 2017: estimates from the Global Burden of Disease Study 2017. Ann Transl Med. 2020 Mar;8(6):299. doi: 10.21037/atm.2020.02.175.

Yates M, Oliveira CB, Galloway JB, Maher CG. Defining and measuring imaging appropriateness in low back pain studies: a scoping review. Eur Spine J. 2020 Mar;29(3):519-529. doi: 10.1007/s00586-019-06269-7.

Zileli M, Sharif S, Fornari M, Ramani P, Jian F, Fessler R, Kim SH, Takami T, Shimokawa N, Dechambenoit G, Qureshi M, Konovalov N, Masini M, Osorio-Fonseca E, Sanchez JAS, Bajamal AH, Parthiban J, Sih IM, Alves ÓL, Oertel J, Rasulic L, Costa F, Peul WC, Sharma K, Eldin MM, Ismail NJ, Esene IN, Hossain M, Kalevski S, Hausmann ON, Yaman O, Arif S, Brady Z. History of Spinal Neurosurgery and Spine Societies. Neurospine. 2020 Dec;17(4):675-694. doi: 10.14245/ns.2040622.311.

Reference	Study Design	Total Patients / Sex / Mean Age (Range)	MRI Specifications	MRI Sequence	Disc Characteristics	Disc Patology	MRI to Surgery Time	Surgery	Interpretators	Imaging	True Positive	False Positive	True Negative	False Negative	Sensitifity	Specificity
(Forristall, Marsh and Pay, 1988)[9]	Cohort	Total 32 patient/25 surgery; 25 Male; 7 Female; 45 (22-74)	1.5 Tesla MRI General Electric SIGNA	Axial Sagital T1WI T2WI	Lumbar disc, not specified; 31 level disc	HNP, not specified	-	Open lumbal surgery, not specified	1 Neuroradiologist; 1 Spine orthopedist	MRI	22	0	7	2	92%	100%
(Jackson et al., 1989)[13]	Cohort	59 patients; 33 Male; 26 Female; 39 (18-70)	1.5 Tesla MRI General Electric SIGNA	Axial Sagital T1WI T2WI	VL3-4 VL4-5 VL5-S1; 120 level disc	Protrusion Ekstrusion Sequestrasion	-	Laminectomy-disectomy	2 Neuroradiologists; 2 Spine surgeons	MRI	38	8	53	21	64%	87%
(Bischoff et al., 1993)[4]	Cohort	57 patients; 29 Male; 28 Female; 42 (20-79)	1.5 Tesla MRI General Electric Signa	Axial Sagital T1WI T2WI	Lumbar disc, not specified; 72 level disc	HNP, not specified	-	Laminectomy disectomy	2 Neuroradiologists; 2 Spine Orthopedists	MRI	25	10	27	10	71%	73%
										CT-Mielography	28	9	27	8	77%	75%
(Thornbury et al., 1993)[31]	Cohort	95 patients; 61 Male; 34 Female; 39 (21-72)	1.5 Tesla MRI General Electric SIGNA	Axial Sagital T1WI T2WI	Lumbar disc, not specified; 22 level disc	HNP, not specified	>6 months	Laminectomy-chemnucleolysis	2 Neuroradiologists; 1 Neurosurgeon; 1 Orthopedists	MRI	7	4	7	4	64%	64%
										CT Mielography	8	3	8	3	72.7%	72.7%
(Bernard, 1994)[3]	Cohort	33 patients; 20 Male; 13 Female; 50 (23-74)	1.5 Tesla No brand listed	Axial Sagital T1WI/+C; Sagital T2WI	VL4-5 VL5-S1; 66 level disc	HNP, not specified	Mean 2.8 months; Range 1-9 months	Laminotomy-Laminectomy + Disectomy	2 Neuroradiologists; 1 Spine surgeon	MRI	33	9	11	13	72%	55%
										CT-Discography	34	8	12	12	73.9%	60%
(Pfirrmann et al., 2004)[26]	Cohort	80 patients; 48 Male; 32 Female; 46 (29-83)	1.5 Tesla MRI Symphony Siemens	Sagital T1WI T2WI	Lumbar disc, No brand listed; 94 level disc	HNP, not specified	-	Laminectomy-disectomy	2 Neuroradiologists; 3 Spine orthopedists	MRI	62	2	27	3	95%	93%
(Chawalparit et al., 2006)[5]	Cohort	123 patients; 61 Male; 62 Female; 42 (21-60)	1.5 Tesla MRI Philip ASC II	Axial Sagital T1WI T2WI	Lumbar disc, not specified; 33 level disc	HNP, not specified	-	Open lumbal surgery, not specified	3 Neuroradiologists; 2 Spine surgeons	MRI	19	3	7	4	83%	70%
(Parmar et al., 2018)[23]	Cohort	72 patients; 33 Male; 39 Female; 47 (31-75)	1.5 Tesla MRI Siemens Magnetom	Axial Sagital T1WI T2WI	Lumbar Disc, not specified; 72 level disc	HNP, not specified	>1 year	Endoscopic Spine Surgery	1 Neuroradiologist; 2 Spine surgeon	MRI	30	2	25	15	67%	93%

Table 3 above. Table of literature characteristics included in the calculation of meta-analysis statistics

Reference	Study Design	Quality Assessment Form for Cohort Studies								Quality (AHRQ)
		Selection				Comparability	Outcome			
		Representativeness of the exposed cohort	*Selection of the non-exposed cohort*	*Ascertainment of exposure*	*No Endpoint*	*Analysis controlled for confounders*	*Assessment outcome*	*follow-up long enough for outcomes*	*Adequacy of follow-up*	
Bernard *et al.* 1994	Cohort	Truly representative (*)	Same community as the exposed cohort (*)	Surgical record (*)	Yes (*)	The study controls for age, sex, surgical findings (*)	*Independent* (*)	Yes (*)	Complete Follow-up (*)	(*Good*)
Bischoff et al. 1993	Cohort	Truly representative (*)	No description of the derivation of the non exposed cohort	Surgical record (*)	Yes (*)	The study controls for age, sex, surgical findings (*)	*Independent* (*)	Yes (*)	Complete Follow-up (*)	(*Good*)
Chawalparit *et al.* 2006	Cohort	Selected Group	No description of the derivation of the non exposed cohort	Surgical record (*)	Yes (*)	Cohorts are not comparable	*Independent* (*)	Yes (*)	Complete Follow-up (*)	(*Fair*)
Forristal *et al.* 1988	Cohort	Truly representative (*)	No description of the derivation of the non exposed cohort	Surgical record (*)	Yes (*)	The study controls for age, sex, surgical findings (*)	*Independent* (*)	Yes (*)	Complete Follow-up (*)	(*Good*)
Jackson *et al.* 1989	Cohort	Truly representative (*)	No description of the derivation of the non exposed cohort	Surgical record (*)	Yes (*)	The study controls for age, sex, surgical findings (*)	*Independent* (*)	Ya (*)	Complete Follow-up (*)	(*Good*)
Thornbury *et al.* 1993	Cohort	Truly representative (*)	Same community as the exposed cohort (*)	Surgical record (*)	Yes (*)	The study controls for age, sex, surgical findings (*)	*Independent* (*)	Yes (*)	Complete Follow-up (*)	(*Good*)
Pfirmann *et al.* 2004	Cohort	Truly representative (*)	No description of the derivation of the non exposed cohort	Surgical record (*)	Yes (*)	The study controls for age, sex, surgical findings (*)	*Independent* (*)	Yes (*)	Complete Follow-up (*)	(*Good*)
Parmar *et al.* 2018	Cohort	Truly representative (*)	Same community as the exposed cohort (*)	Surgical record (*)	Yes (*)	The study controls for age, sex, surgical findings (*)	*Independent* (*)	Yes (*)	Complete Follow-up (*)	(*Good*)

(*) = Asterisk for the Newcastle-Ottawa Quality Assessment Form for Cohort Studies

Table 5 above. Study quality analysis using the Newcastle-Ottawa Quality Assessment Form for Cohort Studies