



Real-world study for identifying the predictive factors of surgical intervention and the value of magnetic resonance imaging in patients with low back pain

Hui Wang¹, Chang Liu², Zhou Meng³, Wenxian Zhou¹, Tao Chen¹, Kai Zhang⁴, Aimin Wu¹

¹Department of Orthopedics, The Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, Zhejiang Provincial Key Laboratory of Orthopedics, Wenzhou, China; ²Institute for Musculoskeletal Health, The University of Sydney and Sydney Local Health District, Camperdown, Sydney, Australia; ³Department of Medicine, University of Maryland Medical Center Midtown Campus, Baltimore, MD, USA; ⁴Department of Orthopedics, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai Key Laboratory of Orthopedic Implants, Shanghai, China

Contributions: (I) Conception and design: H Wang, K Zhang, A Wu; (II) Administrative support: K Zhang, A Wu; (III) Provision of study materials or patients: H Wang, Z Meng, W Zhou; (IV) Collection and assembly of data: C Liu, Z Meng, T Chen; (V) Data analysis and interpretation: H Wang, W Zhou; (VI) Manuscript writing: All authors; (VII) Final approval of manuscript: All authors.

Correspondence to: Ai-Min Wu. Department of Orthopaedics, The Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University, Zhejiang Provincial Key Laboratory of Orthopaedics, 109 Xueyuan Road, Wenzhou 325027, China. Email: aiminwu@wmu.edu.cn; Kai Zhang. Department of Orthopedics, Shanghai Ninth People's Hospital, Shanghai Jiao Tong University School of Medicine, Shanghai Key Laboratory of Orthopedic Implants, 639 Zhizaoju Road, Shanghai 200011, China. Email: mmc2008@163.com.

Background: Low back pain (LBP) is a prevalent disease and can be disabling. Currently, many patients with LBP with or without radiculopathy commonly undergo magnetic resonance imaging (MRI) for diagnosis and therapeutic assessment, yet the final intervention is mainly centered around nonoperative treatment. This study's aim was to identify the predictive factors of surgical treatment and the value of MRI in patients with LBP with or without radiculopathy.

Methods: The study included a training cohort that consisted of 461 patients with MRI from January 2014 to December 2018. Demographic characteristics and MRI findings were collected from our medical records. We developed and validated 2 nomograms to predict the possibility of receiving surgical treatment in LBP patients, based on multivariable logistic regression analysis. The performance of the 2 nomograms was assessed in terms of their calibration, discrimination, and clinical usefulness. An independent validation cohort containing 163 patients was comparatively analyzed.

Results: The baseline model incorporated 6 clinicopathological variables, while the MRI model consisted of 9 variables including several MRI findings. Internal validation revealed the good performance of the 2 nomograms in discrimination and calibration, with a concordance index (C-index) of 0.799 (95% CI: 0.743–0.855) for the baseline model and 0.834 (95% CI: 0.783–0.884) for the MRI model, which showed that the addition of MRI findings to the nomogram failed to achieve better prognostic value (Z statistic = -1.509; P=0.131). Application of the 2 models in the validation cohort also showed good discrimination (baseline model: C-index 0.75, 95% CI: 0.671–0.829; MRI model: C-index 0.777, 95% CI: 0.696–0.857) and calibration. No significant predictive benefit was found in the MRI model in the validation cohort (Z statistic = -0.588; P=0.557).

Conclusions: This study showed that clinical demographic characteristics provide good prognostic value to determine whether LBP patients with or without radiculopathy require surgical treatment. The addition of MRI findings yielded no significantly incremental prognostic value.

Keywords: Low back pain (LBP); magnetic resonance imaging (MRI); nomogram; prediction.

Submitted Jun 02, 2021. Accepted for publication Nov 04, 2021.

doi: 10.21037/qims-21-584

View this article at: <https://dx.doi.org/10.21037/qims-21-584>

Introduction

Low back pain (LBP) is one of the most common reasons for people visiting the clinic or being absent from work. It is reported that the global life disability increased by 54% between 1990 and 2015 due to LBP (1). According to demographic statistics, at least 80% of the population experience LBP at some point in their lives. Among them, approximately 10% experience chronic and long-term persistent pain, leading to deteriorative physical disability and low quality of life (2). Although there is a high prevalence of LBP, its treatment is mainly focused on nonsurgical intervention. For LBP patients who are suspected of developing vertebral body lesions or nerve root impairment, advanced imaging examinations such as lumbar radiography, computerized tomography (CT), and magnetic resonance imaging (MRI) are highly recommended (3), as they can provide detailed anatomic resolution and pathologic information for surgeons to make accurate diagnosis and treatment decisions. Nevertheless, diagnostic imaging is not commonly applied for LBP patients, since the main type of LBP is nonspecific pain and can be relieved after physical therapy or pharmacological treatment. Moreover, patients who are informed of abnormal imaging findings can easily develop an increased sense of dissatisfaction, which may result in redundant complaints and overtreatment (4). Furthermore, frequent radiological examinations can also incur various side effects, including excessive radiation exposure and medical consumption. This not only damages the patient's physical body but also increases the socioeconomic burden placed upon them (5). Research has also reported that the vast majority of people with LBP do not display any imaging abnormalities in the spine, while abnormalities could occur in some asymptomatic people (6-8), suggesting that there may be no significant correlation between abnormal imaging findings and clinical symptoms. Therefore, routine imaging tests are not commonly recommended for patients with LBP.

The application of diagnostic imaging can be understood in several situations, including poor improvement of symptoms after 6 weeks of pharmaceutical administration or rehabilitation exercise, or the detection of any red flags found in morphologic natural history or physical examination, such as unexplained weight loss, history of

cancer, significant trauma, cauda equina syndrome, and severe neurologic compromise (9). However, retrospective analyses of clinical practices have revealed that excessive imaging examinations are often performed in LBP patients without the above characteristics. In addition, for some LBP patients with indicators, imaging abnormalities related to the symptoms can be identified, yet the final treatments are still focused on nonsurgical treatment including physical therapy or pharmacological treatment. Therefore, for these patients, imaging examinations are considered to have limited value because of their lack of effect in the overall therapeutic process (10).

Although various guidelines have stated the clinical implications of image examination in a series of diseases, a consensus has not thus far been reached regarding the practicability of imaging examinations for LBP patients. To date, the demographic characteristics of patients whose imaging findings display abnormality of the lumbar area for whom the final treatments remain unaffected have not been clearly studied. An explicit understanding of the above population is critical for the rational application of lumbar imaging. The aims of this study were as follows: (I) to compare the clinical demographic characteristics and imaging findings in LBP patients with or without surgical therapy after imaging; (II) to create and validate 2 models that included clinical baseline data and MRI findings, respectively; and (III) to develop nomograms for surgical prediction. We present the following article in accordance with the TRIPOD reporting checklist (available at <https://dx.doi.org/10.21037/qims-21-584>).

Methods

Study design and participants

This retrospective study was conducted at the Second Affiliated Hospital and Yuying Children's Hospital of Wenzhou Medical University. The training cohort comprised LBP patients (18–80 years) with or without radiculopathy who attended the hospital between January 2014 and December 2018 according to the medical records. The exclusion criteria were as follows: (I) no imaging was performed during the episode of LBP, (II) the patient had

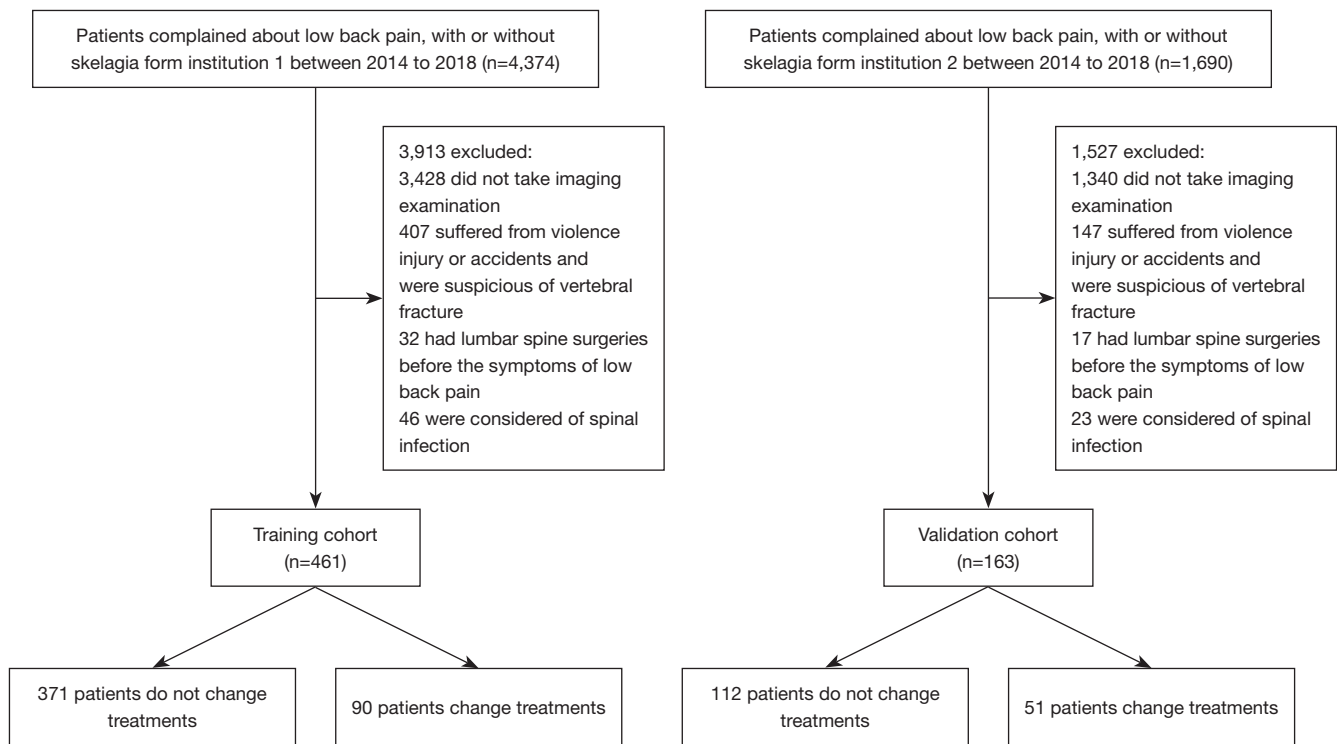


Figure 1 Study flow chart of participants.

experienced a direct or indirect attack on the lumbar spine and was suspected of lumbar vertebral fracture, and (III) the patient underwent previous lumbar surgery and had residual LBP. Patients with unknown survival information or surgical treatment were also excluded from this study. In total, 461 patients were included in the training cohort. At the same time, an independent validation cohort comprising 163 patients was enrolled at the First Affiliated Hospital of Wenzhou Medical University based on the same inclusion and exclusion criteria. All included patients were diagnosed and treated by the experienced senior surgeons of spine surgery from the 2 institutions, according to the *Diagnosis and Treatment of Low Back Pain* by Chou *et al.* (11). Pharmacological and nonpharmacological therapies were prescribed for the initial administration of LBP, and MRI was applied to detect severe lumbar disorders or minor alleviation of back pain after a 6-week course of pharmacological therapy. Surgical therapies were decided upon both by surgeons and patients and were based on the patients' clinical characteristics and MRI findings. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the institutional review board of our hospital and individual

consent for this retrospective analysis was waived (*Figure 1*).

Data collection and cohort definition

To evaluate the rate of receiving surgical treatment in LBP patients, the therapeutic processes within 1 year after imaging were collected and analyzed. Demographic and clinical data, along with MRI findings, were collected from the institutional picture archiving and communication system (PACS; Toshiba, Tokyo, Japan) medical records. Patients' clinical parameters included gender, age, duration of pain, leg pain, restricted waist movement, lameness, restricted mobility, and straight leg raise test. All the participants were divided into surgery and no-surgery groups based on the final therapeutic methods. The surgery group referred to participants who had undergone MRI within 1 year, and the therapeutic processes had changed from the initial medical treatment or rehabilitation treatment to surgical treatment. The no-surgery group referred to participants who had persisted with pharmacotherapy or physical therapy during the period of LBP within 1 year after MRI.

The 10-item Oswestry disability index (ODI) has been widely developed to assess pain-related disability in people

with acute, subacute, or chronic LBP (12). The ODI includes 1 item on pain and 9 items on activities of daily living, in which each item is rated with a score from 0 (high functioning) to 5 (low functioning) on a 6-point ordinal scale. The 5 levels of disability were calculated based on a percentage score of the ODI ranging from 0 to 100: minimal disability (0 to 20%), moderate disability (21% to 40%), severe disability (41% to 60%), house bound (61% to 80%), and bed bound or exaggerated symptoms (81% to 100%) (12).

The 11-point numeric rating scale (NRS) is the most common scale for measuring acute and chronic pain, as it can provide significantly discriminative power for describing pain intensity in LBP patients (13). In this study, both ODI and NRS scores were measured to describe low back and leg pain at the time after MRI imaging.

Furthermore, to analyze the difference of MRI findings in the surgery and no-surgery groups, the lumbar MRI of each participant was investigated and interpreted by the radiologists on duty at the time of the examination. In addition, all the imaging information was made available to both the physicians and patients for the following treatment decisions. In this study, 2 experienced radiologists who were blinded to the patients' clinical information and treatment progresses were invited to record the altered morphology by using nomenclature and classification of lumbar disc abnormalities, according to the Lumbar Disc Nomenclature: Version 2.0 by Fardon *et al.* (14). When there were disagreements, another veteran radiologist was invited to compare the 2 disparate records and reach a consensus on the final interpretation.

Model design and evaluation

For the construction of clinical prediction models, clinical demographic characteristics and MRI findings were tested in univariable logistic regression models. Factors showing association in univariable comparisons ($P < 0.20$) were included in multivariable logistic regression models for the further selection of final models by using a stepwise forward procedure. Finally, 2 multivariable logistic regression models, including baseline and MRI models, were developed and validated to predict the possibility of receiving surgical treatment. The baseline model included the clinical baseline data consisting of age, leg pain, leg weakness, lower limb hypoesthesia, straight leg raise test, and NRS; meanwhile, the MRI model included disc protrusion, nerve root impairment, spondylolisthesis, spinal stenosis, and compression fracture in combination with clinical predictor variables.

After identifying the possible predictors, we conducted further analyses to validate the 2 models. Receiver operating characteristic (ROC) curve was used to illustrate the diagnostic performance of the 2 models designed, and the areas under the curve (AUC) of the ROC in different models were calculated to determine whether they were significantly different. Calibration curves were adopted to assess the calibration of the 2 models, and bootstrap resampling was used to calculate a relatively corrected concordance index (C-index) (15,16). In addition, the logistic regression formulae developed in the training dataset were applied to all validation cohort participants, and the total points for each participant were calculated to perform the logistic regression analysis. Moreover, both baseline and MRI nomograms were constructed and verified to predict the possibility of receiving surgical treatments (17,18). A nomogram is a reliable tool and is designed to create a simple intuitive graph of a statistical predictive model to assess the risk of a given clinical event. Nomograms can provide approximate complicated calculation quickly without a computer or calculator. In this study, the predictive factors identified by the multivariable logistic analysis were applied to construct the baseline and MRI nomograms. The nomograms were constructed to quantify and visualize the influence of the different predictive factors on horizontal lines of different lengths, with a longer horizontal line indicating a stronger influence. Furthermore, a series of points were visualized on each horizontal line, which reflected the influence of each factor. By adding the points associated with each factor, we were able to calculate the anticipated magnitude of risk on the bottom of the nomograms.

Clinical use

Decision curve analysis (DCA) is a statistical method that is designed to evaluate the performance of the models in supporting clinical decisions and to decide which can lead to the best decision (19). In this study, DCA was designed to evaluate the clinical utility of the constructed models by calculating the net benefits (NBs) at different risk threshold probabilities in each model (20). The NB is a simple measure for assessing the potential clinical effects of risk models in DCA by combining the value of false positives and true positives into a number that can represent the NBs after intervention. We introduced 2 default strategies, including "All" and "None", to interpret NB properly. The 2 default strategies represented patients who were managed with or without the use of a model, respectively.

Table 1 Characteristics of patients in the training and validation cohorts

Parameters	Training cohort			Validation cohort		
	No surgery group (n=371)	Surgery group (n=90)	P value ^a	No surgery group (n=112)	Surgery group (n=51)	P value ^b
Gender (male), n (%)	197 (53.1)	55 (61.1)	0.211	55 (49.1)	25 (49.0)	0.824
Age, mean (SD), years	49.26 (15.41)	53.70 (16.27)	0.016*	51.59 (14.83)	56.75 (16.51)	0.049*
Duration of pain, mean [IQR], weeks	8 [3–48]	16 [4–48]	0.021*	12 [5–24]	16 [4–48]	0.547
Leg pain, n (%)	188 (50.7)	69 (76.7)	<0.001*	62 (55.4)	41 (80.4)	0.002*
Restricted waist movement, n (%)	26 (7.0)	26 (28.9)	<0.001*	23 (20.5)	18 (35.3)	0.044*
Leg numbness, n (%)	65 (17.5)	42 (46.7)	<0.001*	44 (39.3)	31 (60.8)	0.011*
Leg weakness, n (%)	9 (2.4)	14 (15.6)	<0.001*	2 (1.8)	5 (9.8)	0.031*
Lower limb hypoesthesia, n (%)	7 (1.9)	10 (11.1)	<0.001*	20 (17.9)	25 (49.0)	<0.001*
Lameness, n (%)	12 (1.9)	9 (10.0)	0.018*	12 (10.7)	12 (23.5)	0.032*
Restricted mobility, n (%)	40 (10.8)	38 (42.2)	<0.001*	20 (17.9)	16 (31.4)	0.054
Straight leg raise test, n (%)	39 (10.5)	30 (33.3)	<0.001*	48 (42.9)	23 (45.1)	0.789
ODI, mean (SD)	45.65 (12.83)	55.67 (12.60)	<0.001*	41.09 (9.9)	48.27 (11.82)	<0.001*
NRS, mean (SD)	5.12 (0.96)	5.90 (1.32)	<0.001*	4.32 (1.25)	4.78 (1.25)	0.03*
Imaging finding						
Disk bulge, n (%)	113 (30.5)	12 (13.3)	0.002*	48 (42.9)	18 (35.3)	0.362
Disk protrusion, n (%)	163 (43.9)	51 (56.7)	0.04*	55 (49.1)	28 (54.9)	0.493
Disk extrusion, n (%)	31 (8.4)	16 (17.8)	0.014*	6 (5.4)	10 (19.6)	0.005*
Nerve root impairment, n (%)	30 (8.1)	24 (26.7)	<0.001*	7 (6.3)	10 (19.6)	0.01*
Spondylolisthesis, n (%)	27 (7.3)	17 (18.9)	0.002*	13 (11.6)	9 (17.6)	0.295
Spinal stenosis, n (%)	29 (7.8)	26 (28.9)	<0.001*	6 (5.4)	14 (27.5)	<0.001*
Schmorl's node, n (%)	29 (7.8)	3 (3.3)	0.133	7 (6.3)	5 (9.8)	0.519
Modic change, n (%)	20 (5.4)	7 (7.8)	0.539	9 (8.0)	5 (9.8)	0.766
Compression fracture, n (%)	8 (2.2)	5 (5.6)	0.164	3 (2.7)	3 (5.9)	0.378
Intraspinal tumors, n (%)	6 (1.6)	5 (5.6)	0.07	2 (1.8)	2 (3.9)	0.590

^a, comparison of demographic characteristics and MRI imaging of surgery and no surgery group in the training cohort; ^b, comparison of demographic characteristics and MRI imaging of surgery and no surgery group in the validation cohort; *, P <0.05. SD, standard deviation; ODI, Oswestry disability index; NRS, numerical rating scale.

Statistical analysis

The χ^2 test or 2-tailed Fisher's exact test was used to evaluate categorical variables, and the Student's *t*-test (for normally distributed variables) or Mann-Whitney U test (for variables that were not normally distributed) were applied to analyze continuous variables. Moreover, mean and standard deviation were calculated to express the result of continuous variables. In all analyses, a P value ≤ 0.05 was considered statistically significant, and all reported P values were

2 sided. All analyses were performed using SPSS version 24.0 software (IBM Corp., Armonk, NY, USA) and R x64 4.0.0 software (available at: <https://www.R-project.org>).

Results

Participant demographics

From January 2014 to December 2018, 461 patients who underwent MRI imaging were included in the training

cohort, and another 163 patients meeting the same inclusion criteria we enrolled as the external validation cohort. Participant demographics and MRI findings in the 2 cohorts are described in *Table 1*.

Comparison of the MRI findings in the surgery and no-surgery groups

In the training cohort, the MRI findings displayed more significant disc protrusion and disc extrusion in the surgery group than in the no-surgery group (*Table 1*). Participants in the surgery group were more likely to have severe nerve root impairment, spondylolisthesis, and spinal stenosis than were those in the no-surgery group. However, several imaging abnormalities, including Schmorl's node, Modic change, compression fracture, and intraspinal tumors, were no significantly difference between the surgery group and no-surgery group. Moreover, patients with nerve root impairment were more likely to have leg numbness (37%, $P=0.01$) and positive straight leg raise test (31.5%, $P<0.001$), but there were no differences in leg weakness (9.3%, $P=0.23$) or restricted mobility (25.9%, $P=0.06$). No association was found between the clinical symptoms and spondylolisthesis as indicated by imaging findings in LBP patients. Participants with or without spondylolisthesis were similar in restricted mobility and waist movement, straight leg raise test, leg numbness, and weakness.

Demographic characteristics and MRI findings in LBP patients with or without radiculopathy

In LBP patients, radiculopathy is a secondary compression or impairment of the nerve root leading to paresthesia, weakness, or diminished deep tendon reflexes in nerve root distribution (11). In this study, the characteristics of radiculopathy included leg numbness, leg weakness, or lower-limb hypoesthesia. Both demographic characteristics and MRI findings in LBP patients with or without radiculopathy were analyzed and compared (*Table 2*). Our study showed that compared to patients without radiculopathy, LBP patients with radiculopathy in the training and validation cohorts had higher proportion in leg pain, restricted waist movement, lameness, restricted mobility, straight leg raise test, and ODI and NRS scores, indicated severe physical disability and compromised life quality in LBP patients with radiculopathy. The proportion of participants undergoing surgical treatments was

significantly higher in patients with radiculopathy in the 2 cohorts ($P<0.001$). However, when compared with MRI findings, only nerve root impairment and spinal stenosis were identified as significant difference in patients with or without radiculopathy.

Prediction models development and comparison

The results of univariate logistic regression analysis are presented in *Table 3*. In multivariate logistic regression analysis, our study showed that in the training cohort, age [odds ratio (OR) =1.020; $P=0.042$], leg pain (OR =1.956, $P=0.045$), leg weakness (OR =3.215, $P=0.023$), low-limb hypoesthesia (OR =3.618, $P=0.033$), straight leg raise test (OR =2.248, $P=0.019$), and NRS (OR =1.433, $P=0.022$) were independent predictors in the multivariate baseline model, indicated that those factors are associated with the possibility of taking surgical intervention in the baseline model. Among them, leg weakness, low-limb hypoesthesia, and straight leg raise test had higher OR values and were positively associated with surgical intervention. When incorporated with MRI findings, leg weakness (OR =3.728, $P=0.021$), straight leg raise test (OR =2.18, $P=0.034$), NRS (OR =1.621, $P=0.007$), disc protrusion (OR =2.327, $P=0.04$), nerve root impairment (OR =2.554, $P=0.02$), spondylolisthesis (OR =3.961, $P=0.004$), spinal stenosis (OR =2.865, $P=0.009$), compression fracture (OR =6.589, $P=0.019$), and intraspinal tumors (OR =16.971, $P<0.001$) were independent predictors in the multivariate MRI model. Among them, MRI findings including disc protrusion, nerve root impairment, spondylolisthesis, spinal stenosis, compression fracture, and intraspinal tumors had higher OR values, indicated those factors are positively associated with the possibility of taking surgical intervention in MRI model (*Table 4*).

To further evaluate the predictive performance of the baseline model and the developed MRI model, ROC curves were plotted, and the AUC was calculated by the trapezoidal rule. Compared with the baseline model, the AUC increased from 79.9% (95% CI: 74–85%) to 83.4% (95% CI: 78–88%) in the MRI model in the training cohort ($Z=-1.509$; $P=0.131$; *Figure 2A*). In the validation cohort, the AUC increased from 75% (95% CI: 67.1–82.9%) to 77.7% (95% CI: 69.6–85.7%) in the MRI model ($Z=-0.588$; $P=0.557$; *Figure 2B*). Although a slightly higher AUC was observed in the MRI model, integration of the MRI findings into the prediction model did not elicit significantly improved prediction performance in the

Table 2 Demographic characteristics and MRI findings in LBP patients with/without radiculopathy

Parameters	Training cohort			Validation cohort		
	LBP (n=340)	LBP with radiculopathy (n=121)	P value ^a	LBP (n=77)	LBP with radiculopathy (n=86)	P value ^b
Gender (male), n (%)	180 (52.9)	72 (59.5)	0.213	35 (45.5)	45 (52.3)	0.381
Age, mean (SD), years	49.37 (15.99)	52.23 (14.54)	0.085	50.35 (16.97)	55.76 (13.67)	0.026
Duration of pain, mean [IQR], weeks	8 [2–48]	12 [8–48]	0.267	12 [4–24]	12 [8–31]	0.556
Leg pain, n (%)	153 (45)	104 (86.0)	<0.001*	31 (40.3)	72 (83.7)	<0.001*
Restricted waist movement, n (%)	27 (7.9)	25 (20.7)	<0.001*	11 (14.3)	30 (34.9)	0.002*
Lameness, n (%)	9 (2.6)	12 (9.9)	0.001*	5 (6.5)	19 (22.1)	0.005*
Restricted mobility, n (%)	37 (10.9)	41 (33.9)	<0.001*	3 (3.9)	33 (38.4)	<0.001*
straight leg raise test, n (%)	32 (9.4)	37 (30.6)	<0.001*	20 (26.0)	51 (59.3)	<0.001*
ODI, mean (SD)	44.02 (12.51)	57.69 (10.24)	<0.001*	39.74 (11.24)	46.56 (9.87)	<0.001*
NRS, mean (SD)	5.07 (1.014)	5.85 (1.062)	<0.001*	4.26 (1.39)	4.65 (1.11)	0.048*
Imaging finding						
Disk bulge, n (%)	93 (27.4)	32 (26.4)	0.847	28 (36.4)	38 (44.2)	0.310
Disk protrusion, n (%)	152 (44.7)	62 (51.2)	0.216	36 (46.8)	47 (54.7)	0.314
Disk extrusion, n (%)	32 (9.4)	15 (12.4)	0.351	5 (6.5)	11 (12.8)	0.177
Nerve root impairment, n (%)	32 (9.4)	22 (18.2)	0.01*	4 (5.2)	13 (15.1)	0.039*
Spondylolisthesis, n (%)	31 ()	13 (10.7)	0.601	8 (10.4)	14 (16.3)	0.272
Spinal stenosis, n (%)	32 (9.4)	23 (19.0)	0.005*	5 (6.5)	15 (17.4)	0.033*
Schmorl's node, n (%)	25 (9.1)	7 (5.8)	0.436	6 (7.8)	6 (7.0)	0.842
Modic change, n (%)	18 (5.3)	9 (7.4)	0.109	9 (11.7)	5 (5.8)	0.181
Compression fracture, n (%)	12 (3.5)	1 (0.8)	0.199	3 (3.9)	3 (3.5)	1.000
Intraspinal tumors, n (%)	7 (2.1)	4 (3.3)	0.490	1 (1.3)	3 (3.5)	0.623
Surgery, n (%)	40 (11.8)	50 (41.3)	<0.001*	11 (14.3)	40 (46.5)	<0.001*

^a, comparison of demographic characteristics and MRI imaging of LBP patients with/without radiculopathy in the training cohort; ^b, comparison of demographic characteristics and MRI imaging of LBP patients with/without radiculopathy in the validation cohort; *, P<0.05. LBP, low back pain; SD, standard deviation; ODI, Oswestry disability index; NRS, numerical rating scale.

training and the validation cohorts.

Construction and validation of nomogram

Both baseline and MRI nomogram were constructed based on the multivariate models in the training cohort to predict the possibility of receiving surgical treatments (Figure 3). Calibration curves demonstrated excellent coherence between the 2 nomogram-predicted risks and the actual outcome in the training and validation cohorts (Figure 4).

Bootstrap validation method was used to analyze the internal validation of the 2 nomogram models. The C-indexes for the prediction of receiving surgical treatments were 0.799 (95% CI: 0.743–0.854) and 0.834 (95% CI: 0.783–0.884) in the baseline model and MRI model (Figure 4A, 4B), respectively, indicating that the 2 nomograms were in good association. The Hosmer-Lemeshow test was used, with the results showing no significant difference between the baseline and MRI models in the training cohort (P=0.06 and P=0.93, respectively), which suggested that there was no departure

Table 3 Univariate logistic regression analysis of the possibility of receiving surgical treatment in the training cohort

Parameters	OR	95% CI	P value
Gender, male vs. female	1.388	0.867–2.221	0.172
Age	1.019	1.003–1.034	0.016
Duration of pain	1.001	0.999–1.004	0.172
Leg pain	3.198	1.884–5.429	<0.001
Restricted waist movement	5.391	2.942–9.877	<0.001
Leg numbness	4.119	2.516–6.745	<0.001
Leg weakness	7.409	3.094–17.741	<0.001
Lower limb hypoesthesia	6.5	2.401–17.595	<0.001
Lameness	3.324	1.355–8.154	0.009
Restricted mobility	6.047	3.554–10.290	<0.001
straight leg raise test	4.256	2.457–7.375	<0.001
ODI	1.062	1.042–1.083	<0.001
NRS	1.923	1.541–2.399	<0.001
Imaging finding			
Disk bulge	0.351	0.184–0.671	0.002
Disk protrusion	1.669	1.049–2.656	0.031
Disk extrusion	2.371	1.234–4.559	0.01
Nerve root impairment	4.133	2.273–7.516	<0.001
Spondylolisthesis	2.967	1.538–5.725	0.001
Spinal stenosis	4.791	2.648–8.667	<0.001
Schmorl's node	0.407	0.121–1.366	0.146
Modic change	1.48	0.606–3.616	0.39
Compression fracture	2.669	0.852–8.363	0.092
Intraspinal tumors	3.578	1.067–12.001	0.039

ODI, Oswestry Disability Index; NRS, Numeric Rating Scales; OR, odds ratio; CI, confidence interval.

from perfect fit.

After successful internal validation, external validation was performed. Good calibration was observed in the validation cohort (*Figure 4C,4D*), and the C-index for the prediction of receiving surgical treatments were 0.75 (95% CI: 0.671–0.829) and 0.777 (95% CI: 0.696–0.857) in the baseline model and MRI model. The Hosmer-Lemeshow test also revealed there to be no departure from the perfect fit of the baseline and MRI models in the validation cohort (P=0.946 and P=0.090, respectively).

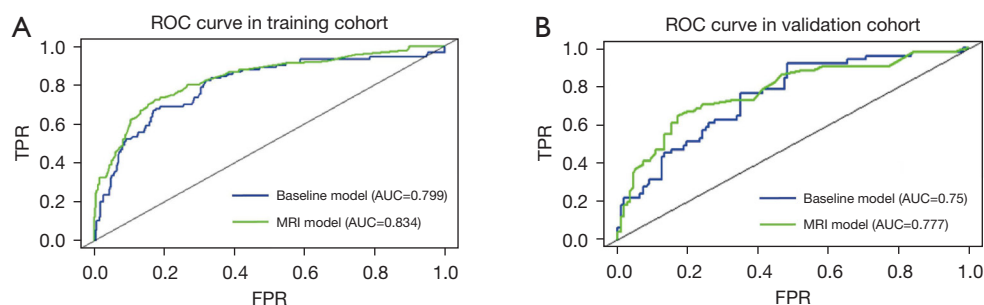
Clinical use

The DCA for the baseline and MRI nomogram is presented in *Figure 5*. The y-axis represents the net benefit, and the x-axis represents the threshold probability. Our study showed that the threshold probability of the decision curve is 10% and the corresponding NB is 0.22 in the baseline model, which indicated that using the baseline model to predict receiving surgical treatment in LBP patients adds more benefit than does the treat-all-patients measure or the treat-none measure when the threshold probability is >10%.

Table 4 Multivariate logistic regression analysis of the possibility of receiving surgical treatment in the baseline and MRI models in training cohort

Parameters	Baseline Model			MRI Model		
	OR	95% CI	P value	OR	95% CI	P value
Age	1.020	1.001–1.041	0.042	NA	NA	NA
Leg pain	1.956	1.015–3.769	0.045	NA	NA	NA
Leg weakness	3.215	1.174–8.809	0.023	3.728	1.222–11.375	0.021
Lower limb hypoesthesia	3.618	1.106–11.836	0.033	NA	NA	NA
straight leg raise test	2.248	1.141–4.429	0.019	2.18	1.062–4.474	0.034
NRS	1.433	1.054–1.949	0.022	1.621	1.144–2.297	0.007
Imaging finding						
Disk protrusion	NA	NA	NA	2.327	1.027–5.219	0.04
Nerve root impairment	NA	NA	NA	2.554	1.162–5.616	0.02
Spondylolisthesis	NA	NA	NA	3.961	1.566–10.018	0.004
Spinal stenosis	NA	NA	NA	2.865	1.308–6.276	0.009
Compression fracture	NA	NA	NA	6.589	1.366–31.799	0.019
Intraspinal tumors	NA	NA	NA	16.971	3.463–83.174	<0.001

ODI, Oswestry Disability Index; NRS, Numeric Rating Scales; OR, odds ratio; CI, confidence interval; N/A, not applicable.

**Figure 2** The ROC curves of the baseline model and MRI models in the training (A) and validation cohort (B). ROC, receiver operating characteristic; MRI, magnetic resonance imaging.

Discussion

In this study, 2 diagnostic nomogram models were developed and validated to predict the possibility of receiving surgical treatments in LBP patients. Our data showed that in the baseline model, age, leg pain, leg weakness, lower limb hypoesthesia, straight leg raise test, and NRS score were positively associated with the possibility of taking surgical intervention in LBP patients. When lumbar MRI findings were included, significantly enhanced model fit and

superior diagnostic accuracy were not apparent compared to the baseline model, suggesting that advanced lumbar MRI examination may not be useful to decide whether LBP patients with or without radiculopathy need to undergo surgical treatment. Therefore, our research may provide a completely new understanding of MRI's clinical practicability in LBP patients.

For most patients with LBP, the pain symptoms can be resolved after pharmacological therapy and rehabilitation exercise. Although some patients still complain

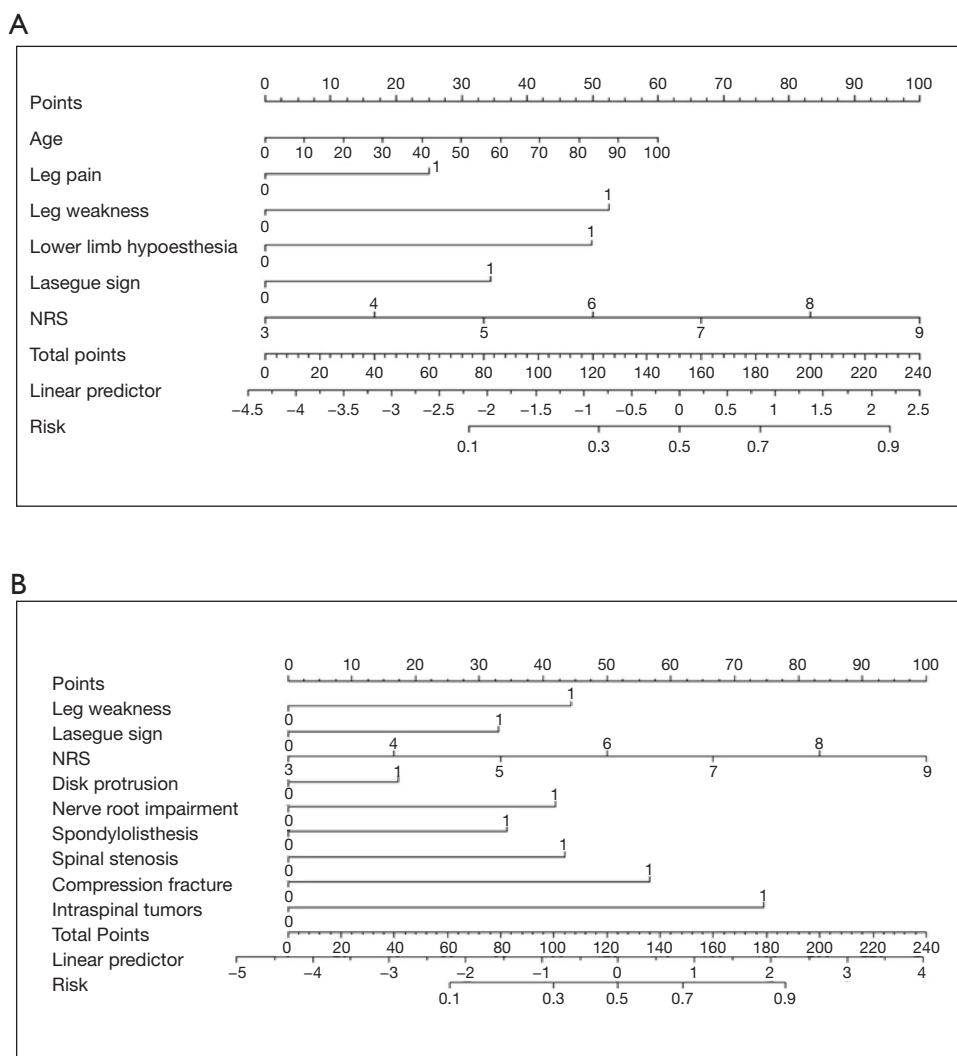


Figure 3 Nomograms for the prediction of receiving surgical treatment in the training cohort. (A) Baseline nomogram; (B) MRI nomogram. MRI, magnetic resonance imaging.

about chronic pain or unrelieved pain relief after the recommended treatments, the diagnostic imaging is only proposed when the final decision-making is changed. Since abnormal imaging findings are detectable in people with or without LBP, imaging findings may not be a good indicator to predict pain episodes (21). Also, patients' demand and expectation for imaging can result in inappropriate imaging examinations, and individuals who are informed of abnormal imaging findings are prone to develop dissatisfaction and impact treatments due to the misconception regarding and fear of bad clinical outcomes (5,22). Therefore, good education for pain management and development in clinician communication are essential, as these can improve

the patients' notions of clinical diagnosis and the treatment plan in the absence of imaging examinations.

For LBP patients with abnormal imaging, it is still unclear whether their adverse symptoms are associated with abnormal lumbar structures, as several morphological disorders, such as disc bulge or mild degeneration, are age-related morphologic changes and are commonly visible in MRI examination. Additionally, several studies have reported that there is no significant value for MRI on clinical assessment and outcome prediction (4,23). Jarvik *et al.* reported no significant difference of Roland score in 2 groups that underwent rapid MRI and plain radiography (24). Carragee *et al.* proposed the use of demographic and clinical

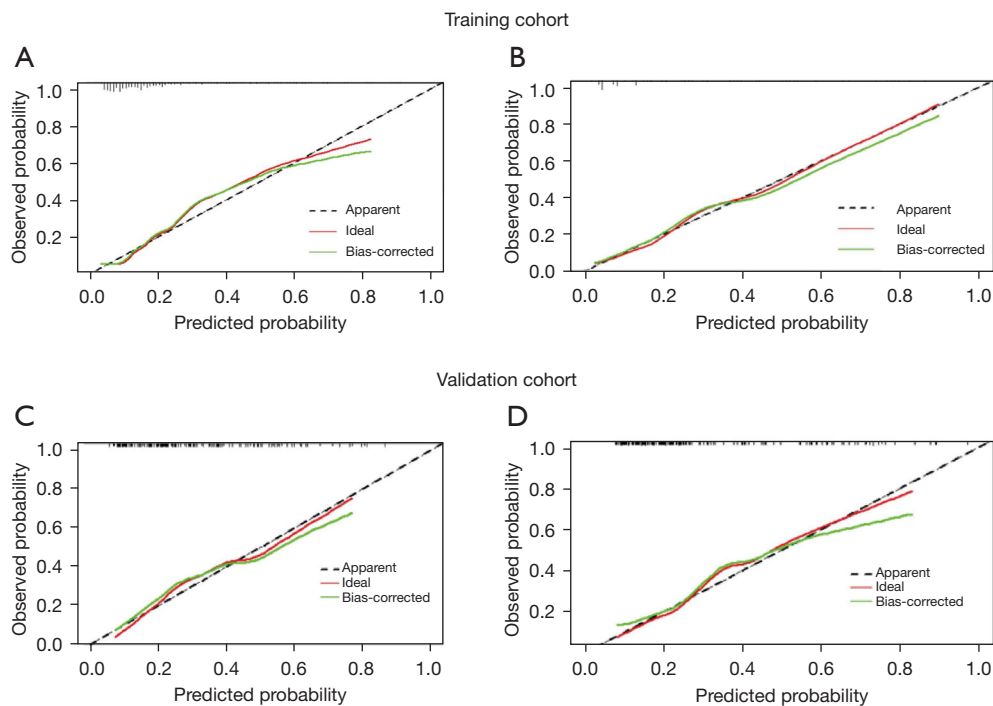


Figure 4 Calibration curves of the baseline and MRI models in the training and MRI cohorts. (A) Calibration curve of the baseline model to predict the possibility of receiving surgical treatments in the training cohort; (B) calibration curve of the MRI model to predict the possibility of receiving surgical treatments in the training cohort; (C) calibration curve of the baseline model to predict the possibility of receiving surgical treatments in the validation cohort; (D) calibration curve of the MRI model to predict the possibility of receiving surgical treatments in the validation cohort. MRI, magnetic resonance imaging.

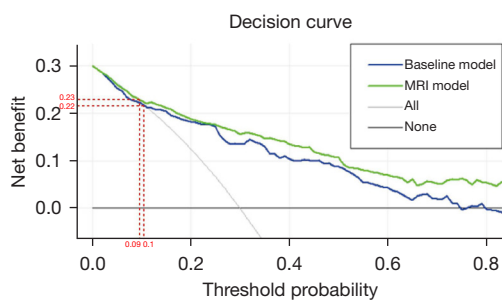


Figure 5 Decision curve analysis for the baseline and MRI models. The y-axis represents the net benefit, and the x-axis represents the threshold probability. The blue line represents the baseline nomogram. The green line represents the MRI model. The thin black line represents the assumption that all patients have received surgical treatment. The bold black line represents the assumption that no patients have received surgical treatment. MRI, magnetic resonance imaging.

characteristics to predict outcomes of conservative treatment and reported that morphometric features of disc herniation and the spinal canal on MRI seem to be useful predictors of operative treatment, with young people with a small ratio of disc hemiarea or remaining canal hemiarea usually achieving good outcomes (23). Furthermore, MRI has not been shown to be useful in assessing the need for surgical fixation in patients with thoracolumbar fracture when compared with plain radiographs (25). Furthermore, a pilot study of MRI imaging in acute lumbar radiculopathy found that acute sciatica recovered after conservative treatment and that the size of disc herniation could also be significantly reduced after conservative management. However, patients with herniated nucleus pulposus (HNP) are inclined to develop severe neurologic symptoms, with little improvement in MRI imaging after conservative treatment (26). Age is

a significant predictive factor for clinical diagnosis, and the characteristic of LBP in young individuals is usually functional pain rather than an indicator of organic disorders. It should be kept in mind that MRI is not commonly recommended for young LBP patients except for those with distressing symptoms or for considering the existence of potential malignant diseases.

Preoperative MRI may be helpful for selecting the optimal operation methods and predict the prognosis of LBP patients after surgical treatment. Kuittinen *et al.* reported that preoperative MRI of the lumbar spine could predict the outcome of LBP patients after surgical treatment in a 2-year follow up period (27). In their study, the additional MRI indicated that the severe lumbar spinal central stenosis could predict lower LBP and leg pain when compared to that in patients with moderate central stenosis after surgical treatment. Grassner *et al.* found that preoperative MRI could precisely identify the injured levels of the spine, which provided evidence to change the planned surgical approach, so as to undertake timely surgical intervention (28). Moreover, the generation of synthetic CT scans from MR imaging has been shown to help reduce workflow complexity, radiation exposure, and costs associated with adjunctive CT scanning in the lumbar spine (29). In our study, we found that LBP patients with more significant disk protrusion, spondylolisthesis, and spinal stenosis in the preoperative MRI were inclined to undergo lumbar interbody fusion, while patients with less severe disk protrusion or extrusion tended to receive percutaneous lumbar discectomy. The above results indicate that preoperative MRI may be helpful for surgeons in identifying the severity of lumbar spine injury and ultimately determining the optimal operation methods for LBP patients.

In summary, although advanced imaging technologies provide detailed anatomic information, abnormal imaging findings in LBP patients may not be associated with the chief complaint. In our study, we discovered that for LBP patients, demographic and clinical characteristics, including age, leg pain, leg weakness, lower limb hypoesthesia, straight leg raise test, and NRS score were positively associated with the possibility of taking surgical intervention in LBP patients, indicated that those factors may serve as promising predictors to decide whether patients need surgical treatment, while the incorporation of MRI findings showed limited incremental prognostic value. The findings of MRI may be advantageous for the selection of optimal operation methods. Therefore, for patients with LBP,

demographic and clinical characteristics have predictive value, and aged LBP patients with leg pain, leg weakness, lower limb hypoesthesia, straight leg raise test, and higher NRS score were positively associated with the possibility of taking surgical intervention and were therefore more likely to undergo surgical treatment.

Limitations

There were some limitations in this study. First, bias attributable to the imbalance between the surgery group and no-surgery group could not be eliminated. However, both internal and external validation was performed to address the association and calibration of the models, and therefore render our results more convincing. Second, we only analyzed the clinical characteristics and MRI findings of the 2 groups in the training and validation cohort, while the clinical outcomes of each patient following surgery or no surgery were not identified. Third, several comorbidities like heart or lung diseases in participants were not traceable and recorded, which might have excluded patients from surgery. For this reason, we studied 2 populations from different institutions, with a total of 624 patients included. Multivariable Cox analysis was employed to reduce potential confounding, and nomograms were constructed to predict the possibility of receiving surgical treatments. Although there were several limitations, this work still provides important insights into the clinical practicability of MRI in LBP patients.

Conclusions

Previous studies have reported the irrational use of advanced imaging examinations for the diagnosis and treatment of LBP patients. Yet, no studies have discussed the potential risk factors of LBP patients who need to undergo surgical therapies. In our study, 2 models were developed and validated to predict the possibility of receiving surgical treatment for LBP patients with or without radiculopathy. Our study showed that LBP patients with leg weakness are prone to receiving surgical treatments after pharmacological treatments, while additional MRI seemed to add limited value to the outcome prediction of LBP patients. Therefore, for LBP patients whose clinical baseline characteristics suggest a low risk for receiving surgical treatment, advanced imaging examinations should not be routinely prescribed. Diagnostic MRI may be helpful to ascertain the optimal surgical method once a patient is considered suitable for surgical treatment.

Our study can help in reducing the unnecessary use of MRI in patients based on their clinical data.

Acknowledgments

Funding: This work has received funding from the Wenzhou Municipal Science and Technology Bureau (Y20190018), the Wenzhou Leading Talent Innovative Project (RX2016004), the Zhejiang Provincial Natural Science Foundation of China (LGF21H060010), and the Zhejiang Provincial Traditional Chinese Medicine Science and Technology Program (2020ZB146). The funders had no role in the design, execution, or writing up of the study.

Footnote

Reporting Checklist: The authors have completed the TRIPOD reporting checklist. Available at <https://dx.doi.org/10.21037/qims-21-584>

Conflicts of Interest: All authors have completed the ICMJE uniform disclosure form (available at <https://dx.doi.org/10.21037/qims-21-584>). The authors have no conflicts of interest to declare.

Ethical Statement: The authors are accountable for all aspects of the work in ensuring that questions related to the accuracy or integrity of any part of the work are appropriately investigated and resolved. The study was conducted in accordance with the Declaration of Helsinki (as revised in 2013). The study was approved by the institutional review board of our hospital and individual consent for this retrospective analysis was waived.

Open Access Statement: This is an Open Access article distributed in accordance with the Creative Commons Attribution-NonCommercial-NoDerivs 4.0 International License (CC BY-NC-ND 4.0), which permits the non-commercial replication and distribution of the article with the strict proviso that no changes or edits are made and the original work is properly cited (including links to both the formal publication through the relevant DOI and the license). See: <https://creativecommons.org/licenses/by-nc-nd/4.0/>.

References

- Hartvigsen J, Hancock MJ, Kongsted A, Louw Q, Ferreira ML, Genevay S, Hoy D, Karppinen J, Pransky G, Sieper J, Smeets RJ, Underwood M; Lancet Low Back Pain Series Working Group. What low back pain is and why we need to pay attention. *Lancet* 2018;391:2356-67.
- Risbud MV, Shapiro IM. Role of cytokines in intervertebral disc degeneration: pain and disc content. *Nat Rev Rheumatol* 2014;10:44-56.
- Ruiz Santiago F, Láinez Ramos-Bossini AJ, Wáng YXJ, López Zúñiga D. The role of radiography in the study of spinal disorders. *Quant Imaging Med Surg* 2020;10:2322-55.
- Modic MT, Obuchowski NA, Ross JS, Brant-Zawadzki MN, Grooff PN, Mazanec DJ, Benzel EC. Acute low back pain and radiculopathy: MR imaging findings and their prognostic role and effect on outcome. *Radiology* 2005;237:597-604.
- Chou R, Fu R, Carrino JA, Deyo RA. Imaging strategies for low-back pain: systematic review and meta-analysis. *Lancet* 2009;373:463-72.
- Jensen MC, Brant-Zawadzki MN, Obuchowski N, Modic MT, Malkasian D, Ross JS. Magnetic resonance imaging of the lumbar spine in people without back pain. *N Engl J Med* 1994;331:69-73.
- De Roo B, Hoste P, Stichelbaut N, Annemans L, Bacher K, Verstraete K. Belgian multicentre study on lumbar spine imaging: Radiation dose and cost analysis; Evaluation of compliance with recommendations for efficient use of medical imaging. *Eur J Radiol* 2020;125:108864.
- Boden SD, Davis DO, Dina TS, Patronas NJ, Wiesel SW. Abnormal magnetic-resonance scans of the lumbar spine in asymptomatic subjects. A prospective investigation. *J Bone Joint Surg Am* 1990;72:403-8.
- Patel ND, Broderick DE, Burns J, Deshmukh TK, Fries IB, Harvey HB, Holly L, Hunt CH, Jagadeesan BD, Kennedy TA, O'Toole JE, Perlmutter JS, Policeni B, Rosenow JM, Schroeder JW, Whitehead MT, Cornelius RS, Corey AS. ACR Appropriateness Criteria Low Back Pain. *J Am Coll Radiol* 2016;13:1069-78.
- Bernstein IA, Malik Q, Carville S, Ward S. Low back pain and sciatica: summary of NICE guidance. *BMJ* 2017;356:i6748.
- Chou R, Qaseem A, Snow V, Casey D, Cross JT Jr, Shekelle P, Owens DK; American Pain Society Low Back Pain Guidelines Panel. Diagnosis and treatment of low back pain: a joint clinical practice guideline from the American College of Physicians and the American Pain Society. *Ann Intern Med* 2007;147:478-91.
- Smeets R, Köke A, Lin CW, Ferreira M, Demoulin C. Measures of function in low back pain/disorders: Low Back Pain Rating Scale (LBPRS), Oswestry Disability

- Index (ODI), Progressive Isoinertial Lifting Evaluation (PILE), Quebec Back Pain Disability Scale (QBPDS), and Roland-Morris Disability Questionnaire (RDQ). *Arthritis Care Res (Hoboken)* 2011;63 Suppl 11:S158-73.
13. Jensen MP, Turner JA, Romano JM. What is the maximum number of levels needed in pain intensity measurement? *Pain* 1994;58:387-92.
 14. Fardon DE, Williams AL, Dohring EJ, Murtagh FR, Gabriel Rothman SL, Sze GK. Lumbar disc nomenclature: version 2.0: Recommendations of the combined task forces of the North American Spine Society, the American Society of Spine Radiology and the American Society of Neuroradiology. *Spine J* 2014;14:2525-45.
 15. Sun F, Ma K, Yang X, Li M, Shi Y, Zhan C, Jiang W, Wang Q. A nomogram to predict prognosis after surgery in early stage non-small cell lung cancer in elderly patients. *Int J Surg* 2017;42:11-6.
 16. Zhang L, Yan L, Niu H, Ma J, Yuan BY, Chen YH, Zhuang Y, Hu Y, Zeng ZC, Xiang ZL. A nomogram to predict prognosis of patients with unresected hepatocellular carcinoma undergoing radiotherapy: a population-based study. *J Cancer* 2019;10:4564-73.
 17. Balachandran VP, Gonen M, Smith JJ, DeMatteo RP. Nomograms in oncology: more than meets the eye. *Lancet Oncol* 2015;16:e173-80.
 18. Dai D, Jin H, Wang X. Nomogram for predicting survival in triple-negative breast cancer patients with histology of infiltrating duct carcinoma: a population-based study. *Am J Cancer Res* 2018;8:1576-85.
 19. Van Calster B, Wynants L, Verbeek JFM, Verbakel JY, Christodoulou E, Vickers AJ, Roobol MJ, Steyerberg EW. Reporting and Interpreting Decision Curve Analysis: A Guide for Investigators. *Eur Urol* 2018;74:796-804.
 20. Huang YQ, Liang CH, He L, Tian J, Liang CS, Chen X, Ma ZL, Liu ZY. Development and Validation of a Radiomics Nomogram for Preoperative Prediction of Lymph Node Metastasis in Colorectal Cancer. *J Clin Oncol* 2016;34:2157-64.
 21. Tonosu J, Oka H, Higashikawa A, Okazaki H, Tanaka S, Matsudaira K. The associations between magnetic resonance imaging findings and low back pain: A 10-year longitudinal analysis. *PLoS One* 2017;12:e0188057.
 22. Franz EW, Bentley JN, Yee PP, Chang KW, Kendall-Thomas J, Park P, Yang LJ. Patient misconceptions concerning lumbar spondylosis diagnosis and treatment. *J Neurosurg Spine* 2015;22:496-502.
 23. Carragee EJ, Kim DH. A prospective analysis of magnetic resonance imaging findings in patients with sciatica and lumbar disc herniation. Correlation of outcomes with disc fragment and canal morphology. *Spine (Phila Pa 1976)* 1997;22:1650-60.
 24. Jarvik JG, Maravilla KR, Haynor DR, Levitz M, Deyo RA. Rapid MR imaging versus plain radiography in patients with low back pain: initial results of a randomized study. *Radiology* 1997;204:447-54.
 25. Rajasekaran S, Vaccaro AR, Kanna RM, Schroeder GD, Oner FC, Vialle L, Chapman J, Dvorak M, Fehlings M, Shetty AP, Schnake K, Maheshwaran A, Kandziora F. The value of CT and MRI in the classification and surgical decision-making among spine surgeons in thoracolumbar spinal injuries. *Eur Spine J* 2017;26:1463-9.
 26. Modic MT, Ross JS, Obuchowski NA, Browning KH, Cianflocco AJ, Mazanec DJ. Contrast-enhanced MR imaging in acute lumbar radiculopathy: a pilot study of the natural history. *Radiology* 1995;195:429-35.
 27. Kuittinen P, Sipola P, Leinonen V, Saari T, Sinikallio S, Savolainen S, Kröger H, Turunen V, Airaksinen O, Aalto T. Preoperative MRI findings predict two-year postoperative clinical outcome in lumbar spinal stenosis. *PLoS One* 2014;9:e106404.
 28. Grassner L, Wutte C, Zimmermann G, Grillhösl A, Schmid K, Weiß T, Maier W, Hauck S, Hollerith T, Vogel M, Bierschneider M, Vastmans J, Thomé C, Gonschorek O, Strowitzki M. Influence of Preoperative Magnetic Resonance Imaging on Surgical Decision Making for Patients with Acute Traumatic Cervical Spinal Cord Injury: A Survey Among Experienced Spine Surgeons. *World Neurosurg* 2019;131:e586-92.
 29. Staartjes VE, Seevinck PR, Vandertop WP, van Stralen M, Schröder ML. Magnetic resonance imaging-based synthetic computed tomography of the lumbar spine for surgical planning: a clinical proof-of-concept. *Neurosurg Focus* 2021;50:E13.

Cite this article as: Wang H, Liu C, Meng Z, Zhou W, Chen T, Zhang K, Wu A. Real-world study for identifying the predictive factors of surgical intervention and the value of magnetic resonance imaging in patients with low back pain. *Quant Imaging Med Surg* 2022;12(3):1830-1843. doi: 10.21037/qims-21-584