Eur Spine J (2018) 27:1058–1066
https://doi.org/10.1007/s00586-017-5305-2

ORIGINAL ARTICLE

Diagnostic accuracy of MR imaging for direct visualization of lumbar pars defect in children and young adults: a systematic review and meta-analysis

Amira Dhouib[1] · Anne Tabard-Fougere[2] · Sylviane Hanquinet[1] · Romain Dayer[2]

Received: 19 May 2017 / Revised: 13 July 2017 / Accepted: 17 September 2017 / Published online: 23 September 2017

Abstract
Purpose The accurate diagnosis of spondylolysis is widely made with CT scan considered as the gold standard. However, CT represents significant radiation exposure particularly substantial in a young and sometimes still growing population. Although the role of MRI in identifying edema/inflammation within the pars as an active lesion is proved, its ability to demonstrate and classify pars fracture line as same as CT is still controversial. This meta-analysis aimed to determine sensitivity and specificity of MRI in the direct visualisation of the pars defect.
Methods The PubMed and Embase databases were systematically searched for relevant studies from the earliest researchable time to December 2016 for cases in which the accuracy of MRI was reported for the diagnosis of spondylolysis in young patients. Two reviewers independently assessed the methodological quality for each selected study using the quality assessment of diagnostic accuracy studies 2 tool. A meta-analysis of the reported sensitivity and specificity of pooled data of selected studies was performed by a systematic review. For each selected study, sensitivity and specificity was recalculated, by considering only direct visualisation of a fracture line of the pars. The hierarchic summary receiver operating characteristic curve was generated to estimate the diagnostic performance of MR imaging. Heterogeneity was also tested.
Results The systematic review identified 4 out of a total of 1300 studies to be included in the meta-analysis. On a per-pars basis (a total of 1122 pars), the pooled sensitivity and specificity of the MRI for the direct diagnosis of a pars defect were 81% (95% CI 54–94%) and 99% (95% CI 98–100%), respectively. A high overall heterogeneity ($I2 = 79.5\%$) was computed with respective high and low heterogeneity on sensitivity ($I2 = 87.9\%$) and specificity ($I2 = 38.4\%$).
Conclusions This meta-analysis demonstrated a high diagnostic performance of MR imaging for the diagnosis of a pars defect in young adults. This technique may be considered as a first-line imaging technique as it helps to avoid exposure to ionising radiation.

Keywords Systematic review and meta-analysis · Diagnostic accuracy · Magnetic resonance imaging · Spondylolysis

Introduction

Spondylolysis is defined as a defect in the pars interarticularis. It most commonly occurs at the L5 vertebral level (95%), and the incidence decreases proceeding cephalad [1]. Spondylolysis is a frequent cause of low back pain in children and young adults, with a reported incidence of 4.4% in the paediatric population [2, 3]. It is also the most significant risk factor for low back pain in high school and college American football players [4].

Spondylolysis is described by some authors as a stress fracture in young people and is thought to be due to repetitive hyperextension and rotational movements of the trunk [5]. Consequently, it is more frequent in young athletes. In

✉ Anne Tabard-Fougere
anne.tabard@hcuge.ch

[1] Pediatric Radiology Unit, Department of Radiology, University Hospital of Geneva, Geneva, Switzerland

[2] Service of Pediatric Orthopaedics, Department of Child and Adolescent, University Hospitals of Geneva, 6 Rue Willy Donzé, 1211 Geneva 14, Switzerland

a series of 100 adolescent athletes with low back pain, 47% were ultimately found to have spondylolysis [6].

Computed tomography (CT) scanning is widely regarded as the gold standard for making an accurate diagnosis of spondylolysis [1, 7–10]. It is considered a good tool for visualizing the osseous anatomy, the cortical integrity and the extent and healing of the fracture [11, 12]. It is notable that the term "fracture" is here employed to describe a cortical discontinuity (defect) of the pars. It does not refer to a traumatic origin.

However, the major disadvantage of a CT scan is significant radiation exposure, which particularly substantial in young, and sometimes still growing populations. The approximate absorbed dose of a lumbosacral spine ct ranges between 16 and 26 mGy for children aged between 6- and 15-years-old based on diagnostic reference levels used in Switzerland. These values are based on national surveys and on values derived from literature [13]. Taking in account the ICRP tissue-weighting factor of bone marrow, effective dose will range between 1.92 and 3.12 mSv. MRI needs longer examination time, which can be an issue in young children. However, most of the patients with a suspicion of spondylolysis are adolescents or young adults.

Although the role of magnetic resonance imaging (MRI) in identifying oedema/inflammation within the pars as an active lesion has been proven, its ability to demonstrate and a pars fracture in the same way as a CT scan remains controversial [14].

Therefore, we performed this systematic review and meta-analysis to summarise knowledge about the sensitivity and specificity of MRI in the diagnosis of spondylolysis in young patients, in terms of direct visualization of the pars defect.

Materials and methods

This systematic review and meta-analysis was conducted according to the published preferred reporting items for systematic reviews and meta-analysis statements [15].

Search strategy

A comprehensive literature search of the PubMed and Embase databases without the use of search filters was conducted in March 2016 by one paediatric radiologist (AD). The search strategy was designed on the basis of the following research question: What is the diagnostic accuracy of MR imaging for the diagnosis of spondylolysis in children and young adults?

According to the recommendations of Pai et al. [16], and with the help of an experienced research librarian, Boolean combinations were defined as follows:

On PubMed: (((((((spondylolysis[MeSH]) OR spondylolyse*[Title/Abstract]) OR pars fracture [Title/Abstract]) OR pars defect [Title/Abstract]) OR pars interarticularis [Title/Abstract]) OR lumbar stress fracture [Title/Abstract]) OR isthmic fracture [Title/Abstract])) AND (("Magnetic Resonance Imaging" [Mesh]) OR ((Imaging, Magnetic Resonance [Title/Abstract] OR MRI Scan [Title/Abstract] OR Scan [Title/Abstract] OR MRI [Title/Abstract] OR Magnetic Resonance [Title/Abstract]))

On Embase: 'spondylolysis'/exp OR 'spondylolysis' OR 'spondylolyse*':ab,ti OR 'pars fracture':ab,ti OR 'pars defect':ab,ti OR 'pars interarticularis':ab,ti OR 'lumbar stress fracture':ab,ti OR 'isthmic fracture':ab,ti AND ('nuclear magnetic resonance imaging'/exp OR 'magnetic resonance imaging':ab,ti OR 'mr imaging':ab,ti OR 'mri':ab,ti)

The last search was performed on March 15, 2016, with no limitation of time. The search was supplemented by manual search of references of retrieved articles to identify any missing articles. A limited updated literature search was performed from March 15, 2016, to December 15, 2016. Two articles were identified, and the full-text were retrieved [17, 18], but they were already excluded because of an inappropriate study design.

Study selection

The two authors (AD and RD) independently selected full-text articles on the basis of the title and abstract of the identified studies according to the inclusion and exclusion criteria defined below. Any disagreement was solved by discussion.

Selected full-text articles were retrieved if they met the following criteria: the purpose of the study was to evaluate the diagnostic accuracy of MRI in lumbar spondylolysis. The patient population had an upper limitation of 30 years of age with clinical suspicion of lumbar spondylolysis. Both MRI and CT scans were performed, with CT used as the gold standard. We excluded studies according to the following exclusion criteria: no original cases of spondylolysis (review articles, commentaries, or letters to the editor), case reports and non-peer-reviewed meeting abstracts or posters. Only English-language articles were considered, and no publication date limitation was imposed.

Date extraction and processing

All included articles underwent a detailed review. One review author (AD) extracted the data, and a second author (RD) checked the following extracted data: author, year of publication, number of included patients, age range of patients, vertebral level of spondylolysis and the total

number of pars interarticularis imaged. Imaging characteristics were also extracted when available. For MRI imaging, the magnetic field strength was noted in addition to the MRI protocol, including the sequences used 2D or 3D sequence type, slice thickness and reference plan orientation for lecture. The accuracy value of the MRI (sensitivity, specificity) for each study was extracted. For CT scans, the type of scanner, section thickness and reference plan for lecture were recorded. Finally, the time between the MRI and CT was noted. No disagreements were found between the authors.

Data quality assessment

We used the standard quality of diagnostic accuracy studies 2 (QUADAS 2) tool to evaluate the methodological quality of the included studies [19].

Eleven criteria in four separate domains (patient selection, index test, reference standard and flow and timing) were used to evaluate the risk of bias and concerns regarding applicability to the search question. The assessment was performed independently by two authors (AD, AT), and disagreements were resolved by consensus. When a discrepancy remained, a third author (RD) was solicited to formulate a unique quality judgement.

Statistical analysis and data synthesis

The primary outcomes of this meta-analysis were the sensitivity and specificity of MRI to detect lumbar spondylolysis based on the CT scans.

For each selected study, sensitivity and specificity was recalculated, by considering only direct visualization of a fracture line (pars defect) on T1-weighted images. Thus, cases with isolated marrow oedema on fluid-sensitive sequences without the presence of a defect were considered as negative. Sensitivity and specificity were considered to be low if they were 50% or less, low to moderate if they were between 51 and 64%, moderate if they were between 65 and 74%, moderate to high if they were between 75 and 84% and high if they were 85% or more [20]. The pooled results and their corresponding 95% confidence intervals (CIs) were computed using a bivariate random-effects model described previously by Reitsma et al. [21] and represented by a forest plot, which illustrates individual studies, pools sensitivity and specificity estimates with 95% CIs. Interstudy heterogeneity was evaluated by computing the $I2$ value using Cochran Q statistics for each forest plot. The $I2$ value evaluated the proportion of the entire variation among studies attributable to heterogeneity rather than to chance, with heterogeneity ranked as low (25%), moderate (50%) and high (75%) [22].

Additionally, a hierarchical summary receiver operating characteristic (HSROC) curve was determined on the basis of estimates of sensitivity and specificity [23], with the calculated area under the ROC curve.

Statistical analysis was performed using the "mada" (meta-analysis of diagnostic accuracy) and "mvmeta" (multivariate and univariate meta-analysis and meta-regression) packages within the R v.3.1.3 software and the RStudio interface [RStudio Team (2015), RStudio: Integrated Development for R. RStudio, Inc., Boston, MA, USA]. A p value less than 0.05 was considered as a statistically significant difference.

Results

Study selection

Details of the search strategy and the study selection process were reported in a flowchart in Fig. 1.

Initially 1300 records were identified: 770 in PubMed and 530 in Embase. Two hundred and fifty-six duplicates were identified and then excluded. After the selection based on the title and abstract, 1020 irrelevant records were excluded. The full-text version of the remaining 19 citations was examined in more detail. Only six studies fulfilled the inclusion criteria. In two cases, discussion was needed with a third reviewer (RD) to reach consensus for eligibility and inclusion. Finally, a total of six eligible studies were included and considered for further analyses. Only four papers were selected for meta-analysis.

Study characteristics and data extraction

Relevant study characteristics are summarised in Table 1. All six studies selected by the systematic review were published between 1993 and 2015. Patients of the included studies were aged between 5 and 30 years. The total number of pars explored was initially impossible to calculate. Indeed, the total number of included pars in the study of Rush et al. [24] was not specified. We contacted the corresponding author to obtain this information.

There was variation in the MRI techniques used. Magnetic field strength varied between 0.5 and 1.5 T. The MRI protocols included variable sequences with variable reference plans of lecture for the detection of the pars defect. Only the T1-weighted sequence was performed in all studies. Slice thickness also varied between 3 and 5 mm. In two studies, a thin-slice 3D sequence was used [25, 26].

A variation among CT scan characteristics was also observed. The reference lecture plan was variable, but in most cases, an axial oblique plan was used. In one study, the plan was not specified [24]. CT scan slice thickness varied between 1.5 and 3.0 mm and not specified in the same two

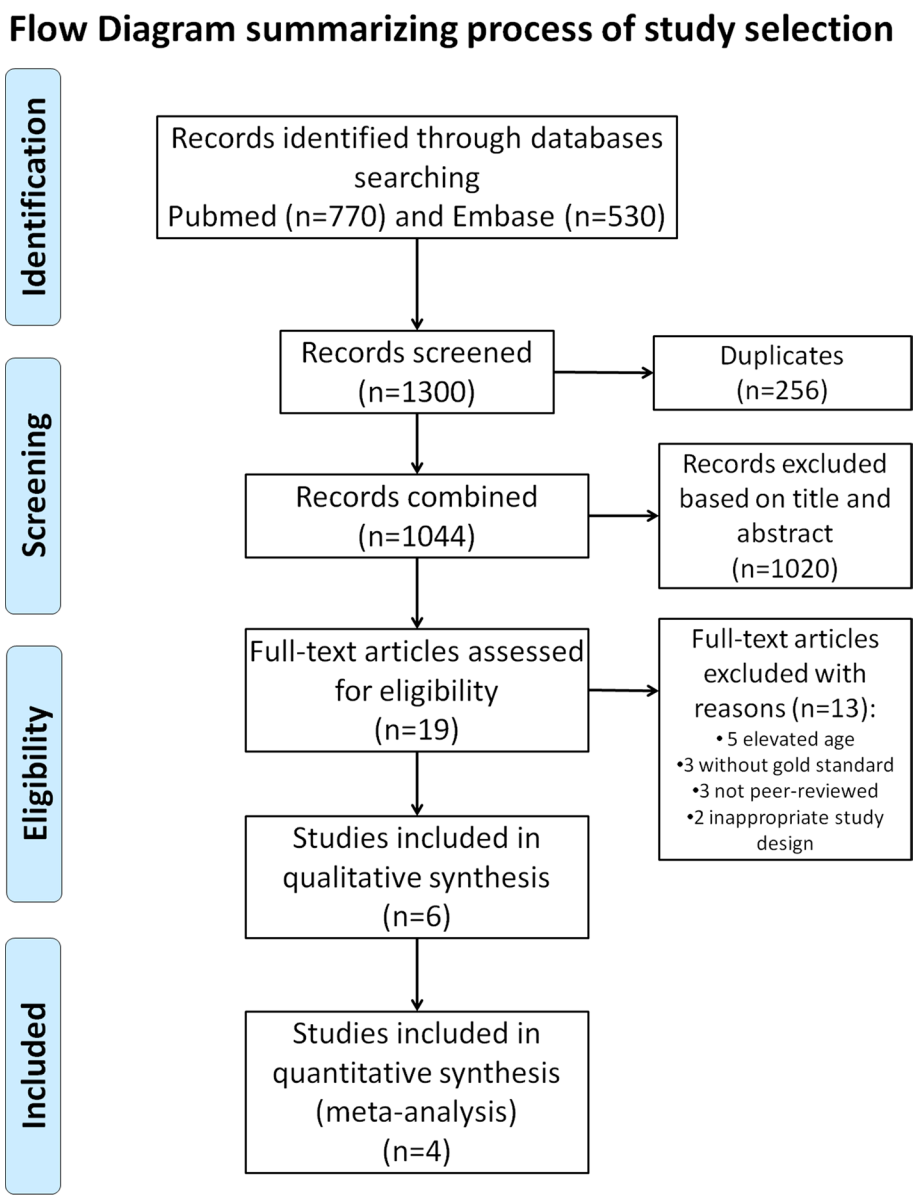

Fig. 1 The flowchart shows the flow of information through the different phases of systematic review toward a meta-analysis

studies. In addition, the number of CT barrettes was only systematically specified in one study [26].

Quality assessment

The distribution of QUADAS 2 scores of the methodological quality, with risk of bias and regarding the applicability of every included study is presented in Table 2.

Three studies were assessed with a low risk of bias and minimal concerns regarding applicability [24–26]. The study of Masci et al. [27] had a high risk in patient selection since only patients having a positive single-photon emission computed tomography (SPECT) were included. The study of Yamaguchi et al. [14] was evaluated with a high risk of bias and serious concerns about applicability due to few details concerning methodologic features for both index test than for gold standard being reported. We also considered a high risk of bias in the study of Yamane et al. [28]. These two studies [14, 28] were excluded from the meta-analysis because of methodological shortcomings.

Data analysis

The results of specificity and sensitivity of the four selected studies included for the meta-analysis are presented in Fig. 2, with a moderate to high combined sensitivity [81% (54 to 94—95% CI)] and a high specificity [99% (98 to 100%—95% CI)]. Sensitivity of MRI to detect pars defect was high for Studies 1 and 2 [25, 27] (> 85%), including a total of 344 pars, moderate to high (84%) for Study 3 [26] including 570

Table 1 Characteristics of selected studies

Year of publ.	First author	Journal	Patients included	Age range	Nb of PI imaged	MRI characteristics						CT characteristics		
						MRI type	Mag. field strength	Seq.	3D seq.	Slice thickness (mm)	Ref. plan orientation	CT barette's number	CT section thickness	Ref. lecture plan
2012	Yamagushi	*J Child Orthop*	9											
1993	Yamane	*JBJS*	79	5–19	81	General Electric	0.5	sag + coro T1 and T2	No	5	coro + sag	NS	2 mm	Transverse oblique
2015	Rush	*J Pediatr Orthop*	26	11–20		Siemens	1	sag T1, sag T2, sag T2 fs	No	4	sag	NS	NS	NS
2006	Masci	*Br J Sports Med*	39	10–30	50	General Electric	1.5	sag T1, sag + ax T2, ax T2 fs, Ax oblique STIR	No	3 and 3.5	sag + ax oblique	NS	3 mm	Transverse oblique
2010	Ganiyusufoglu	*Clinic Radiol*	57	9–30		Siemens	1.5	sag + ax T1, sag + ax T2, sag STIR, VIBE	Yes (vibe)	1.7, 3, 4	sag,	16	1.5	ax oblique, sag, coronal
2005	Campbell	*Skelet Radiol*	72	8–26	294	Siemens	1	sag + ax obl T1, STIR or T2 fs sag, FLASH	Yes (flash)	1.9, 3, 4	ax obl, sag obl	NS	3 mm	ax obl

Mag. magnitude, *Ref* reference, *Nb* number, *publ.* publication, *seq.* sequence, *NS* not specified, *sag* sagittal, *coro* coronal, *obl* oblique, *ax* axial

Table 2 Tabular presentation for QUADAS-2 results

Study	Risk of bias				Applicability concerns		
	Patient selection	Index test	Reference standard	Flow and timing	Patient selection	Index test	Reference standard
Yamaguchi et al. [14]	☹	☹	?	?	☺	?	?
Yamane et al. [28]	☹	☹	☺	☺	☺	☹	☺
Rush et al. [24]	☺	☺	?	☺	☺	☺	?
Masci et al. [27]	☹	☺	☺	?	☺	☺	☺
Ganiyusufoglu et al. [26]	☺	☺	☺	☺	☺	☺	☺
Campbell et al. [25]	☺	☺	☺	☺	☺	☺	☺

☺ Low risk ☹ High risk ? Unclear risk

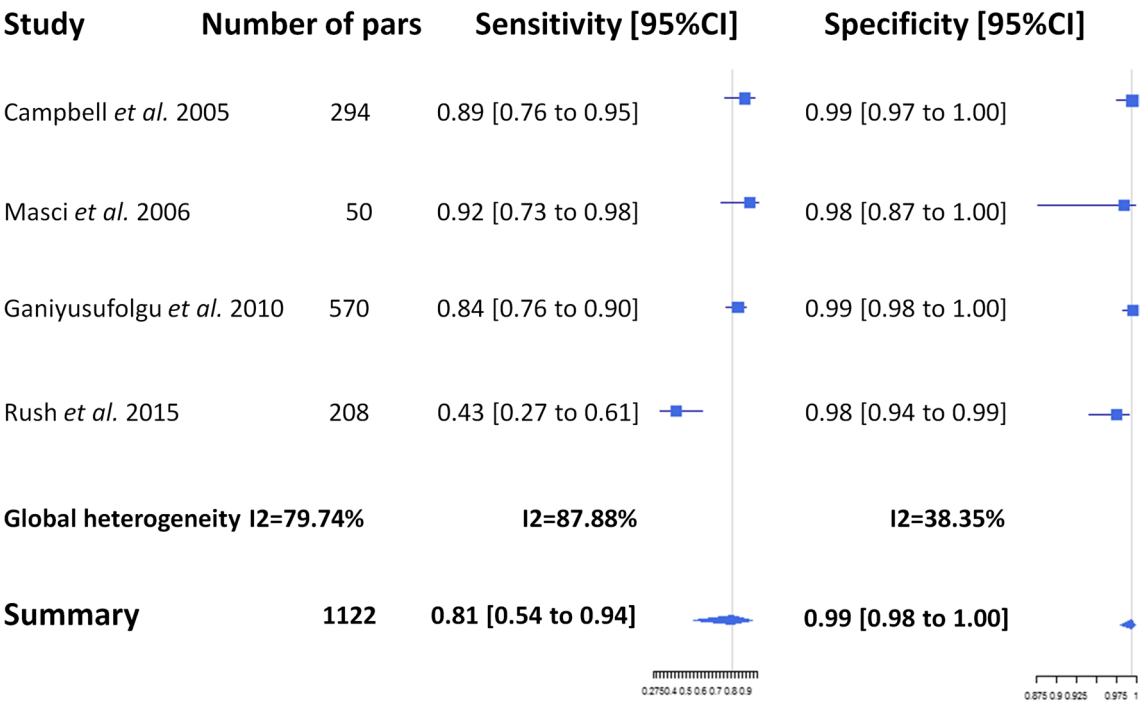

Fig. 2 The forest plot of the meta-analysis shows the results of pooled sensitivity and specificity of MRI imaging for diagnostic accuracy of lumbar spondylolysis among children and adolescents. The point estimates (blue squares) and pooled estimates (blue diamond with vertical line) with 95% CIs of sensitivity or specificity (horizontal lines) from each study are shown. The *I*2 statistics described interstudy heterogeneity

pars and low (43%) for Study 4 [24], including 208 pars. Specificity was high for all evaluated studies (> 85%), including a total of 1122 pars. A high overall heterogeneity ($I2 = 79.7\%$) was detected in these estimates, with a high heterogeneity on sensitivity ($I2 = 87.9\%$) and low heterogeneity on specificity ($I2 = 38.4\%$). Figure 3 illustrates the HSROC curve, which had an area under the curve of 98.7% with a summary estimate of sensitivity of 81.0%, and a summary estimate of specificity of 99.0% (false negative rate = 1.0%), showing the excellent diagnostic accuracy of MRI.

Exploration of the source of heterogeneity

Influence analyses are summarised in Table 3. When Study 4 [24] was excluded, the *I*2 decreased from 87.9 to 0.0%. Combined sensitivity and specificity increased from 81.0 to 86.7%

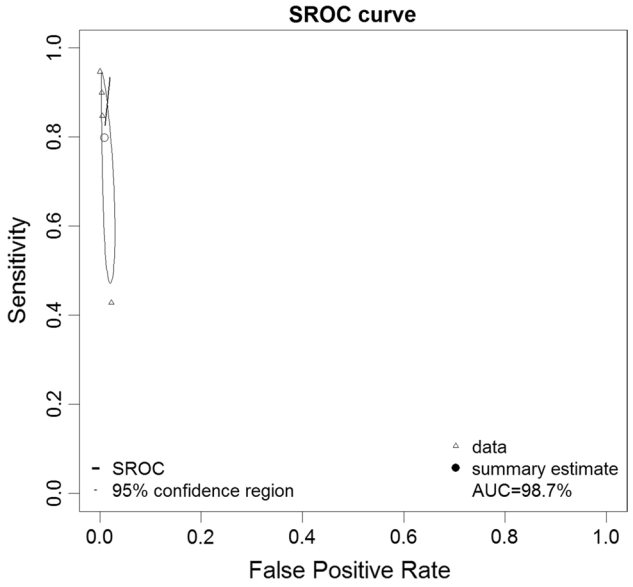

Fig. 3 The summary receiver operating characteristic (SROC) plots of MRI imaging for diagnostic accuracy of lumbar spondylolysis among children and adolescents is represented by the bold line. Each triangle indicates included studies, and a circle represents the summary point, indicating an estimate of sensitivity and specificity, with the fine line representing the 95% confidence region, and *AUC* area under the curve

and from 99.4 to 99.6%, respectively. Heterogeneity on selectivity decreased from 38.3 to 0.0%. This same study decreased the overall heterogeneity from 79.7 to 0.0%. In contrast, when Study 2 [27] was excluded, the *I*2 increased from 38.3 to 58.7%, and combined specificity decreased from 99.4 to 99.3%.

Discussion

Summary of evidence

The present investigation is the first meta-analysis to evaluate the accuracy of MRI for the direct visualization of pars defect in children and young adults.

The systematic review identified six studies that met the inclusion criteria, and four studies that demonstrated moderate to high combined sensitivity (> 0.75%) and high combined specificity (> 0.85%) of MRI for the direct visualization of the fracture in the pars interarticularis were selected for the meta-analysis.

We found moderate to high sensitivity (> 0.75%) of MRI for the diagnosis of lumbar spondylolysis for all included studies except one [24] with low sensitivity (< 50%). Although the MRI protocol was correctly reported, it is notable that only sagittal views were explored with a relatively high slice thickness (4 mm). This could potentially explain the high number of false negatives. These results are in line with results of previous studies [29–32], which judged MRI sensitivity as low. The poor results of these studies are likely due to relatively thick slices, wide inter-slice distances and the inclusion of elderly adults in the cohorts. In the adult population, spondylolysis was most likely to be confused with facet osteoarthritis and other degenerative changes of the posterior elements. Facet hypertrophy may account for hypointense pars and explain the diminished diagnostic accuracy of the MRI in these patients.

The two largest series [25, 26] that account for the most important number of pars interarticularis showed a high sensitivity. These two studies used a thin-section 3D sequence, not routinely used in the exploration of the lumbar spine. From an anatomic point of view, the pars interarticularis is a difficult region to explore. It is oriented obliquely to all three orthogonal planes [33]. In this situation, a thin-section 3D sequence is particularly helpful for direct visualization of the pars with its superior and inferior cortical margins, through multiplanar oblique reconstruction along the axis of the pars. This thin slice thickness could improve the accuracy of MRI in the diagnosis of spondylolysis by deferring the partial volume effects induced by thicker slices.

The calculated sensitivity was valuable only for direct visualization of the pars defect. The purpose of the present meta-analysis was to evaluate the accuracy of MRI in the diagnosis of spondylolysis, with CT scans considered as the gold standard. Thus, we were primarily interested by

Table 3 Influence of analyses on heterogeneity

Study excluded	Overall heterogeneity (*I*2) (%)	Sensitivity (95% CI)	Heterogeneity on sensitivity (*I*2) (%)	Specificity (% CI)	Heterogeneity on sensitivity (*I*2) (%)
None	79.7	81.0 (54.5–93.8)	87.9	99.4 (98.0–99.8)	38.3
Study 1—Campbell et al. [25]	84.0	77.2 (39.8–94.5)	90.6	99.2 (97.0–99.8)	45.9
Study 2—Masci et al. [27]	85.3	76.6 (42.6–93.5)	91.1	99.3 (97.2–99.8)	58.7
Study 3—Ganiyusufoglu et al. [26]	81.7	80.4 (39.5–96.3)	89.6	99.2 (96.8–99.8)	21.3
Study 4—Rush et al. [24]	0.0	86.7 (80.4–91.1)	0.0	99.6 (98.7–99.9)	0.0

Negative values of heterogeneity *I*2 are equal to zero so that *I*2 lies between 0.0 and 100.0% and a value of 0.0% indicates no observed heterogeneity and larger values show increasing heterogeneity [22]

sequences showing the fracture, particularly T1-weighted images. The pars defects appear as an interruption of the cortex and marrow through the pars. This pattern is best seen on T1-weighted images, which allows the greatest contrast between hyperintense bone marrow and the signal void of the bony cortex [29]. Marrow oedema into the pars, shown on fluid-sensitive sequences, was not considered as a spondylolysis sign in this context for two reasons. First, the CT scan, considered as the gold standard, is not able to distinguish active from inactive inflammatory lesions. Second, previous studies have already demonstrated that MRI is effective in detecting spondylolysis activity [25, 34–36]. Some studies have evaluated the effectiveness of MRI in diagnosing active lesions in comparison with SPECT or scintigraphy [25, 27, 34], and opinions have varied on whether SPECT or MRI is the more sensitive tool. Other studies [35, 36] supported the belief that MRI is a useful non-invasive method in the early diagnosis of spondylolysis, the so-called stress reaction, when fracture is not yet visible on CT scans. This pattern was shown as high signal changes on fluid-sensitive sequences. The presence of marrow oedema in the pars or in the pedicle, although suggesting a possible fracture, remains a non-specific pattern when isolated. This type of oedema may also be related to other inflammatory lesions, such as overuse injuries or osteoid osteomas.

The value of MRIs was considered high (> 0.85%) for all evaluated studies. In two evaluated studies [26, 37], all lumbar levels were analysed by both MRI and CT scan. Thus, all true negative patients were considered when specificity was calculated, which increased the study's value.

Limitations

This systematic review has several limitations. First, the search process was limited to English-language publications and did not include grey literature. Second, and perhaps the most important limitation to our work, is the relatively small number of studies that met our inclusion criteria, limiting the statistical power to estimate the diagnostic accuracy of MRI for the diagnosis of lumbar spondylolysis.

Differences in methodological quality according to the QUADAS 2 tool could also be a source of heterogeneity. We found that the study of Masci et al. [27] demonstrated selection bias because only patients with a positive SPECT were explored by both CT and MRI. This could increase the prevalence of spondylolysis in the evaluated population and thus affect the results.

We also observed heterogeneity in the MRI protocols. Sequences were variable, but the T1 weighted sequence was systematically performed in all studies. The plane of lecture was also variable. Thin-slice 3D sequences with oblique reconstruction in the axis of the pars were used for the interpretation of images in only two studies.

The same heterogeneity was observed for the CT scan protocols. The reference plane and slice thickness varied between included studies.

Contribution of MRI in the diagnosis of each subtype of spondylolysis (partial fracture, total fracture and pseudarthrosis) is not discussed in our study due to the small number of pars in each subtype in the included studies.

Conclusion

Results of this meta-analysis revealed moderate to high sensitivity and high specificity of MRI for the diagnosis of lumbar spondylolysis in children and young adults. These results suggest that MRI is highly effective for the diagnosis of spondylolysis. It may be considered as a first-line imaging test, as it helps to avoid exposure to ionising radiation. The use of 3D slices is recommended.

Acknowledgements The authors wish to express their gratitude to the experienced research librarian Mrs. Mafalda Burri for his technical assistance to perform the search strategy of the comprehensive literature.

Compliance with ethical standards

Funding source No funding was secured for this study. The authors have no financial relationships relevant to this article to disclose.

Conflict of interest The authors have no conflicts of interest to disclose.

References

1. Grogan JP, Hemminghytt S, Williams AL et al (1982) Spondylolysis studied with computed tomography. Radiology 145:737–742
2. Ralston S, Weir M (1998) Suspecting lumbar spondylolysis in adolescent low back pain. Clin Pediatr (Phila) 37:287–293
3. Fredrickson BE, Baker D, McHolick WJ et al (1984) The natural history of spondylolysis and spondylolisthesis. J Bone Jt Surg Br 66:699–707
4. Iwamoto J, Abe H, Tsukimura Y, Wakano K (2004) Relationship between radiographic abnormalities of lumbar spine and incidence of low back pain in high school and college football players: a prospective study. Am J Sports Med 32:781–786
5. Sakai T, Yamada H, Nakamura T et al (2006) Lumbar spinal disorders in patients with athetoid cerebral palsy. Spine (Phila Pa 1976) 31:E66–E70
6. Micheli LJ, Wood R (1995) Back pain in young athletes. Significant differences from adults in causes and patterns. Arch Pediatr Adolesc Med 149:15–18
7. Blanda J, Bethem D, Moats W, Lew M (1993) Defects of pars interarticularis in athletes: a protocol for nonoperative treatment. J Spinal Disord 6:406–411

8. Miller SF, Congeni J, Swanson K (2004) Long-term functional and anatomical follow-up of early detected spondylolysis in young athletes. Am J Sports Med 32:928–933
9. Herman MJ, Pizzutillo PD (2005) Spondylolysis and spondylolisthesis in the child and adolescent: a new classification. Clin Orthop Relat Res 434(5):46–54
10. Langston JW, Gavant ML (1985) "Incomplete ring" sign: a simple method for CT detection of spondylolysis. J Comput Assist Tomogr 9:728–729
11. Harvey CJ, Richenberg JL, Saifuddin A, Wolman RL (1998) The radiological investigation of lumbar spondylolysis. Clin Radiol 53:723–728
12. Congeni J, McCulloch J, Swanson K (1997) Lumbar spondylolysis. A study of natural progression in athletes. Am J Sports Med 25:248–253
13. Galanski M, Nagel HD, Stamm G (2007) Results of a federation inquiry 2005/2006: pediatric CT X-ray practice in Germany (article in German). RoFo 179:1110–1111
14. Yamaguchi KT, Skaggs DL, Acevedo DC et al (2012) Spondylolysis is frequently missed by MRI in adolescents with back pain. J Child Orthop 6:237–240
15. Moher D, Liberati A, Tetzlaff J et al (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. PLoS One 6:e1000097
16. Pai M, McCulloch M, Gorman JD et al (2004) Systematic reviews and meta-analyses: an illustrated, step-by-step guide. Natl Med J India 17:86–95
17. Delavan JA, Stence NV, Mirsky DM et al (2016) Confidence in assessment of lumbar spondylolysis using three-dimensional volumetric T2-weighted MRI compared with limited field of view, decreased-dose CT. Sport Heal A Multidiscip Approach 8:364–371
18. Ang EC, Robertson AF, Malara FA et al (2016) Diagnostic accuracy of 3-T magnetic resonance imaging with 3D T1 VIBE versus computer tomography in pars stress fracture of the lumbar spine. Skelet Radiol 45:1533–1540
19. Whiting PF, Rutjes AWS, Westwood ME et al (2011) QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. Ann Intern Med 155:529
20. Schneiders AG, Sullivan SJ, Hendrick PA et al (2012) The ability of clinical tests to diagnose stress fractures: a systematic review and meta-analysis. J Orthop Sport Phys Ther 42:760–771
21. Reitsma JB, Glas AS, Rutjes AWS et al (2005) Bivariate analysis of sensitivity and specificity produces informative summary measures in diagnostic reviews. J Clin Epidemiol 58:982–990
22. Higgins JPT, Thompson SG, Deeks JJ, Altman DG (2003) Measuring inconsistency in meta-analyses. BMJ 327:557–560
23. Rutter CM, Gatsonis CA (2001) A hierarchical regression approach to meta-analysis of diagnostic test accuracy evaluations. Stat Med 20:2865–2884
24. Rush JK, Astur N, Scott S et al (2015) Use of magnetic resonance imaging in the evaluation of spondylolysis. J Pediatr Orthop 35:271–275
25. Campbell RSD, Grainger AJ, Hide IG et al (2005) Juvenile spondylolysis: a comparative analysis of CT, SPECT and MRI. Skelet Radiol 34:63–73
26. Ganiyusufoglu AK, Onat L, Karatoprak O et al (2010) Diagnostic accuracy of magnetic resonance imaging versus computed tomography in stress fractures of the lumbar spine. Clin Radiol 65:902–907
27. Masci L, Pike J, Malara F et al (2006) Use of the one-legged hyperextension test and magnetic resonance imaging in the diagnosis of active spondylolysis. Br J Sports Med 40:940–946 **(discussion 946)**
28. Yamane T, Yoshida T, Mimatsu K (1993) Early diagnosis of lumbar spondylolysis by MRI. J Bone Jt Surg Br 75:764–768
29. Johnson D, Farnum G, Latchaw R, Erba S (1989) MR imaging of the pars interarticularis. Am J Roentgenol 152:327–332
30. Grenier N, Kressel HY, Schiebler ML, Grossman RI (1989) Isthmic spondylolysis of the lumbar spine: MR imaging at 1.5 T. Radiology 170:489–493
31. Saifuddin A, Burnett SJD (1997) The value of lumbar spine MRI in the assessment of the pars interarticularis. Clin Radiol 52:666–671
32. Ulmer JL, Mathews VP, Elster AD et al (1997) MR imaging of lumbar spondylolysis: the importance of ancillary observations. Am J Roentgenol 169:233–239
33. Campbell RS, Grainger AJ (1999) Optimization of MRI pulse sequences to visualize the normal pars interarticularis. Clin Radiol 54:63–68
34. Ishibashi Y, Okamura Y, Otsuka H et al (2002) Comparison of scintigraphy and magnetic resonance imaging for stress injuries of bone. Clin J Sport Med 12:79–84
35. Sairyo K, Katoh S, Takata Y et al (2006) MRI signal changes of the pedicle as an indicator for early diagnosis of spondylolysis in children and adolescents: a clinical and biomechanical study. Spine (Phila Pa 1976) 31:206–211
36. Kobayashi A, Kobayashi T, Kato K et al (2013) Diagnosis of radiographically occult lumbar spondylolysis in young athletes by magnetic resonance imaging. Am J Sports Med 41:169–176
37. Murray CJL, Vos T, Lozano R et al (2012) Disability-adjusted life years (DALYs) for 291 diseases and injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of Disease Study 2010. Lancet 380:2197–2223