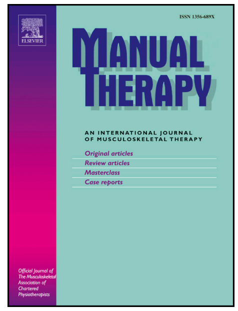
Diagnostic utility of patient history and physical examination data to detect spondylolysis and spondylolisthesis in athletes with low back pain: A systematic review

Grødahl, Linn Helen J.; Fawcett, Louise; Nazareth, Madeleine; Smith, Richard; Spencer, Simon; Heneghan, Nicola; Rushton, Alison

DOI:
10.1016/j.math.2016.03.011

Accepted Manuscript



Diagnostic utility of patient history and physical examination data to detect spondylolysis and spondylolisthesis in athletes with low back pain: A systematic review

Linn Helen J. Grødahl, Louise Fawcett, Madeleine Nazareth, Richard Smith, Simon Spencer, Nicola Heneghan, Alison Rushton

PII: S1356-689X(16)30001-7

DOI: 10.1016/j.math.2016.03.011

Reference: YMATH 1842

To appear in: *Manual Therapy*

Received Date: 24 August 2015

Revised Date: 22 March 2016

Accepted Date: 24 March 2016

Please cite this article as: Grødahl LHJ, Fawcett L, Nazareth M, Smith R, Spencer S, Heneghan N, Rushton A, Diagnostic utility of patient history and physical examination data to detect spondylolysis and spondylolisthesis in athletes with low back pain: A systematic review, *Manual Therapy* (2016), doi: 10.1016/j.math.2016.03.011.

Title: Diagnostic utility of patient history and physical examination data to detect spondylolysis and spondylolisthesis in athletes with low back pain: a systematic review

Linn Helen J. Grødahl [1]*, Louise Fawcett, Madeleine Nazareth, Richard Smith, Simon Spencer, Nicola Heneghan, Alison Rushton

* Corresponding author: lhj.grodahl@gmail.com, Tel; +47 93466301, Postal address; Nils Bays vei 86, 0855 Oslo, Norway

School of Sport, Exercise and Rehabilitation Sciences
College of Life and Environmental Sciences
University of Birmingham
Edgbaston
Birmingham
B15 2TT
UK

[1] Present address: Nils Bays vei 86, 0855 Oslo, Norway.

Abstract

Background

In adolescent athletes, low back pain has a 1-year prevalence of 57% and causes include spondylolysis and spondylolisthesis. An accurate diagnosis enables healing, prevention of progression and return to sport.

Objective

To evaluate the diagnostic utility of patient history and physical examination data to identify spondylolysis and/or spondylolisthesis in athletes.

Design

Systematic review was undertaken according to published guidelines, and reported in line with PRISMA.

Method

Key databases were searched up to 13/11/15. Inclusion criteria: athletic population with LBP, patient history and/or physical examination accuracy data for spondylolysis and/or spondylolisthesis, any study design including raw data. Two reviewers independently assessed risk of bias (ROB) using QUADAS-2. A data extraction sheet was pre-designed. Pooling of data and investigation for heterogeneity enabled a qualitative synthesis of data across studies.

Results

Of the eight included studies, two were assessed as low ROB, one of which also had no concerns regarding applicability. Age (<20 years) demonstrated 81% sensitivity and 44% specificity and gender (male) 73% sensitivity and 57% specificity for spondylolysis. Difficulty falling asleep, waking up because of pain, pain worse with sitting and walking all have sensitivity >75% for spondylolisthesis. Step-deformity palpation demonstrated 60-88% sensitivity and 87-100% specificity for spondylolisthesis. The one-legged hyperextension test was not supported for spondylolysis (sensitivity 50-73%, specificity 0-87%).

Conclusion

No recommendations can be made utilising patient history data. Based on one low ROB study, step deformity palpation may be useful in diagnosing spondylolisthesis. No physical tests demonstrated diagnostic utility for spondylolysis. Further research is required.

INTRODUCTION

Low back pain (LBP) in adolescent athletes (aged 12-20 years) has a 1 year prevalence of up to 57% (Schmidt et al., 2014), compared to the age matched broad population (10-19 years) 1 year prevalence of 23% (Hoy et al., 2012). In the adult population, disc pathology and degenerative changes are predominantly associated with LBP, whereas athletic adolescents are more predisposed to posterior element derangements, including spondylolysis and spondylolisthesis (Micheli and Wood, 1995). Spondylolysis is an osseous defect of the pars interarticularis of a vertebral arch (Haun and Kettner, 2005); and spondylolisthesis is a translation of a vertebral body on the adjacent vertebra, most often referred to as a listhesis in the anterior direction (Haun and Kettner, 2005).

In the general population, spondylolysis is present in 4.4% of asymptomatic children and by adulthood in 6% (Fredrickson et al., 1984). Occurrence of symptomatic spondylolisthesis into adulthood has been reported as 5% (Beutler et al., 2003). The male:female ratio is 2:1 (Beutler et al., 2003, Lonstein, 1999). The prevalence of spondylolysis in the athletic population is 13.90% (Rossi and Dragoni, 2001), higher percentages are seen in sports like; diving 40.35% (Rossi and Dragoni, 2001), throwing sports 27% (Soler and Calderon, 2000), sailing 17.18% and gymnastics 16.64% (Rossi and Dragoni, 2001). Progression to spondylolisthesis has been reported as 47.5% (Rossi and Dragoni, 2001) and has been associated with mechanical stress related to certain sports involving repetitive lumbar hyperextension (Jackson et al., 1976). The progression of listhesis is seen to be greater in adolescence with 7% slippage and reduced to 2% slippage by the 5th decade of life (Beutler et al., 2003).

Establishing an accurate diagnosis to enable healing and prevention of progression to non-union of the pars interarticularis is the primary management goal for athletes with spondylolysis (Iwamoto et al., 2010). Higher healing rates have been seen if spondylolysis is detected early (Fujii et al., 2004, Morita et al., 1995, Saraste, 1986). A recent systematic review concluded that no clinical test possessed the diagnostic utility (the diagnostic usefulness of a test) to diagnose spondylolysis, but that the lumbar spinous palpation test demonstrated diagnostic utility for diagnosing spondylolisthesis (Alqarni et al., 2015), with specificity 87-100% and sensitivity 60-88%. The review included a general, non-athletic population in their eligibility criteria; but two of the included studies (Gregg et al., 2009, Masci et al., 2006) investigated a sporting population for spondylolysis. Clinical tests that can distinguish spondylolysis from other causes of LBP in athletes have not been identified (Alqarni et al., 2015, Kujala et al., 1999). However, patient history data are strong contributors to establishing an accurate diagnosis (Peterson et al., 1992); through clinical reasoning processes (Rushton and Lindsay, 2010), and Alqarni et al (2015) only explored physical data, searching to February 1st 2014. An updated review including patient history data is therefore required

Objective

To identify and evaluate the diagnostic utility of patient history and physical examination data to identify spondylolysis and/or spondylolisthesis in athletes.

METHODOLOGY

Methods

A systematic review was conducted according to a pre-defined protocol designed according to The Cochrane Handbook for Diagnostic Test Accuracy studies (Bossuyt et al., 2013, Bossuyt and Leeflang, 2008, de Vet et al., 2008), the Centre for Reviews and Dissemination (CRD, 2009) and the Preferred Reporting Items for Systematic reviews and Meta-Analysis guidelines (Moher et al., 2009).

Search strategy

Two reviewers (XX, XX) independently searched key bibliographic databases: MEDLINE, Cochrane Library, AMED, CINAHL, Sport Discus, Pub Med Central and Web of Science. Databases were searched from date of inception to 13th November 2015. A third reviewer (XX) mediated any disagreements. All three reviewers attended a meeting with a research assistant where the search strategies for the main databases were discussed. The following terms and combinations of them, were used: low back pain, spondylolysis, spondylolisthesis, stress fracture, pars interarticularis, stability, range of motion, test, diagnosis, diagnostic test, signs, symptoms, patient history, physical examination, accuracy, sensitivity, specificity, reliability, validity, athletes and sport. Terms were searched for as text words and database subject headings, covering synonyms and related terms. Box 1 details the MEDLINE search strategy. Screening reference lists of included studies and relevant publications augmented the search.

Box 1: MEDLINE search strategy

1. spondylolisthesis.mp. or exp Spondylolisthesis/
2. spondylolysis.mp. or exp Spondylolysis/
3. stress fracture.mp. or exp Fractures, Stress/
4. pars interarticularis.mp.
5. 1 or 2 or 3 or 4
6. physical examination.mp. or exp Physical Examination/
7. physical test.mp.
8. clinical test.mp.
9. Diagnosis/ or Diagnosis, Differential/ or Diagnos*.mp.
10. palpation.mp. or Palpation/
11. symptom.mp. or exp Symptom Assessment/
12. (stability or instability).mp. [mp=title, abstract, original title, name of substance word, subject heading word, keyword heading word, protocol supplementary concept word, rare disease supplementary concept word, unique identifier]
13. patient history.mp.
14. accuracy.mp. or exp "Sensitivity and Specificity"/
15. exp "Reproducibility of Results"/ or reliability.mp.

16. exp Athletic Injuries/ or extension related stress injury.mp.
17. low back pain.mp. or Back Pain/ or exp Low Back Pain/ or exp Lumbar Vertebrae/
18. "range of motion".mp. or exp "Range of Motion, Articular"/
19. 6 or 7 or 8 or 10 or 11 or 12 or 13
20. 9 or 14 or 15
21. 9 or 16 or 17
22. 5 and 19 and 20 and 21

Eligibility criteria

The title and abstract of identified studies were screened by two reviewers (XX, XX) for eligibility using pre-specified inclusion criteria. Retrieved full texts were screened by the same two reviewers, and a third reviewer mediated any disagreement (XX). Inclusion criteria:

- Any study design using primary diagnostic accuracy data;
- Population with LBP with/without radiculopathy presenting with suspected spondylolysis and/or spondylolisthesis. An initial scoping search revealed few studies focused only to an athletic/young population. Therefore no age restriction was applied for study eligibility but the young/athletic population (aged 11-30 years and engaged in sport activities on a regular basis) was the focus of the analysis.
- Study investigating patient history and/or physical examination data, including specificity, sensitivity, likelihood ratios, and predictive values or presenting the raw data needed for calculation of these values.

Studies that did not compare patient history and/or physical examination data against diagnostic imaging (plain radiograph, magnetic resonance imaging, computed tomography), or where full texts were not available in English, were excluded.

Risk of bias

Two reviewers (XX,XX) independently conducted the risk of bias (ROB) assessment using the Quality Assessment of Diagnostic Accuracy Studies 2 (QUADAS-2)-tool (Whiting et al., 2011), a development of the original tool. The QUADAS-2 has been utilised in recent systematic reviews (Hegedus et al., 2012, Schwieterman et al., 2013) and comprises the same four domains: patient selection, index test, reference standard and flow, and timing. In addition, it rates applicability to the review question in three areas: patient selection, index test and reference standard. In line with the review objective, applicability was related to participants aged 11-30 years and/or engaged in sport activities on a regular basis. The ROB assessment enabled a judgment of ‘high’, ‘low’ or ‘unclear’ on individual items and a summary judgement of ‘at risk or low risk’. In line with the QUADAS guideline, a study is rated ‘at risk’ if it has one or more ‘unclear’ and/or ‘high’ judgements. The applicability assessment enabled a judgement of ‘with concerns’ or ‘no concerns’ (young/athletic population) regarding applicability. In line with the QUADAS guideline a study is rated ‘with concern’ if it has one or more ‘unclear’ and/or ‘low’ judgements. To ensure inter-agreement an initial ‘training’ discussion and analysis of the individual items of the QUADAS-2 was conducted with the two reviewers to agree implementation. In line with Whiting et al. (2011), review specific guidance was developed and the two reviewers conducted a training ROB assessment on 2 studies in an associated area where discrepancies were resolved through discussion. A third reviewer (XX) mediated any disagreements.

Data extraction and data items

Study characteristics and diagnostic accuracy data were extracted by one reviewer (XX) using a pre-designed data extraction sheet. The extraction sheet covered five areas: the first section contained information on the author and publication details, the second covered the studies' method sections (aim of study, study design, method of recruitment, eligibility criteria, ethical approval). The third covered participants (description, geography, setting, number, age, gender, ethnicity, diagnosis, stage of illness, other). The fourth covered data regarding the diagnostic tests (sensitivity, specificity, predictive values, likelihood ratios and other), and the fifth section was 2x2 contingency tables for the diagnostic tests. Data were audited by the second reviewer (XX). A third reviewer (XX) mediated any disagreements. Authors were contacted for additional data where necessary. Study characteristics and diagnostic data were collated and presented in tabular form for further analysis.

Summary measures

Sensitivity, specificity, predictive values (PV) and likelihood ratios (LR) are presented as summary measures. In cases where only raw data were presented, the sensitivity, specificity, predictive values and likelihood ratios were calculated according to the formulae of Couglin et al., (1992) and Akobeng (2007b) by one reviewer (XX) and audited by the second (XX). Level of accuracy of sensitivity and specificity was graded as low (<50%), low/moderate (51-64%), moderate (65-74%), moderate/high (75-84%) and high (>85%) (Schneiders et al., 2012). The discriminatory properties of the test was graded using positive (+) and negative (-) LR as: conclusive evidence (LR+ >10 and LR- <0.1), strong diagnostic evidence (LR+ 5-10 and LR- 0.1-0.2), weak diagnostic evidence (LR+ 2-5 and LR- 0.2-0.5), negligible evidence (LR+ 1-2 and LR- 0.5-1) (Jaeschke et al., 1994). Strength of agreement in reliability was graded according to Landis and Koch (1997) as: 0 poor, 0–0.20 slight, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 substantial and 0.81–1.00 almost perfect.

<u>Synthesis of result</u>

Pooling of diagnostic accuracy data and investigation for heterogeneity are the main aims for a meta-analysis (CRD, 2009). Heterogeneity was explored to evaluate if the studies were suitable for combining in an analysis. Study design, population, comparable diagnostic data and reference standard were considered for clinical comparison (Lijmer et al., 2002). No sub-group analyses were planned, as from the initial scoping search, it was not anticipated that sufficient studies would be available. Studies of high ROB tend to over-estimate diagnostic performance of a test (Lijmer et al., 2002, Rutjes et al., 2006), and were therefore excluded from a possible meta-analysis. *A priori*, it was decided that if meta-analyses were not clinically or statistically meaningful based on the criteria stated, a descriptive synthesis would be carried out.

RESULTS

Study identification

The searches identified 1512 studies. The screening of title and abstract resulted in 18 studies retrieved for full-text assessment. Ten studies did not meet the eligibility criteria, leaving 8 studies included in the analysis (Figure 1). Disagreements of study eligibility were resolved by discussion between the two reviewers (xx, xx).

Study description

Table 1 summarises the characteristics of included studies. Five studies investigated spondylolisthesis (Ahn and Jhun, 2015, Collaer et al., 2006, Ferrari et al., 2014, Kalpakcioglu et al., 2009, Moller et al., 2000). Two of the studies investigated patient history data (Kalpakcioglu et al., 2009, Moller et al., 2000), and all studies investigated physical examination data. No studies specifically investigated a young/athletic population.

Three studies investigated spondylolysis (Gregg et al., 2009, Masci et al., 2006, Sundell et al., 2013). Two studies investigated a population engaged in regular sporting activity aged <30 years (Masci et al., 2006, Sundell et al., 2013). One study (Gregg et al., 2009) failed to report mean age, but reported that 65.8% participants were aged <20 years and 70.2% of the participants were engaged in regular sporting activity. One study investigated patient history data (Gregg et al., 2009), and all 3 studies investigated physical examination data (Gregg et al., 2009, Masci et al., 2006, Sundell et al., 2013).

Figure 1. Flow diagram for study identification (from Moher et al, 2009)

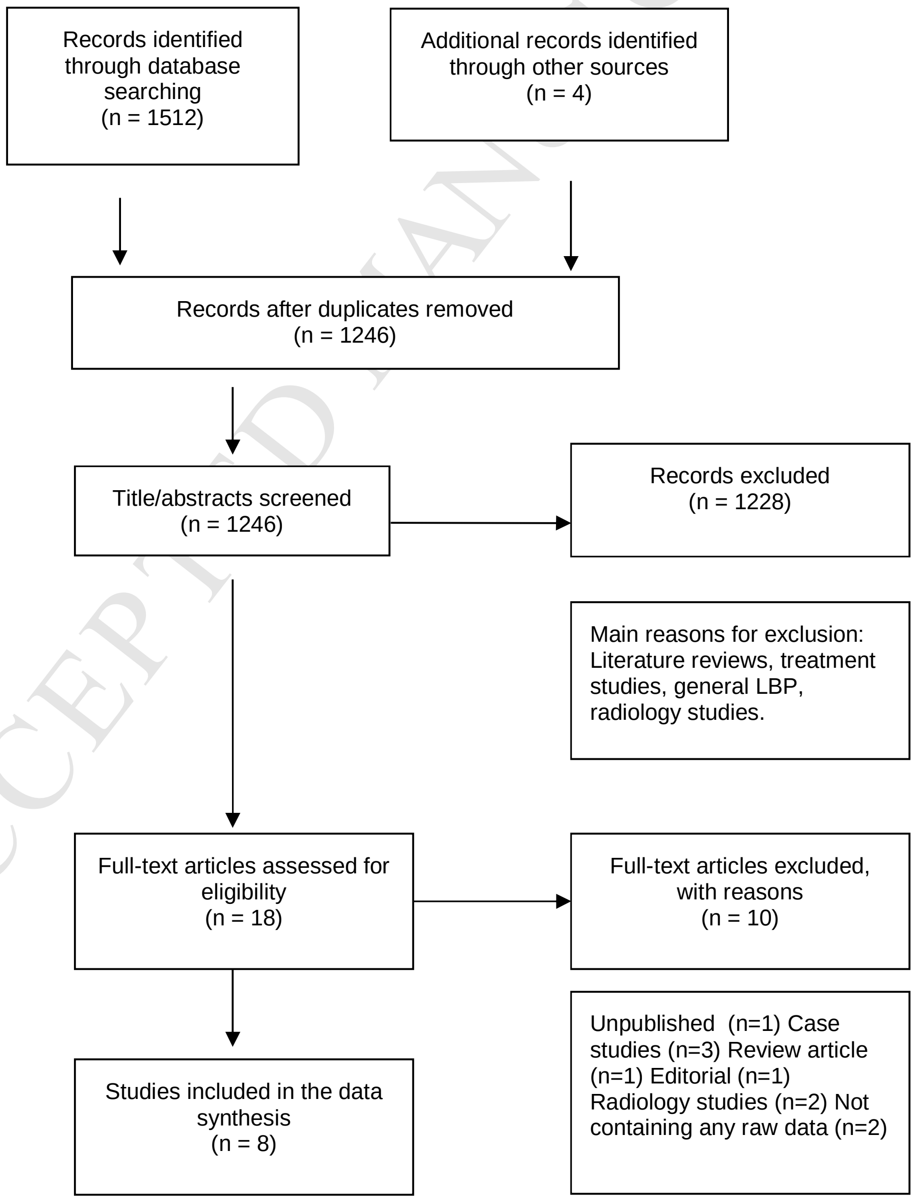


Table 1. Characteristics of included studies

Author, (year), Country	**Type of study**	**Pathology**	**Setting**	**Inclusion/Exclusion criteria**	**Population** (Number, gender, age)	**Outcome measures**	**Reference standard**	**Risk of bias**
Ahn et al., (2015)	Prospective cohort design	Spondylo-litshesis	PMC	LBP/lumbar radicular pain. Exclusion criteria: Contraindication for radiology, pregnancy, history of lumbar spine surgery, difficulty standing, unable to flex and extend the spine.	N = 86 65 women 31 men Mean age: 52.8 8 (± 13.9)	Step-deformity palpation/ Low midline sill sign (inspection and palpation)	Lumbar lateral radiography	At risk
Collaer et al., (2009) USA	Prospective cohort design	Spondylo-litshesis	H/SM	LBP and/or radiculopathy, age >16 years, no history of thoracic, lumbar, and sacral surgery, a same-day standing lateral lumbar radiograph	N = 30 15 women 15 men Mean age: 40 (±15)	Static palpation of step deformity	Lumbar lateral radiography	Low risk
Ferrari et al., (2014) Italy	Non - experimental	Spondylo-litshesis	PC	> 18 years, LBP with/without referred pain, diagnosis of spondylolisthesis confirmed by radiographs, computed tomography or magnetic resonance imaging (MRI ability in spoken and written Italian. Exclusion criteria: any previous lumbar surgery, systemic diseases (inflammatory, infectious, cancer, etc.), neuromuscular disorders, or cognitive deficits.	N = 119 67 women 52 men Mean age: 45.4 (±14.65)	ASLR, PLE, PIT, AM	Radiographs, MRI or CT	At risk
Gregg et al., (2009) New Zealand	Retrospective non-experimental	Spondylo-lysis	SM	All patients with LBP referred for a SPECT scan to confirm suspected diagnosis of spondylolysis over a 2-year period.	N = 82 39 women 43 men Mean age: data missing	Age at bone scan (Greater or less than 20 years old), gender, injured Period (Greater or less than 3 months), onset of symptoms (Sudden or Gradual), Sports participation (Yes or No), OLHET	SPECT scan	At risk
Kalpakcioglu et al., (2009) Turkey	Retrospective cohort design	Spondylo-litshesis	H	LBP and radiological diagnosis of spondylolisthesis (all patients except control group).	N = 130	Step deformity (palpation), Lumbar flexion, Lumbar extension, Lumbar lateral flexion, Lumbar rotation, SLR,	Antero-posterior, lateral,	At risk

				Exclusion criteria: Patients with inflammatory, infectious, metabolic, tumoral, toxic systemic diseases, radiological findings of severe degenerative changes, and those who had received physical therapy and rehabilitation or surgical treatment.	113 women 17 men Mean: 54.8 (SD not reported)	ASLR, Femoral stretch test, Achilles reflex, Patellar reflex, Loss of strength, Sensorial change, Weak/ dropping abdominal wall. Paravertebral muscle hypertrophy, Paravertebral muscle spasm, Increase in lumbar lordosis, Lumbar scoliosis, Signs of slipping (inspection). Hamstring muscle spasm , Contracting hamstring muscle , Z posture, Gait disorder , Walking distance <250 m	oblique and lateral flexion/extension radiograph	
Masci et al., (2006) Australia	Prospective cohort design	Spondylo-lysis	SM	Aged 10–30 years, engaged in regular activity, symptoms of LBP < 6 months, provisional diagnosis of active spondylolysis, referred for bone scintigraphy (with SPECT)/ computed tomography as the initial investigation. Exclusion criteria: contraindication to MRI and a recent history of bone scintigraphic evidence of active spondylolysis (within the preceding 12 months).	N = 71 Female/male ratio: not stated Age 10-30 Mean age not reported	OLHET	Bone scintigraphy with SPECT/CT	Low risk
Möller et al., (2000) Sweden	Cross-sectional clinical	Spondylo-litshesis	H	Lumbar isthmic spondylolisthesis of all grades with at least 1 year of low back pain or sciatica and severely restricted functional ability in patients aged 18–55 years. Exclusion criteria: Patients with mild symptoms, previous spine surgery, or alcohol or drug abuse.	N = 111 54 females and 57 males Mean age: 39 (SD = not reported)	EHL-reduced power, Positive SLR, Crossed SLR, Femoral stretch test, Lateral flexion, Hamstring tightness, Straight leg raise. Achilles reflex. Patellar reflex, Sensorial change, lumbosacral tenderness, sacro-iliac joint test positive. Mean age, mean age at onset of symptoms, women, men, earlier period in H due to LBP, analgesics, wake up during sleep due to pain, pain with coughing, worse at sitting, bladder dysfunction, bowel dysfunction, sexual dysfunction, sciatica, bilateral sciatica, organic pain drawing.	Radiographically confirmed and referred for surgery. Participants with sciatica examined with MRI or myelography	At risk

Sundell et al., (2013) Sweden	Case- series	Spondylo-lysis	PC	Adolescents 13–20 years, ≥ 6 h of sports participation/week and > 3 weeks of LBP, hindering their ADL or physical activity. Exclusion criteria: condition that could affect the results, metal placed in the body that could disturb the MRI investigation.	N = 25 11 females and 14 males Mean age: 15.3 (SD not reported)	OLHET, The prone back extension with fixed pelvis test, The coin test , The percussion test with reflex hammer, The rocking test, The sacrum nutation test , The HOOK test , MCI control test	MRI and CT	Low risk

PMC – Pain Management Clinic H – Hospital, SM – Sports medicine clinic, PC, Physiotherapy clinic, AM - Aberrant movement, OLHET – One-legged hyperextension test, ASLR – Active straight leg raise, PIT - Prone instability test, PLE - Passive lumbar extension test, EHL - extensor hallucis longus, SLR - straight leg raise, ASLR - Active SLR, CT - Computerised tomography, SPECT - Single-photon emission computerized tomography.

Risk of bias assessment

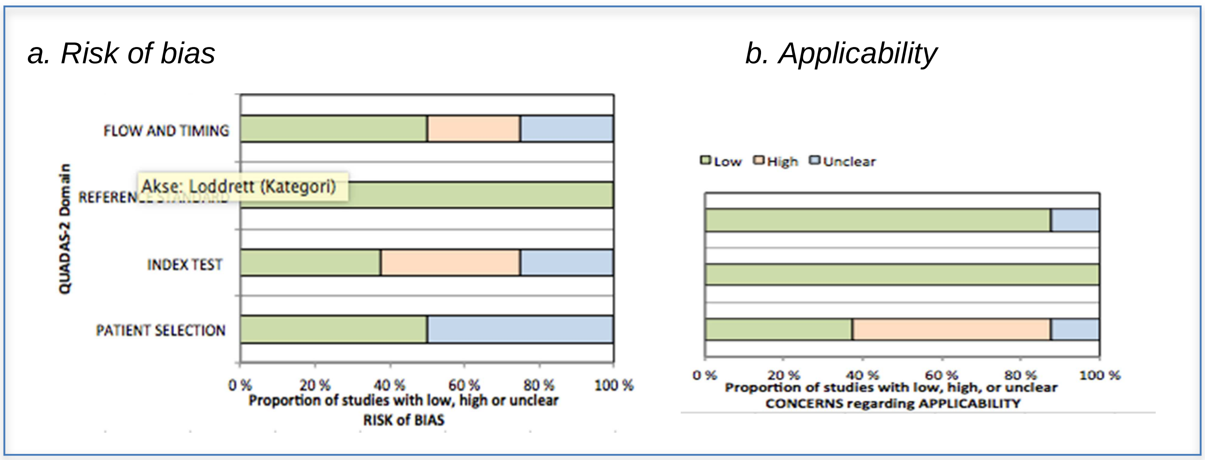
Two studies were assessed as low ROB (Collaer et al., 2006, Masci et al., 2006), but only 1 study had in addition no concerns with regards to applicability (Masci et al., 2006) (Table 2). Patient selection procedure, blinding to the results of the reference standard, and poor reporting of methodology, were the main concerns for ROB (Figure 2). Two further studies were assessed as low concern for applicability (Gregg et al., 2009, Sundell et al., 2013), the remaining studies main reason for concern was the high mean age. The reference standard varied across the 7 studies (Table 1). There was complete agreement between the two reviewers through discussion for the ROB and applicability assessment.

Table 2. Risk of bias assessment of included studies using QUADAS-2

Study	RISK OF BIAS				Summary	APPLICABILITY CONCERNS			Summary
	PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD	FLOW AND TIMING		PATIENT SELECTION	INDEX TEST	REFERENCE STANDARD	
Ahn et al, 2015	?	?	☺	☺	At risk	☹	☺	☺	With concern
Collaer et al, 2009	☺	☺	☺	☺	At risk Low risk	☹	☺	☺	With concern
Ferrari et al, 2014	☺	☹	☺	☹	At risk	☹	☺	☺	With concern
Gregg et al, 2009	?	?	?	☹	At risk	☺	☺	☺	With concern
Kalpakcioglu et al, 2009	?	☹	☺	?	At risk	?	☺	?	No concern
Masci et al, 2006	☺	☺	☺	☺	Low risk	☺	☺	☺	With concern
Möller et al, 2000	☺	☹	☺	?	At risk	☹	☺	☺	No concern
									With

Sundell et al, 2013	?	☺	☺	☺	At risk	☺	☺	☺	concern No concern

Figure 2. Proportion of studies assessed as low, high or unclear ROB and / or applicability



Synthesis of results

The main limitation for performing a meta-analysis was heterogeneity, based on:

- Limited numbers of studies (3 for spondylolysis and 5 for spondylolisthesis)
- Difference in study design case-series/retrospective cohort/prospective cohort/non-experimental/cross-sectional
- ROB assessment
- Difference in reference standard utilised
- Physical examination and patients history data utilised

Since a meta-analysis was not possible, the diagnostic accuracy data is presented as summary measures (sensitivity, specificity, predictive values and likelihood ratios) in tabular form.

Patient history data

Spondylolysis

One study (at ROB, with no concerns of applicability) assessed patient history data (Gregg et al., 2009). Age, gender, sport participation, injury period and onset of symptom data were assessed (Table 3). Sensitivity values were moderate (73%) for gender (male), moderate/high (81%) for age and sudden onset, and high for participation in sport (85%). However, the specificity values were low for age (44%) and sport participation (34%), and low-moderate for gender (male) (57%).

Table 3: Patient history data for spondylolysis

Author (year)	**Patient history**	**Diagnostic data**		
		Present in % of participants	Sensitivity	Specificity
Gregg et al., (2009)	Age at bone scan (<20 years old)	81.8	80.77	43.64
	Gender (Male)	73.1	73.08	57.4
	Injured Period (<3 months)	65.4	65.38	54.55
	Onset of symptoms (Sudden)	70.8	87.5	50.91
	Sports participation (Yes)	84.6	84.62	33.93

*Numbers in blue: calculations made by reviewers based on raw data presented in the original study.

Spondylolisthesis

Two studies (at ROB, with concerns of applicability) investigated patient history data for spondylolisthesis (Kalpakcioglu et al., 2009, Moller et al., 2000). ROB in both studies related to blinding. In addition, Kalpakcioglu et al (2009) did not report if all participants were included in data analyses, and Möller et al (2000) lacked clarity around which reference standard was utilised. A range of patient history data was investigated (Table 4). Sciatica was the only symptom investigated in both studies with low/moderate sensitivity 61% (Kalpakcioglu et al., 2009) and 68% (Moller et al., 2000) and low specificity 27% (Kalpakcioglu et al., 2009). Some reported symptoms from the Moller et al. (2000) study demonstrated moderate/high sensitivity: difficulty falling asleep (75%), waking up because of pain (81%), and pain worse with sitting (85%, high) and with walking (80%), but specificity was not investigated.

Table 4: Patient history data for spondylolisthesis

Author (year)	**Patient history**	**Diagnostic data**		
		Present in % of participants	Sensitivity	Specificity
Kalpakcioglu et al, (2009)	Pain localised to the low back	23	23	76.67
	Sciatca	61	61	26.67
	In gluteal region or backside of femur	16	16	96.67

Möller et al (2000)	Reduced physical condition	71	63.37	-
	Reduced mental condition	42	40.54	
	Difficulty falling asleep	75	74.77	
	Symptom free periods	8	8.11	
	Wake up during sleep due to pain,	81	81.08	
	Earlier period in hospital due to LBP	20	18.98	
	Pain with coughing,	39	37.84	
	Worse at sitting,	85	84.68	
	Worse at walking	81	80.18	
	Bladder dysfunction,	12	11.71	
	Bowel dysfunction,	13	12.61	
	Sexual dysfunction,	21	17.12	
	Analgesics,	62	62.26	
	Sciatica,	69	67.57	
	Bilateral sciatica,	28	27.03	
	Organic pain drawing	72	70.27	

Physical examination data

Spondylolysis

Three studies investigated physical examination data (Gregg et al., 2009, Masci et al., 2006, Sundell et al., 2013). Only one test was investigated in all studies, the one-legged hyperextension test with low to moderate sensitivity (50-73%) and low to high specificity (0-87%). No other test demonstrated collated sensitivity and specificity to a moderate level (Table 5).

Table 5: Diagnostic test accuracy data for spondylolysis

AUTHOR (YEAR)	PHYSICAL TEST	DIAGNOSTIC DATA					
		Sensitivity %	**Specificity %**	**+ ve LR**	**-ve LR**	**+ve PPV %**	**-ve PPV %**
Gregg et al., (2009)	OLHET	73	17.2	-	-	-	-
Masci et al., (2006)	OLHET	L side: 50 R side: 55	*(67.6)* 32.4 (45.2) 45.4	0.74 1.01	1.54 0.98	40.48 53.85	41.38 46.88
Sundell et al., (2013)	OLHET	61.54	0	0.62	-	40	0
	Prone back extension with fixed pelvis test	46.15	33.33	0.69	1.62	42.86	36.36
	Coin test	84.62	16.67	1.02	0.92	52.38	50
	Percussion test with reflex hammer	38.76	50	0.77	1.12	45.45	42.86
	Rocking test	69.23	25	0.92	1.23	50	42.86

Sacrum nutation test	23	58.33	0.55	1.32	37.50	41.18
HOOK test	46.15	75	1.85	0.72	66.67	56.25
MCI control test	69.23	50	1.38	0.62	60	60

*Numbers in blue: calculations made by reviewers based on raw data presented in the original study.
+ve – positive, -ve – negative, LR – likelihood ratio, PPV – predictive values, OLHET – One-legged hyperextension test.

Spondylolisthesis

Five studies investigated physical examination data (Ahn and Jhun., 2015; Collaer et al., 2006; Ferrari et al., 2014; Kalpakcioglu et al., 2009; Möller et al., 2000). Those at ROB did not have proper blinding and 2 studies did not include all participants in data analyses (Ferrari et al., 2014, Kalpakcioglu et al., 2009). All studies had concerns regarding applicability related to the age of the population.

Variants of step deformity palpation was investigated in 3 studies with sensitivity ranging from low/moderate 60% (Collaer et al., 2006), to moderate 81.3% (Ahn and Jhun, 2015) and high 88% (Kalpakcioglu et al., 2009). Specificity was high in all 3 studies: 87% (Collaer et al., 2006), 89.1% (Ahn and Jhun, 2015) and 100% (Kalpakcioglu et al., 2009). Active straight leg raise was investigated in two studies with low/moderate 64% (Ferrari et al., 2014) to high sensitivity 87% (Kalpakcioglu et al., 2009) and low 45% (Ferrari et al., 2014) to moderate/high specificity 77% (Kalpakcioglu et al., 2009). Paravertebral muscle hypertrophy had moderate sensitivity 65% and specificity 70% (Kalpakcioglu et al., 2009). Lumbar extension showed moderate/high sensitivity 79% and moderate specificity 67% (Kalpakcioglu et al., 2009). No other test demonstrated collated sensitivity and specificity to a moderate level. The passive lumbar extension test exhibited the best ability to predict the probability of a positive dynamic radiograph; more specifically for the participants with a positive passive lumbar extension, 84% were positive, decreasing to 54% in the group with negative passive lumbar extension.

Table 6. Diagnostic test accuracy data for spondylolisthesis

Author (Year)	**Physical test**	**Diagnostic data**						
		Sensitivity %	**Specificity %**	**+ ve LR**	**-ve LR**	**+ve PPV %**	**-ve PPV %**	**Other**
Ahn and Juhn, (2015)	Step deformity (palpation) *Authors name for the test: Low midline sill sign*	81.3	89.1	7.43	0.21	78.8	90.5	
Collaer et al., (2006)	Step deformity (palpation)	60	87	4.68	0.46	37.5	94.44	**Post-test probability** + ve 30 % - ve 5%
Ferrari et al., (2014)	ASLR	64	45	1.16	0.80	69	40	
	PIT	44	45	0.80	1.24	43	32	
	PLE	43	86	3.07	0.66	85	46	
	AM	41	77	1.78	0.76	76	42	
Kalpakcioglu et al., (2009)	Gait disorder	5	93.33	0.75	1.02	71.43	22.76	5
	Weak/dropping abdominal wall	99	40	1.65	0.03	84.62	92.31	99
	Paravertebral m. hypertrophy	65	70	2.17	0.50	87.84	37.50	65
	Paravertebral m. spasm	87	13.33	1.0	0.98	76.99	23.53	87
	Increase in lumbar lordosis	58	63	1.58	0.66	84.06	31.15	58
	Lumbar scoliosis	4	96.67	1.20	0.99	80	23.20	4
	Signs of slipping (inspection)	21	100	-	0.29	100	27.52	21
	Step deformity (palpation)	88	100	-	0.12	100	71.43	88
	Hamstring muscle spasm	27	96.67	8.10	0.76	96.43	28.43	27
	Contracting hamstring muscle	1	90	0.10	1.10	25	21.43	1
	Z posture	2	100	-	0.98	100	23.44	2
	Lumbar flexion	19	3.33	0.20	24.3	39.58	1.22	19
	Lumbar extension	79	66.67	2.37	0.31	88.76	48.78	79
	Lumbar lateral flexion	46	83.33	2.76	0.65	90.20	31.65	46
	Lumbar rotation	10	96.67	3.0	0.93	90.91	24.37	10
	Straight leg raise	10	90	1.0	1.0	76.92	23.08	10
	ASLR	87	76.67	3.73	0.17	92.55	63.89	87
	Femoral stretch test	14	96.67	4.20	0.89	93.33	25.22	14
	Achilles reflex	13	93.33	1.95	0.93	86.67	24.35	13
	Patellar reflex	8	96.67	2.40	0.95	88.89	23.97	8
	Loss of strength	1	96.67	0.30	1.02	50	22.66	1
	Sensorial change	2	100	-	0.98	100	23.44	2
	Walking distance <250 m	74	60	1.85	0.43	86.05	40.91	74
Möller et al., (2000)	EHL-reduced power	6.31	-	0.06	-	-	-	6
	Positive SLR	11.71		0.12				12
	Crossed SLR	0		-				0
	Femoral stretch test	1.80		0.02				2
	Lateral flexion	42.34		0.42				46
	Hamstring tightness	20.72		0.21				22
	Achilles reflex	5.41		0.05				6
	Patellar reflex	4.50		0.05				5
	Sensorial change	23.42		0.43				22
	Lumbosacral tenderness	66.67		0.67				68
	Sacro-iliac joint test positive.	6.31		0.06				7

*Numbers in blue: calculations made by reviewers based on raw data presented in the original study.

+ve – positive, -ve – negative, LR – likelihood ratio, PPV – predictive values, SLR – straight leg raise, ASLR – active straight leg raise, EHL – extensor halluces longus, PIT – prone instability test, PLE – passive lumbar extension, AM – aberrant movement.

DISCUSSION

The purpose of this review was to investigate the diagnostic utility of patient history and physical examination data to enable diagnosis of spondylolysis and spondylolisthesis in a young athletic population. It is considered an update of the Alqarni et al. (2015) study. In addition to updating the physical examination data with two additional studies, it includes an evaluation of patient history data, which is important in determining clinical diagnosis.

Spondylolysis

Gregg et al. (2009) found gender (male) to have moderate sensitivity. Epidemiological data support this finding with a reported male-to-female ratio of 3:1 for spondylolysis (Kalichman et al., 2009). Spondylolysis is suggested to have its onset in childhood and adolescence (Haun and Kettner, 2005) and Gregg et al. (2009) found age <20 to have moderate sensitivity. Sensitivity values were also moderate/high for sudden onset and high for participation in sport. Despite these promising findings, specificity values were low to low/moderate and the study did not report key methodological information, leaving it at ROB (Lijmer et al., 1999, Rutjes et al., 2006).

The one-legged hyperextension test has been suggested to have diagnostic utility for spondylolysis (Jackson et al., 1981, Micheli and Wood, 1995). Gregg et al. (2009) found moderate sensitivity for the one-legged hyperextension test, but the result was only available in 44 of the 87 participants questioning the validity of findings from a study at ROB. The low sensitivity value and low/moderate to low specificity values from one study with low ROB (Masci et al., 2006), confirms low diagnostic utility of the one-legged hyperextension test for spondylolysis. Sundell et al. (2013), a study not included in the recent review from Alqarni et al. (2015), recruited a young, athletic population, but the one-legged hyperextension test did not show any diagnostic utility in this population. Several hypotheses may explain this limited diagnostic utility of the one-legged hyperextension test. For example, extension will decrease the intervertebral foramina width (Fujiwara et al., 2001) and can affect neurological tissue, the one-legged hyperextension test has been used to diagnose sciatica due to herniated disc (Poiraudeau et al., 2001); and even though coupling of the lumbar spine is disputed (Cook and Showalter, 2004, Harrison et al., 1998), there is a similarity to the extension-rotation test suggested to assist in diagnosing facet joint dysfunction (Laslett et al., 2006). Sundell et al. (2013) included a range of tests not previously studied, but none of the included tests showed both sensitivity and specificity to a moderate level.

Overall for spondylolysis, no patient history or physical examination data have diagnostic utility to inform clinical practice. A well designed low ROB study is required to further investigate the diagnostic utility of patient history data: age, sudden onset of pain and participation in sport are aspects worth bringing forward because of its high sensitivity. A study may also usefully explore the diagnostic utility of clusters of patient history and physical examination data.

Spondylolisthesis

Difficulty falling asleep, waking up because of pain, and pain worse with sitting and walking have shown to have moderate/high to high sensitivity (Moller et al., 2000). However, pain worse with sitting and prolonged weight bearing positions are also features present with lumbar disc pathology (Chan et al., 2013). Specificity values need to be presented before these findings can be integrated into the patient history for diagnosing spondylolisthesis. For implementation in the young athletic population, where acute and repetitive injuries dominate, further studies are required. In the Moller et al. (2000) study, 71% participants had reduced physical condition and LBP for >1 year, and the mean age of participants in the included studies was 39 years (Moller et al., 2000) to 54.4 years (Kalpakcioglu et al., 2009) questioning its relevance to a young athletic population based on pathogenesis of the different types of spondylolisthesis. Degenerative spondylolisthesis follows degeneration of disc or facet and is often accompanied by symptoms of spinal stenosis and nerve root compression (Ulmer et al., 1994) which is rarely seen in a population under 40 years (Kalichman et al., 2009).

Step-deformity palpation was investigated in 2 studies, of low (Collaer et al., 2006) and high ROB (Kalpakcioglu et al., 2009), and was the only test with diagnostic values with moderate-high sensitivity and high specificity for spondylolisthesis. The low midline sill sign is described by the authors (Ahn and Jhun, 2015) as a new test for spondylolisthesis, but when comparing the description of performance of the test is it clearly a variant of step-deformity palpation. Kalpakcioglu et al., (2009) does not provide a description of how the test was performed, but Collaer et al., (2006) describes the test with the patients standing and palpation is performed by keeping a firm pressure and sliding the fingertips from the upper lumbar region to the sacrum palpating for the absence or presence of a step deformity. The low midline sill sign test (Ahn and Jhun, 2015) consists of inspection and palpation. Inspection is considered positive when a sill like the capital "L" is outlined in the lumbar spine. The palpation is considered positive when the upper spinous process in displaced anterior to the lower one and a sill like a capital "L" is palpated on the midline. Collaer et al. (2006), investigated inter-rater reliability of step-deformity palpation and found a kappa value of 0.179, 0.394, and 0.314 showing slight to fair agreement (Landis and Koch, 1977). The result is supported by other studies, with findings of a general low inter-agreement, for palpation of the spinous processes of the lumbar spine (Haneline and Young, 2009, Stovall and Kumar, 2010). Adding in the vast range of different palpation techniques utilised in clinical practice (Billis et al., 2003, Harlick et al., 2007, McKenzie and Taylor, 1997, Merz et al., 2013, Robinson et al., 2009) challenges clinicians to perform the tests as described to get the same diagnostic value. In summary, the step-deformity palpation is showing moderate to high diagnostic value in these individual studies, but the lack of uniformity in test description and performance will only promote a low inter-rater reliability and challenge the diagnostic usefulness for clinicians.

Hamstring tightness is advocated as a diagnostic sign for spondylolisthesis (Koerner and Radcliff, 2013, Stanitski, 2006). None of the studies investigating hamstring tightness described the test, but two commonly used methods described in literature are the straight leg raise and the finger to ground distance in forward bending (Goeken and Hof, 1993). The patient population in these studies had a high reported incidence of sciatica (61-69%) so it cannot be excluded that this had an influence on the results owing to involvement of neurological tissue (Rebain et al., 2002). Hamstring tightness has also been an observed phenomenon in non-specific LBP (Esola et al., 1996, Halbertsma et al., 2001, McClure et al., 1997, McHugh et al., 1998), but no significant correlation has been found between hamstring tightness and LBP (Hellsing, 1988, Nourbakhsh and Arab, 2002). Paravertebral muscle spasm and hypertrophy is promoted as a diagnostic sign for spondylolisthesis (McNeely et al., 2003), but with low specificity values (<14%) it is at risk of producing a high number of false positives (Akobeng, 2007a). The often-advocated signs of hamstring tightness and paravertebral muscle symptoms cannot therefore be justified, and should not be relied on in a clinical setting.

Based on current evidence, no patient history data has diagnostic utility for spondylolisthesis in athletes. A well-designed study presenting specificity values will help determine the diagnostic value of pain patterns aggravated by prolonged sitting and standing, pain worse at night and waking up form pain. Step deformity palpation has shown diagnostic utility in the form of the diagnostic data presented, but as the advance in evidence based practice and the reinforcement of clinical reasoning in the last couple of decades has shown us, we cannot rely solely on one clinical test to make a diagnosis. For both spondylolysis and spondylolisthesis, this review, in line with the conclusion of the Alqarni et al. (2015) review, accentuates the potential of utilising a more cluster-based approach in the diagnostic process. Looking more towards pooling of signs, symptoms and physical examination tests is a useful focus for future research.

Strengths and limitations

The strengths of this review are that it robustly synthesises the existing diagnostic data, highlights the need for methodologically stronger studies and provides a clear direction for future research. However, only a small number of studies were identified which made it impossible to calculate pooled estimates of the data. The poorly reported methodology in many studies left them at ROB and this is a reoccurring weakness particularly in studies investigating physical examination data (Lijmer et al., 1999, Rutjes et al., 2006). Future studies need to strengthen the methodological quality, especially around patient recruitment, blinding, description of the diagnostic tests and reporting patient flow and timing, to have any confidence in results. Being a reasonably new development, the QUADAS-2 requires further development, particularly as studies with unclear ROB components are managed the same as if they were high ROB.

CONCLUSION

Based on current evidence, no patient history or physical examination data has the diagnostic utility to identify spondylolysis in athletes. The commonly utilised one-legged hyperextension test does not possess diagnostic utility, and is not recommended. Conclusions are however limited by risk of bias. There is currently no evidence to support patient history data that can be utilised to identify spondylolisthesis in athletes. Step-deformity palpation demonstrates diagnostic utility in the general population, but no studies have investigated its diagnostic utility in a young/athletic population.

Well-designed low risk of bias studies using a sporting population are required to investigate the diagnostic utility of patient history and physical examination data individually and in clusters. The data support a focus on both patient history and physical examination data rather than the current emphasis on physical testing

REFERENCE LIST

Ahn, K. and Jhun, H.J. (2015) New physical examination tests for lumbar spondylolisthesis and instability: low midline sill sign and interspinous gap change during lumbar flexion-extension motion. **BMC Musculoskelet Disorder**, 16.

Alqarni, A.M., Schneiders, A.G., Cook, C.E., et al. (2015) Clinical tests to diagnose lumbar spondylolysis and spondylolisthesis: A systematic review. **Physical Therapy in Sport**, (0).

Beutler, W.J., Fredrickson, B.E., Murtland, A., et al. (2003) The natural history of spondylolysis and spondylolisthesis: 45-year follow-up evaluation. **Spine (Phila Pa 1976)**, 28: (10): 1027-1035; discussion 1035.

Billis, E.V., Foster, N.E. and Wright, C.C. (2003) Reproducibility and repeatability: errors of three groups of physiotherapists in locating spinal levels by palpation. **Manual Therapy**, 8: (4): 223-232.

Bossuyt, Davenport C, Deeks J, et al. (2013) Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy Version 0.9. . **The Cochrane Collaboration**, Chapter 11:Interpreting results and drawing conclusions.

Bossuyt and Leeflang, M. (2008) Chapter 6: Developing Criteria for Including Studies. **Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy**, Version 0.4 [updated September 2008]. (The Cochrane Collaboration, 2008.).

Chan, A.Y.P., Ford, J.J., McMeeken, J.M., et al. (2013) Preliminary evidence for the features of non-reducible discogenic low back pain: survey of an international physiotherapy expert panel with the Delphi technique. **Physiotherapy**, 99: (3): 212-220.

Collaer, J.W., McKeough, M. and Boissonnault, W.C. (2006) Lumbar Isthmic Spondylolisthesis Detection with Palpation: Interrater Reliability and Concurrent Criterion-Related Validity. **Journal of Manual & Manipulative Therapy (Journal of Manual & Manipulative Therapy)**, 14: (1): 22-29.

Cook, C. and Showalter, C. (2004) A survey on the importance of lumbar coupling biomechanics in physiotherapy practice. **Manual Therapy**, 9: (3): 164-172.

CRD (2009) Systematic Reviews: CRD's guidance for undertaking reviews in healthcare. **Centre for Reviews and Dissemination**, (Chapter 2).

de Vet, H., Eisinga A, Riphagen II, et al. (2008) Chapter 7: Searching for Studies. **Cochrane Handbook for Systematic Reviews of Diagnostic Test Accuracy**, Version 0.4 [updated September 2008]: (The Cochrane Collaboration, 2008.).

Esola, M.A., McClure, P.W., Fitzgerald, G.K., et al. (1996) Analysis of lumbar spine and hip motion during forward bending in subjects with and without a history of low back pain. **Spine (Phila Pa 1976)**, 21: (1): 71-78.

Ferrari, S., Vanti, C., Piccarreta, R., et al. (2014) Pain, Disability, and Diagnostic Accuracy of Clinical Instability and Endurance Tests in Subjects With Lumbar Spondylolisthesis. **Journal of Manipulative and Physiological Therapeutics**, (0).

Fredrickson, B.E., Baker, D., McHolick, W.J., et al. (1984) The natural history of spondylolysis and spondylolisthesis. **Journal of Bone and Joint Surgery Am**, 66: (5): 699-707.

Fujii, K., Katoh, S., Sairyo, K., et al. (2004) Union of defects in the pars interarticularis of the lumbar spine in children and adolescents. The radiological outcome after conservative treatment. **Journal of Bone and Joint Surgery Br**, 86: (2): 225-231.

Fujiwara, A., An, H.S., Lim, T.-H., et al. (2001) Morphologic Changes in the Lumbar Intervertebral Foramen Due to Flexion-Extension, Lateral Bending, and Axial Rotation: An In Vitro Anatomic and Biomechanical Study. **Spine (Phila Pa 1976)**, 26: (8): 876-882.

Goeken, L.N. and Hof, A.L. (1993) Instrumental straight-leg raising: results in healthy subjects. **Archives of Physical Medicine and Rehabilitation**, 74: (2): 194-203.

Gregg, C.D., Dean, S. and Schneiders, A.G. (2009) Variables associated with active spondylolysis. **Physical Therapy in Sport**, 10: (4): 121-124.

Halbertsma, J.P., Goeken, L.N., Hof, A.L., et al. (2001) Extensibility and stiffness of the hamstrings in patients with nonspecific low back pain. **Archives of Physical Medicine and Rehabilitation**, 82: (2): 232-238.

Haneline, M.T. and Young, M. (2009) A review of intraexaminer and interexaminer reliability of static spinal palpation: a literature synthesis. **Journal of Manipulative and Physiological Therapeutics**, 32: (5): 379-386.

Harlick, J.C., Milosavljevic, S. and Milburn, P.D. (2007) Palpation identification of spinous processes in the lumbar spine. **Manual Therapy**, 12: (1): 56-62.

Harrison, D.E., Harrison, D.D. and Troyanovich, S.J. (1998) Three-dimensional spinal coupling mechanics: Part I. A review of the literature. **Journal of Manipulative and Physiological Therapeutics**, 21: (2): 101-113.

Haun, D.W. and Kettner, N.W. (2005) Spondylolysis and spondylolisthesis: a narrative review of etiology, diagnosis, and conservative management. **Journal of Chiropractic Medicine**, 4: (4): 206-217.

Hegedus, E.J., Goode, A.P., Cook, C.E., et al. (2012) Which physical examination tests provide clinicians with the most value when examining the shoulder? Update of a systematic review with meta-analysis of individual tests. **British Journal of Sports Medicine**, 46: (14): 964-978.

Hellsing, A.L. (1988) Tightness of hamstring- and psoas major muscles. A prospective study of back pain in young men during their military service. **Upsala Journal of Medical Sciences**, 93: (3): 267-276.

Hoy, D., Bain, C., Williams, G., et al. (2012) A systematic review of the global prevalence of low back pain. **Arthritis & Rheumatism**, 64: (6): 2028-2037.

Iwamoto, J., Sato, Y., Takeda, T., et al. (2010) Return to sports activity by athletes after treatment of spondylolysis. **World journal of orthopedics**, 1: (1): 26.

Jackson, D.W., Wiltse, L.L. and Cirincoine, R.J. (1976) Spondylolysis in the female gymnast. **Clinical Orthopaedics and Related Research**, (117): 68-73.

Jackson, D.W., Wiltse, L.L., Dingeman, R.D., et al. (1981) Stress reactions involving the pars interarticularis in young athletes. **American Journal of Sports Medicine**, 9: (5): 304-312.

Jaeschke, R., Guyatt, G. and Sackett, D.L. (1994) Users' guides to the medical literature. III. How to use an article about a diagnostic test. A. Are the results of the study valid? Evidence-Based Medicine Working Group. **Jama**, 271: (5): 389-391.

Kalichman, L., Kim, D.H., Li, L., et al. (2009) Spondylolysis and spondylolisthesis: prevalence and association with low back pain in the adult community-based population. **Spine (Phila Pa 1976)**, 34: (2): 199-205.

Kalpakcioglu, B., Altinbilek, T. and Senel, K. (2009) Determination of spondylolisthesis in low back pain by clinical evaluation. **Journal of Back and Musculoskeletal Rehabilitation**, 22: (1): 27-32.

Koerner, J. and Radcliff, K. (2013) Spondylolysis in the Athlete. **Operative Techniques in Sports Medicine**, 21: (3): 177-184.

Kujala, U.M., Kinnunen, J., Helenius, P., et al. (1999) Prolonged low-back pain in young athletes: a prospective case series study of findings and prognosis. **European Spine Journal**, 8: (6): 480-484.

Landis, J.R. and Koch, G.G. (1977) The measurement of observer agreement for categorical data. **biometrics**, 159-174.

Laslett, M., McDonald, B., Aprill, C.N., et al. (2006) Clinical predictors of screening lumbar zygapophyseal joint blocks: development of clinical prediction rules. **Spine J**, 6: (4): 370-379.

Lijmer, J.G., Bossuyt, P.M. and Heisterkamp, S.H. (2002) Exploring sources of heterogeneity in systematic reviews of diagnostic tests. **Statistics in Medicine**, 21: (11): 1525-1537.

Lijmer, J.G., Mol, B., Heisterkamp, S., et al. (1999) EMpirical evidence of design-related bias in studies of diagnostic tests. **Jama**, 282: (11): 1061-1066.

Lonstein, J.E. (1999) Spondylolisthesis in children. Cause, natural history, and management. **Spine (Phila Pa 1976)**, 24: (24): 2640-2648.

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. **British journal of sports medicine**, 40: (11): 940-946.

McClure, P.W., Esola, M., Schreier, R., et al. (1997) Kinematic Analysis of Lumbar and Hip Motion While Rising From a Forward, Flexed Position in Patients With and Without a History of Low Back Pain. **Spine (Phila Pa 1976)**, 22: (5): 552-558.

McHugh, M.P., Kremenic, I., Fox, M.B., et al. (1998) The role of mechanical and neural restraints to joint range of motion during passive stretch. **Medicine and Science in Sports and Exercise**, 30: (6): 928-932.

McKenzie, A.M. and Taylor, N.F. (1997) Can Physiotherapists Locate Lumbar Spinal Levels by Palpation? **Physiotherapy**, 83: (5): 235-239.

McNeely, M.L., Torrance, G. and Magee, D.J. (2003) A systematic review of physiotherapy for spondylolysis and spondylolisthesis. **Manual Therapy**, 8: (2): 80-91.

Merz, O., Wolf, U., Robert, M., et al. (2013) Validity of palpation techniques for the identification of the spinous process L5. **Manual Therapy**, 18: (4): 333-338.

Micheli, L.J. and Wood, R. (1995) Back pain in young athletes. Significant differences from adults in causes and patterns. **Archives of Pediatric and Adolescent Medicine**, 149: (1): 15-18.

Moher, D., Liberati, A., Tetzlaff, J., et al. (2009) Preferred Reporting Items for Systematic Reviews and Meta-Analyses: The PRISMA Statement. **Journal of Clinical Epidemiology**, 62: (10): 1006-1012.

Moller, H., Sundin, A. and Hedlund, R. (2000) Symptoms, signs, and functional disability in adult spondylolisthesis. **Spine (Phila Pa 1976)**, 25: (6): 683-689; discussion 690.

Morita, T., Ikata, T., Katoh, S., et al. (1995) Lumbar spondylolysis in children and adolescents. **Journal of Bone and Joint Surgery Br**, 77: (4): 620-625.

Nourbakhsh, M.R. and Arab, A.M. (2002) Relationship between mechanical factors and incidence of low back pain. **Journal of Orthopaedic Sports Physical Therapy**, 32: (9): 447-460.

Peterson, M.C., Holbrook, J.H., Von Hales, D., et al. (1992) Contributions of the history, physical examination, and laboratory investigation in making medical diagnoses. **Western journal of medicine**, 156: (2): 163.

Poiraudeau, S., Foltz, V., Drape, J.L., et al. (2001) Value of the bell test and the hyperextension test for diagnosis in sciatica associated with disc herniation: comparison with Lasegue's sign and the crossed Lasegue's sign. **Rheumatology (Oxford)**, 40: (4): 460-466.

Rebain, R., Baxter, G.D. and McDonough, S. (2002) A systematic review of the passive straight leg raising test as a diagnostic aid for low back pain (1989 to 2000). **Spine (Phila Pa 1976)**, 27: (17): E388-395.

Robinson, R., Robinson, H.S., Bjørke, G., et al. (2009) Reliability and validity of a palpation technique for identifying the spinous processes of C7 and L5. **Manual Therapy**, 14: (4): 409-414.

Rossi, F. and Dragoni, S. (2001) The prevalence of spondylolysis and spondylolisthesis in symptomatic elite athletes: radiographic findings. **Radiography**, 7: (1): 37-42.

Rutjes, A.W., Reitsma, J.B., Di Nisio, M., et al. (2006) Evidence of bias and variation in diagnostic accuracy studies. **Canadian Medical Association Journal**, 174: (4): 469-476.

Saraste, H. (1986) Symptoms in relation to the level of spondylolysis. **International Orthopaedics**, 10: (3): 183-185.

Schmidt, C.P., Zwingenberger, S., Walther, A., et al. (2014) Prevalence of low back pain in adolescent athletes - an epidemiological investigation. **International Journal of Sports Medicine**, 35: (8): 684-689.

Schneiders, A.G., Sullivan, S.J., Hendrick, P.A., et al. (2012) The ability of clinical tests to diagnose stress fractures: a systematic review and meta-analysis. **Journal of Orthopaedic Sports Physical Therapy**, 42: (9): 760-771.

Schwieterman, B., Haas, D., Columber, K., et al. (2013) Diagnostic accuracy of physical examination tests of the ankle/foot complex: a systematic review. **International Journal of Sports Physical Therapy**, 8: (4): 416-426.

Soler, T. and Calderon, C. (2000) The prevalence of spondylolysis in the Spanish elite athlete. **American Journal of Sports Medicine**, 28: (1): 57-62.

Stanitski, C.L. (2006) Spondylolysis and Spondylolisthesis in Athletes. **Operative Techniques in Sports Medicine**, 14: (3): 141-146.

Stovall, B.A. and Kumar, S. (2010) Reliability of bony anatomic landmark asymmetry assessment in the lumbopelvic region: application to osteopathic medical education. **Journal of the American Osteopathic Association**, 110: (11): 667-674.

Sundell, C.G., Jonsson, H., Adin, L., et al. (2013) Clinical examination, spondylolysis and adolescent athletes. **International Journal of Sports Medicine**, 34: (3): 263-267.

Ulmer, J.L., Elster, A.D., Mathews, V.P., et al. (1994) Distinction between degenerative and isthmic spondylolisthesis on sagittal MR images: importance of increased anteroposterior diameter of the spinal canal ("wide canal sign"). **American Journal of Roentgenology**, 163: (2): 411-416.

Whiting, Rutjes, A.W., Westwood, M.E., et al. (2011) QUADAS-2: a revised tool for the quality assessment of diagnostic accuracy studies. **Annals of Internal Medicine**, 155: (8): 529-536.