Journal of Anatomy

J. Anat. (2020) **236**, pp3–20 doi: 10.1111/joa.13086

REVIEW ARTICLE

Changes in the macroscopic morphology of hip muscles in low back pain

Mohammadreza Pourahmadi,[1] Mohammad Asadi,[1] Jan Dommerholt[2] and Ali Yeganeh[3]

[1]*Department of Physiotherapy, School of Rehabilitation Sciences, Iran University of Medical Sciences, Tehran, Iran*
[2]*Bethesda Physiocare, Inc., Bethesda, MD, USA*
[3]*Trauma and Injury Research Center, Rasoul Akram General Hospital, Iran University of Medical Sciences, Tehran, Iran*

Abstract

Low back pain is a major health issue affecting the lumbopelvic muscles. Morphological changes in hip muscles, such as alterations in the muscle cross-sectional area and muscle volume, may occur in patients with low back pain. This systematic review was conducted to investigate whether patients with low back pain have macroscopic changes in their hip muscle morphology compared with asymptomatic, healthy individuals, based on current evidence. The electronic databases of PubMed/Medline, Ovid, Scopus, Embase®, and Google Scholar were searched from the inception to August 31, 2018. We only included full texts of original studies regarding macroscopic morphological alterations, including atrophy and fat infiltration, in hip muscles of patients with low back pain compared with asymptomatic controls. The quality of the included studies was determined using an assessment tool based on the Newcastle–Ottawa Scale. The scale was modified for the purposes of this study. Sixteen comparative observational studies were found eligible to be included in this review. Eleven were classified as high quality and four as moderate quality. The morphological changes in the psoas major, gluteus maximus, gluteus medius, gluteus minimus, and piriformis muscles were assessed in the primary studies. All selected studies were considered B level of evidence studies. The strength of conclusions for the psoas major, gluteal, and piriformis muscles was moderate. The results revealed that there is substantial controversy about the morphological changes in hip muscles in patients with low back pain; however, the majority of high-quality studies concluded that atrophy of hip muscles is evident in patients with low back pain. The psoas major muscle was the most commonly investigated hip muscle for morphological changes. Major methodological limitations of the included studies were identified and discussed. The present systematic review does not include a formal meta-analysis because of very significant differences in the primary studies in terms of study populations and methodologies. Finally, in clinical practice, it is recommended that physical therapists develop exercise programs to improve hip muscle function in patients with low back pain.

Key words: cross-sectional area; hip muscles; low back pain; morphological changes; review.

Introduction

Low back pain (LBP) is the main musculoskeletal disorder responsible for disability worldwide (Hoy et al. 2014; Maher et al. 2017; Hartvigsen et al. 2018; Oliveira et al. 2018), with 84% of people in all age groups expected to experience LBP during their lifetime (Balagué et al. 2012; Ostelo, 2018). A systematic review showed that approximately 12% of the general population are affected by LBP (Hoy et al. 2012). Moreover, there is strong evidence that LBP may result in significant health and socioeconomic problems, such as work absenteeism and high costs, for both patients and society (Saragiotto et al. 2016). Trunk muscle degeneration is a common feature in patients with LBP (Parkkola et al. 1993; Danneels et al. 2000). Macroscopic trunk muscle degeneration is characterized by a decrease in the cross-sectional area (CSA; Fortin & Macedo, 2013), smaller functional CSA (FCSA), also referred to as lean muscle mass (Pourahmadi et al. 2016), an altered FCSA/CSA ratio (Pourahmadi et al. 2016), and an increase in the amount of fat content of the lumbar erector spinae muscles (Mengiardi et al. 2006; Yanik et al. 2013). Fortin & Macedo (2013) reported that the erector spinae muscles are significantly smaller in patients with chronic LBP than in control patients. Goubert

Correspondence
Mohammadreza Pourahmadi, Department of Physiotherapy, School of Rehabilitation Sciences, Iran University of Medical Sciences, Madadkaran All., Shahnazari St., Madar Sq., Mirdamad Blvd., PO Box 4391-15875, Tehran 1545913187, Iran. E: pourahmadipt@gmail.com

Accepted for publication *10 August 2019*
Article published online *1 September 2019*

et al. (2016) identified multifidus and erector spinae muscle atrophy in patients with chronic LBP.

On the other hand, deficits in hip muscle endurance, strength, and motor control have been identified in individuals with LBP, yet it is unknown whether these deficits are a cause or an effect of LBP (Amabile et al. 2017). The hip muscles can balance the forces applied directly to the pelvis by the trunk muscles. The postural function of some deep hip muscles (e.g. psoas major) has been shown to maintain the lumbar lordosis and stabilize the lumbar spine, sacroiliac, and hip joints during sitting as well as standing, walking and running (Arbanas et al. 2012). Mayoux-Benhamou et al. (1994) mentioned that the CSA of the psoas major muscle is correlated with lumbar lordosis; when the psoas major muscle is atrophied, the lumbar curvature is accentuated. Moreover, abnormal changes in lumbar lordosis alter muscle activity and stress patterns, leading to the development of LBP (Kendall et al. 2005). Insufficiency of other hip muscles has also been shown to be associated with the development of LBP (Lee & Kim, 2015). Neumann (2017) alleged that hypofunction of the gluteal muscles would cause instability in the lumbopelvic region. Decreased gluteus medius muscle activity has been shown to be a predictor of the presence of LBP (Cooper et al. 2016).

Muscle inhibition and atrophy are catastrophic consequences frequently observed in the context of pain (Rantanen et al. 1993; Goubert et al. 2016). Muscle inhibition is also evident in muscle groups distant from the site of pain (Suter & McMorland, 2002). Previous studies provided evidence that pain-related nerve inhibition reduces muscle activity to prevent the motion of the painful area and tissue damage (Rantanen et al. 1993). Falla & Farina (2008) reported that pain can influence muscle structures by compromising muscle contraction, strength, force production, and motor unit activity. Comprised muscle function caused by pain can ultimately result in altered muscle structure (Falla & Farina, 2008). Insights into whether structural muscle alterations happen and how hip muscles specifically change in patients with LBP are important for the prevention and management of LBP. Hence, this study was carried out to review the published literature critically and to evaluate the macroscopic morphological changes in hip muscles in the presence of LBP. To date, no relevant reviews have been published on this topic. In addition, systematic reviews can be very useful decision-making tools because they objectively summarize large volumes of research evidence and identify critical research gaps.

Materials and methods

Search strategy

This systematic review was performed in accordance with the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA; Moher et al. 2009) and the Cochrane group guideline recommendations (Higgins & Green, 2011). A review protocol was registered on PROSPERO (International Prospective Register of Systematic Reviews; http://www.crd.york.ac.uk/PROSPERO/; identification no. CRD42018108800). Ethical approval and patient consent were not required for this systematic review. Wide electronic search strategies were constructed with the combined keywords including *muscle morphology*, *hip, back pain*, and *comparative study* to search English-language human studies in the peer-reviewed literature investigating hip muscle morphology in adult patients (≥ 18 years) with LBP. No restrictions were imposed on assessment instruments for hip muscle morphology, techniques or position of participants during testing.

PubMed/Medline (NLM), Ovid, Scopus, Embase®, and Google Scholar databases were searched from inception to 31 August 2018. Medical Subject Heading (MeSH) terms were recognized and all trees of these MeSH terms were searched in the biomedical literature engines. Therefore, the syntax of this review was a combination of MeSH terms and free-text words. The Boolean operator 'OR' was used to include synonyms and the Boolean operator 'AND' was used to combine and narrow down the searches. To retrieve all the possible variations of a specific root word, wildcards and truncations were used as well. The strategy was slightly modified for searches of other databases. Details of search syntax for PubMed/Medline (NLM) databases are provided in Supporting Information Appendix S1. To retrieve only comparative observational studies, a modified version of the syntax developed by Furlan et al. (2006) was used. Citation tracking and reference lists scanning of the included studies and relevant reviews were searched for eligible studies. Manual search of keywords via the internet was also performed. We did not review content from file sources that were from mainstream publishers (e.g. Wiley, ScienceDirect, Sage, BMJ, and BioMed Central), as we expected these to be captured in our broader search strategy.

Eligibility criteria

At the completion of the search, all references were imported into the EndNote referencing software (version X8; Thomson Reuters, New York, NY, USA) and duplicates were removed. Titles and abstracts of all primary articles that met the search strategy were scrutinized by two reviewers (M.P. and M.A.) to determine studies eligible for inclusion. In the absence of sufficient information in the title and abstract of an article, a full-text evaluation was undertaken. The same two reviewers then evaluated the full text of potentially relevant non-duplicated papers. All disagreements were resolved by discussion between the reviewers. Where there was no consensus, a third reviewer (A.Y.) acted as arbitrator. Studies were screened for selection according to the review objectives and Participants, Diagnosis, Comparison, Outcomes, and Study design criteria (Pourahmadi et al. 2019):

- Participants and Diagnosis: Studies in which participants were adult patients (≥ 18 years) of both genders with LBP. LBP was defined as pain or discomfort on the posterior aspect of the trunk from the lower margin of the 12th ribs to the lower gluteal folds with or without pain referred into one or both lower limbs that lasts for at least 1 day (Hoy et al. 2014; Pourahmadi et al. 2018b).
- Comparison: Studies were included in which the participants were adult individuals without symptoms of LBP, who had never undergone lumbar spine surgery. Studies were not included in this review when the unaffected side was only considered as a control.

- Outcomes: Macroscopic morphological changes in hip muscles included the assessment of CSA, FCSA, fat infiltration, fat deposit, muscle density, and muscle volume. No restriction was placed on the instruments and techniques used to measure the macroscopic morphological changes, position of participants, test sites, and target muscles. Microscopic changes such as changes in fiber distribution were not assessed in this review.
- Study design: Comparative observational studies published in peer-reviewed journals with full text available in English. Results obtained from theses/dissertations, conference proceedings, abstracts, policy documents, commercial documents, and websites were excluded.

Quality assessment of eligible studies

The quality of each primary study was assessed with the Newcastle–Ottawa Scale (NOS) for comparative observational studies (Stang, 2010). The NOS is recommended by the Cochrane Non-Randomized Studies Methods Working Group to assess the quality of observational studies. The original scale, which is very comprehensive, is based on the following three subscales: Selection (4 items), Comparability (1 item), and Outcome or Exposure (3 items; Griffin et al. 2012). Considering the purposes of this review, the modified version of the NOS was used (Griffin et al. 2012) for the reliability and validity assessment of the degree of muscle morphology utilized in the study. The psychometric properties of the measurement of muscle morphology were deemed essential in determining the overall external validity of the study. Additionally, aspects of the statistical analysis in the original reports were also evaluated, including sample size justifications and appropriateness and clarity of the statistical analysis method presentations (Griffin et al. 2012). Differences in age, gender, and physical activity were considered to investigate the comparability subscale of the NOS (Griffin et al. 2012). Finally, a total score of 3 or less was considered poor, scores of 4–6 were considered moderate, and scores of 7–10 high quality (Yong et al. 2018). Unacceptable bias was defined as a zero score in any of the NOS subscales.

The quality assessment was conducted independently by two reviewers (M.P. and M.A.). The level of inter-rater agreement was measured with Cohen's kappa coefficient using a method developed for comparing the level of agreement with categorical data along with their respective 95% confidence intervals (κ 0–0.20 = poor agreement; 0.21–0.40 = fair agreement; 0.41–0.60 = moderate agreement; 0.61–0.80 = good agreement; and 0.81–1 = very good agreement; Pourahmadi et al. 2018b). Any discrepancies were resolved through discussion. When no consensus was reached, a third reviewer (A.Y.) acted as arbitrator. The quality assessment score was not decisive for inclusion in this study but was taken into account while presenting the results.

Level of evidence and strength of conclusion

The level of evidence applied to each study was based on the 2005 classification system of the Dutch Institute for Healthcare Improvement [CBO; Meeus & Gebruers, 2016; Table 1]. In addition, the strength of conclusion was determined for each hip muscle by considering the level of evidence of the included studies and the consistency of the reported results (Meeus & Gebruers, 2016). The strength of conclusion was classified according to De Meulemeester et al. (2017) as: (i) high, (ii) moderate, (iii) low, and (iv) very low (Table 2).

Data extraction

Data extraction was carried out by two reviewers (M.P. and M.A.). A customized data extraction form was developed for the outcomes of interest, including morphometric changes in hip muscles. The data extraction form was a Microsoft Excel spreadsheet (Microsoft, Redmond, Washington, DC, USA) designed according to the Cochrane meta-analysis guidelines and adjusted to the needs of this review. The following information was documented for each paper which met the inclusion criteria: first author's name, year of publication, location, participant characteristics, definition of patient group, specific measurement techniques employed, type and name of outcome measure(s), test site and participant position, study results, and any other relevant details.

A meta-analysis was not conducted, because the original studies were highly heterogeneous in terms of LBP sub-classification and methodologically different in relation to measurement of the morphological changes in muscles. Hence, this review focused only on a descriptive and qualitative synthesis of the searched studies.

Results

Identification of studies

A total of 674 publications were detected in the initial literature search: 671 in the electronic databases (D'Hooge et al. 2012; Arbanas et al. 2013; Singh et al. 2016) and three through other sources (i.e. reference lists of relevant papers, reviews, and manual search of keywords). After screening the abstracts and analyzing the eligibility of the full-text papers, 16 studies remained and were included in our qualitative analysis (Parkkola et al. 1993; Dangaria & Naesh, 1998; Danneels et al. 2000; Kamaz et al. 2007; Hides et al. 2008; Stewart et al. 2010; D'Hooge et al. 2012, 2013; Arbanas et al. 2013; Gildea et al. 2013; Abbas et al. 2016; Hyun et al. 2016; Singh et al. 2016; Skorupska et al. 2016; Amabile et al. 2017; Sions et al. 2017). Fifteen full-text studies were excluded as they did not recruit healthy controls or patients with LBP or had an ineligible design (Barker et al. 2004; Ranson et al. 2006; Kang et al. 2007; Kalichman et al. 2010; Bouche et al. 2011; Kim et al. 2011; Sanchis-Moysi

Table 1 Levels of evidence (adopted from Meeus & Gebruers, 2016)

Level	Intervention
A1	Systematic review and meta-analyses based on a minimum of 2 independent conducted studies of evidence level A2
A2	Randomized controlled trials: double blinded, with sound methodology and sufficient sample size
B	Comparative studies, but lacking the quality criteria of A2 (including cohort studies, case-control studies, and randomized controlled trials of moderate quality or insufficient sample size)
C	Non-comparative studies
D	Expert opinion

Table 2 Strength of conclusion (adopted from Meeus & Gebruers, 2016)

Level	Conclusion based on
1	One study of evidence level A1 or at least 2 independently conducted studies of evidence level A2
2	One study of evidence level A2 or at least 2 independently conducted studies of evidence level B
3	One study of evidence level B or C or conflicting evidence (inconsistent results)
4	Expert opinion

et al. 2011; Akgul et al. 2013; Joseph et al. 2015; Bhadresha et al. 2016; Jeon et al. 2016; Salah El-din Mahmoud et al. 2016; Verla et al. 2016; Lee et al. 2017; Sasaki et al. 2017) The procedure is displayed in a flow chart (Fig. 1).

Quality assessment

The result of the quality assessment is presented in Table 3. The level of inter-rater agreement of quality assessment was good ($\kappa = 0.61 \pm 0.25$). In terms of quality assessment, a median NOS score of 7 [interquartile range (IQR) = 6–8] indicated a high methodological quality of the included studies. More specifically, among 16 studies, 11 studies were of high quality according to the NOS scale (Parkkola et al. 1993; Dangaria & Naesh, 1998; Danneels et al. 2000; Kamaz et al. 2007; Stewart et al. 2010; D'Hooge et al. 2012, 2013; Arbanas et al. 2013; Gildea et al. 2013; Abbas et al. 2016; Amabile et al. 2017) and five studies were rated as moderate quality (Hides et al. 2008; Hyun et al. 2016; Singh et al. 2016; Skorupska et al. 2016; Sions et al. 2017). Sample size justification, case–control matching for physical activity, and selection of control group were the criteria that most frequently, were not met. The percentage of studies that met each modified NOS item is shown in Table 3.

Level of evidence and strength of conclusion

Relevant studies were located and categorized as presented in Table 4. Following the analysis of the level of evidence, all selected studies were classified as level of evidence 'B'. In addition, the strength of conclusion was moderate for the psoas major, gluteus maximus, gluteus medius, gluteus minimus, and piriformis muscles (Table 4).

Description of participant characteristics

Table 5 provides a summary of the number of participants recruited, along with their health status, gender, and age. A total of 1218 participants were originally recruited in the 16 studies. Among these 1218 participants, 627 (51%) had LBP and 591 (49%) were asymptomatic. Dangaria & Naesh (1998) did not specify the gender of their participants. Nevertheless, male and female participants made up approximately 41 and 59%, respectively, of the total sample. Two studies enrolled only female participants (Kamaz et al. 2007; Amabile et al. 2017) and two studies included only male participants (Hides et al. 2008; Stewart et al. 2010); other studies assessed the morphological changes in hip muscles in both males and females (Parkkola et al. 1993; Danneels et al. 2000; D'Hooge et al. 2012, 2013; Arbanas

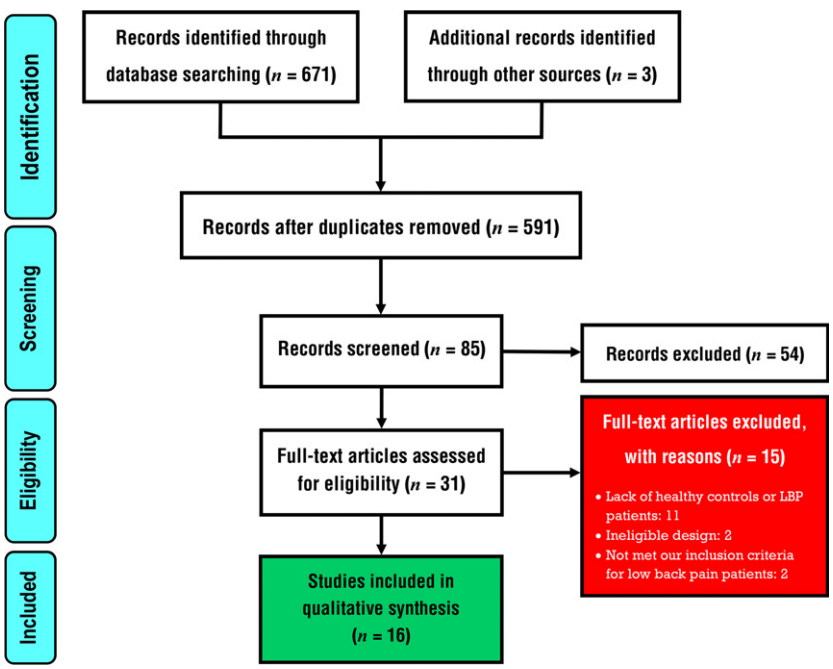

Fig. 1 PRISMA diagram showing the flow of studies through phases of the review.

Table 3 Quality assessment for each included study

	Selection				Comparability		Outcome measurement		Statistical analysis		
Study	Patient definition[1]	Representativeness of patients[2]	Selection of controls[3]	Definition of controls[4]	Age and gender[5]	Physical activity[6]	Reliability of outcome measure[7]	Validity of outcome measure[8]	Sample size[9]	Statistical method[10]	Total score
Amabile et al. (2017)	1	1	0	1	1	0	1	1	0	1	7[11]
Sions et al. (2017)	1	0	1	0	1	0	1	1	0	1	6
Hyun et al. (2016)	1	0	0	1	1	0	1	1	0	1	6
Abbas et al. (2016)	1	1	1	1	1	0	1	1	1	1	9[11]
Skorupska et al. (2016)	1	0	0	1	0	0	1	1	0	1	5
Singh et al. (2016)	1	1	0	1	1	0	0	1	0	1	6
Gildea et al. (2013)	1	0	1	0	1	1	1	1	0	1	7[11]
D'Hooge et al. (2013)	1	1	0	1	1	1	1	1	0	1	8[11]
Arbanas et al. (2013)	1	1	1	1	1	0	0	1	1	1	8[11]
D'Hooge et al. (2012)	1	1	0	1	1	1	0	1	0	1	7[11]
Stewart et al. (2010)	0	1	1	1	0	1	1	1	0	1	7[11]
Hides et al. (2008)	0	1	1	0	0	1	1	1	0	1	6
Kamaz et al. (2007)	1	1	1	1	1	1	1	1	0	1	9[11]
Danneels et al. (2000)	1	1	1	1	1	1	1	1	0	1	9[11]
Dangaria & Naesh (1998)	1	1	1	1	1	0	0	1	0	1	7[11]
Parkkola et al. (1993)	1	1	1	1	1	1	0	1	0	1	8[11]
Percentage of articles meeting each NOS item	87.5%	75%	62.5%	81%	81%	50%	69%	100%	12.5%	100%	

[1]The inclusion/exclusion criteria are clearly defined.
[2]How were cases selected? (e.g. random sample).
[3]Controls were derived from the same community as patients.
[4]Controls defined as individuals with no history of low back pain during at least the past 6 months.
[5]The patient group and control group were adequately matched for age and gender.
[6]The patient group and control group were adequately matched for physical activity.
[7]The measure of muscle morphological changes has documented reliability.
[8]The measure of muscle morphological changes has documented validity.
[9]The sample size was justified.
[10]The statistical analysis was clearly presented and was appropriate.
[11]Indicates high-quality studies.

Table 4 Level of evidence of each included study and strength of conclusion for each hip muscle

Muscle	Study	Level of evidence	Strength of conclusion
Psoas Major	Sions et al. (2017)	B	2 (moderate)
	Hyun et al. (2016)	B	
	Abbas et al. (2016)	B	
	Singh et al. (2016)	B	
	Gildea et al. (2013)	B	
	D'Hooge et al. (2013)	B	
	Arbanas et al. (2013)	B	
	D'Hooge et al. (2012)	B	
	Stewart et al. (2010)	B	
	Hides et al. (2008)	B	
	Kamaz et al. (2007)	B	
	Danneels et al. (2000)	B	
	Dangaria & Naesh (1998)	B	
	Parkkola et al. (1993)	B	
Gluteus maximus	Amabile et al. (2017)	B	2 (moderate)
	Kamaz et al. (2007)	B	
Gluteus medius	Skorupska et al. (2016)	B	2 (moderate)
Gluteus minimus	Skorupska et al. (2016)	B	2 (moderate)
Piriformis	Skorupska et al. (2016)	B	2 (moderate)

et al. 2013; Abbas et al. 2016; Hyun et al. 2016; Singh et al. 2016; Skorupska et al. 2016).

The mean and standard deviation (SD) of age of the participants was equal to 46.78 $\pm$ 11.68 years old at baseline. Sions et al. (2017) did not provide the SD of age for their sample size. The majority of the included studies (seven studies; 44%) enrolled patients with chronic LBP (Parkkola et al. 1993; Danneels et al. 2000; Kamaz et al. 2007; Arbanas et al. 2013; Singh et al. 2016; Amabile et al. 2017; Sions et al. 2017). Two studies recruited patients with unilateral recurrent non-specific LBP (D'Hooge et al. 2012, 2013), whereas one study included patients with low back and leg pain (Skorupska et al. 2016). One study included patients with degenerative lumbar kyphosis (Hyun et al. 2016). One study enrolled patients with degenerative lumbar spinal stenosis (Abbas et al. 2016), and one study included patients with unilateral sciatica caused by disc herniation (Dangaria & Naesh, 1998). Three studies did not provide detailed information on the LBP subclassification (Hides et al. 2008; Stewart et al. 2010; Gildea et al. 2013). Participants in one study by D'Hooge et al. (2012) seem to be the same as those in another study also by D'Hooge et al. (2013), as similar demographic characteristics were reported in both studies. Finally, a calculation of sample size was performed in two studies (Arbanas et al. 2013; Abbas et al. 2016).

Methodology considerations and outcome measures

Three studies (~ 19%) were conducted in Belgium (Danneels et al. 2000; D'Hooge et al. 2012, 2013). Another three studies originated from Australia (Hides et al. 2008; Stewart et al. 2010; Gildea et al. 2013), and the remaining studies were from the USA (Amabile et al. 2017; Sions et al. 2017), Poland (Skorupska et al. 2016), Croatia (Arbanas et al. 2013), Finland (Parkkola et al. 1993), Israel (Abbas et al. 2016), Brunei (Dangaria & Naesh, 1998), India (Singh et al. 2016), Turkey (Kamaz et al. 2007), and Korea (Hyun et al. 2016). Eleven studies used magnetic resonance imaging (MRI) to assess the morphological changes in the hip muscles (Parkkola et al. 1993; Dangaria & Naesh, 1998; Hides et al. 2008; Stewart et al. 2010; D'Hooge et al. 2012, 2013; Arbanas et al. 2013; Gildea et al. 2013; Singh et al. 2016; Skorupska et al. 2016; Sions et al. 2017). Four studies utilized computed tomography (CT) scans (Danneels et al. 2000; Kamaz et al. 2007; Abbas et al. 2016; Amabile et al. 2017). One study used both MRI and CT scan techniques (Hyun et al. 2016; Table 5).

Fourteen comparative observational studies assessed muscle CSA (Parkkola et al. 1993; Dangaria & Naesh, 1998; Danneels et al. 2000; Kamaz et al. 2007; Hides et al. 2008; Stewart et al. 2010; D'Hooge et al. 2012, 2013; Arbanas et al. 2013; Gildea et al. 2013; Abbas et al. 2016; Singh et al. 2016; Amabile et al. 2017; Sions et al. 2017) and four studies (Danneels et al. 2000; D'Hooge et al. 2012; Arbanas et al. 2013; Hyun et al. 2016; Sions et al. 2017) also investigated the fatty degenerative changes in hip muscles. Three-dimensional muscle volume calculation was performed in the study by Skorupska et al. (2016). Two studies evaluated the CSA of the psoas major at five lumbar spine levels (L_1–L_5; Gildea et al. 2013; Hyun et al. 2016), and other studies examined the CSA of the psoas major muscle at different lumbar spine levels (Parkkola et al. 1993; Dangaria & Naesh, 1998; Danneels et al. 2000; Kamaz et al. 2007; Hides et al. 2008; Stewart et al. 2010; D'Hooge et al. 2012, 2013; Arbanas et al. 2013; Abbas et al. 2016; Singh et al. 2016; Sions et al. 2017). The L_3 vertebra and L_3–L_4 intervertebral disc level were the most common sites for psoas major muscle measurements.

The reliability of measurements techniques used in the included studies was moderate to excellent [intra-class correlation coefficients (ICCs) 0.58–0.99], according to the scale developed by Bland & Altman (1999) (ICCs $\leq$ 0.20 poor, 0.21–0.40 fair, 0.41–0.60 moderate, 0.61–0.80 good, and 0.81–1.00 excellent; Table 5). The included studies reported contrasting results in terms of differences in the psoas major muscle CSA between patients with LBP and asymptomatic participants (Parkkola et al. 1993; Dangaria & Naesh, 1998; Danneels et al. 2000; Kamaz et al. 2007; Hides et al. 2008; Stewart et al. 2010; D'Hooge et al. 2012, 2013; Arbanas et al. 2013; Abbas et al. 2016; Hyun et al. 2016; Singh et al. 2016). Hyun et al. (2016), D'Hooge et al. (2013, 2012), Hides et al. (2008), and Danneels et al. (2000) stated that no significant differences were found for the psoas major muscle CSA between the two groups at any lumbar spine level. D'Hooge et al. (2012) reported that the there were no

Table 5 Description of studies assessing morphological changes in hip muscles in patients with LBP

No.	Study characteristics	Participant characteristics	Evaluation technique	Outcome measure	Target muscle, test site, and position of subjects	Reliability of measurement technique	Main findings
1	Amabile et al. (2017)Journal: *PLoS One*	Case: 36 female CLBP patients, Mean age: 51.6 (8.7) yearsControl: 32 female patients without a history of LBP, Mean age: 51.3 (8.6) years	Axial CT scan	CSA	Gluteus maximusAt the level of the apex of the coccyx, bilateralParticipants position was not clearly presented	Inter-rater reliability; ICC: 0.94	Mean normalized CSA was significantly smaller in the case group than in the control group ($P < 0.05$)
2	Sions et al. (2017) Journal: *Journal of Orthopaedic and Sports Physical Therapy*	Case: 53 patients with LBP (22 ♂, 31 ♀), Mean age: 69.9 (68.1, 71.7) yearsControl: 49 participants without a history of LBP (15 ♂, 34 ♀), Mean age: 72.2 (70.3, 74.1) years	Axial T1-weightedMRI	CSA, muscle-to-fat indices, and relative muscle CSA, i.e. CSA void of fat	Psoas majorAt the level of L2 through L5 vertebral body, bilateralParticipants position was not clearly presented	Cited a previous study (Sions et al. 2015) Intra-rater reliability; ICC: 0.58–0.99Inter-rater reliability; ICC: 0.75–0.97	Patients with LBP had decreased psoas major muscle-to-fat index ($P = 0.003$) and relative muscle CSA ($P = 0.003$) compared to healthy participants
3	Hyun et al. (2016) Journal: *Clinical Spine Surgery*	Case: 20 patients with degenerative lumbar kyphosis (1 ♂, 19 ♀), Mean age: 61.5 (6.3) yearsControl: 20 healthy volunteers (1 ♂, 19 ♀), Mean age: 60.5 (8.5) years	Axial T2-weightedMRI and axial CT scan	Lumbar muscularity (*CSA of muscle – VB ratio × 100*) degree of fatty change (*mean signal intensity of muscle – subcutaneous fat ratio × 100*)	Psoas major At the inferior endplate of L_1, L_2, L_3, L_4, and L_5, bilateral Participants position was not clearly presented	MRI measurements: Intra-rater reliability; ICC: 0.86–0.91, Inter-rater reliability; ICC: 0.79–0.89CT scan measurements: Intra-rater reliability; ICC: 0.78, Inter-rater reliability; ICC: 0.65	Muscularity and fatty degeneration in the psoas major were not significantly different between the case and control groups ($P > 0.05$)
4	Abbas et al. (2016) Journal: *BMC Musculoskeletal Disorders*	Case: 165 patients with degenerative lumbar spinal stenosis (80 ♂, 85 ♀), Age range: 40–88 yearsControl: 180 healthy volunteers (90 ♂, 90 ♀), Age range: 40–99 years	Axial CT scan	Density, CSA	Psoas major At the middle part of L_3 vertebral body, bilateral Participants position was not clearly presented	Intra-rater reliability; ICC: 0.92–0.94Inter-rater reliability; ICC: 0.89–0.92	Mean psoas major density was significantly higher in the case group than in the control group ($P \leq 0.01$). Furthermore, mean CSA was only significantly greater in male patients compared with healthy males ($P = 0.04$)

(continued)

Table 5 (continued)

No.	Study characteristics	Participant characteristics	Evaluation technique	Outcome measure	Target muscle, test site, and position of subjects	Reliability of measurement technique	Main findings
5	Skorupska et al. (2016)Journal: *PLoS One*	Case: 71 patients with LBLP (27 ♂, 44 ♀), Mean age: 47.7 (8.4) yearsControl: 29 healthy volunteers (10 ♂, 19 ♀), Mean age: 47.6 (9.9) years	Axial T2-weightedMRI	Volume (muscles images without adipose tissue infiltration were manually segmented in 3-D medical images)	Piriformis, gluteus minimus, gluteus medius, and gluteus maximus The area from the lumbar spine down to pelvic and upper thigh muscles, bilateral Participants position was supine	Inter-rater reliability; ICC: 0.85–0.90	More than 50% of LBLP patients showed a smaller volume for all pelvic muscles for the symptomatic side, both left and right, except for the gluteus medius (gluteus maximus ($P < 0.001$), gluteus minimus ($P < 0.01$) and piriformis ($P < 0.05$). Furthermore, significant differences were found between symptomatic and non-symptomatic sides of the LBLP group for the gluteus maximus ($P < 0.001$), gluteus minimus ($P < 0.001$) and piriformis ($P < 0.005$)
6	Singh et al. (2016) Journal: *European Spine Journal*	Case: 50 patients with CLBP (26 ♂, 24 ♀), Mean age: 33.54 (8.33) yearsControl: 15 healthy volunteers (9 ♂, 6 ♀), Mean age: 28.93 (7.24) years	Axial T2-weightedMRI	CSA	Psoas major At L_3–L_4, L_4–L_5, and L_5–S_1 intervertebral disc levels, bilateral Participants position was supine with a pillow positioned underneath the knees	Not reported	Mean psoas major CSA was significantly smaller in the case group than in the control group at the L_3–L_4 intervertebral disc level ($P \leq 0.048$). However, no significant difference was found in psoas major CSA between the two groups at L_4–L_5 and L_5–S_1 intervertebral disc levels ($P \geq 0.07$)

(continued)

Table 5 (continued)

No.	Study characteristics	Participant characteristics	Evaluation technique	Outcome measure	Target muscle, test site, and position of subjects	Reliability of measurement technique	Main findings
7	Gildea et al. (2013) Journal: *Journal of Orthopaedic and Sports Physical Therapy*	Case 1: 13 ballet dancer patients with LBP only (4 ♂, 9 ♀), Mean age: 24 (3) yearsCase 2: 10 ballet dancer patients with LBP and hip pain (5 ♂, 5 ♀), Mean age: 25 (5) yearsControl: 8 healthy ballet dancers (5 ♂, 3 ♀), Mean age: 22 (3) years	Axial T2-weightedMRI	CSA	Psoas major At each lumbar body level, bilateral Participants position was supine with their hips and knees resting in slight flexion on a wedge	Cited two previous studies (Hides et al. 1995, 2007)	The size of the psoas major muscle was the same in ballet dancers with and without LBP and in those with hip-region pain and LBP ($P = 0.55$). The psoas major muscle was larger in male dancers than in female dancers ($P < 0.0001$). Finally, the psoas major muscle size increased with the number of years of professional dancing
8	D'Hooge et al. (2013)Journal: *The Clinical Journal of Pain*	Case: 13 patients with unilateral, recurrent, mechanical, nonspecific LBP (6 ♂, 7 ♀), Mean age: 32.09 (11.52) yearsControl: 13 healthy volunteers (6 ♂, 7 ♀), Mean age: 32.13 (10.57) years	Axial T2-weighted muscle functional MRI	CSA	Psoas major At the upper and the lower endplates of L_4, bilateral Participants position was supine with a foam wedge supporting the legs	Intra-rater reliability; ICC: 0.92 (0.70–0.98)	No significant difference was observed between the two groups after rest for 30 min and after prone trunk extension ($P \geq 0.09$). In addition, no difference was observed between painful and non-painful sides for both groups ($P \geq 0.24$). Only the case group showed significant mean CSA difference between rest and exercise at the lower endplates of L_4 ($P = 0.005$)

(continued)

Table 5 (continued)

No.	Study characteristics	Participant characteristics	Evaluation technique	Outcome measure	Target muscle, test site, and position of subjects	Reliability of measurement technique	Main findings
9	Arbanas et al. (2013)Journal: *European Spine Journal*	Case: 42 patients with CLBP (18 ♂, 24 ♀), Mean age: 51.1 (14.5) yearsControl : 49 healthy volunteers (24 ♂, 25 ♀), Mean age: 55.1 (17.1) years	Axial T2-weightedMRI	CSA and fat infiltration	Psoas major At L_3–L_4, L_4–L_5, and L_5–S_1 intervertebral disc levels, bilateral Participants position was not clearly presented	Not clearly presented	CLBP patients had bigger psoas major CSA than controls at L_3–L_4 and L_4–L_5 levels ($P < 0.05$). Patients with degenerative changes in the lumbar spine had smaller psoas major CSA compared to the patients without apparent changes at L_3–L_4 and L_4–L_5 levels ($P < 0.05$). At L_4–L_5 and L_5–S_1 levels, patients with Modic changes in the lumbar vertebral bodies demonstrated smaller CSA of the psoas major compared to the patients without Modic changes ($P < 0.05$). However, psoas major CSA in the patients with degenerative changes in lumbar spine and Modic changes was still bigger than the one of the controls. Analysis of fat infiltration in psoas major showed that there was no significant difference between the two groups ($P = 0.80$)
10	D'Hooge et al. (2012)Journal: *Manual Therapy*	Case: 13 patients with unilateral, recurrent, mechanical, non-specific LBP (6 ♂, 7 ♀), Mean age: 32.09 (11.52) yearsControl: 13 healthy volunteers (6 ♂, 7 ♀), Mean age: 32.13 (10.57) years	Axial T1-weightedMRI	Total muscle CSA, lean muscle CSA, fat CSA, muscle-fat-index	Psoas major At the upper endplate of L_3 and L_4 and lower endplate of L_4, bilateral Participants position was supine with a foam wedge supporting the legs (~ 30° hip flexion)	Not reported	No significant differences were found for total, lean muscle, and fat CSA between the case and control groups. However, muscle-fat-index was increased bilaterally at L_4 upper and lower endplates ($P < 0.05$)

(continued)

Table 5 (continued)

No.	Study characteristics	Participant characteristics	Evaluation technique	Outcome measure	Target muscle, test site, and position of subjects	Reliability of measurement technique	Main findings
11	Stewart et al. (2010) Journal: *The British Journal of Sports Medicine*	Case: 17 elite male footballers with a history of LBP, Mean age: not clearly presentedControl: 13 elite male cricketers without LBP, Mean age: not clearly presented	MRI	CSA	Psoas major At L_2–L_3, L_3–L_4, L_4–L_5, and L_5–S_1 intervertebral disc levels, bilateral Participants position was supine with the knees resting on a small foam wedge	Cited two previous studies (Hides et al. 2006, 2008)	The psoas major was larger on the side of the dominant kicking leg at all four vertebral levels measured (P = 0.01). Participants who reported LBP had larger psoas major muscles than the control groups (P = 0.04)
12	Hides et al. (2008) Journal: *British Journal of Sports Medicine*	Case: 8 elite male cricketers with a history of LBP, Mean age: not clearly presentedControl: 13 elite male cricketers without LBP, Mean age: not clearly presented	MRI	CSA	Psoas major At the L_3–L_4 intervertebral disc level, bilateral Participants position was supine with the hips and knees resting on a foam wedge	Cited two previous studies (Hides et al. 1995, 2006)	No significant differences were found for psoas major CSA between the case and control groups and also between the sides contralateral and ipsilateral to the hand dominance in both groups (P = 0.95)
13	Kamaz et al. (2007) Journal: *Diagnostic and Interventional Radiology*	Case: 36 female patients with CLBP, Mean age: 43.2 (6.9) yearsControl: 34 healthy females, Mean age: 44.4 (6.9) years	Axial CT scan	CSA	Psoas major and gluteus maximusAt the L_4 upper and lower endplates, bilateralParticipants position was prone	Accuracy coefficient: 0.68–0.99	In the case group psoas major CSA was smaller than in the control group (P = 0.04) at the L_4 endplate. However, no significant difference was found between the two groups in gluteus maximus CSA (P = 0.50)
14	Danneels et al. (2000)Journal: *European Spine Journal*	Case: 32 patients with CLBP (17 ♂, 15 ♀), Mean age: 37.34 (9.78) yearsControl: 23 healthy volunteers (13 ♂, 10 ♀), Mean age: 36.91 (10.26) years	Axial CT scan	CSA, fat deposit	Psoas major At the upper endplate of L_3 and the upper and lower endplate of L_4, bilateral Participants position was prone	Inter-rater reliability; ICC: 0.81–0.92	No significant differences were found for psoas major CSA with and without fat at the three different levels between the case and control groups ($P \geq 0.25$)

(continued)

Table 5 (continued)

No.	Study characteristics	Participant characteristics	Evaluation technique	Outcome measure	Target muscle, test site, and position of subjects	Reliability of measurement technique	Main findings
15	Dangaria & Naesh (1998)Journal: *Spine*	Case: 25 patients with unilateral sciatica caused by disc herniation, Mean age: not clearly presentedControl: 15 healthy volunteers, Mean age: not clearly presented	MRI	CSA	Psoas major At the L_3–L_4, L_4–L_5, and L_5–S_1 intervertebral disc levels, bilateral Participants position was not clearly presented	Not reported	In the case group, significant reduction in CSA of the psoas major was observed at the level and the site of the disc herniation ($P < 0.05$). Insignificant difference in the CSA of the psoas major was observed in the control group ($P > 0.05$)
16	Parkkola et al. (1993)Journal: *Spine*	Case 1: 38 patients with moderate CLBP (20 ♂, 18 ♀), Age range: 30–47Case 2: 10 patients with severe CLBP (4 ♂, 6 ♀), Age range: 30-47Control: 60 healthy volunteers (33 ♂, 27 ♀), Age range: 30–47	Axial T2-weightedMRI	CSA	Psoas major At the L_4–L_5 level, bilateral Participants position was not clearly presented	Not reported	Mean psoas major CSA of the patients was smaller than those of the volunteers ($P < 0.05$)

Studies are listed in time-ordered manner from most recent to oldest by year of publication.
CLBP, chronic low back pain; CSA, cross-sectional area; CT, computed tomography; ICC, intra-class correlation coefficient; LBLP, low back with leg pain; LBP, low back pain; MRI, magnetic resonance imaging; NR, not reported; VB, vertebral body; y, year.

statistically significant differences between psoas major muscle FCSA and fat CSA in patients with LBP and asymptomatic participants. Hides et al. (2008) also found no significant difference for the psoas major muscle CSA between the sides contralateral and ipsilateral to the hand dominance in elite cricketers with and without LBP. Moreover, Hyun et al. (2016) mentioned that lumbar muscularity index was not significantly different (Table 5) in the psoas major muscle between patients with LBP and asymptomatic participants. Gildea et al. (2013) reported that the size of the psoas major muscle was the same in ballet dancers with and without LBP and in those with LBP and hip-region pain. They also indicated that the size of the psoas major muscle was related to the number of years of professional dancing (Gildea et al. 2013; Table 5).

In contrast, several studies reported significant reductions in the psoas major muscle CSA of patients with LBP compared with asymptomatic participants (Parkkola et al. 1993; Dangaria & Naesh, 1998; Kamaz et al. 2007; Singh et al. 2016; Sions et al. 2017). Singh et al. (2016) indicated that the mean psoas major muscle CSA was significantly smaller in the LBP group than in the control group at the L_3–L_4 intervertebral disc level; however, no significant difference was found in the psoas major muscle CSA between two groups at the L_4–L_5 and L_5–S_1 intervertebral disc levels (Singh et al. 2016; Table 5). Furthermore, Sions et al. (2017) stated that the psoas major muscle-to-fat index decreased in older patients with LBP compared with LBP-free peers.

Finally, a few studies indicated that patients with LBP had a greater psoas major muscle CSA than controls (Stewart et al. 2010; Arbanas et al. 2013; Abbas et al. 2016). Stewart et al. (2010) showed that elite male footballers with LBP had a greater psoas major muscle CSA on the side of the dominant kicking leg at the L_2–L_3 to L_5–S_1 intervertebral disc levels. Furthermore, one study reported an increased muscle-fat-index of the psoas major muscle in patients with LBP (D'Hooge et al. 2012). Abbas et al. (2016) assessed the psoas major muscle density in patients with degenerative lumbar spinal stenosis and noted that the mean psoas major muscle density was significantly higher in patients compared to controls.

The CSA of the gluteus maximus muscle was investigated in two studies (Kamaz et al. 2007; Amabile et al. 2017), but the results were not consistent. Amabile et al. (2017) reported that the mean normalized gluteus maximus muscle CSA was significantly smaller in female patients with CLBP than in asymptomatic female participants, whereas Kamaz et al. (2007) found no significant difference. One study assessed the volume of the gluteus maximus muscle and found a significantly smaller gluteus maximus muscle volume in the experimental group than in healthy controls (Skorupska et al. 2016).

The volume of the gluteus medius, gluteus minimus, and piriformis muscles were evaluated in a study by Skorupska et al. (2016). Gluteus minimus and piriformis muscles atrophy were confirmed in patients with low back and leg pain, but no significant difference was observed for the gluteus medius muscle between patients with low back and leg pain and asymptomatic participants (Skorupska et al. 2016). Skorupska et al. (2016) also showed that the volume of the gluteus maximus, gluteus minimus, and piriformis muscles was significantly smaller in the symptomatic side vs. the non-symptomatic side of the low back and leg pain group.

During our search across the selected databases, no comparative observational study was detected in which morphological changes in other hip muscles were assessed in patients with LBP besides the psoas major, gluteal, and piriformis muscles. Table 5 summarizes the characteristics and main findings of the included studies.

Discussion

Our systematic review is the first qualitative study that assessed the macroscopic morphological changes in hip muscles in patients with LBP. The level of evidence of each included study was determined as level B. The results of this study revealed that there is substantial controversy about the morphological changes in the psoas major and gluteus maximus muscles in patients with LBP. A formal meta-analysis was not performed due to very significant differences in the included studies in terms of study populations and methodology.

It is generally accepted that the density and CSA of muscles reflect the muscles' physical function and performance of people with LBP (Keller et al. 1999; Käser et al. 2001). Muscle CSA is related to the force in various muscles, and therefore provides an indication of the muscle's force generation capacity (Maughan et al. 1983). Conditions of muscle such as density, CSA size, and fatty infiltration can be attained via medical imaging modalities providing non-invasive, direct, reliable, and quantitative information (Abbas et al. 2016). CT scan and MRI have been employed for measuring CSA and the degeneration rate of muscles in patients with muscular dystrophy (Abbas et al. 2016). The findings obtained from direct and objective examination of muscles will contribute to the explanation of the pathogenesis of LBP, as well as its diagnosis and treatment (Akima et al. 2000, 2001; Danneels et al. 2000; Kader et al. 2000).

Fourteen of 16 included studies (87.5%) assessed the morphological changes in the psoas major muscle in patients with LBP. The psoas major muscle is a powerful flexor of the hip joint and, because of its line of action, it is also a weak medial rotator and adductor of the femur (Ward, 1999). However, some anatomists claim that, in the anatomical position the psoas major muscle flexes the hip joint with no rotational component (Skyrme et al. 1999). With the hip joint in the abducted position, the psoas major muscle produces flexion, adduction, and lateral rotation of the femur at the hip joint (Skyrme et al. 1999). This muscle is the

largest muscle in cross section at the lower levels of the lumbar spine (McGill et al. 1988). Nachemson (1966, 1968) indicated that the psoas major muscle is active during upright standing, lifting, and forward bending. These findings prompted the inference that the psoas major muscle may function as a lumbar spine stabilizer (Sajko & Stuber, 2009). Other studies proposed various roles for the psoas major muscle with respect to lumbar spine stability and movement, including being a flexor and lateral flexor of the lumbar spine on the pelvis, a stabilizer of the lumbar spine and the hip joint, power source for bipedal walking and running, and controller of the lumbar lordosis while supporting difficult lumbar loads (Andersson et al. 1997; Sajko & Stuber, 2009).

Patients with LBP have been found to have a smaller psoas major muscle CSA and more fatty infiltrations localized to the suspected pathological spinal level and symptomatic side, although these findings were not consistent in all studies. This discrepancy may be explained by the differences in the chronicity of LBP and study population. Dangaria & Naesh (1998) demonstrated that there is a significant relationship between a reduction in the CSA of the psoas major muscle and the duration of continuous sciatica of the affected leg (Spearman's $\rho = 0.8$; $P = 0.05$). In addition, 46% of the included studies did not control the physical activity level as a potential confounding variable (Goldman et al. 2016) between the patients with LBP and asymptomatic participants. Therefore, the results of some included studies (Dangaria & Naesh, 1998; Arbanas et al. 2013; Abbas et al. 2016; Hyun et al. 2016; Singh et al. 2016) could be confounded by this potentially important factor. Stewart et al. (2010) and Hides et al. (2008) did not minimize the effect of age as another potential confounding variable in their studies, and the results were not presented after adjustment for age. Goldman et al. (2016) stipulated that age may confound the estimates of muscle CSA.

Pourahmadi et al. (2018a) showed that patients with chronic non-specific LBP had limited lumbar spine and hip joints sagittal plane angles, and smaller angular velocity compared with asymptomatic individuals during a functional task. Patients with LBP usually limit their movements at the lumbar spine and adjacent joints as a protective strategy to avoid pain progression in their affected area (Pourahmadi et al. 2018a). Decreased range of motion of the lumbar spine may be due to pain and resulting muscle inhibition. Muscle inhibition might cause disuse muscle atrophy and a decrease in muscle performance capabilities (Ross et al. 2002). When muscles are degenerated because of immobilization or decreased movement, they undergo a variety of histologic changes (Boonyarom & Inui, 2006). Decreased muscle size and an increased infiltration of muscle by fat and connective tissue are two characteristic features of muscle atrophy (Boonyarom & Inui, 2006). Parkkola et al. (1993) reported that the amount of fat infiltration in the muscles is related to the degree of muscle atrophy. Kamaz et al. (2007) also found atrophic psoas major muscles, but the atrophy was not related to the side of symptoms. They mentioned that the side of clinical symptoms may change over time in patients with chronic LBP. Both sides are often affected in the majority of patients (Kamaz et al. 2007). There is evidence suggesting that fat infiltration, fibrosis, slow-to-fast muscle fiber transition, and muscle fiber atrophy are prominent features of sub-acute and chronic LBP (Hodges & Danneels, 2019).

Stewart et al. (2010) demonstrated that the psoas major muscle was larger on the side of the dominant kicking leg at the lumbar levels in elite footballers. They argued that the difference in psoas major muscle CSA observed between the sides is likely related to the increased muscular demands and relatively large forces generated by the psoas major muscle during repetitive kicking (Stewart et al. 2010). Moreover, some studies showed an increased CSA of the psoas major muscle in patients with LBP than in asymptomatic individuals (Arbanas et al. 2013; Abbas et al. 2016). Considering the stabilizing function of the psoas major muscle together with the back extensor muscles, hypertrophy of the psoas major muscle might be due to its increased activity, with increased levels of instability associated with degenerative disorders of the lumbar spine (Zhao et al. 2005; Arbanas et al. 2013). Hides et al. (2007) assessed the effects of prolonged bed rest on the psoas major muscle in healthy male participants and reported an increase in the CSA of the psoas major muscle. They attributed this change to the increase in the muscle tone and the possibility of maintaining a flexed trunk position during bed rest by the participants, resulting in shortening the psoas major muscle and increasing its CSA (Hides et al. 2007).

In this review, three of 16 included studies ($\sim$ 19%) investigated the morphological changes in the gluteus maximus muscle in patients with LBP. The gluteus maximus muscle is the largest, thickest, and most powerful muscle in the gluteal region of the body (Taylor et al. 2015). It is the most superficial of the three gluteal muscles and is considered to be important for both functional and sport activities such as jogging, running, and lifting (Contreras et al. 2015). The gluteus maximus muscle is quadrilateral, with its fasciculi directed downward and outward obliquely at a 45° angle from the pelvis to the buttocks (Taylor et al. 2015). This muscle is aligned and leveraged to extend, laterally rotate, and assist in abduction of the hip joint (Hollman et al. 2013). The gluteus maximus muscle is also functionally coupled with the back extensor muscles to perform lifting from full flexion (Clark et al. 2003). While arising from full trunk flexion into extension, most movement occurs at the hip joint and is accomplished by the gluteus maximus and hamstring muscles during the first 50% of the movement cycle (Amabile et al. 2017).

Two studies showed morphological changes in gluteus maximus muscle in patients with LBP (Skorupska et al. 2016; Amabile et al. 2017). A comparative study revealed that the

gluteus maximus muscle showed fatigue faster in women with chronic LBP than in a group of healthy controls during a sustained back extension endurance test (Kankaanpää et al. 1998). Nadler et al. (2000) assessed the side-to-side symmetry of the gluteus maximus muscle strength in collegiate athletes and reported a significant difference in side-to-side symmetry of the gluteus maximus muscle function in female participants who had LBP or lower extremity pain. Amabile et al. (2017) showed a correlation between the decreased CSA of the gluteus maximus muscle with the development of LBP and suggested investigating the role of the gluteus maximus muscle in LBP and the nature of its atrophy in future studies. Skorupska et al. (2016) offered two possible explanations for the presence of atrophy of the symptomatic hip muscles observed in patients with low back and leg pain. First, a neurogenic type atrophy caused by nerve compression may induce metabolic changes in the sympathetic nervous system, promoting increased metabolic activity of the musculoskeletal system and vasoconstriction, and eventually resulting in muscle atrophy (MacIntyre et al. 1995; Skorupska et al. 2016). Secondly, atrophy of the pelvis muscles can be observed because of a patient's unwillingness to use the symptomatic leg or because of improper functioning of the trunk and pelvis stabilizers, which is one of the main contributors to the development of chronic LBP (Hodges & Richardson, 1996; Nelson-Wong et al. 2008). However, Kamaz et al. (2007) did not find any significant difference in the gluteus maximus muscle CSA between patients with chronic LBP and asymptomatic controls. The insignificant results were attributed to individual differences (Kamaz et al. 2007).

In the current systematic review, the majority of the included studies (12 studies) did not include any power analysis to calculate the number of participants needed to prevent type II statistical errors (false-negative results). Therefore, the generalizability of the results of these studies is limited due to low external validity. In two comparative observational studies (Hides et al. 2008; Stewart et al. 2010) the patients with LBP were not subclassified Because LBP is a heterogeneous disorder, it is necessary to provide a homogeneous sample by recruiting specific and well-defined groups of patients with LBP in order to assess better and more precisely the muscles' morphological changes. Additionally, most of the included studies had some methodological weaknesses, such as a lack of a proper sample size calculation, or no case-control matching for common potential confounding variables. Future research studies should consider the limitations of the previous studies in order to improve the quality of the results in this field.

Like other studies, there were some limitations in the current systematic review. First, only studies published in peer-review journals were included, and therefore a publication bias may have occurred. Secondly, there is a possibility of language bias, as only those full-text studies published in English were included in this review. Finally, a meta-analysis was not performed due to the high methodological heterogeneity of the selected studies.

Conclusion

In this study, we attempted to provide a comprehensive qualitative synthesis of previously published literature regarding the macroscopic morphological changes in hip muscles in subjects with LBP. Among 16 comparative observational studies included in this review, 11 were considered to be of high quality. The current systematic review indicated that morphological changes in hip muscles could happen in patients with LBP, but the results were not consistent across the reported studies. The majority of high-quality studies demonstrated, however, that the CSA and volume of hip muscles, such as the gluteus maximus, gluteus minimus, piriformis, and psoas major muscles, generally decrease in patients with LBP. A meta-analysis was not conducted as part of this systematic review because there were very significant differences in the selected studies in terms of study populations and methodology. Further high-quality research could assess the effects of general and specific physical therapy exercises in normalizing the morphological changes in hip muscles in patients with LBP. Finally, it is suggested that physical therapists plan their treatment strategies accordingly to improve and normalize hip muscle function in patients with LBP.

Conflict of interest

The authors declare that they have no conflict of interest.

References

Abbas J, Slon V, May H, et al. (2016) Paraspinal muscles density: a marker for degenerative lumbar spinal stenosis? *BMC Musculoskelet Disord* **17**, 422.

Akgul O, Gulkesen A, Akgol G, et al. (2013) MR-defined fat infiltration of the lumbar paravertebral muscles differs between non-radiographic axial spondyloarthritis and established ankylosing spondylitis. *Mod Rheumatol* **23**, 811–816.

Akima H, Kuno S-y, Takahashi H, et al. (2000) The use of magnetic resonance images to investigate the influence of recruitment on the relationship between torque and cross-sectional area in human muscle. *Eur J Appl Physiol* **83**, 475–480.

Akima H, Kubo K, Imai M, et al. (2001) Inactivity and muscle: effect of resistance training during bed rest on muscle size in the lower limb. *Acta Physiol Scand* **172**, 269–278.

Amabile AH, Bolte JH, Richter SD (2017) Atrophy of gluteus maximus among women with a history of chronic low back pain. *PLoS One* **12**, e0177008.

Andersson EA, Nilsson J, Thorstensson A (1997) Intramuscular EMG from the hip flexor muscles during human locomotion. *Acta Physiol Scand* **161**, 361–370.

Arbanas J, Starčević-Klasan G, Malnar D (2012) Composition of the psoas major muscle regarding its complex function. *Med Flum* **48**, 123–130.

Arbanas J, Pavlovic I, Marijancic V, et al. (2013) MRI features of the psoas major muscle in patients with low back pain. *Eur Spine J* **22**, 1965–1971.

Balagué F, Mannion AF, Pellisé F, et al. (2012) Non-specific low back pain. *Lancet* **379**, 482–491.

Barker KL, Shamley DR, Jackson D (2004) Changes in the cross-sectional area of multifidus and psoas in patients with unilateral back pain: the relationship to pain and disability. *Spine* **29**, E515–E519.

Bhadresha A, Lawrence OJ, McCarthy MJ (2016) A comparison of magnetic resonance imaging muscle Fat content in the lumbar paraspinal muscles with patient-reported outcome measures in patients with lumbar degenerative disk disease and focal disk prolapse. *Global Spine J* **6**, 401–410.

Bland JM, Altman DG (1999) Measuring agreement in method comparison studies. *Stat Methods Med Res* **8**, 135–160.

Boonyarom O, Inui K (2006) Atrophy and hypertrophy of skeletal muscles: structural and functional aspects. *Acta Physiol* **188**, 77–89.

Bouche KG, Vanovermeire O, Stevens VK, et al. (2011) Computed tomographic analysis of the quality of trunk muscles in asymptomatic and symptomatic lumbar discectomy patients. *BMC Musculoskelet Disord* **12**, 65.

Clark BC, Manini TM, Ploutz-Snyder LL (2003) Derecruitment of the lumbar musculature with fatiguing trunk extension exercise. *Spine* **28**, 282–287.

Contreras B, Vigotsky AD, Schoenfeld BJ, et al. (2015) A comparison of gluteus maximus, biceps femoris, and vastus lateralis electromyographic activity in the back squat and barbell hip thrust exercises. *J Appl Biomech* **31**, 452–458.

Cooper NA, Scavo KM, Strickland KJ, et al. (2016) Prevalence of gluteus medius weakness in people with chronic low back pain compared to healthy controls. *Eur Spine J* **25**, 1258–1265.

Dangaria TR, Naesh O (1998) Changes in cross-sectional area of psoas major muscle in unilateral sciatica caused by disc herniation. *Spine* **23**, 928–931.

Danneels L, Vanderstraeten G, Cambier DC, et al. (2000) CT imaging of trunk muscles in chronic low back pain patients and healthy control subjects. *Eur Spine J* **9**, 266–272.

De Meulemeester K, Calders P, De Pauw R, et al. (2017) Morphological and physiological differences in the upper trapezius muscle in patients with work-related trapezius myalgia compared to healthy controls: A systematic review. *Musculoskelet Sci Pract* **29**, 43–51.

D'Hooge R, Cagnie B, Crombez G, et al. (2012) Increased intramuscular fatty infiltration without differences in lumbar muscle cross-sectional area during remission of unilateral recurrent low back pain. *Manual Ther*, **17**, 584–588.

D'Hooge R, Cagnie B, Crombez G, et al. (2013) Lumbar muscle dysfunction during remission of unilateral recurrent nonspecific low-back pain: evaluation with muscle functional MRI. *Clin J Pain*, **29**, 187–194.

Falla D, Farina D (2008) Neuromuscular adaptation in experimental and clinical neck pain. *J Electromyogr Kinesiol* **18**, 255–261.

Fortin M, Macedo LG (2013) Multifidus and paraspinal muscle group cross-sectional areas of patients with low back pain and control patients: a systematic review with a focus on blinding. *Phys Ther* **93**, 873–888.

Furlan AD, Irvin E, Bombardier C (2006) Limited search strategies were effective in finding relevant nonrandomized studies. *J Clin Epidemiol* **59**, 1303–1311.

Gildea JE, Hides JA, Hodges PW (2013) Size and symmetry of trunk muscles in ballet dancers with and without low back pain. *J Orthop Sports Phys Ther* **43**, 525–533.

Goldman LH, Tang K, Facchetti L, et al. (2016) Role of thigh muscle cross-sectional area and strength in progression of knee cartilage degeneration over 48 months – data from the Osteoarthritis Initiative. *Osteoarthritis Cartilage* **24**, 2082–2091.

Goubert D, Oosterwijck JV, Meeus M, et al. (2016) Structural changes in lumbar muscles in non-specific low back pain: a systematic review. *Pain Physician* **19**, E985–E1000.

Griffin DW, Harmon D, Kennedy N (2012) Do patients with chronic low back pain have an altered level and/or pattern of physical activity compared to healthy individuals? A systematic review of the literature. *Physiotherapy* **98**, 13–23.

Hartvigsen J, Hancock MJ, Kongsted A, et al. (2018) What low back pain is and why we need to pay attention. *Lancet* **391**, 2356–2367.

Hides JA, Richardson CA, Jull GA (1995) Magnetic resonance imaging and ultrasonography of the lumbar multifidus muscle. Comparison of two different modalities. *Spine* **20**, 54–58.

Hides J, Wilson S, Stanton W, et al. (2006) An MRI investigation into the function of the transversus abdominis muscle during 'drawing-in' of the abdominal wall. *Spine* **31**, E175–E178.

Hides JA, Belavý DL, Stanton W, et al. (2007) Magnetic resonance imaging assessment of trunk muscles during prolonged bed rest. *Spine* **32**, 1687–1692.

Hides JA, Stanton WR, Freke M, et al. (2008) MRI study of the size, symmetry and function of the trunk muscles among elite cricketers with and without low back pain. *Br J Sports Med* **42**, 809–813.

Higgins J, Green S (2011) Cochrane Handbook for Systematic Reviews of Interventions, version 5.1.0 (updated March 2011). The Cochrane Collaboration. Available from handbook.cochrane.org.

Hodges PW, Danneels L (2019) Changes in structure and function of the back muscles in low back pain: different time points, observations, and mechanisms. *J Orthop Sports Phys Ther* **49**, 464–476.

Hodges PW, Richardson CA (1996) Inefficient muscular stabilization of the lumbar spine associated with low back pain: a motor control evaluation of transversus abdominis. *Spine* **21**, 2640–2650.

Hollman JH, Hohl JM, Kraft JL, et al. (2013) Modulation of frontal-plane knee kinematics by hip-extensor strength and gluteus maximus recruitment during a jump-landing task in healthy women. *J Sport Rehabil* **22**, 184–190.

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

Hoy D, March L, Brooks P, et al. (2014) The global burden of low back pain: estimates from the Global Burden of Disease 2010 study. *Ann Rheum Dis* **73**, 968–974.

Hyun S-J, Bae C-W, Lee S-H, et al. (2016) Fatty degeneration of the paraspinal muscle in patients with degenerative lumbar kyphosis: a new evaluation method of quantitative digital analysis using MRI and CT scan. *Clin Spine Surg* **29**, 441–447.

Jeon I-c, Kwon O-y, Weon J-h, et al. (2016) Comparison of psoas major muscle thickness measured by sonography during active straight leg raising in subjects with and without uncontrolled lumbopelvic rotation. *Manual Ther* **21**, 165–169.

Joseph LH, Hussain RI, Naicker AS, et al. (2015) Pattern of changes in local and global muscle thickness among individuals with sacroiliac joint dysfunction. *Hong Kong Physiother J* **33**, 28–33.

Kader D, Wardlaw D, Smith F (2000) Correlation between the MRI changes in the lumbar multifidus muscles and leg pain. *Clin Radiol* **55**, 145–149.

Kalichman L, Hodges P, Li L, et al. (2010) Changes in paraspinal muscles and their association with low back pain and spinal degeneration: CT study. *Eur Spine J* **19**, 1136–1144.

Kamaz M, Kiresi D, Oguz H, et al. (2007) CT measurement of trunk muscle areas in patients with chronic low back pain. *Diagn Interv Radiol* **13**, 144.

Kang C, Shin M, Kim S, et al. (2007) MRI of paraspinal muscles in lumbar degenerative kyphosis patients and control patients with chronic low back pain. *Clin Radiol* **62**, 479–486.

Kankaanpää M, Taimela S, Laaksonen D, et al. (1998) Back and hip extensor fatigability in chronic low back pain patients and controls. *Arch Phys Med Rehabil* **79**, 412–417.

Käser L, Mannion AF, Rhyner A, et al. (2001) Active therapy for chronic low back pain: part 2. Effects on paraspinal muscle cross-sectional area, fiber type size, and distribution. *Spine* **26**, 909–919.

Keller A, Johansen JG, Hellesnes J, et al. (1999) Predictors of isokinetic back muscle strength in patients with low back pain. *Spine* **24**, 275–280.

Kendall FP, McCreary EK, Provance PG, et al. (2005) *Muscles, Testing and Function: With Posture and Pain*. Baltimore: Lippincott Williams and Wilkins.

Kim WH, Lee S-H, Lee DY (2011) Changes in the cross-sectional area of multifidus and psoas in unilateral sciatica caused by lumbar disc herniation. *J Korean Neurosurg Soc* **50**, 201.

Lee SW, Kim SY (2015) Effects of hip exercises for chronic low-back pain patients with lumbar instability. *J Phys Ther Sci* **27**, 345–348.

Lee SH, Park SW, Kim YB, et al. (2017) The fatty degeneration of lumbar paraspinal muscles on computed tomography scan according to age and disc level. *Spine J* **17**, 81–87.

MacIntyre DL, Reid WD, McKenzie DC (1995) Delayed muscle soreness. The inflammatory response to muscle injury and its clinical implications. *Sports Med* **20**, 24–40.

Maher C, Underwood M, Buchbinder R (2017) Non-specific low back pain. *Lancet* **389**, 736–747.

Maughan R, Watson JS, Weir J (1983) Strength and cross-sectional area of human skeletal muscle. *J Physiol* **338**, 37–49.

Mayoux-Benhamou M, Revel M, Vallee C, et al. (1994) Longus colli has a postural function on cervical curvature. *Surg Radiol Anat* **16**, 367–371.

McGill S, Patt N, Norman R (1988) Measurement of the trunk musculature of active males using CT scan radiography: implications for force and moment generating capacity about the L4 L5 joint. *J Biomech* **21**, 329–341.

Meeus M, Gebruers N (2016) *Health Literacy: From Reference to Review*. Den Haag: Acco Leuven.

Mengiardi B, Schmid MR, Boos N, et al. (2006) Fat content of lumbar paraspinal muscles in patients with chronic low back pain and in asymptomatic volunteers: quantification with MR spectroscopy. *Radiology* **240**, 786–792.

Moher D, Liberati A, Tetzlaff J, et al. (2009) Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. *Ann Intern Med* **151**, 264–269.

Nachemson A (1966) Electromyographic studies on the vertebral portion of the psoas muscle: with special reference to its stabilizing function of the lumbar spine. *Acta Orthop Scand* **37**, 177–190.

Nachemson A (1968) The possible importance of the psoas muscle for stabilization of the lumbar spine. *Acta Orthop Scand* **39**, 47–57.

Nadler SF, Malanga GA, DePrince M, et al. (2000) The relationship between lower extremity injury, low back pain, and hip muscle strength in male and female collegiate athletes. *Clin J Sport Med* **10**, 89–97.

Nelson-Wong E, Gregory DE, Winter DA, et al. (2008) Gluteus medius muscle activation patterns as a predictor of low back pain during standing. *Clin Biomech Elsevier Ltd* **23**, 545–553.

Neumann DA (2017) *Kinesiology of the Musculoskeletal System: Foundations for Physical Rehabilitation.)*. St. Louis: Mosby.

Oliveira CB, Maher CG, Pinto RZ, et al. (2018) Clinical practice guidelines for the management of non-specific low back pain in primary care: an updated overview. *Eur Spine J* **27**, 2791–2803.

Ostelo R (2018) Lancet series: the 'Magnum Opus' regarding the evidence on low back pain. *Spine* **43**, 1239–1240.

Parkkola R, Rytökoski U, Kormano M (1993) Magnetic resonance imaging of the discs and trunk muscles in patients with chronic low back pain and healthy control subjects. *Spine* **18**, 830–836.

Pourahmadi MR, Taghipour M, Ebrahimi Takamjani I, et al. (2016) Motor control exercise for symptomatic lumbar disc herniation: protocol for a systematic review and meta-analysis. *BMJ Open* **6**, e012426.

Pourahmadi MR, Ebrahimi Takamjani I, Jaberzadeh S, et al. (2018a) Test-retest reliability of sit-to-stand and stand-to-sit analysis in people with and without chronic non-specific low back pain. *Musculoskelet Sci Pract* **35**, 95–104.

Pourahmadi MR, Mohsenifar H, Dariush M, et al. (2018b) Effectiveness of mobilization with movement (Mulligan concept techniques) on low back pain: a systematic review. *Clin Rehabil* **32**, 1289–1298.

Pourahmadi MR, Ebrahimi Takamjani I, Jaberzadeh S, et al. (2019) Kinematics of the spine during sit-to-stand movement using motion analysis systems: a systematic review of literature. *J Sport Rehabil* **28**, 77–93.

Ranson CA, Burnett AF, Kerslake R, et al. (2006) An investigation into the use of MR imaging to determine the functional cross sectional area of lumbar paraspinal muscles. *Eur Spine J* **15**, 764–773.

Rantanen J, Hurme M, Falck B, et al. (1993) The lumbar multifidus muscle five years after surgery for a lumbar intervertebral disc herniation. *Spine* **18**, 568–574.

Ross MD, Hooten S, Moore D (2002) Lower leg girth and ankle plantar-flexor endurance after anterior cruciate ligament reconstruction. *J Sport Rehabil* **11**, 128–138.

Sajko S, Stuber K (2009) Psoas Major: a case report and review of its anatomy, biomechanics, and clinical implications. *J Can Chiropr Assoc* **53**, 311.

Salah El-din Mahmoud W, Yousef A, Manssor E, et al. (2016) The relationship between pain and functional disability with morphological changes in psoas major in discogenic low back pain patients. *Int J Ther Rehabil* **23**, 363–370.

Sanchis-Moysi J, Idoate F, Izquierdo M, et al. (2011) Iliopsoas and gluteal muscles are asymmetric in tennis players but not in soccer players. *PLoS One* **6**, e22858.

Saragiotto BT, Maher CG, Yamato TP, et al. (2016) Motor control exercise for chronic non-specific low-back pain. *Cochrane Library* (1), CD012004.

Sasaki T, Yoshimura N, Hashizume H, et al. (2017) MRI-defined paraspinal muscle morphology in Japanese population: the Wakayama Spine Study. *PLoS One* **12**, e0187765.

Singh R, Yadav SK, Sood S, et al. (2016) Magnetic resonance imaging of lumbar trunk parameters in chronic low backache patients and healthy population: a comparative study. *Eur Spine J* **25**, 2864–2872.

Sions JM, Smith AC, Hicks GE, et al. (2015) Trunk muscle size and composition assessment in older adults with chronic low back pain: an intra-examiner and inter-examiner reliability study. *Pain Med* **17**, 1436–1446.

Sions JM, Elliott JM, Pohlig RT, et al. (2017) Trunk muscle characteristics of the multifidi, erector spinae, psoas, and quadratus lumborum in older adults with and without chronic low back pain. *J Orthop Sports Phys Ther* **47**, 173–179.

Skorupska E, Keczmer P, Łochowski RM, et al. (2016) Reliability of MR-based volumetric 3-D analysis of pelvic muscles among subjects with low back with leg pain and healthy volunteers. *PLoS One* **11**, e0159587.

Skyrme A, Cahill D, Marsh H, et al. (1999) Psoas major and its controversial rotational action. *Clin Anat* **12**, 264–265.

Stang A (2010) Critical evaluation of the Newcastle-Ottawa scale for the assessment of the quality of nonrandomized studies in meta-analyses. *Eur J Epidemiol* **25**, 603–605.

Stewart S, Stanton W, Wilson S, et al. (2010) Consistency in size and asymmetry of the psoas major muscle among elite footballers. *Br J Sports Med* **44**, 1173–1177.

Suter E, McMorland G (2002) Decrease in elbow flexor inhibition after cervical spine manipulation in patients with chronic neck pain. *Clin Biomech Elsevier Ltd* **17**, 541–544.

Taylor V, Guttmann GD, Reeves RE (2015) A variant accessory muscle of the gluteus maximus. *Int J Anat Var* 8.

Verla T, Adogwa O, Elsamadicy A, et al. (2016) Effects of psoas muscle thickness on outcomes of lumbar fusion surgery. *World Neurosurg* **87**, 283–289.

Ward AB (1999) Botulinum toxin type A treatment of hip and thigh spasticity: a technique for injection of psoas major muscle. *Eur J Neurol* **6**, s91–s93.

Yanik B, Keyik B, Conkbayir I (2013) Fatty degeneration of multifidus muscle in patients with chronic low back pain and in asymptomatic volunteers: quantification with chemical shift magnetic resonance imaging. *Skeletal Radiol* **42**, 771–778.

Yong W, Sanguankeo A, Upala S (2018) Association between primary Sjögren's syndrome, cardiovascular and cerebrovascular disease: a systematic review and meta-analysis. *Clin Exp Rheumatol* **36**, S190–S197.

Zhao F, Pollintine P, Hole BD, et al. (2005) Discogenic origins of spinal instability. *Spine* **30**, 2621–2630.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Search strategy for PubMed/Medline (NLM).