The Spine Journal 22 (2022) 1205–1221

Systematic Review/Meta-Analysis

Microscopic changes in the spinal extensor musculature in people with chronic spinal pain: a systematic review

Shilpa Purushotham, MBBS, MSc, FHEA, (PhD)[a,b], Rob S. Stephenson, BSc, PhD[b], Andy Sanderson, BSc, PhD[c], Deepa Abichandani, BSc, MPT, MSc[d], Carolyn Greig, BSc, MSc, PhD, FTPS[e,f,g], Adrian Gardner, BM, PhD, MRCS, FRCS (T&O)[h], Deborah Falla, BPhty (Hons), PhD[a,*]

[a] *Centre of Precision Rehabilitation for Spinal Pain (CPR Spine), School of Sports, Exercise and Rehabilitation Sciences, University of Birmingham, 142 Edgbaston Park Rd, Birmingham, B15 2TT, UK*

[b] *Anatomy Department, Birmingham Medical School, College of Medical and Dental Sciences, University of Birmingham, Vincent Drive, Edgbaston, Birmingham, B15 2TT, UK*

[c] *Department of Sport and Exercise Sciences, Musculoskeletal Science and Sports Medicine Research Centre, Manchester Metropolitan University, Oxford Road, Manchester, M15 6BH, UK*

[d] *Division of Physiotherapy, Institute of Health and Social Care, London South Bank University, Borough Road, London, SE1 0AA, UK*

[e] *School of Sport, Exercise and Rehabilitation Sciences, University of Birmingham, 142 Edgbaston Park Rd, Birmingham, B15 2TT, UK*

[f] *MRC-Versus Arthritis Centre for Musculoskeletal Ageing and Health, University of Birmingham, Edgbaston, Birmingham, B15 2TT, UK*

[g] *National Institute for Health Research, Birmingham Biomedical Research Centre at University Hospitals Birmingham NHS Foundation Trust, Birmingham, B15 2TH, UK*

[h] *Spine Unit, The Royal Orthopaedic Hospital NHS Foundation Trust, Bristol Road South, Northfield, Birmingham, UK*

Received 6 August 2021; revised 27 January 2022; accepted 31 January 2022

Abstract

BACKGROUND CONTEXT: Chronic spinal pain is one the most common musculoskeletal disorders. Previous studies have observed microscopic structural changes in the spinal extensor muscles in people with chronic spinal pain. This systematic review synthesizes and analyzes all the existing evidence of muscle microscopic changes in people with chronic spinal pain.

PURPOSE: To assess the microscopy of spinal extensor muscles including the fiber type composition, the area occupied by fiber types, fiber size/cross sectional area (CSA), and narrow diameter (ND) in people with and without chronic spinal pain. Further, to compare these outcome measures across different regions of the spine in people with chronic neck, thoracic and low back pain.

STUDY DESIGN: Systematic review with meta-analysis.

METHODS: MEDLINE (Ovid Interface), Embase, PubMed, CINAHL Plus, and Web of Science were searched from inception to October 2020. Key journals, conference proceedings, grey literature and hand searching of reference lists from eligible studies were also searched. Two independent reviewers were involved in the selection process. Only studies examining the muscle microscopy of the spinal extensor muscles (erector spinae [ES] and/or multifidus [MF]) between people with and without chronic spinal pain were selected. The risk of bias from the studies was assessed using modified Newcastle Ottawa Scale and the level of evidence was established using the GRADE approach. Data were synthesized based on homogeneity on the methodology and

FDA device/drug status: Not applicable.

Author disclosures: ***SP:*** Nothing to disclose. ***RSS:*** Nothing to disclose. ***AS:*** Nothing to disclose. ***DA:*** Nothing to disclose. ***CG:*** Nothing to disclose. ***AG:*** Nothing to disclose. ***DF:*** Nothing to disclose.

*Corresponding author. Deborah Falla, BPhty (Hons), PhD, Centre of Precision Rehabilitation for Spinal Pain (CPR Spine), School of Sports, Exercise and Rehabilitation Sciences, University of Birmingham, Birmingham, B15 2TT, UK

E-mail address: d.falla@bham.ac.uk (D. Falla).

https://doi.org/10.1016/j.spinee.2022.01.023

outcome measures of the studies for ES and MF muscles and only four studies were eligible for analysis.

RESULTS: All the five studies included were related to chronic low back pain (CLBP). Meta-analysis (inverse variance method for random effect to calculate mean difference and 95% CI) was performed for the ES fiber type composition by numbers for both type I and type II fibers (I^2=43% and 0% respectively indicating homogeneity of studies) and showed no difference between the people with and without CLBP with an overall effect estimate Z= 1.49 (p=.14) and Z=1.06 (p=.29) respectively. Meta-analysis was performed for ES fiber CSA for both type I and type II fibers (I^2=0 for both) and showed no difference between people with and without CLBP with an overall effect estimate Z=0.08 (p=.43) and Z=0.75 (p=.45) respectively. Analysis was not performed for ES area occupied by fiber types and ND due to heterogeneity of studies and lack of evidence respectively. Similarly, meta-analysis was not performed for MF fiber type composition by numbers due to heterogeneity of studies. MF analysis for area occupied by fiber type, fiber CSA and ND did not yield sufficient evidence.

CONCLUSIONS: For the ES muscle, there was no difference in fiber type composition and fiber CSA between people with and without CLBP and no conclusions could be drawn for ND for the ES. For the MF, no conclusions could be drawn for any of the muscle microscopy outcome measures. Overall, the quality of evidence is very low and there is very low evidence that there are no differences in microscopic muscle features between people with and without CLBP.

Keywords: Chronic low back pain; Cross-section area; Erector spinae; Fiber type composition; Fiber size; Multifidus; Muscle; Narrow diameter; Spinal extensors

Introduction

As per the 2016 Global Burden of Diseases, low back pain (LBP) is the most common musculoskeletal disorder and is a leading cause of disability [1]. The lifetime prevalence of LBP is up to 84% [2] and the proportion of people who experience persistence of pain for at least 1 year or more after the first episode is estimated to be 25% to 60%, resulting in significant health, economic and societal impacts [3,4]. Approximately 11 to 12% of those affected by chronic low back pain (CLBP) experience severe disability [3,5]. Some of the more specific causes of LBP include trauma, spondylolisthesis, structural deformities, inflammation disorders and malignancy [6,7]. However, in most cases, no specific nociceptive or pathological source can be identified, hence the term nonspecific LBP [5,8]. Neck pain is the second most common spinal-related pain with up to 70% of people experiencing at least one episode in their lifetime [3,9]. Thoracic pain is also common and can be as disabling as LBP and neck pain, although the epidemiology of thoracic pain is not as well documented [10].

Structural changes have been observed in the spinal extensor musculature of people experiencing spinal pain compared to asymptomatic people [11−15], these changes are thought to underlie many of the functional deficits that chronic pain patients present with (eg, less strength and endurance, faster fatigability) [16−21]. Some of the well-noted macroscopic changes (as seen in MRI and CT scans) in both cervical and lumbar muscles of people with spinal pain which have been evaluated in high quality systematic reviews include fatty infiltration, and changes in muscle cross-sectional area [22−26]. Microscopic variations of spinal muscles have also been examined, including changes in muscle fiber type composition, fiber diameter and fiber cross-sectional area. A number of primary studies have been conducted in this area and there are two systematic reviews which both examined microscopic changes of the lumbar muscles [24,27]. However, there are some important limitations of these reviews; Cagnie et al [27] did not evaluate the quality of studies, and the population of interest was not limited to a living population (ie, cadaveric studies were included) which questions the histological reliability of tissues along with unknown medical history. The review by Goubert et al [24] did not include studies clearly defining the LBP population and the searches in both reviews were restricted to only two databases. Moreover, no previous systematic review has been conducted to consider microscopic changes of the spinal muscles in people with either chronic neck or thoracic pain.

The current systematic review was conducted to overcome these limitations in order to gain a thorough understanding on the microscopic changes that occur in the spinal extensor muscles in people with chronic spinal-related pain in comparison with asymptomatic individuals. This knowledge is critical to the development of targeted therapies for management of people with chronic spinal pain. Therefore, the specific objectives were as follows:

1. To assess the muscle fiber type composition, area/proportion of muscle fiber types (including fiber size and diameter) of the spinal extensor muscles of people experiencing chronic spinal pain, compared with asymptomatic individuals.

2. To compare the muscle fiber type composition and area/proportion of muscle fiber types (including fiber size and diameter) of the spinal extensor muscles across different regions of the spine in people with chronic neck, thoracic or low back pain.

Methodology

The protocol for the systematic review was designed using PRISMA-P guidelines (Preferred Reporting Items for Systematic Reviews and Meta-Analysis Protocols) [28] and the Cochrane handbook for systematic reviews [29–31]. The protocol was registered under PROSPERO (CRD42020198087) and published [32]. This systematic review is reported in line with the PRISMA statement [28] as presented in Appendix 1.

A minor amendment to the eligibility criteria was made from the published protocol. Initially, the criteria as described in the protocol was nonspecific chronic spinal pain in a living population. As the screening results did not yield adequate studies with a focus on nonspecific chronic spinal pain, the inclusion criteria were broadened to include chronic spinal pain associated with spinal pathology.

Eligibility criteria

The PICO(S) framework was adopted to develop the eligibility criteria for inclusion and exclusion of the primary studies [33]. PICO(S) was chosen as it has a good balance between sensitivity and specificity [34]. Since no interventional methods were required for this systematic review, Intervention (I) was not included in the eligibility criteria.

Inclusion criteria

- Population (P): Studies including adults ≥18 years of age experiencing chronic spinal pain. Spinal pain was considered chronic if persisting >3 months as per NICE guidelines [35]. Spinal regions considered for this review include cervical, thoracic and lumbar regions.
- Comparator (C): Studies comparing the microscopic changes in the spinal extensor musculature (cervical, thoracic or lumbar region) between people with chronic spinal pain and healthy pain-free controls; or studies on people with chronic spinal pain comparing between different extensor muscles of the spine, or the same muscle in different regions of the spine.
- Outcome (O): Muscle fiber type composition, area occupied by different fiber types, fiber diameter or narrow diameter, and fiber size or cross-sectional area (CSA) of the intrinsic muscles of the spine, including, but not limited to, the erector spinae (ES) and multifidus (MF).
- Study design (S): Observational studies constituted the highest level of evidence for this review, as ascertained by scoping searches. For example, case-control studies and cohort studies.

Exclusion criteria

- Studies including chronic spinal pain attributed to trauma, fractures, surgery, deformity such as scoliosis, inflammatory disorders, infection, or malignancy.
- Studies examining the effect of exercises, or other interventions, on microscopic changes in the spinal muscles.
- Studies not in English.
- Studies where the full text is not available and authors were uncontactable.

Information sources

Electronic databases MEDLINE (Ovid Interface), Embase, PubMed, CINAHL Plus, and Web of Science were searched for relevant studies. Searches were conducted from database inception until October 16, 2020. Moreover, hand searching was conducted for relevant journals including Journal of Anatomy, The Spine Journal, European Spine Journal, and the Clinical Journal of Pain. Anticipation of publication bias was assessed by searching thoroughly for unpublished literature from conference proceedings from 2018 to 2020, including the annual meeting of the Society for Back Pain Research and the World Institute of Pain. Grey literature was searched from Open Grey and British National bibliography databases. Hand searching of reference lists from eligible studies was also completed to ensure that no relevant studies were missed.

Search strategy

The search strategies were developed using free-text or MESH terms applying PICOS criteria for specific databases. The keywords for the search strategy were developed initially for MEDLINE (Ovid) database which included MESH terms for a comprehensive search. This search strategy was adapted in different databases to meet the variations in keywords, MESH headings and syntax where necessary whilst retaining the consistency of search and search terms [36]. The search strategies are presented in Appendix 2.

Study selection

The lead author (SP) conducted the searches from all the information sources. The full results of all the searches including the abstracts and citations of these potentially relevant studies obtained from the comprehensive literature search were imported into EndNote X9 (Clarivate Analytics) data management software.

Following the removal of duplicates [37], two reviewers (S.P./D.A.) independently screened all the potential eligible studies to assess eligibility for inclusion in this review. The selection process followed the best practice guidelines as suggested in the Cochrane Back Review Group [30]. In the first stage, the title and abstract of all nonduplicate studies obtained from search results were screened and clearly ineligible studies were excluded. All potentially eligible

studies, and any studies where eligibility was unclear following title and abstract screening were further screened for eligibility by reading the full text. A screening form was used for full text screening to objectively assess all the eligible studies based on the inclusion and exclusion criteria. Following screening, all studies were classified into three groups: yes (eligible), no (ineligible), maybe (unclear). The studies that were identified as unclear by the independent reviewers, were then discussed for clarification. The two reviewers had agreement of 97.7% on the screening, when the interrater reliability was measured using Cohen's kappa tool (0.57- moderate agreement) [38,39]. Any disagreement between reviewers at each stage or if the reviewers were unable to reach consensus, then the third reviewer (DF) was consulted to reach a decision.

Data extraction

Following the full-text screening process, all relevant information was extracted from each included study by the lead author (SP) using a standardized data extraction form which was developed based on the Cochrane data extraction template and guided by the objectives and inclusion criteria of the review [31,40]. This form was verified by the third reviewer and piloted on a small number of studies to confirm the completeness of data extraction before applying to all the included studies. The data extracted was subsequently checked by a second reviewer (AS) for thoroughness and accuracy. Any discrepancies between the reviewers were discussed and resolved. The third reviewer (DF) was consulted for any further queries and to determine the relevant data for analysis.

Data extraction items are summarized in Appendix 3. Data from some studies were presented only in graphs, and therefore the corresponding authors for these included studies were contacted via email to obtain the raw data for accuracy. A reminder email was sent after 2 weeks, allowing an additional 2 weeks for their response. If no response was received, the graphical information from the publications was converted to numerical values using 'WebPlotDigitizer 4.4' software (https://automeris.io/WebPlotDigitizer/) [41].

Risk of bias assessment

The methodological quality of the included studies was assessed by two reviewers (SP and AS) independently using the Newcastle-Ottawa Scale (NOS) [42]. As there is no consensus on the optimal study quality or risk of bias (RoB) tools for observational studies [43], the NOS was chosen because it is validated, adaptable, and quick to complete [44,45]. Since all of the included studies were case-control studies, the NOS scale was modified to match the design of the evidence (Appendix 4). During the RoB assessment, stars were awarded by each reviewer across three main domains: selection (maximum 4*), comparability (maximum 2*), and outcome (maximum 3*) [42].

Data analysis

The data synthesized in this review were influenced by the methodology and the outcome measures. The homogeneity of studies was assessed based on certain factors such as muscle biopsy site and the muscle tissue sampled, population considered for the patient and control group, age group, pain duration, and the outcome measures.

Meta-analysis was performed when two or more studies measured the same outcome for a given muscle as per the Cochrane guidelines for systematic reviews [46]. Two studies reported their findings in graphical presentation [13,47]. The raw data was unavailable for these two studies despite contacting the authors and hence data were extracted from the graphs. One study presented results separately for males and females and in subgroups within the LBP group or control group, and these results were combined for meta-analysis [13]. Similarly, as there was variation in how each study examined the type II fibers, it was decided to pool all the different subtypes under one umbrella as type II fibers for the meta-analysis on discussion with the third reviewer (DF). Subtypes of type II fibers were combined within each study and presented collectively as type II fibers for all the outcome measures in the meta-analysis.

Meta-analysis was performed using Review Manager 5.4 software (v.5.3 Cochrane Collaboration) [48] for continuous data outcomes (data type), inverse variance statistical method (statistical details used: mean difference, total and subtotals, 95% CI) using random effect analysis measure. Mean difference, CI, I^2, and p value were calculated. The $p<.05$ was considered significant. Statistical heterogeneity was analyzed using the I^2 statistic with 25%, 50%, or 75% indicating low, moderate and high heterogeneity between the studies respectively as indicated in the protocol [32].

Confidence in cumulative evidence

The Grading of Recommendations Assessment, Development, and Evaluation (GRADE) approach was used to assess the quality of evidence obtained from our review as per the Cochrane recommendations for systematic reviews [49,50]. The GRADE tool was applied across each outcome measure for both ES and MF for five determinants: RoB, inconsistency, imprecision, indirectness and publication bias [51–55]. Since all the studies in our review are observational studies, the GRADE score starts with a low rating [49,51]. RoB was assessed for individual studies using the NOS scale; whereas, inconsistency and imprecision was assessed across the studies for heterogeneity, and CI and sample size, respectively [31]. Indirectness was assessed for individual studies by checking how directly the outcome measure was assessed and if they met the PICO criteria of this review [31]. Publication bias was scored moderate

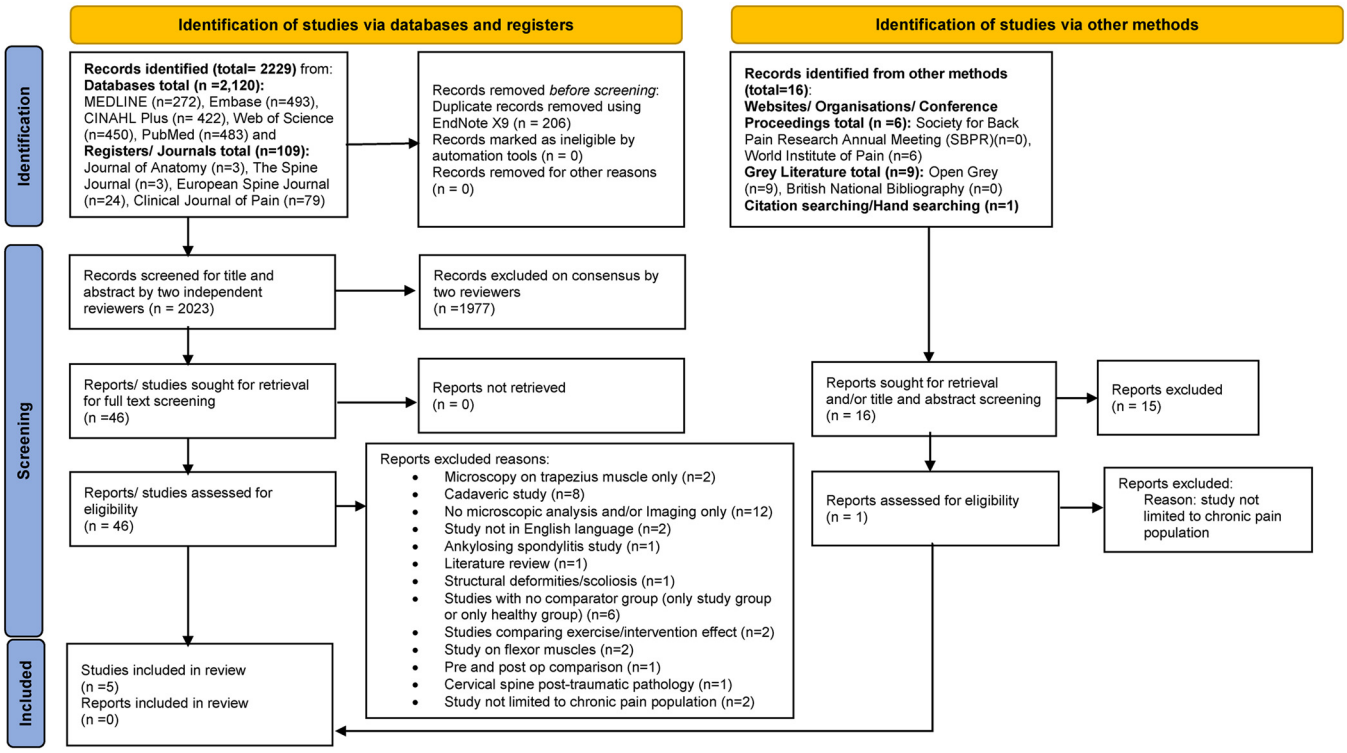

Fig. 1. *PRISMA flow chart of study selection.* PRISMA 2020 flow diagram for new systematic reviews which included searches of databases, registers, and other sources (adapted from ([28]a)).

throughout as the search were limited to English language accounting for inclusion bias [31].

Results

Study selection

Following all searches (as shown in Fig. 1) including databases (n=2,120), key journals (n=109), conference proceedings (n=6), grey literature (n=9), and hand searches (n=1), a total of 2,245 potential studies and reports were exported to EndNote X9. After removing duplicates, the remaining 2,039 studies (2,023 from databases and registers and 16 from other methods) were screened for title and abstract and 1,993 records were excluded (1,977 from databases and registers and 15 from other methods). Full text screening was completed for the 47 remaining studies (46 from databases and registers and one from other methods) and following discussion between the two reviewers, 15 studies were potentially included clearly excluding the remaining (reasons stated in Fig. 1). Ten studies were further excluded for various reasons including: studies on a healthy population only, studies not limited to a chronic pain population i.e. participants with acute pain included, studies with no healthy comparators and interventional studies. Five studies [12,13,47,56,57] matched the inclusion criteria and objectives and were included in the systematic review.

No included studies examined people with chronic neck pain or thoracic pain, and therefore, the results of this systematic review solely relate to CLBP. Additionally, no study compared microscopic changes in muscles in different regions of the spine and therefore it was not possible to assess the secondary objective.

Study characteristics

The individual study characteristics of the included studies are summarized in Table 1.

Population

The study population sample size varied from 20 to 64 whereas the control group sample size varied from 6 to 32 across studies. The overall age of the participants ranged from 18 to 61 years in the CLBP pain group and 17 to 60 years in the control group. One study [47] considered a control group population with acute fracture of lower lumbar spine (at L3, L4, L5) from whom intraoperative biopsy samples were taken within 48 hours of trauma. Further, the lower age limit of one of the participants from control group of this study was 17 years but as this was the only limitation, the study was considered for inclusion on discussion with the third reviewer (DF). Two studies [13,57] considered the outcomes for male and females separately whereas one study [12] included only males in their study and two

Table 1
Summary of characteristics of individual studies included in the systematic review

Study/ Author	Participant no. & characteristics		Methodology			Results				
	Study group	Control group	Muscle (s)	Vertebral level	Biopsy Technique	Fiber type composition by numbers in %	Area occupied by each fiber type in % (RCSA)	Fiber size or CSA (μm^2)	Narrow diameter (μm)	Any otherrelevant
[56]	20 NSCLBP (M:F 10:10)	18 Asymptomatic (M:F 9:9)	ES MF (right side)	L4	Ultrasound guided biopsy	**ES** in NSCLBP- no difference in % type I fibers (p=.0978) but a significant decrease in % type IIx fibers (p=.0019). **MF**- no difference b/n groups	**ES** in NSCLBP −no difference in RCSA of type I fiber (p=.0596) but significant lower RSCA of type IIx fibers (p=.0441). **MF**- no difference b/n the groups	**ES and MF**- no difference b/n groups	-	-
[12]	35 NSCLBP (Males only)	32 Asymptomatic (males only)	ES (left side)	L3 (L3 and T10 described in reference study)	Percutaneous biopsy using conchtome technique	No difference b/n groups for % type I fibers	No difference b/n groups for % area occupied by type I fibers.	No b/n group differences in type I or type II fiber size	No b/n group differences in type I or type II narrow diameter	Type I significantly larger size than type II (p<.01) in those with CLBP. No significant atrophy of any fiber type in those with CLBP compared to controls
[57]	21 LBP for spinal surgery (failed conservative) 79% first surgery 21 pairs (M12, F9 pairs matched with controls)	21 Asymptomatic	MF b/l (superficial) from CLBP group. ES (left side) from control group	L3 or L4 level for CLBP group L3 level for control group	Punch biopsy using Tilley-Henckel forceps (semi-open technique)	Significantly greater % type I in controls (p=.0001). No gender difference for % type I fibers across both the groups (p=.944). Type IIB % significantly higher compared to type I fibers in CLBP only (p=.0001). No difference in % type IIA b/n CLBP and controls groups (p=.134). Type IIC (>1%) significantly higher in CLBP group (p=.049).	Type IIA area significantly higher in men (p=.047) across both groups. No differences in % area of type I (p=0.065) and type IIB (p=.885) b/n the genders. Significantly less % area of type I in CLBP group (p=.0004) and significantly greater % area of type IIB in CLBP group (p=.0001) compared to controls	No difference in mean size of each fiber types b/n CLBP and control groups for a given gender. Men had significantly larger fibers than women (p<.006) in both groups (in relation to body size). Size ratio I:II higher in women in both groups. (p=.018). No difference in I:II size ratio b/n CLBP and control (p=.167)	-	Patients <1 year LBP had double the quantity of type IIC than those with LBP> 1 year. Pathological fibers incidence- no difference b/n CLBP and controls. Based on presence or absence of neurological symptoms- no difference. Based on first and second surgery- no difference.
[47]	26 CLBP (spine instability for surgery) (M10, F16)	6 Fracture L4-L5 (<48 hours for surgery)	ES MF (left side)	L4-L5 L5-S1	Intraoperative samples	**ES-** Significant increase of type I in CLBP group (p<.05). Significant decrease in type IIA fibers in study group (p<.01). No difference b/n groups for type IIX and IIC fibers. **MF**- Significant decrease in the % of type IIA fibers in	-	-	-	-

Table 1 (***Continued***)

Study/ Author	Participant no. & characteristics		Methodology			Results				
	Study group	Control group	Muscle (s)	Vertebral level	Biopsy Technique	Fiber type composition by numbers in %	Area occupied by each fiber type in % (RCSA)	Fiber size or CSA (μm^2)	Narrow diameter (μm)	Any otherrelevant
						CLBP group (p<.05). Significant increase in the % of type IIX fibers in CLBP group (p<.05). No difference b/n groups for the % of type I and IIC fibers.				
[13]	64 CLBP (IVD disorders for surgery) (M33, F31) Divided based on physical activity: LPA, MPA, HPA	17 Asymptomatic (M9, F8)	MF	L4-L5 (MRI, CT to confirm level)	Intraoperative samples for study group	% type II greater than type I in whole CLBP group. No difference b/n LPA, MPA and HPA for the % of type II fibers. Significantly increased % of type I in control group compared to CLBP group (p<.05)	No change in the % area of type I fibers in females compared to males. No difference b/n males and females for type I and type II % area in CLBP groups (p>.05)	-	Significantly increased diameter in both type I and II fibers in control group for both genders compared to CLBP (p<.05). No difference b/n LPA, MPA and HPA (p>.05). Significantly larger type I and II fibers in men in both CLBP and control group (p<.05).	Pathological fibers -significantly less in control group (p<.05) No difference b/n males and females in both CLBP and control groups. No differences b/n LPA, MPA and HPA (p>.05).

RCSA, relative cross-sectional area; CSA, cross-sectional area; NSCLBP, nonspecific chronic low back pain; CLBP, chronic low back pain; LBP, low back pain; ES, erector spinae; MF, multifidus, M, males; F, females; IVD, intervertebral disc; LPA, low physical activity; MPA, medium physical activity; HPA, high physical activity.

studies [47,56] presented their results combined for both males and females.

Two studies [12,56] compared nonspecific CLBP to healthy controls and three studies [13,47,57] compared CLBP attributed to a pathology with healthy controls. The causes of specific CLBP included ischemic and degenerative spondylolisthesis, intervertebral disc degeneration including disc herniation and stenosis, facet joint degeneration, or degenerative scoliosis.

Biopsy site and muscle samples

Two studies [47,56] compared muscle microscopy of both ES and MF muscles separately, whereas Mazis et al [13] examined only MF, and Crossman et al [12] examined only the ES. Mannion et al examined different muscles for comparison, MF in the CLBP group and ES in the control group [57]. Muscle samples were either stained by H and E stains [12,13], staining for mATPase after alkaline and acid preincubations [47,57] or by immunofluorescent antibodies [56]. Stained sections were viewed under the microscope [13,47] or from photographs taken from microscopic cameras [12,56,57] and analyzed using various imaging software program.

For all studies, biopsies were taken between L3 and L5 but one study [47] also included samples taken at the L5 −S1 vertebral level. Crossman et al [12] did not explicitly state the biopsy details but referred to a previous paper from the same research team [58] which indicated that biopsies were taken from T10 and L3 levels.

Risk of bias

All the five studies were found to be of poor quality presenting with a high RoB following the NOS assessment [42]. The RoB scores were affected in all three domains but mainly in selection and outcome domains as seen in Table 2. In the selection domain, overall studies scored low for selection of controls as the recruitment of participants were poorly explained or selected from a particular section of population. This was followed by lack of independent validation and clear representativeness of cases. In the outcome domain, there was lack of clarity on whether the same method of ascertainment was applied for cases and controls. Two studies showed total lack of comparability of the study groups based on age, gender, and BMI.

Data analysis

The homogeneity of the studies for data analysis was based on:

- Muscle biopsy site (L3−L5/S1) and the muscle tissue (ES or MF) sampled for the study.
- Population considered: Study group with CLBP (specific or nonspecific) and a control group as the healthy population.
- Age group: Included adults with age ranging from 17 to 61 years.
- Outcome measures: Fiber type composition by numbers and by area in percentage, fiber size/ CSA and fiber diameter or narrow diameter.

Pain related factors could not be considered as a measure of homogeneity as all the studies did not consistently measure pain features. However, meta-analysis was carried out where there was sufficient homogeneity in the outcome measures for a chosen muscle.

Meta-analysis

Four studies [12,13,47,56] were considered for meta-analysis as these studies measured at least one of the desired outcomes on either the ES or MF. The study by Mannion et al [57] compared different muscles MF and ES in CLBP population and asymptomatic controls respectively and hence could not be considered for meta-analysis or narrative interpretation of results as it would be inappropriate to compare two different muscles to report differences in CLBP and a control group. This study [57] was not considered in the analysis of any of the outcome measures, however, it was considered for reporting on gender differences in the discussion.

Table 2
Summary of risk of bias assessment by two reviewers (R1 and R2) independently and overall risk of bias for each study on agreement between the reviewers

Studies	Selection domain		Comparability domain		Outcome domain		Independent NOS assessment		Agreed combined NOS assessment
	R1	R2	R1	R2	R1	R2	R1	R2	
[56]	**	**	-	-	***	***	Poor	Poor	Poor
[12]	*	*	*	*	-	-	Poor	Poor	Poor
[57]	**	**	**	**	*	*	Poor	Poor	Poor
[47]	**	***	-	-	**	**	Poor	Poor	Poor
[13]	*	*	**	**	*	*	Poor	Poor	Poor

Table 3
Studies examining ES fiber type composition by numbers in percentage and the results of the individual studies comparing people with CLBP to controls

Fiber type	Study	Muscle	CLBP compared to Controls		
			Male	Female	Combined
Type I	[56]	ES			=
	[12]	ES	=		
	[47]	ES			↑
Type IIA	[56]	ES			=
	[47]	ES			↓
Type IIAX	[56]	ES			=
Type IIX (IIB)	[56]	ES			↓
	[47]	ES			=
Type IIC	[47]	ES			=

↑ significantly greater in CLBP compared to controls, ↓ significantly less in CLBP compared to controls, = Comparable between groups, grey cells- not examined.

Table 4
Studies examining area occupied by each fiber type in the ES in percentage and their results comparing people with CLBP to controls

Fiber type	Study	Muscle	CLBP compared to Controls		
			Male	Female	Combined
Type I	[56]	ES			=
	[12]	ES	=		
Type IIA	[56]	ES			=
Type IIAX	[56]	ES			=
Type IIX (IIB)	[56]	ES			↓

↑ significantly greater in CLBP compared to controls, ↓ significantly less in CLBP compared to controls, = Comparable between groups, gray cells- not examined.

Analysis of outcome measures: Erector Spinae muscle (ES): Fiber type composition by numbers in percentage

For the ES, type I and type II fiber composition by numbers (in %) were examined in three studies [12,47,56] and the individual study results are presented in Table 3. Two studies [12,56] show comparable results for type I fibers between the CLBP and control group, whereas Matejka et al [47] reported a significant increase in type I fibers in the CLBP population. Only two studies [47,56] examined type II fibers and reported comparable results across the fiber subtypes except for a significant decrease in type IIA and type IIX fibers reported in Matejka et al [47] and Agten et al [56], respectively.

For the percentage of type I fibers, three studies including a total of 81 participants with CLBP and 56 controls were analyzed. The data from individual studies were pooled for meta-analysis as shown in the forest plot (Fig. 2). An I^2=43% indicates sufficient homogeneity of these studies. The 95% CI of the pooled result crosses the line of no effect and the effect estimate showed no statistically significant difference (p=.14) between the CLBP and control groups. Similarly, a meta-analysis was conducted on the percentage of type II fibers. Two studies with a total of 46 and 24 participants for the CLBP and control group respectively were analyzed and the results are presented in Fig. 3. An I^2=0% indicates homogeneity of these studies. The 95% CI of the pooled result crosses the line of no effect and the effect estimate shows no statistically significant difference between the two groups (p=.29).

The GRADE assessment (Table 7) revealed that there was high RoB indicating poor quality of studies but there was consistency, precision and directness among the studies and publication bias scored moderate. Overall, the quality of evidence was low, which indicates that there were no significant differences in ES fiber type composition between people with and without CLBP.

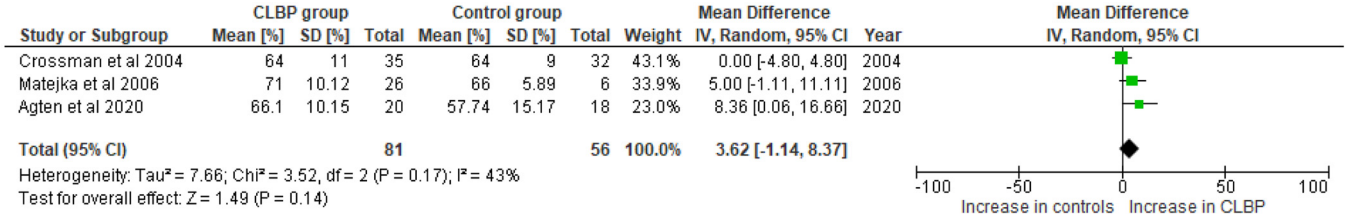

Fig. 2. Forest plot demonstrating the meta-analysis of the percentage of type I fibers in the ES in people with CLBP vs. controls. CLBP, Chronic low back pain; SD, standard deviation; IV, inverse variance; CI, confidence interval.

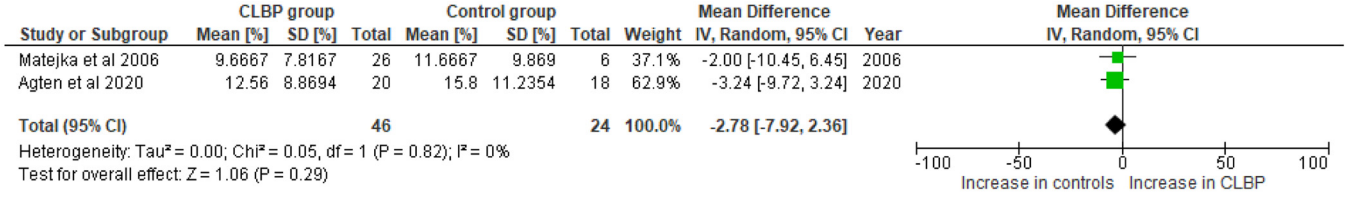

Fig. 3. Forest plot demonstrating the meta-analysis of the percentage of type II fiber in the ES in people with CLBP vs. controls. CLBP, Chronic low back pain; SD, standard deviation; IV, inverse variance; CI, confidence interval.

Table 5
Studies examining the ES fiber size or CSA (μm2) and their results comparing people with CLBP to controls

Fiber type	Study	Muscle	CLBP compared to Controls		
			Male	Female	Combined
Type I	[56]	ES			=
	[12]	ES	=		
Type II	[12]	ES	=		
Type IIA	[56]	ES			=
Type IIAX	[56]	ES			=
Type IIX (IIB)	[56]	ES			=
Ratio I/II	[12]	ES	=		
Mean Fiber area (MFA)	[12]	ES	=		

= Comparable between groups, grey cells- not examined.

Table 6
Studies examining the narrow diameter (ND) (in μm) and their results comparing people with CLBP to controls.

Fiber type	Study	Muscle	CLBP compared to Controls		
			Male	Female	Combined
Type I	[12]	ES	=		
Type II	[12]	ES	=		
ND ratio I/II	[12]	ES	=		
Mean Fiber ND	[12]	ES	=		

= Comparable between groups, grey cells- not examined.

Area occupied by each fiber type in percentage

The area occupied by each fiber type was examined in two studies [12,56] and the individual study results are shown in Table 4. For percentage area occupied by type I, the individual study results showed no significant difference between the CLBP and control group. The area occupied by type II fibers was reported only by Agten et al [56] and the results are comparable between the groups for type IIA and type IIAX, whereas, there was a significant decrease in the percentage area occupied by type IIX fibers in those with CLBP.

Meta-analysis was considered for the area occupied by type I fibers where two studies with a total of 55 and 50 participants for the CLBP and control group respectively were analyzed. Homogeneity was examined between the studies by Agten et al [56] and Crossman et al [12], however, the studies were found to be highly heterogeneous (I^2=86%), so no further meta-analysis were conducted. With regard to type II fibers, there were not sufficient studies to perform meta-analysis or analyze the results based on direction of effect.

Overall, the GRADE assessment (Table 7) indicated poor quality of studies with high RoB, inconsistency and moderate publication bias, though there was directness and precision of studies to some extent. To summarize, analyzing by the direction of effect, there was very low quality evidence for no significant difference between people with and without CLBP for the area occupied by different fiber types in the ES.

Fiber size or CSA

Fiber CSA was measured in two studies [12,56] and the individual study results are presented in Table 5. The CSA of both type I and type II fibers were examined by both studies and there was no significant difference between the CLBP and control population. In addition, the size ratio of I:II fibers was reported by Crossman et al [12] which showed the size of type I fibers were significantly greater than type II fibers in people with CLBP. Meta-analysis was conducted for both type I and type II fibers independently.

For both type I and type II fibers, two studies were analyzed with a total of 55 and 50 participants with and without CLBP, respectively. The study results for type I fiber were pooled for meta-analysis and are presented in Fig. 4. The 95% CI of the pooled result crossed the line of no effect and the effect estimate was not statistically significant (p=.43) between the CLBP and control groups. I^2=0% indi-

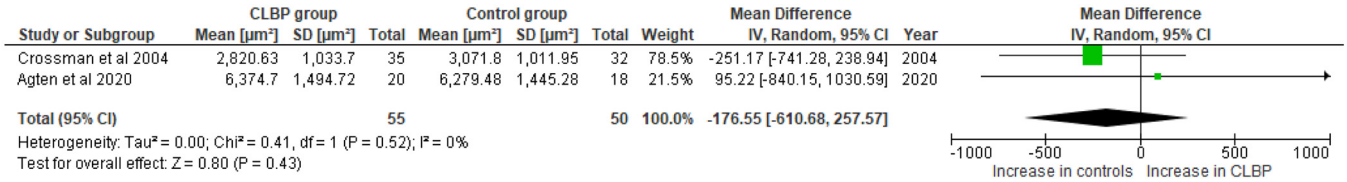

Fig. 4. Forest plot demonstrating the meta-analysis of type I fiber size or CSA in ES (in μm^2) in people with and without CLBP. CLBP, Chronic low back pain; CSA, cross-sectional area; SD, standard deviation; IV, inverse variance; CI, confidence interval.

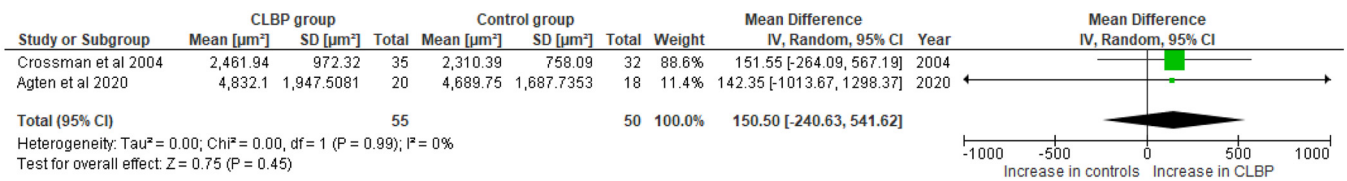

Fig. 5. Forest plot demonstrating the meta-analysis of type II fiber size or CSA in ES (in μm^2) in people with and without CLBP. CLBP, Chronic low back pain; CSA, cross-sectional area; SD, standard deviation; IV, inverse variance; CI, confidence interval.

Table 7
Shows the summary of the studies included and the overall results for each outcome measure in the ES in people with CLBP compared to a control group. The table also presents the GRADE assessment scores and the overall quality of evidence for each outcome

Outcome measures		Erector spinae muscle						GRADE assessment scores	GRADE levels of certainty/ quality of evidence
		Studies/ authors	No. of participants		Overall outcomes (CLBP compared to controls) Analysis				
			CLBP group	Control group	↑	↓	↔		
Fiber type composition by numbers in %	**Type I**	[56] [12] [47]	81	56			✓	ROB- high (-) Inconsistency- no (+) Imprecision- no (+) Indirectness- no (+) Publication bias- moderate (-)	Low
	Type II	[56] [47]	46	24			✓		
Area occupied by each fiber type in %	**Type I**	[56] [12]	55	50			✓	ROB- high (-) Inconsistency- yes (-) Imprecision −no (+) Indirectness- no (+) Publication bias- moderate (-)	Very low
	Type II	[56]	20	18		IIX only	✓		
Fiber size or CSA (μm^2)	**Type I**	[56] [12]	55	50			✓	ROB- high (-) Inconsistency- no (+) Imprecision- yes (-) Indirectness- no (+) Publication bias- moderate (-)	Very low
	Type II	[56] [12]	55	50			✓		
Narrow diameter (μm)	**Type I**	[12]	35	32			✓	ROB- high (-) Inconsistency # (-) Imprecision # (-) Indirectness- yes (-) Publication bias- moderate (-)	Very low
	Type II	[12]	35	32			✓		

- not enough evidence to address the particular domain.

cated homogeneity of studies. Similarly, meta-analysis was performed on type II fiber size/CSA and the results are presented in Fig. 5. The 95% CI of the pooled result crossed the line of no effect and the effect estimate was not statistically significant (p=.45) between the people with CLBP and controls. I^2=0% indicates sufficient homogeneity of studies.

The GRADE assessment (Table 7) demonstrated poor quality of studies (high RoB), lack of precision and moderate level of publication bias. There was however good consistency in the studies and a fair overall level of directness. To summarize, there is very low level quality of evidence which indicates that was there are no significant differences between people with and without CLBP for fiber size/CSA of the ES.

Narrow diameter

Fiber diameter/ narrow diameter was measured in only one study for type I and type II fibers [12]. The results of this study are comparable and show no significant difference between the CLBP and control groups (Table 6). However, it has to be noted that the study by Crossman et al [12] only included males. Meta-analysis could not be done for narrow diameter as there was an insufficient number of studies examining this outcome measure.

The GRADE assessment (Table 7) was incomplete from precision and consistency perspective. Besides, there was high RoB, moderate publication bias and lacked directness of the study. Hence, there is very low quality of evidence and no conclusion can be drawn for the comparison of narrow diameter between people with or without CLBP for the ES.

Table 8
Studies examining the MF fiber type composition by numbers in percentage and the results of the individual studies comparing people with CLBP to controls.

Fiber type	Study	Muscle	CLBP compared to Controls		
			Male	Female	Combined
Type I	[56]	MF			=
	[47]	MF			=
	[13]	MF			↓
Type II	[13]	MF			↑
Type IIA	[56]	MF			=
	[47]	MF			↓
Type IIAX	[56]	MF			=
Type IIX (IIB)	[56]	MF			=
	[47]	MF			↑
Type IIC	[47]	MF			=

↑ significantly greater in CLBP compared to controls, ↓ significantly less in CLBP compared to controls, = Comparable between groups, gray cells- not examined.

Table 9
Study examining area occupied by each fiber type in the MF muscle in percentage and their results comparing people with CLBP to controls

Fiber type	Study	Muscles	CLBP compared to Controls		
			Male	Female	Combined
Type I	[56]	MF			=
Type IIA	[56]	MF			=
Type IIAX	[56]	MF			=
Type IIX (IIB)	[56]	MF			=

= Comparable between groups, gray cells- not examined.

Table 10
Showing the study examining the MF muscle fiber size or CSA (μm^2) and the results comparing the CLBP with healthy controls

Fiber type	Study	Muscles	CLBP compared to Controls		
			Male	Female	Combined
Type I	[56]	MF			=
Type IIA	[56]	MF			=
Type IIAX	[56]	MF			=
Type IIX (IIB)	[56]	MF			=

= Comparable between groups, gray cells- not examined.

Multifidus

Fiber type composition by numbers in percentage

Type I and type II fiber composition were examined in three studies [13,47,56] and the individual study results are presented in Table 8. For type I fibers, two studies [47,56] showed no significant difference between groups whereas one study [13] reported a significant decrease in the proportion of type I fibers in the MF in people with CLBP compared to controls. For type II fibers, Mazis et al (2009) [13] demonstrated a significantly higher proportion of type II fibers in people with CLBP. In contrast, two studies examined type IIA fibers which showed either comparable results [56] or a significantly lower proportion [47] in people with CLBP compared to controls. Likewise, for type IIX fibers, Agten et al [56] reported comparable results between groups whereas, Matejka et al [47] reported a significantly higher proportion in people with CLBP compared to controls. For type IIAX and type IIC fibers which were assessed by Agten et al [56] and Matejka et al [47] respectively, no significant difference was observed between groups.

For both type I and type II fibers, three studies including a total of 110 participants with CLBP and 41 controls were analyzed. Meta-analysis was considered for each fiber type separately (type I and type II fibers) but very high heterogeneity was present between the studies (I^2=86% and 84%, respectively). Since the results did not meet the threshold of homogeneity stated in the protocol for meta-analysis (I^2<50%) [32], no further analysis was conducted.

The GRADE assessment (Table 12) showed poor quality of results with high RoB, and inconsistency among the studies with moderate publication bias though the studies score well on precision and directness. To summarize, there is very low quality of evidence and due to conflicting evidence, no conclusions can be drawn for the comparison of fiber type composition by numbers in people with and without CLBP for the MF muscle.

Table 11
Showing studies examining the narrow diameter (ND) (μm) in MF muscle and their results comparing the CLBP with healthy controls

Fiber type	Study	Muscles	CLBP compared to Controls		
			Male	Female	Combined
Type I	[13]	MF	↓	↓	
Type II	[13]	MF	↓	↓	

↑ significantly greater in CLBP compared to controls, ↓ significantly less in CLBP compared to controls, = Comparable between groups, gray cells- not examined.

Area occupied by each fiber type in percentage

The area occupied by each fiber type in the MF was examined by only one study [56] and the results showed no significant difference between the people with and without CLBP (Table 9).

The GRADE assessment (Table 12) revealed poor quality in terms of RoB, moderate publication bias and no sufficient evidence to comment on consistency and precision domain. This outcome measure only scores for directness of the study. Hence, there is very low quality of evidence and no conclusions can be drawn between people with and without CLBP for comparison of the area occupied by fiber types in the MF muscle.

Fiber size or CSA

MF fiber CSA for all fiber types was measured by only one study [56] and the results from this study presented no significant difference between the people with or without CLBP (Table 10).

The GRADE assessment (Table 12) was very similar to the previous outcome and revealed high RoB and moderate publication bias. This outcome measure scores only on directness of the study and there was not enough evidence to comment on the consistency and precision domains. Hence, there is very low quality of evidence and no conclusions can be drawn for the comparison of fiber size/CSA between people with and without CLBP for the MF muscle.

Narrow diameter

Fiber diameter/ narrow diameter was measured in only one study [13]. The results of this study demonstrated that narrow diameter is significantly less for both type I and type II fibers in both males and females in people with CLBP compared to controls (Table 11). In addition, this study [13] and the study conducted by Mannion et al [57] (which examined different muscles; MF in people with

Table 12
Studies included and the overall results for each outcome measure in the MF muscle in people with CLBP group compared to a healthy control group. The table also presents the GRADE assessment scores and the overall quality of evidence for each outcome

Outcome measures		Multifidus muscle						GRADE assessment scores	GRADE levels of certainty/ quality of evidence
		Studies/ authors	No. of participants		Overall outcome analysis (CLBP compared to Controls)				
			CLBP group	Control group	↑	↓	↔		
Fiber type composition by numbers in %	**Type I**	[56] [47] [13]	110	41			✓	ROB- high (-) Inconsistency - yes (-) Imprecision- no (+) Indirectness- no (+) Publication bias- moderate (-)	Very low
	Type II	[56] [47] [13]	110	41			✓		
Area occupied by each fiber type in %	**Type I**	[56]	20	18			✓	ROB- high (-) Inconsistency # (-) Imprecision # (-) Indirectness- no (+) Publication bias- moderate (-)	Very low
	Type II	[56]	20	18			✓		
Fiber size or CSA (μm^2)	**Type I**	[56]	20	18			✓	ROB- high (-) Inconsistency # (-) Imprecision # (-) Indirectness - no (+) Publication bias- moderate (-)	Very low
	Type II	[56]	20	18			✓		
Narrow diameter (μm)	**Type I**	[13]	64	17		✓		ROB- high (-) Inconsistency # (-) Imprecision # (-) Indirectness- yes (-) Publication bias- moderate (-)	Very low
	Type II	[13]	64	17		✓			

- not enough evidence to address the particular domain.

CLBP and ES in controls) reported that narrow diameter in females was significantly less compared to males in both the CLBP group and control group.

The GRADE assessment (Table 12) for this outcome showed poor quality with high RoB and moderate publication bias. This study lacked directness and there is insufficient evidence to comment on consistency and precision domains. Hence, there is very low quality of evidence and no conclusions can be drawn for narrow diameter when comparing between people with and without CLBP for the MF muscle.

Discussion

This systematic review assessed whether there were any differences in the microscopic features of the ES and MF muscles in people experiencing chronic spinal pain compared to controls. Only primary studies in the lumbar region were included as there was no previous research conducted in the cervical or thoracic region that met our eligibility criteria. Thus, this compromised the secondary objective to compare the same muscle across different regions of the spine and so the focus of the review moved from chronic spinal pain to only CLBP.

The meta-analysis results for ES demonstrated no significant difference between people with and without CLBP in terms of fiber type composition by numbers and fiber size/ CSA. However, there was a slight tendency for an increase in the number of type I fibers and an increase in size/CSA of type II fibers in people with CLBP compared to controls, although this was not statistically significant. The results for the area occupied by each fiber type showed no difference between people with and without CLBP, and the results for fiber diameter/narrow diameter were inconclusive due to inadequate evidence. The overall confidence in the quality of evidence was very low for the ES.

The results for the MF muscle were inconclusive due to heterogeneity of studies measuring fiber type composition, and there was inadequate evidence for other outcome measures including percentage area occupied by fiber types, fiber size/CSA, and fiber diameter/ narrow diameter. The overall confidence in the quality of evidence was also very low for the MF muscle across all outcomes.

The fiber type composition in the ES matched with the independent results from the two included studies of the review [12,56]. Whereas, some previous studies have shown a significant increase in type I and significant decrease in particularly type IIA fibers in people with CLBP [47,59], but Shahidi et al [59] only investigated

chronic degenerative spine pathology and did not have a pain-free comparator group. Another study by Mannion and colleagues [57] contradicts these findings, but this study attempted to compare two separate groups using two different muscles (ES and MF). The systematic review conducted by Cagnie et al [27] presented conflicting evidence for the same outcome measure but this review included only two studies, one of which was a cadaveric study on the MF [60] and the other assessed and compared the ES between people with LBP patients and healthy controls [12]. Similarly, the systematic review by Goubert et al [24] reported no differences in the fiber type of paraspinal muscles between the groups, but this result was based on just one study [12].

For fiber type composition of the MF, the study by Agten et al [56] showed comparable results between groups, but a cadaveric study by Bajek et al [61] and the study by Mazis et al [13] contradicted each other as the former concluded that there is a significant increase in type I fibers and the latter a significantly lower proportion of type I fibers in people with CLBP. Rantanen et al [62] compared MF biopsies pre- and postsurgery for lumbar disc herniation and noted selective atrophy of type II fibers. These results are also supported by a cadaveric study in the neck region which report a significant increase in type I fibers in MF when compared with a flexor muscle of the neck [63]. Most previous research that compared different extensor muscles in people with LBP due to acute disorders noted no significant differences in the fiber type distribution among the spinal extensor muscles [64–66]. This implies that although we could not draw conclusions for fiber distribution in the MF muscle, the direction of effect may be similar for both ES and MF muscle.

Previous research which examined the spinal extensor muscles in either a healthy population or cadavers demonstrated significantly more type I fibers compared to disc herniation patients, likely due to the fact that these extensor muscles primarily provide postural stability [61,67,68]. Research shows reduced use of skeletal muscles (due to pain and immobility/severe deconditioning) eventually leads to conversion of type I to type II fibers [69,70]. A longer duration of pain is associated with significantly reduced number of type I fibers and higher proportion type II fibers [71]. On the other hand, with aging, there can be a conversion of fibers to type I fibers or there is higher proportion of type I fibers due to preferential atrophy of functional type II fibers [71–73]. By considering these points when interpreting the findings of the current review, the lack of significant findings between groups may have been influenced by variability in pain intensity/duration and age, especially given the small sample sizes. To note, not all of the studies reported pain intensity or duration for the CLBP group. As there was insufficient data available on confounding factors to further analyze their effects, it is evident that it is important to consider the implications of these factors in future research. An additional consideration is that the subtypes (intermediate fibres) of type II fibers could not be examined due to insufficient published evidence. Knowledge of the number of intermediate fibers (type I/IIA, type IIA/IIX, and type IIC) in people with CLBP could support the existence of ongoing fiber type conversion.

ES fiber size showed a tendency to be smaller for type I and larger for type II fibers in people with CLBP although this was not statistically significant. The data from the two studies that did examine fiber size had a wide variation in the numerical values for each fiber type CSA [12,56], This is most likely attributed to the different methodology used in these studies to calculate CSA that is, Agten et al [56] measured the CSA on the photographs taken from fluorescent microscopy, whereas, mean fiber size was quantified by the narrow diameter measurement in the study by Crossman et al [12]. Similarly, for the MF muscle, no conclusions could be drawn for fiber size and narrow diameter, however the only study measuring fiber size [56] showed comparable results between people with and without CLBP, and the only study measuring narrow diameter [13] showed a decrease in people with CLBP.

With regard to male and female differences in muscle microscopic features, some studies have shown that females have significantly smaller fiber diameter in their spinal extensor muscles compared to males for both type I and type II fibers [13,57]. On the other hand, the proportional area occupied by type I fibers has been shown to be greater in females compared to males [27] irrespective of pain duration/ intensity [57,71,74,75]. In addition to these outcomes, some studies have observed the presence of pathological changes such as moth-eaten fibers, core fibers, target fibers and angulated fibers in patients with chronic back pain [13,62,76].

Although the individual studies have suggested microscopic changes in the muscle fibers and possible fiber type conversions when people have chronic pain, this review gives comparable results for limited outcome measures (for fiber type composition and size for the ES) and is inconclusive for most of the outcome measures. Nonetheless, to date, it still remains unclear whether or not any of the microscopic structural changes in spinal musculature are a cause or consequence of pain [24,27] and this can only be addressed in future longitudinal studies.

Study limitations

This review failed to address the second objective of the study, that is, to compare the muscle microscopic features in different regions of the spine despite a comprehensive search. This is due to the strict eligibility criteria that eliminated all the cadaveric studies. However, this decision was justified as cadaveric muscle samples may not be pain-free samples without a clearly documented medical history. Further, this review considers studies that examined people with either specific or nonspecific pain introducing heterogeneity in the study group. This however was a necessity as there was insufficient evidence on nonspecific pain alone. This review also revealed the need for future research

investigating people with neck or thoracic pain and as a result, our review included only studies on people with CLBP and even in this population, there were very few studies.

This review is limited to microscopic findings since there are several systematic reviews on macroscopic changes of the spinal muscles [22–24,26]. Despite these limitations, this systematic review is comprehensive and gives a clear picture of the existing evidence as well as its inadequacies.

Future research

There is a lot of scope for additional research examining microscopic changes within the spinal extensor muscles. To date, there is no study comparing microscopic changes in the thoracic or cervical region in people with and without chronic pain. Additionally, no study has examined microscopic changes within the ES or MF across different regions of the spine (ie, painful versus non-painful regions) in people with chronic pain. There are also very few studies which have analyzed the muscle composition and fiber characteristics of the MF muscle in people with CLBP. Future research is also required to examine the spinal extensor musculature in people with specific conditions for example, intravertebral disc herniation, and to compare muscle microscopic changes at the affected segment compared to unaffected segments and the asymptomatic side. Further, extending the research to examine correlations with age, pain intensity and duration should be considered in future studies. Eventually, studies examining the correlation between macroscopic and microscopic changes in the spinal extensor muscles could be conducted to explore the pathogenesis of chronic pain.

Conclusion

This review found no significant difference between the people with and without CLBP for fiber type composition (both by numbers and proportional area occupied by fibers) and fiber CSA/size in the ES. No conclusion could be drawn for fiber diameter/narrow diameter in the ES due to insufficient evidence. For the MF muscle, no conclusion could be drawn for all of the outcome measures including muscle fiber composition, fiber size and narrow diameter due to heterogeneity of studies and inadequate evidence. The confidence in the overall quality of evidence is very low and supports no difference in the muscle microscopic features of ES and MF between people with and without CLBP. Further research is needed to clarify whether or not there are differences in microscopic morphology in spinal extensor muscles across different regions of spine in people with and without chronic spinal pain.

Supporting information:

Appendix 1: PRISMA 2020 Checklist

Appendix 2: Database searches

Appendix 3: Comprehensive data extraction table used to extract data from eligible studies

Appendix 4: Newcastle Ottawa Scale (NOS) for case-control studies- adaptation and justification

Declarations of Competing Interests

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

The authors have not declared a specific grant for this research from any funding agency in the public, commercial or not-for-profit sectors.

Supplementary materials

Supplementary material associated with this article can be found in the online version at https://doi.org/10.1016/j.spinee.2022.01.023.

References

[1] Hay SI, Abajobir AA, Abate KH, Abbafati C, Abbas KM, Abd-Allah F, et al. Global, regional, and national disability-adjusted life-years (DALYs) for 333 diseases and injuries and healthy life expectancy (HALE) for 195 countries and territories, 1990–2016: a systematic analysis for the Global Burden of Disease Study 2016. Lancet North Am Ed 2017;390(10100):1260–344.

[2] Thiese MS, Hegmann KT, Wood EM, Garg A, Moore JS, Kapellusch J, et al. Prevalence of low back pain by anatomic location and intensity in an occupational population. BMC Musculoskelet Dis 2014;15:283.

[3] Manchikanti L, Singh V, Datta S, Cohen SP, Hirsch JA. Comprehensive review of epidemiology, scope, and impact of spinal pain. Pain Phys 2009;12(4):E35–70.

[4] WHO. World Health Organization (2015). World Report on Ageing and Health, Geneva. World Health Organization; 2015.

[5] Balagué F, Mannion AF, Pellisé F, Cedraschi C. Non-specific low back pain. Lancet North Am Ed 2012;379(9814):482–91.

[6] de Campos TF. Low back pain and sciatica in over 16s: assessment and management NICE Guideline [NG59]. J Physiother 2017;63(2):120.

[7] Ehrlich GE. Back pain. J Rheumatol Suppl 2003;67:26–31.

[8] Hartvigsen J, Hancock MJ, Kongsted A, Louw Q, Ferreira ML, Genevay S, et al. What low back pain is and why we need to pay attention. Lancet North Am Ed 2018;391(10137):2356–67.

[9] Fejer R, Kyvik KO, Hartvigsen J. The prevalence of neck pain in the world population: a systematic critical review of the literature. Eur Spine J 2006;15(6):834–48.

[10] Briggs AM, Smith AJ, Straker LM, Bragge P. Thoracic spine pain in the general population: prevalence, incidence and associated factors in children, adolescents and adults. A systematic review. BMC Musculoskelet Dis 2009;10(1):77.

[11] Chan ST, Fung PK, Ng NY, Ngan TL, Chong MY, Tang CN, et al. Dynamic changes of elasticity, cross-sectional area, and fat infiltration of multifidus at different postures in men with chronic low back pain. Spine J 2012;12(5):381–8.

[12] Crossman K, Mahon M, Watson PJ, Oldham JA, Cooper RG. Chronic low back pain-associated paraspinal muscle dysfunction is not the result of a constitutionally determined "adverse" fiber-type composition. Spine 2004;29(6):628–34.

[13] Mazis N, Papachristou DJ, Zouboulis P, Tyllianakis M, Scopa CD, Megas P. The effect of different physical activity levels on

muscle fiber size and type distribution of lumbar multifidus. A biopsy study on low back pain patient groups and healthy control subjects. Eur J Phys Rehabil Med 2009;45(4):459–67.
[14] Schomacher J, Falla D. 'Function and structure of the deep cervical extensor muscles in patients with neck pain'. Man Ther 2013;18 (5):360–6.
[15] Wallwork TL, Stanton WR, Freke M, Hides JA. The effect of chronic low back pain on size and contraction of the lumbar multifidus muscle. Man Ther 2009;14(5):496–500.
[16] Falla D. Unravelling the complexity of muscle impairment in chronic neck pain. Man Ther 2004;9(3):125–33.
[17] Jorgensen K, Nicolaisen T. Trunk extensor endurance: determination and relation to low-back trouble. Ergonomics 1987;30 (2):259–67.
[18] Kankaanpaa M, Taimela S, Laaksonen D, Hanninen O, Airaksinen O. Back and hip extensor fatigability in chronic low back pain patients and controls. Arch Phys Med Rehabil 1998;79(4):412–7.
[19] Mannion AF. Fibre type characteristics and function of the human paraspinal muscles: normal values and changes in association with low back pain. J Electromyogr Kinesiol 1999;9(6):363–77.
[20] Mannion AF, Taimela S, Müntener M, Dvorak J. Active therapy for chronic low back pain: part 1. Effects on back muscle activation, fatigability, and strength. Spine 2001;26(8):897–908.
[21] Roy SH, De Luca CJ, Casavant DA. Lumbar muscle fatigue and chronic lower back pain. Spine 1989;14(9):992–1001.
[22] De Pauw R, Coppieters I, Kregel J, De Meulemeester K, Danneels L, Cagnie B. Does muscle morphology change in chronic neck pain patients? – A systematic review. Man Ther 2015;22:42–9.
[23] Farrell SF, Smith AD, Hancock MJ, Webb AL, Sterling M. Cervical spine findings on MRI in people with neck pain compared with pain-free controls: a systematic review and meta-analysis. J Magn Reson Imaging 2019;49(6):1638–54.
[24] Goubert D, Van Oosterwijck J, Meeus M, Danneels L. Structural changes of lumbar muscles in non-specific low back pain. Pain Phys 2016;19(7):E985–99.
[25] Owers DS, Perriman DM, Smith PN, Neeman T, Webb AL. Evidence for cervical muscle morphometric changes on magnetic resonance images after whiplash: a systematic review and meta-analysis. Injury 2018;49(2):165–76.
[26] Ranger TA, Cicuttini FM, Jensen TS, Peiris WL, Hussain SM, Fairley J, et al. Are the size and composition of the paraspinal muscles associated with low back pain? A systematic review. Spine J 2017;17 (11):1729–48.
[27] Cagnie B, Dhooge F, Schumacher C, De Meulemeester K, Petrovic M, van Oosterwijck J, et al. Fiber typing of the erector spinae and multifidus muscles in healthy controls and back pain patients: a systematic literature review. J Manipulative Physiol Ther 2015;38 (9):653–63.
[28] Page MJ, Moher D, Bossuyt PM, Boutron I, Hoffmann TC, Mulrow CD, et al. PRISMA 2020 explanation and elaboration: updated guidance and exemplars for reporting systematic reviews. BMJ 2021:372.
[29] Cumpston M, Li T, Page MJ, Chandler J, Welch VA, Higgins JP, et al. Updated guidance for trusted systematic reviews: a new edition of the Cochrane Handbook for Systematic Reviews of Interventions. Cochrane Database Syst Rev 2019;10:ED000142.
[30] Furlan AD, Pennick V, Bombardier C, van Tulder M. 2009 updated method guidelines for systematic reviews in the Cochrane Back Review Group. Spine 2009;34(18):1929–41.
[31] Higgins, JP, Thomas, J, Chandler, J, Cumpston, M, Li, T, Page, MJ et al. (2019a) *Cochrane Handbook for Systematic Reviews of Interventions*.John Wiley & Sons.
[32] Purushotham S, Stephenson RS, Sanderson A, Falla D. Microscopic changes in the spinal extensor musculature in patients experiencing chronic spinal pain: protocol for a systematic review. BMJ open 2021;11(2):e042729.
[33] Moher D, Shamseer L, Clarke M, Ghersi D, Liberati A, Petticrew M, et al. Preferred reporting items for systematic review and meta-analysis protocols (PRISMA-P) 2015 statement. Syst Rev 2015;4(1):1–9.
[34] Methley AM, Campbell S, Chew-Graham C, McNally R, Cheraghi-Sohi S. PICO, PICOS and SPIDER: a comparison study of specificity and sensitivity in three search tools for qualitative systematic reviews. BMC Health Serv Res 2014;14(1):1–10.
[35] NICE, g. (2018) 'National Institute For Health And Care Excellence Guideline scope Chronic pain: assessment and management'.
[36] Lefebvre C, Glanville J, Briscoe S, Littlewood A, Marshall C, Metzendorf MI, et al. Searching for and selecting studies. Cochrane Handbook Syst Rev Interv 2019:67–107.
[37] Bramer WM, Giustini D, de Jonge GB, Holland L, Bekhuis T. De-duplication of database search results for systematic reviews in EndNote. J Med Library Assoc: JMLA 2016;104(3):240.
[38] Edwards P, Clarke M, DiGuiseppi C, Pratap S, Roberts I, Wentz R. Identification of randomized controlled trials in systematic reviews: accuracy and reliability of screening records. Stat Med 2002;21 (11):1635–40.
[39] Landis JR, Koch GG. An application of hierarchical kappa-type statistics in the assessment of majority agreement among multiple observers. Biometrics 1977:363–74.
[40] Vucic K, Kadic AJ, Puljak L. Survey of Cochrane protocols found methods for data extraction from figures not mentioned or unclear. J Clin Epidemiol 2015;68(10):1161–4.
[41] Drevon D, Fursa SR, Malcolm AL. Intercoder reliability and validity of WebPlotDigitizer in extracting graphed data. Behav Modif 2017;41(2):323–39.
[42] Wells G, Shea B, O'Connell D, Robertson J, Peterson J, Welch V, Losos M, Tugwell P, et al. (2015) 'The Newcastle-Ottawa Scale (NOS) for Assessing the Quality of Nonrandomized Studies in Meta-Analysis'.
[43] Seehra J, Pandis N, Koletsi D, Fleming PS. Use of quality assessment tools in systematic reviews was varied and inconsistent. J Clin Epidemiol 2016;69:179–184e5.
[44] Hootman JM, Driban JB, Sitler MR, Harris KP, Cattano NM. Reliability and validity of three quality rating instruments for systematic reviews of observational studies. Res Synthesis Methods 2011;2(2):110–8.
[45] Luchini C, Stubbs B, Solmi M, Veronese N. Assessing the quality of studies in meta-analyses: advantages and limitations of the Newcastle Ottawa Scale. World J Meta-Anal 2017;5(4):80–4.
[46] Higgins, JP, Thomas, J, Chandler, J, Cumpston, M, Li, T, Page, MJ and Welch, VA. (2019b) Cochrane Handbook for Systematic Reviews of Interventions version 6.2 (updated February 2021). Cochrane, 2021.Available at: www.training.cochrane.org/handbook.
[47] Matejka J, Miloslava Z, Koudela K, Pavelka T. Changes of muscular fibre types in erector spinae and multifidus muscles in the unstable lumbar spine. J Back Musculoskelet Rehabil 2006;19(1):1–5.
[48] Ryan R. Cochrane Consumers and Communication Review Group: data synthesis and analysis. Cochrane Consumers and Communication Review Group; data synthesis and analysis'; 2013 http://cccrg.cochrane.org. Accessed date June 2021.
[49] Balshem H, Helfand M, Schünemann HJ, Oxman AD, Kunz R, Brozek J, et al. GRADE guidelines: 3. Rating the quality of evidence. J Clin Epidemiol 2011;64(4):401–6.
[50] Ghezzi-Kopel K. Cochrane interactive learning. J Med Library Assoc 2018;106(4). Available at: https://training.cochrane.org/interactive-learning/module-7-interpreting-findings. (Accessed June, 2021).
[51] Guyatt, GH, Oxman, AD, Vist, G, Kunz, R, Brozek, J, Alonso-Coello, P, et al. (2011e) 'GRADE guidelines: 4. Rating the quality of evidence—study limitations (risk of bias)', *J Clin Epidemiol*,64(4), pp. 407-415.
[52] Guyatt, GH, Oxman, AD, Kunz, R, Woodcock, J, Brozek, J, Helfand, M, et al. (2011c) 'GRADE guidelines: 7. Rating the quality of evidence—inconsistency', *J Clin Epidemiol*,64(12), pp. 1294-1302.

[53] Guyatt, GH, Oxman, AD, Kunz, R, Brozek, J, Alonso-Coello, P, Rind, D, et al. (2011a) 'GRADE guidelines 6. Rating the quality of evidence—imprecision', *J Clin Epidemiol,*64(12), pp. 1283-1293.

[54] Guyatt, GH, Oxman, AD, Kunz, R, Woodcock, J, Brozek, J, Helfand, M, et al. (2011b) 'GRADE guidelines: 8. Rating the quality of evidence—indirectness', *J Clin Epidemiol,*64(12), pp. 1303-1310.

[55] Guyatt, GH, Oxman, AD, Montori, V, Vist, G, Kunz, R, Brozek, J, et al. (2011d) 'GRADE guidelines: 5. Rating the quality of evidence—publication bias', *J Clin Epidemiol,*64(12), pp. 1277-1282.

[56] Agten A, Stevens S, Verbrugghe J, Timmermans A, Vandenabeele F. Biopsy samples from the erector spinae of persons with nonspecific chronic low back pain display a decrease in glycolytic muscle fibers. Spine J 2020;20(2):199–206.

[57] Mannion, AF, Weber, BR, Dvorak, J, Grob, D and Muntener, M. (1997b) 'Fibre type characteristics of the lumbar paraspinal muscles in normal healthy subjects and in patients with low back pain', *J Orthop Res,*15(6), pp. 881-887.

[58] Mannion AF, Dumas GA, Stevenson JM, Cooper RG. The influence of muscle fiber size and type distribution on electromyographic measures of back muscle fatigability. Spine 1998;23(5):576–84.

[59] Shahidi B, Hubbard JC, Gibbons MC, Ruoss S, Zlomislic V, Allen RT, et al. Lumbar multifidus muscle degenerates in individuals with chronic degenerative lumbar spine pathology. J Orthop Res 2017;35(12):2700–6.

[60] Mattila M, Hurme M, Alaranta H, Paljärvi L, Kalimo H, Falck B, et al. The multifidus muscle in patients with lumbar disc herniation. A histochemical and morphometric analysis of intraoperative biopsies. Spine 1986;11(7):732–8.

[61] Bajek S, Bobinac D, Bajek G, Vranic TS, Lah B, Dragojevic DM. Muscle fiber type distribution in multifidus muscle in cases of lumbar disc herniation. Acta Med Okayama 2000;54(6):235–41.

[62] Rantanen J, Hurme M, Falck B, Alaranta H, Nykvist F, Lehto M, et al. The lumbar multifidus muscle five years after surgery for a lumbar intervertebral disc herniation. Spine 1993;18(5):568–74.

[63] Boyd-Clark LC, Briggs CA, Galea MP. Comparative histochemical composition of muscle fibres in a pre- and a postvertebral muscle of the cervical spine. J Anat 2001;199(Pt 6):709–16.

[64] Bagnall K, Ford D, McFadden K, Greenhill B, Raso V. The histochemical composition of human vertebral muscle. Spine 1984;9(5):470–3.

[65] Ford D, Bagnall K, McFadden K, Greenhill B, Raso J. Analysis of vertebral muscle obtained during surgery for correction of a lumbar disc disorder. Cells Tissues Organs 1983;116(2):152–7.

[66] Regev GJ, Kim CW, Thacker BE, Tomiya A, Garfin SR, Ward SR, et al. Regional Myosin Heavy Chains Distribution in Selected Paraspinal Muscles. Spine (Phila Pa 1976) 2010;35(13):1265–70.

[67] Jorgensen K, Nicholaisen T, Kato M. Muscle fiber distribution, capillary density, and enzymatic activities in the lumbar paravertebral muscles of young men: Significance for isometric endurance. Spine 1993;18(11):1439–50.

[68] Sirca A, Kostevc V. The fibre type composition of thoracic and lumbar paravertebral muscles in man. J Anat 1985;141:131.

[69] Pette D, Staron RS. Mammalian skeletal muscle fiber type transitions. Int Rev Cytol 1997;170:143–223.

[70] Scott W, Stevens J, Binder−Macleod SA. Human skeletal muscle fiber type classifications. Phys Ther 2001;81(11):1810–6.

[71] Mannion AF, Kaser L, Weber E, Rhyner A, Dvorak J, Muntener M. Influence of age and duration of symptoms on fibre type distribution and size of the back muscles in chronic low back pain patients. Eur Spine J 2000;9(4):273–81.

[72] Kamen G, Sison SV, Du C, Patten C. Motor unit discharge behavior in older adults during maximal-effort contractions. J Appl Physiol 1995;79(6):1908–13.

[73] Roos MR, Rice CL, Vandervoort AA. Age-related changes in motor unit function. Muscle Nerve 1997;20(6):679–90.

[74] Mannion, AF, Dumas, GA, Cooper, RG, Espinosa, FJ, Faris, MW and Stevenson, JM. (1997a) 'Muscle fibre size and type distribution in thoracic and lumbar regions of erector spinae in healthy subjects without low back pain: Normal values and sex differences', *J Anat,*190, pp. 505-513.

[75] Thorstensson A, Carlson H. Fibre types in human lumbar back muscles. Acta Physiol Scand 1987;131(2):195–202.

[76] Kaser L, Mannion AF, Rhyner A, Weber E, Dvorak J, Muntener M. Active therapy for chronic low back pain Part 2. Effects on paraspinal muscle cross-sectional area, fiber type size, and distribution. Spine 2001;26(8):909–19.