Received: 6 June 2023 | Accepted: 20 December 2023

DOI: 10.1111/cpf.12870

Clinical Physiology and Functional Imaging

REVIEW ARTICLE

Is musculoskeletal pain associated with increased muscle stiffness? Evidence map and critical appraisal of muscle measurements using shear wave elastography

Andreas Haueise[1] | Guillaume Le Sant[2,3] | Angelika Eisele-Metzger[4,5] | Angela V. Dieterich[1]

[1]Faculty of Health, Security, Society, Furtwangen University, Furtwangen, Germany

[2]CHU Nantes, Movement–Interactions–Performance, MIP, Nantes Université, Nantes, France

[3]School of Physiotherapy, IFM3R, St-Sebastien/Loire, France

[4]Institute for Evidence in Medicine, Medical Center, Faculty of Medicine, University of Freiburg, Freiburg, Germany

[5]Cochrane Germany, Cochrane Germany Foundation, Freiburg, Germany

Correspondence
Angela V. Dieterich, Faculty of Health, Security, Society, Furtwangen University, Konrad-Goldmann-Straße 7, 79100 Freiburg i. Br., Germany.
Email: dita@hs-furtwangen.de

Funding information
Ministerium für Wissenschaft, Forschung und Kunst Baden-Württemberg,
Grant/Award Number: Innovative Projekte 2020; Baden-Württemberg Ministry of Science, Research and Arts, Germany

Abstract

Introduction and Aims: Approximately 21% of the world's population suffers from musculoskeletal conditions, often associated with sensations of stiff muscles. Targeted therapy requires knowing whether typically involved muscles are objectively stiffer compared to asymptomatic individuals. Muscle stiffness is quantified using ultrasound shear wave elastography (SWE). Publications on SWE-based comparisons of muscle stiffness between individuals with and without musculoskeletal pain are increasing rapidly. This work reviewed and mapped the existing evidence regarding objectively measured muscle stiffness in musculoskeletal pain conditions and surveyed current methods of applying SWE to measure muscle stiffness.

Methods: A systematic search was conducted in PubMed and CINAHL using the keywords "muscle stiffness", "shear wave elastography", "pain", "asymptomatic controls" and synonyms. The search was supplemented by a hand search using Google Scholar. Included articles were critically appraised with the AXIS tool, supplemented by items related to SWE methods. Results were visually mapped and narratively described.

Results: Thirty of 137 identified articles were included. High-quality evidence was missing. The results comprise studies reporting lower stiffness in symptomatic participants, no differences between groups and higher stiffness in symptomatic individuals. Results differed between pain conditions and muscles, and also between studies that examined the same muscle(s) and pathology. The methods of the application of SWE were inconsistent and the reporting was often incomplete.

Conclusions: Existing evidence regarding the objective stiffness of muscles in musculoskeletal pain conditions is conflicting. Methodological differences may explain most of the inconsistencies between findings. Methodological standards for SWE measurements of muscles are urgently required.

KEYWORDS
imaging methods, muscle, stiffness, musculoskeletal pain, shear wave elastography

1 | INTRODUCTION

Approximately 21% of the world's population suffers from musculoskeletal conditions (World Health Organization, 2022). Often, patients perceive muscles in the painful body region as stiff. The sensation of stiff muscles is a typical symptom of chronic neck pain, osteoarthritis, rheumatoid arthritis, temporomandibular joint dysfunction and other pathologies (Alfuraih et al., 2020; Dor & Kalichman, 2017; Olchowy et al., 2020; Takasawa et al., 2015).

Several studies reported that perceived symptoms of pain and stiffness do not correlate with measurements of muscle stiffness (Akagi & Kusama, 2015; Kolding et al., 2018; Proulx et al., 2023; Sawada et al., 2020). Other studies identified a correlation, e.g. the large migraine study by Hvedstrup et al. (2020) demonstrated a low, but significant correlation between neck muscle stiffness and pressure pain thresholds in one of the three investigated subgroups, the participants with migraine and ictal neck pain (Hvedstrup et al., 2020). A clear differentiation between pain, stiffness, and tenderness may be difficult (Maigne et al., 2012), but a definite characterization of the impairment is important to determine appropriate therapeutic interventions. When muscles are found to be objectively stiffer, therapeutic interventions should focus on reducing muscle stiffness. If the muscles are more sensitive, interventions should aim at reducing sensitivity rather than hardness.

Ultrasound shear wave elastography (SWE) provides quantitative measures of tissue stiffness for precise locations in superficial and deep tissues at relatively low costs (Hug et al., 2015). Due to these unique advantages, SWE has been recommended as the superior elastography method for quantifying tissue stiffness (Ferraioli et al., 2022). However, the specific properties of muscle tissue require careful consideration of the imaging and measuring methods (Bernabei et al., 2020). First, muscle anisotropy and the unit of measurement (Davis et al., 2019; Ferraioli et al., 2022; Gennisson et al., 2010). SWE systems measure the velocity of the propagating shear waves that have been induced by so-called "push beams." Results are provided in shear wave speed (m/s) or Young's modulus (kPa) (Shiina et al., 2015). The anisotropy of muscle tissue (Gennisson et al., 2010) does not comply with the assumptions of Young's modulus. Measurements of muscle stiffness should be reported as shear modulus, which requires recalculating the system-provided results (Brandenburg et al., 2014; Hug et al., 2015). Second, the scanning direction, shear wave propagation is better along muscle fibres than across, and longitudinal scanning yields higher measures of muscle stiffness (Ewertsen et al., 2018; Gennisson et al., 2010; Wang et al., 2022c). SWE measurements of muscle stiffness have been validated in the longitudinal scanning direction (Eby et al., 2013). Third, the measured location and area of a muscle (Ewertsen et al., 2016). SWE has been developed to support the diagnosis of cancerous or fibrotic tissue within organs of relatively homogeneous structure, such as the liver (Shiina et al., 2015). The measurement tools of several SWE systems provide small measurement boxes in which tissue stiffness at a specific location is precisely quantified for comparison with other tissue locations (Bota et al., 2011). Typically, muscle stiffness presents as inhomogeneous (Davis et al., 2019). The manual placement of small measurement boxes within inhomogeneous tissue entails a risk of bias; a measurement box may be placed where the measure appears representative or appropriate, confirming expectations. In addition, the measurements from small boxes may not be representative of a much larger muscle. Last, SWE measurements rely on the successful and valid detection of the speed of the propagating shear waves (Yavuz et al., 2015). Often, shear wave speed cannot be tracked throughout the complete region of interest. Depending on the system, black areas within the elastogram (Friede et al., 2020; Klauser et al., 2022) or a separate quality map (Barr et al., 2015; Yavuz et al., 2015) indicate image regions of insufficient tracking quality. Trustworthy measurements require transparency regarding the sufficient control of the quality of shear wave tracking (Barr et al., 2015).

Publications on studies that examined muscle stiffness via SWE in musculoskeletal pain conditions are increasing rapidly and report conflicting results (Dieterich et al., 2020; Lin et al., 2022; Taş et al., 2018). We therefore aimed to review the current evidence of objectively increased muscle stiffness with musculoskeletal pain conditions. The specific aims of our work were (i) to describe and map the current literature on objectively measured muscle stiffness in individuals with musculoskeletal pain compared to asymptomatic individuals using SWE and (ii) to compare SWE methods and identify inconsistencies in the use and reporting of measurements by SWE.

2 | METHODS

As far as applicable the reporting of this study is based on the Preferred Reporting Items for Systematic Reviews and Meta-Analyses (PRISMA) 2020 guidelines (Page et al., 2021), complemented by methodological guidance for creating evidence maps (Miake-Lye et al., 2016; Schmucker et al., 2013; Snilstveit et al., 2016; White et al., 2020).

2.1 | Information sources and search strategy

A systematic search of articles was performed using PubMed and CINAHL, supplemented by a hand search using Google Scholar and the bibliographies of the included articles. The date of the last search was 8 December 2022. The following search strategy was used in PubMed: (*muscle stiffness OR elasticity OR hardness OR*

tenderness) *AND* (*pain OR headache OR arthritis OR TMD OR ITBS*) *AND* (*shear wave elastography*) *AND* (*healthy OR asymptomatic OR pain-free OR volunteers*). No filters were used concerning the publication date or the article type. The results were limited to humans and articles written in English or German.

2.2 | Eligibility criteria

Original research that compared the stiffness of defined muscles in defined samples with and without pathology using the same SWE methods was eligible. To be eligible, studies had to include a pain-free group acting as control. Since the study question "Is musculoskeletal pain associated with increased muscle stiffness?" refers to musculoskeletal pain in general, all conditions associated with musculoskeletal pain in different body regions (e.g. chronic neck pain, osteoarthritis and headache) were included. Only studies using longitudinal measurements were included in this review. Studies were excluded if they measured solely the stiffness of fascia or tendinous tissues (e.g. Achilles tendon or plantar fascia) but not any muscle. Conditions that were characterized by biomechanical deviations, e.g. scoliosis, were excluded if musculoskeletal pain was not an inclusion criterion for the group with pathology. Case studies, published abstracts without full text, nonpublished work (e.g. theses) and secondary literature were excluded.

2.3 | Study selection

Two researchers independently screened all titles and abstracts. Potentially relevant full texts were retrieved and independently assessed for eligibility by the same researchers. Discrepancies at each stage were discussed and resolved during consensus meetings.

2.4 | Data extraction

Data extraction was conducted independently by two reviewers using a standardized spreadsheet. Disagreements between reviewers were resolved through discussion involving a third reviewer if needed. The extracted data were (a) sample characteristics (sex, age, group descriptions, sample sizes), (b) pain condition (diagnostic and inclusion criteria, duration), (c) muscles assessed with SWE, (d) measurement procedures (system and unit of measurement, measurement position, data processing, measurement repetitions, reliability analyses and quality control), (e) the results, for each examined muscle separately and (f) the authors' conclusions. In studies that aimed to investigate the effects of interventions, only the baseline measurements and their methods were included in the current review.

2.5 | Critical appraisal of study quality

Although the appraisal of methodological quality is not mandatory in evidence maps (Schmucker et al., 2013), assessing the quality of the included studies is crucial for confidence in their results. The AXIS tool has been proposed for the appraisal of cross-sectional studies (Downes et al., 2016). This tool facilitates a comprehensive review of study quality, targeting the most common aspects that tend to bear risk of bias in cross-sectional designs. Briefly, the tool is composed of 20 items divided into five categories "Introduction," "Methods," "Results," "Discussion" and "Others." In the "Methods" section, reviewers are specifically instructed to conduct a thorough assessment of the methodology's quality. Reviewers are provided with guiding questions and explanatory help texts to aid in their critical assessment of each item. The AXIS tool does not provide a numerical score for grading the quality of publications. This allows the user to prioritize items depending on their relevance to the research question (Downes et al., 2016). In this review, we used the quality categories proposed by Raynaud et al. (2021), as presented in Table 1. As this work refers only to the cross-sectional comparison of muscle stiffness between individuals with a pain condition (cases) and asymptomatic individuals (controls), we used the AXIS tool to assess all included studies. Two researchers individually assessed the study quality and resolved discrepancies during consensus meetings.

In current guidelines on elastography, muscles play a minor role (Săftoiu et al., 2019; Shiina et al., 2015). Expert reviews indicate open questions regarding the interpretation of SWE images (Davis et al., 2019; Ferraioli et al., 2022). Based on methodological literature (Davis et al., 2019; Ferraioli et al., 2022; Javed et al., 2022; Lin et al., 2017; Rominger et al., 2018) and the authors' experience, 11 supplementary items were developed to specify aspects relevant to Items 9 and 11 of the AXIS tool (parameters' measurement and description of methods,

TABLE 1 Judgement criteria for the critical appraisal of included studies using the AXIS tool according to Raynaud et al. (2021).

Low quality High risk of bias	Moderate quality Moderate risk of bias	High quality Low risk of bias
If the study design or population representativeness or selection process or data description or parameters' measurements have not been adequately addressed, or if five or more items have not been adequately addressed	If exactly three or four items or the representativeness of the measurement box have not been adequately addressed	If three or fewer items have not been adequately addressed

respectively) and to compare the SWE methods between studies. The supplementary items targeted the replicability of a study by sufficiently reported information, the control of the image and the measurement quality, the reliability and the representativeness of the measurements, data processing and the statistical analysis. The SWE-specific checklist including justifications for each item is presented in Table 2.

2.6 | Visual evidence mapping

To map the existing evidence visually, we created a 4 × 5 tabulated bubble chart using MATLAB R2021a (MathWorks). Each study is represented by a bubble. The size of the bubble provides information on the study's sample size; the colour of the bubble informs about the study quality according to the adapted AXIS assessment. The bubble chart includes four body regions (head and neck; shoulder, upper extremity and thorax; lumbar region; lower extremity) and five categories of study results (stiffness of the symptomatic group, compared to the nonsymptomatic group):

1. Lower stiffness of the symptomatic group (Symptomatic < Asymptomatic).
2. Mixed results including lower and equal stiffness of the symptomatic group (Symptomatic ≤ Asymptomatic).
3. Equal stiffness in both groups (Symptomatic = Asymptomatic).
4. Mixed results including higher and equal stiffness of the symptomatic group (Symptomatic ≥ Asymptomatic).
5. Higher stiffness of the symptomatic group (Symptomatic > Asymptomatic).

The categories "Symptomatic ≤ Asymptomatic" and "Symptomatic ≥ Asymptomatic" include studies, in which more than one muscle has been examined with inconsistent results between muscles or measurement conditions.

Furthermore, a body chart was created to compile the reported stiffness for each examined muscle. Since the included studies reported muscle stiffness in different measurement units, that is, Young's modulus (E), shear modulus (μ) or shear wave velocity (V_s), the shear modulus was calculated based on the following formulae for comparability.

TABLE 2 Item descriptions and justifications for the supplementary items to the AXIS tool.

Item description	Justification
Was the measurement system (including transducer properties) reported?	Enables the study to be repeated, specific bandwidths are recommended for SWE (Ferraioli et al., 2022) and can influence the numeric outcome (Rominger et al., 2018). Major differences across the technology of different manufacturers (Javed et al., 2022).
Were the system settings (e.g. presets, dynamic range and depth) reported to an extent that enables repetition? (e.g. figure with information)	Enables the study to be repeated and to understand the settings behind the measurements. Presets may influence the SWV values (Rominger et al., 2018).
Were measures undertaken (and reported) to ensure sufficient image quality?	Image quality influences measurement outcomes. Areas without elastography values may distort the measurements. The probe orientation influences the measurement outcomes, specifically in muscle. Artefacts from movement or structural boundaries may distort the measurements (Davis et al., 2019; Ferraioli et al., 2022; Lin et al., 2017).
Were the measurements repeated more than once (technical replicates)?	Improves reliability, reduces unwanted motion artefacts and provides more robust measurements (Davis et al., 2019).
Were reliability analyses conducted?	The literature reports major differences in reliability between muscles and elastography systems (Davis et al., 2019; Javed et al., 2022). Reference to a reliability study that has not been conducted under comparable conditions is questionable.
Was image analysis fully computed?	Reduces bias by a standardized processing procedure for all images.
Was image analysis blinded to group allocation?	Reduces bias by not knowing in which images higher stiffness is expected.
Were the measurement boxes placed in a standardized way/ standardized location?	Reduces bias since boxes cannot be placed where they appear appropriate.
Was fascia included in the measurement box?	Fascia is typically much stiffer than muscle tissue. Such an influence on stiffness measurements must be declared (Ferraioli et al., 2022).
Was the size of the measurement box stated and justified?	Enables study replication and allows for judging one aspect of the representativeness of the measured area.
Were the placement and size of the measurement boxes representative of the muscle?	Depending on the size of the examined muscle, measurement boxes should be large or many to provide as representative measurements as possible.

Abbreviation: SWE, shear wave elastography; SWV, shear wave velocity.

$$\mu = \rho V_s^2,$$

$$E \approx 3\mu,$$

where ρ is the tissue density. Since ρ is assumed to be 1000 kg/m^{-3} in muscular tissue (Hug et al., 2015) the shear modulus in kPa is the square of the shear wave velocity.

3 | RESULTS

3.1 | Selected studies

After the removal of duplicates, a total of 137 articles were screened (130 identified through the systematic search, seven identified by searching additional sources), 30 of which were ultimately included in this review. The search and selection process is shown in Figure 1.

3.2 | Study characteristics

The included articles examined the stiffness of 35 different muscles in 16 musculoskeletal pain conditions. The outcome measure of muscle stiffness was reported using the unit shear modulus in kPa, Young's modulus in kPa or shear wave velocity in m/s. We summarize the different outcome units in the following report using the term "stiffness." The most frequently assessed muscles were the upper part of the trapezius muscle (11 articles, 36.6%) followed by the lumbar multifidus muscle (six articles, 20.0%). Chronic low back pain was the most frequent pain condition (seven articles, 23.3%). All included muscles with their reported range of stiffness are presented in Figure 3. A total of 1943 individuals have been assessed in the included studies, 989 symptomatic and 954 asymptomatic individuals. The sex distribution within study samples was not always declared, but based on the studies with clear information, more women were included. Ultrasound devices from six manufacturers have been used, predominantly Aixplorer systems (SuperSonic Image, Aix-en-Provence) (19 articles, 63.3%). All but two studies used only linear transducers. Hashimoto et al. (2022) used only a convex transducer (Hashimoto et al., 2022), while Sedlackova et al. (2021) used both, a linear and a convex probe (Sedlackova et al., 2021). An overview of the extracted data is presented in Table 3.

3.3 | Study quality

AXIS items that were often not met were the representativeness of the sample (Item 5), the selection process (Item 6) and the justification of the sample size (Item 3). Also, the outcome measurements (Item 9) and the description of methods (Item 11) were often not met due to missing reliability analyses, reporting of Young's modulus and missing information regarding the use of SWE. Moreover, the discussion of study limitations was often missing or limited to the study design without considering the influence of methodological decisions. A colour-coded overview of the results of the AXIS appraisal can be found in Figure S1.

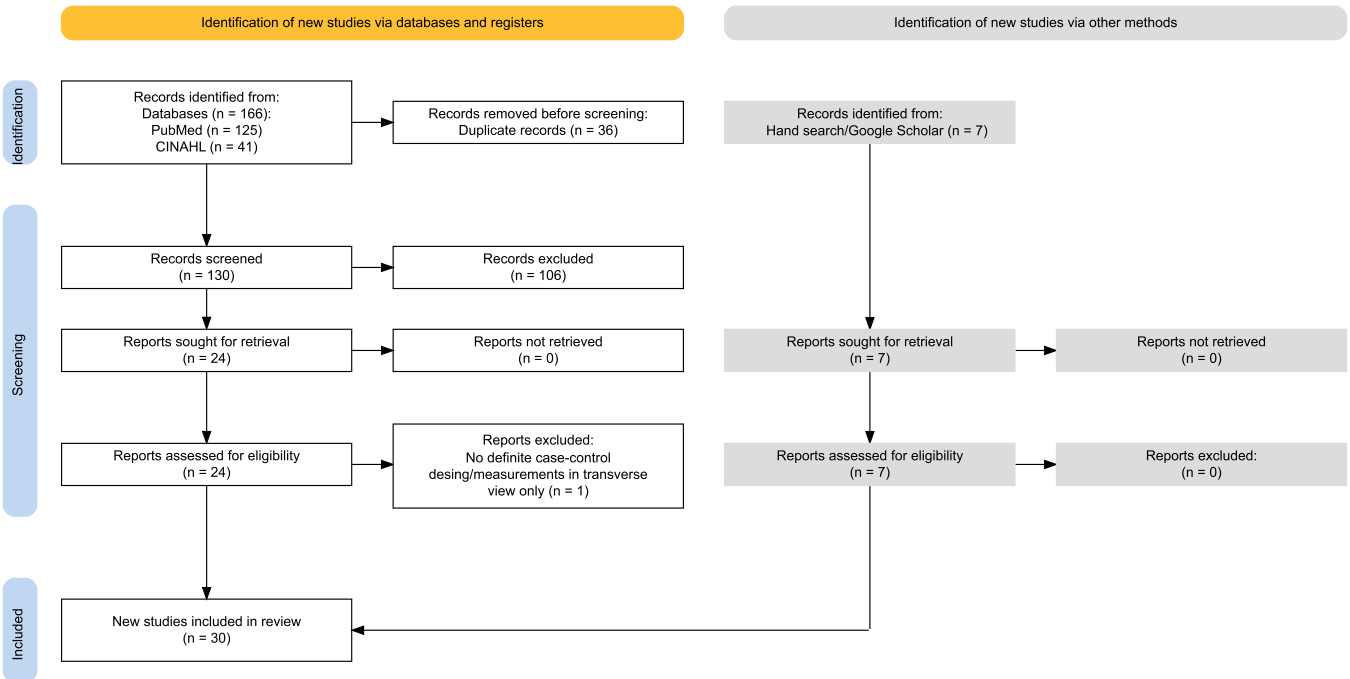

FIGURE 1 PRISMA flowchart with reasons for exclusion. Created with the R package and Shiny app for PRISMA 2020-compliant flow diagrams by (Haddaway et al., 2022). PRISMA, Preferred Reporting Items for Systematic Reviews and Meta-Analyses.

TABLE 3 Main characteristics of included studies.

Reference	Pathology	Sample		Muscle(s) assessed	Measurement position	Measurement repetitions	Results (stiffness in relaxation)
		N symptomatic (sex, age in years)	*N* asymptomatic (sex, age in years)				
Alfuraih et al. (2020)	Three groups of RA: • Newly diagnosed untreated (new RA) • Diagnosed over 1 year and in remission (remission RA) • Diagnosed over 1 year and active (active RA)	New RA: 29 (21 females), age: 56.8; Remission RA: 33 (14 females), age: 65.9; Active RA: 18 (14 females), age: 60.9	40 (28 females), age: 61.5	M. vastus lateralis M. rectus femoris M. vastus medialis M. vastus intermedius M. biceps femoris M. semitendinosus M. semimembranosus M. biceps brachii	All muscles were tested in the most relaxed position. For the quadriceps, the participants were supine on a flat bed. The hamstrings, they were prone on a flat bed with the knees flexed at 90° and rested on a wall. For M. biceps brachii, they were supine and the elbow was flexed at 90° with the forearm rested on the body and the hand in supination.	Three per muscle	No statistically significant differences between RA patients and healthy controls.
Aljinović et al. (2020)	Whiplash injury	75 (43 females), age: 43.1	75 (41 females), age: 46.5	M. trapezius M. splenius capitis M. sternocleidomastoideus	The trapezius and splenius capitis muscles were examined in a seated position with relaxed shoulder girdle and the arms in supination resting on the thighs, the head slightly flexed forward. The sternocleidomastoid muscle was examined in a supine position with the chin up.	Three per measurement box/six per muscle	The trapezius muscle was significantly stiffer in patients with a whiplash injury than in healthy controls.
Botanlioglu et al. (2013)	Patellofemoral pain syndrome	11 (all females), age: 30.8	22 (11 females), age: 28.8	M. vastus lateralis M. vastus medialis obliquus	Sitting with the hip in 90° flexion, knees extended and the ankle in a neutral position, a towel under the popliteal fossa. Measurements were taken in neutral hip position and in 30° hip abduction.	Three per muscle and hip position	No statistically significant differences during relaxation.
Botanlıoğlu et al. (2019)	Patellofemoral pain syndrome	40 (20 females), age: 46.5	40 (20 females), age: 46.5	M. vastus lateralis M. vastus medialis obliquus	Not described	Three per muscle	No statistically significant differences during relaxation for M. vastus medialis obliquus. M. vastus lateralis was stiffer in the symptomatic group.

TABLE 3 (Continued)

Reference	Pathology	Sample		Muscle(s) assessed	Measurement position	Measurement repetitions	Results (stiffness in relaxation)
		N symptomatic (sex, age in years)	*N* asymptomatic (sex, age in years)				
Dieterich et al. (2020)	Chronic nonspecific neck pain	20 (all females), age: 52.5	18 (all females), age: 48.5	M. trapezius M. splenius capitis M. semispinalis capitis M. semispinalis cervicis M. multifidus cervicis	Participants were positioned in upright sitting within a multicervical unit (BTE Technologies)	Three per condition	No statistically significant differences between groups across all muscle layers and tasks.
Ertekin et al. (2021)	Myofascial pain syndrome	30 (all females), age: 20.3	30 (all females), age: 20.9	M. trapezius	Sitting position with the hands on the knees in full resting position and the head upright and in the midline.	At least 10 repetitions	The stiffness in the symptomatic group was significantly higher than in asymptomatic controls.
Friede et al. (2020)	Iliotibial band syndrome	14 (7 females), age: 32.6	14 (7 females), age: 27.2	M. tensor fasciae latae M. gluteus maximus	Subjects lay supine on an examination bed with their backs slightly raised and knees rested on a support cushion (hip angle 140°−150°, knee angle ~90°).	Three per muscle	Stiffness of M. tensor fasciae latae was significantly lower in symptomatic subjects. No information on statistical differences for M. gluteus maximus.
Gao et al. (2020)	Low back pain	20 (10 females), age: 29	9 (4 females), age: 27	M. iliocostalis lumborum	Measured in relaxation when the participant was in a neutral prone position.	Twice per condition	The stiffness of symptomatic dysfunctional muscles was significantly higher than in muscles with normal function and in healthy controls.
Gökşen et al. (2021)	Knee osteoarthritis	31 (21 females), age: 54.6	10 (no further information given)	M. rectus femoris M. biceps femoris	The rectus femoris was measured in supine on the distal one-third between SIAS and patella. The biceps femoris long head was measured in prone on the distal one-fourth between the SIPS and the fibular head.	Three per muscles	No statistically significant differences between the groups.
Hashimoto et al. (2022)	Low back pain + piriformis syndrome	28 (all males), age: 39.3	30 (all males), no further information	M. piriformis	Sidelying with the hip joint bent at 40° and the knee joint bent at 90° to relax the piriformis muscle and the participant.	Not reported	Significantly stiffer muscles in symptomatic subjects.

(Continues)

TABLE 3 (Continued)

Reference	Pathology	Sample		Muscle(s) assessed	Measurement position	Measurement repetitions	Results (stiffness in relaxation)
		N symptomatic (sex, age in years)	*N* asymptomatic (sex, age in years)				
Hvedstrup et al. (2020)	Migraine (with ictal neck pain)	48 (44 females) without ictal neck pain, age: 42; 52 (46 females) with ictal neck pain, age: 44	46 (43 females), age: 42	M. trapezius M. splenius capitis M. semispinalis capitis	Ultrasound SWE examination was performed with individuals sitting relaxed in a comfortable chair, similar to previous studies.	Eight per location	The neck muscles of subjects with ictal neck pain were significantly stiffer than those of subjects without ictal neck pain and of healthy controls. No significant differences in trapezius stiffness across the groups.
Itoigawa et al. (2022)	Posterior shoulder pain	22 (all males), age: 20	25 (all males), no further information was given	M. infraspinatus M. supraspinatus	The participants were in a relaxed state with the shoulder at 0° abduction and neutral rotation when the SSPM and ISPM were measured.	Two per muscle	Significantly higher stiffness of infraspinatus muscle in symptomatic subjects. No differences for the supraspinatus muscle.
Klauser et al. (2022)	Iliotibial band syndrome	14 (7 females), age: 32.6	14 (7 females), age: 26.1	M. tensor fasciae latae M. gluteus maximus	To obtain SWE images, participants lay relaxed supine on an examination bed with their backs slightly raised and knees rested on a support cushion (hip angle 140°–150°, knee angle ~90°	Three per muscle	Significantly lower stiffness of the tensor fasciae latae muscle in symptomatic subjects. No statistical differences for M. gluteus maximus.
Kolding et al. (2018)	Tension-type headache	17 (10 females), age: 33.7	29 (20 females), age: 32.6	M. masseter M. sternocleidomastoideus M. trapezius	The participants were placed in a chair with an adjustable headrest to ensure relaxation of the muscles of the head and neck throughout the examination.	Ten per location	No statistically significant differences between the symptomatic and asymptomatic subjects.
Koppenhaver et al. (2020)	Low back pain	60 (24 females), age: 32.2	60 (34 females), age: 31	M. multifidus lumbalis M. erector spinae lumbalis	Participants were positioned prone on a plinth with pillows placed under their pelvis and ankles to increase comfort and reduce lumbar lordosis.	Three per muscle	Significantly higher stiffness values for both muscles in symptomatic subjects.

TABLE 3 (Continued)

Reference	Pathology	Sample		Muscle(s) assessed	Measurement position	Measurement repetitions	Results (stiffness in relaxation)
		N symptomatic (sex, age in years)	*N* asymptomatic (sex, age in years)				
Leong et al. (2016)	Rotator cuff tendinopathy	26 (all males), age: 23.6	17 (all males), age: 21.7	M. trapezius	Sitting upright on a stool with the head in a neutral position. For the passive tasks, the participants were asked to stay fully relaxed with their elbow flexed at 90° and their forearm in pronation. The arm was first rested on the hip at 0° of shoulder abduction and was then passively positioned at 30° and 60° of shoulder abduction.	Three per position and task	Significantly higher stiffness in symptomatic subjects in 0° abduction. A trend towards higher stiffness in 30° abduction, with no statistically significant differences in 60° abduction.
Li et al. (2021)	Knee osteoarthritis	50 (28 females), age: 56.9	50 (32 females), age: 56.2	M. semimembranosus M. semitendinosus M. biceps femoris	Before scanning, all participants were placed in a prone position on the scanning bed and were asked to relax and be comfortable for 5 min.	Three per location	Significantly higher stiffness in knee OA patients than in asymptomatic controls.
Lin et al. (2022)	Cervicogenic headache	19 (10 females), age: 27.5	20 (10 females), age: 24.6	M. trapezius M. splenius capitis M. semispinalis capitis M. semispinalis cervicis	Subjects were lying prone with their shoulders in abduction and external rotation, hands overlapped under the forehead and the chest supported by a pillow to position the neck in neutral.	Three per muscle	Significantly higher stiffness of all muscles in symptomatic subjects.
Masaki et al. (2017)	Low back pain	9 (8 females), age: 27.5	23 (15 females), age: 32.7	M. multifidus lumbalis M. erector spinae lumbalis	In the prone position.	Two per muscle	Significantly higher stiffness of M. multifidus in symptomatic subjects. No statistical differences in the erector spinae muscle.
Masaki et al. (2019)	Low back pain	23 (all females), age: 74.3	Healthy 19 (all females), age: 72.4 LBP history 16 (all females), age: 70.3	M. multifidus lumbalis M. erector spinae lumbalis	In the prone position.	One per location	The multifidus muscle was significantly stiffer in the LBP history group than in the healthy control group. No information on differences between the LBP and control groups. No significant differences across groups for the erector spinae muscle.

(Continues)

TABLE 3 (Continued)

Reference	Pathology	Sample		Muscle(s) assessed	Measurement position	Measurement repetitions	Results (stiffness in relaxation)
		N symptomatic (sex, age in years)	*N* asymptomatic (sex, age in years)				
Murillo et al. (2019)	Low back pain	15 (8 females), age: 29.4	15 (7 females), age: 26.7	M. multifidus lumbalis (superficial and deep)	Participants were positioned in prone with a rolled towel placed under their abdomen to minimize the lumbar lordosis.	Two per muscle	Significantly higher stiffness in the superficial multifidus in the symptomatic group than in the asymptomatic group. No differences in the deep multifidus between the groups.
Pinto et al. (2022)	Low back pain	78 (46 females), age: 46	73 (47 females), age: 48	M. multifidus lumbalis	In the prone position at rest.	Three per muscle	No statistical differences between the groups.
Sedlackova et al. (2021)	Cervicogenic headache	23 (20 females), age: 36.4	23 (did not differ in sex, age or BMI)	M. sternocleidomastoideus M. trapezius	In a relaxed supine position with their hands placed beside their thighs.	Six per location	Significantly higher stiffness of M. sternocleidomastoideus in symptomatic subjects. No differences in the trapezius muscle.
Takashima et al. (2017)	Temporomandibular disorder	26 (all females), 13 with limited mouth opening, age: 26.9 13 without limited mouth opening, age 30.5	24 (all females), age: 29	M. masseter	In a sitting position with the head held in a natural position.	One per location	Significantly higher stiffness in symptomatic subjects than in asymptomatic subjects.
Taş et al. (2018)	Chronic neck pain	35 (27 females), age: 35.6	35 (24 females), age: 35.2	M. trapezius M. sternocleidomastoideus M. levator scapulae M. splenius capitis	Trapezius and splenius capitis muscles were measured in prone with the arms relaxed at the sides and the wrists pronated, forehead on a U-shaped pillow. The sternocleidomastoid muscle was measured in supine with a soft pillow underneath the neck.	Three per muscle	Significantly higher stiffness of the trapezius, sternocleidomastoideus and levator scapulae muscles in symptomatic subjects. No differences in the splenius capitis muscle.
Valera-Calero et al. (2021)	Myofascial pain syndrome	19, age: 21.5	34, age: 21.9	M. trapezius	In the prone position.	Two per location	Significantly higher stiffness of control points in symptomatic subjects. No differences between trigger points.

TABLE 3 (Continued)

Reference	Pathology	Sample		Muscle(s) assessed	Measurement position	Measurement repetitions	Results (stiffness in relaxation)
		N symptomatic (sex, age in years)	*N* asymptomatic (sex, age in years)				
Wang, Cui, et al. (2022)	Pincer-type femoroacetabular impingement	33 (22 females), age: 39	37 (24 females), age: 39	M. iliocapsularis	Supine with their hip joint in external rotation.	Three per muscle and hip position	Significantly lower stiffness in symptomatic subjects than in asymptomatic subjects.
Wang, Liu, et al. (2022)	Ankylosing spondylitis	30 (3 females), age: 29.5	27 (4 females), age: 27.5	M. multifidus lumbalis	In a prone position and a state of the lumbar multifidus relaxation.	One per muscle and condition	Significantly higher stiffness in symptomatic subjects than in asymptomatic subjects.
Wolff et al. (2022)	Idiopathic chronic neck pain	18 (11 females), age: 34.4	18 (11 females), age: 32.2	M. sternocleidomastoideus M. trapezius	The Trapezius muscle was measured in prone, the face resting in a hollow headrest, arms resting along the sides. M. sternocleidomastoideus was measured in supine, the head resting at the same level as the body, arms resting along the sides.	Two per muscle	No differences between the groups for both muscles in relaxation.
Zhang et al. (2022)	Medial tibial stress syndrome	23 (all males), age: 21.8	20 symptomless runners (all males), age: 22.2; 20 controls (all males), age: 23.6	M. tibialis anterior M. extensor digitorum longus M. peroneus longus M. soleus M. gastrocnemius lateralis M. gastrocnemius medialis M. flexor digitorum longus M. tibialis posterior	The subjects were asked to lie prone on the examination table and bend their knee (90°) so that the calf was in the most relaxed state.	Five per muscle	Significantly higher stiffness of all but the extensor digitorum longus muscles in symptomatic subjects compared to controls and symptomless runners.

Abbreviations: BMI, body mass index; ISPM, infraspinatus muscle; LBP, lower back pain; OA, osteoarthritis; RA, rheumatoid arthritis; SSPM, supraspinatus muscle.

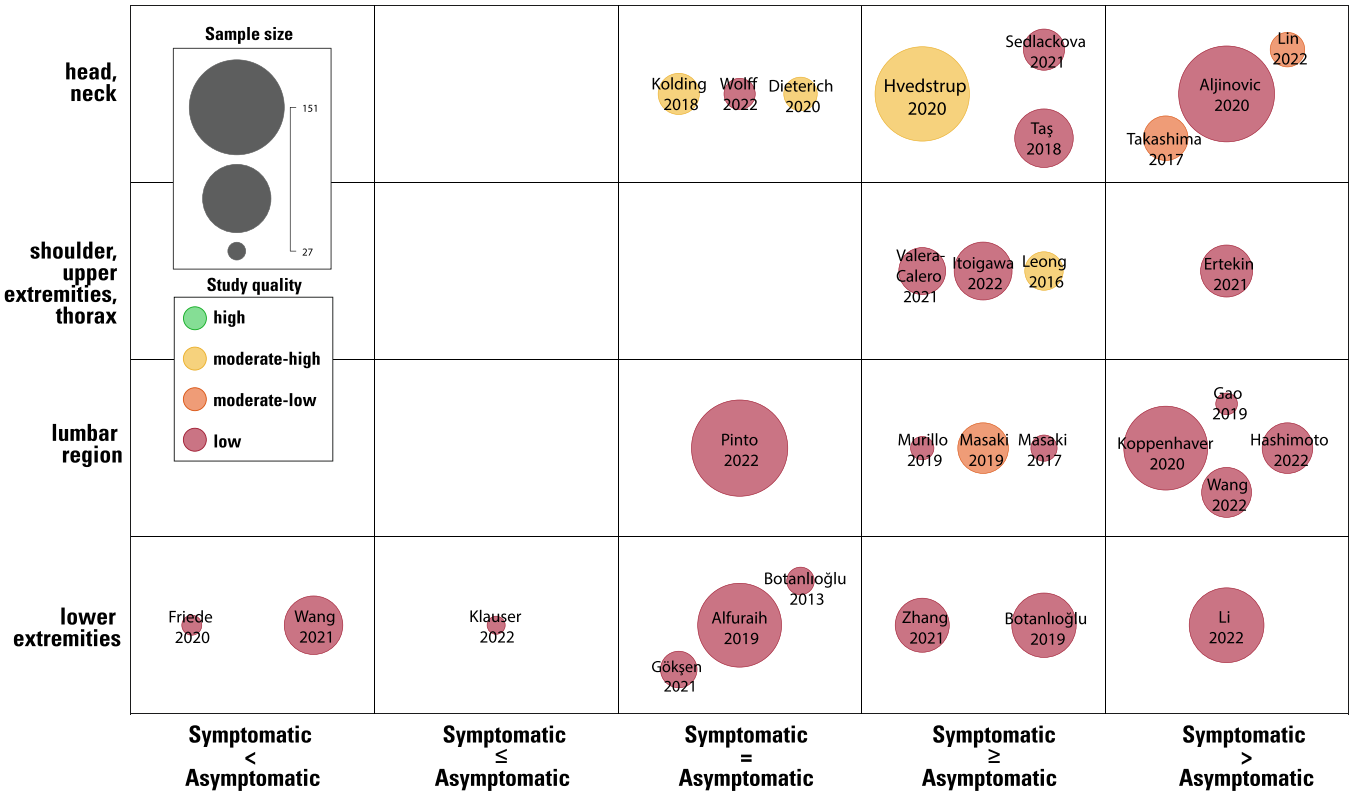

FIGURE 2 Evidence map of the included studies sorted by the examined body region and by five result categories. The signs <, = and > refer to the stiffness of the measured muscles in the symptomatic, compared to the asymptomatic group. The second (≤) and fourth category (≥) include inconsistent findings between muscles or measurement conditions.

3.4 | Evidence map and map of muscle stiffness

An overview of study results is given in the evidence map (Figure 2) and a narrative description of each study is provided in Table 3. Figure 3 displays the reported stiffness values in a body chart. Lower stiffness in the symptomatic group was found in two studies (6.6%) examining the iliotibial band syndrome (Friede et al., 2020) and pincer-type femoroactebabular impingement (Wang, Cui, et al., 2022). Lower or equal stiffness in the symptomatic group was found in one study (3.3%) that examined also the iliotibial band syndrome (Klauser et al., 2022). Equal stiffness between groups was found in seven studies (23.3%). The studies investigated idiopathic/non-specific chronic neck pain (Dieterich et al., 2020; Wolff et al., 2022), tension-type headache (Kolding et al., 2018), low back pain (Pinto et al., 2022), patellofemoral pain syndrome (Botanlioglu et al., 2013), knee osteoarthritis (Gökşen et al., 2021) and rheumatoid arthritis (Alfuraih et al., 2020). No difference or higher stiffness in symptomatic individuals was identified in eleven studies (36.6%) that investigated chronic neck pain (Taş et al., 2018), migraine (Hvedstrup et al., 2020), cervicogenic headache (Sedlackova et al., 2021), myofascial pain syndrome in the neck region (Valera-Calero et al., 2021), rotator cuff tendinopathy (Leong et al., 2016), posterior shoulder pain (Itoigawa et al., 2022), low back pain (Masaki et al., 2017; Masaki et al., 2019; Murillo et al., 2019), patellofemoral pain syndrome (Botanlıoğlu et al., 2019) and medial tibial stress syndrome (Zhang et al., 2022). Higher stiffness in the symptomatic group was reported by nine studies (30.0%) that investigated whiplash injury (Aljinović et al., 2020), cervicogenic headache (Lin et al., 2022), temporomandibular disorder (Takashima et al., 2017), myofascial pain syndrome in the neck region (Ertekin et al., 2021), low back pain (Gao et al., 2020; Koppenhaver et al., 2020; Wang, Liu, et al., 2022), piriformis syndrome (Hashimoto et al., 2022) and knee osteoarthritis (Li et al., 2021).

Most studies were judged low quality. The four studies (13.3%) of moderate-high quality reported equal or equal and higher stiffness in the symptomatic group (Dieterich et al., 2020; Hvedstrup et al., 2020; Kolding et al., 2018; Leong et al., 2016). The three studies (10.0%) of moderate-low quality reported equal and higher or only higher stiffness in the symptomatic group (Lin et al., 2022; Masaki et al., 2019; Takashima et al., 2017).

3.5 | Critical appraisal of SWE methods (supplementary items)

Regarding set-up and data collection procedures, all articles provided information about the ultrasound device and the transducer used for data collection. The type and the version of the elastography system were often not stated. Only eight study reports (26.6%) provided written information on the scanning parameters and settings of the

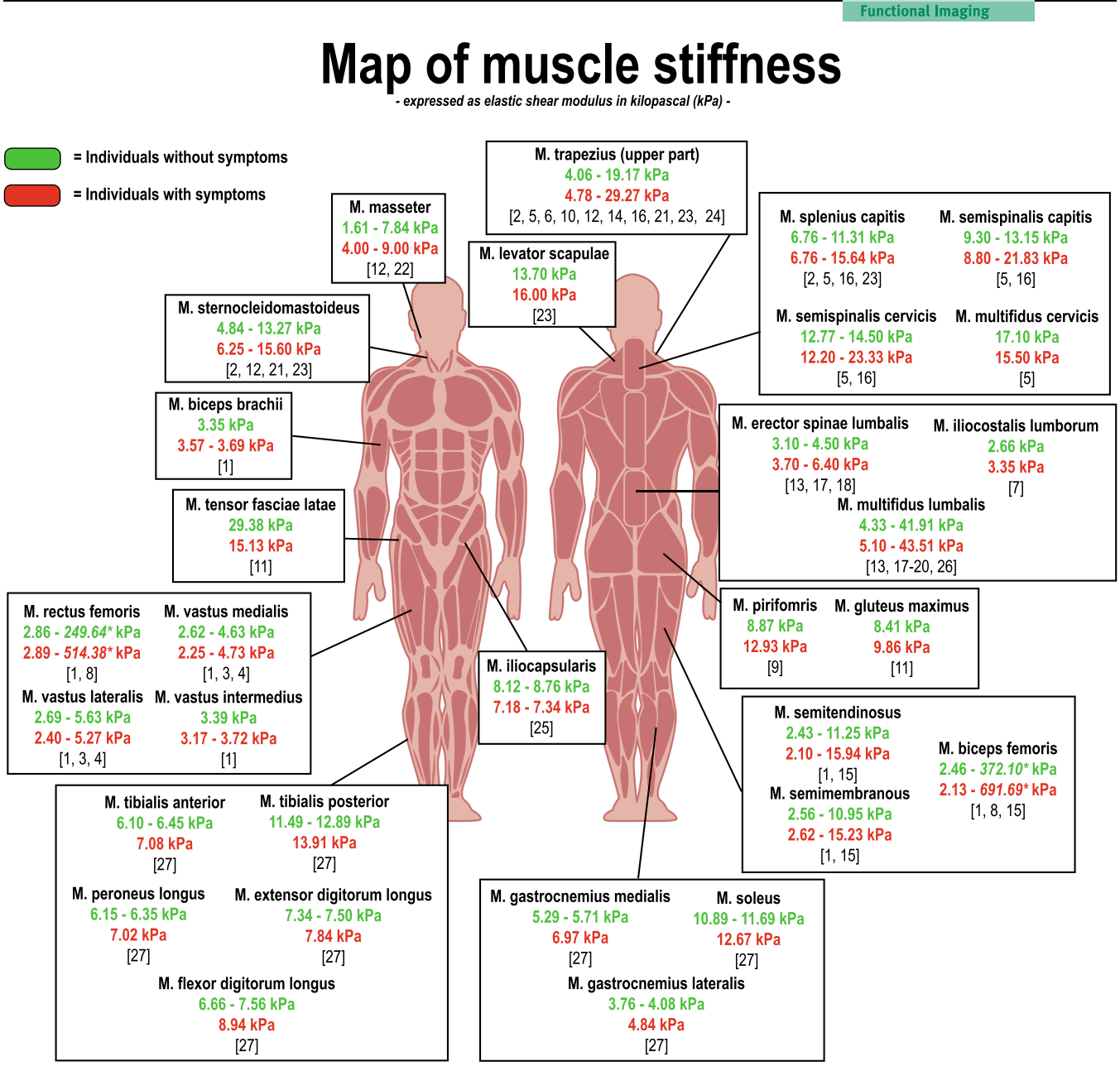

FIGURE 3 Body chart mapping the (range of) shear modulus of all examined muscles, for which absolute values were provided by the authors. *Confidence in these values is low due to them exceeding technical capabilities of elastography systems and probably the physiological range of muscle stiffness. Perhaps the values were reported in the wrong unit or the decimal point was incorrect. References: [1] (Alfuraih et al., 2020), [2] (Aljinović et al., 2020), [3] (Botanlioglu et al., 2013), [4] (Botanlıoğlu et al., 2019), [5] (Dieterich et al., 2020), [6] (Ertekin et al., 2021), [7] (Gao et al., 2020), [8] (Gökşen et al., 2021), [9] (Hashimoto et al., 2022), [10] (Hvedstrup et al., 2020), [11] (Klauser et al., 2022), [12] (Kolding et al., 2018), [13] (Koppenhaver et al., 2020), [14] (Leong et al., 2016), [15] (Li et al., 2021), [16] (Lin et al., 2022), [17] (Masaki et al., 2017), [18] (Masaki et al., 2019), [19] (Murillo et al., 2019), [20] (Pinto et al., 2022), [21] (Sedlackova et al., 2021), [22] (Takashima et al., 2017), [23] (Taş et al., 2018), [24] (Valera-Calero et al., 2021), [25] (Wang, Cui, et al., 2022), [26] (Wang, Liu, et al., 2022) and [27] (Zhang et al., 2022).

SWE system (Dieterich et al., 2020; Friede et al., 2020; Hashimoto et al., 2022; Klauser et al., 2022; Masaki et al., 2017; Masaki et al., 2019; Taş et al., 2018; Wolff et al., 2022). In six further reports (20.0%), the settings can be obtained from the published images (Botanlıoğlu et al., 2019; Itoigawa et al., 2022; Li et al., 2021; Lin et al., 2022; Sedlackova et al., 2021; Takashima et al., 2017). Solely five articles (16.6%) reported precisely how sufficient image quality was monitored and ascertained (Dieterich et al., 2020; Gao et al., 2020; Koppenhaver et al., 2020; Li et al., 2021; Zhang et al., 2022) and six articles (20.0%) provided imprecise information on quality assurance (Friede et al., 2020; Hashimoto et al., 2022; Klauser et al., 2022; Murillo et al., 2019; Valera-Calero et al., 2021; Wang, Liu, et al., 2022).

Thirteen studies (43.3%) included a reliability analysis (Alfuraih et al., 2020; Aljinović et al., 2020; Dieterich et al., 2020; Ertekin et al., 2021; Gao et al., 2020; Klauser et al., 2022; Koppenhaver et al., 2020;

Leong et al., 2016; Li et al., 2021; Masaki et al., 2017; Valera-Calero et al., 2021; Wang, Cui, et al., 2022; Zhang et al., 2022). Two studies (6.6%) conducted a fully computed and therefore reliable image analysis (Dieterich et al., 2020; Wolff et al., 2022). Twenty-six studies (86.6%) used repeated measurements while three studies (10.0%) used only a single repetition per measurement and subject for the statistical analysis (Masaki et al., 2019; Takashima et al., 2017; Wang, Liu, et al., 2022). The number of measurement repetitions was not reported by Hashimoto et al. (2022).

Regarding data analysis, two studies (6.6%) conducted a fully computed image analysis (Dieterich et al., 2020; Wolff et al., 2022), which does not distinguish group allocation; eight articles (26.6%) reported that the examiner conducting data processing was blinded to group allocation (Aljinović et al., 2020; Ertekin et al., 2021; Hvedstrup et al., 2020; Itoigawa et al., 2022; Kolding et al., 2018; Masaki et al., 2017; Masaki et al., 2019; Valera-Calero et al., 2021). In only seven studies (23.3%), a standardized location of the measurement box(es) was stated (Dieterich et al., 2020; Gao et al., 2020; Masaki et al., 2017; Masaki et al., 2019; Valera-Calero et al., 2021; Wang, Cui, et al., 2022; Zhang et al., 2022), while this remained unclear in seven articles (23.3%) (Friede et al., 2020; Hashimoto et al., 2022; Klauser et al., 2022; Leong et al., 2016; Pinto et al., 2022; Takashima et al., 2017; Wang, Liu, et al., 2022). The size of the measurement box was standardized in 17 studies (56.6%) (Alfuraih et al., 2020; Dieterich et al., 2020; Friede et al., 2020; Gao et al., 2020; Hashimoto et al., 2022; Klauser et al., 2022; Koppenhaver et al., 2020; Leong et al., 2016; Li et al., 2021; Lin et al., 2022; Masaki et al., 2017; Pinto et al., 2022; Sedlackova et al., 2021; Taş et al., 2018; Valera-Calero et al., 2021; Wang, Cui, et al., 2022; Zhang et al., 2022). Only in two studies (6.6%), the measurement box included the majority of the visible muscle (Friede et al., 2020; Gao et al., 2020). In eight reports (26.6%), the size of the measurement box was deemed sufficiently representative of the scanned muscle (Dieterich et al., 2020; Hashimoto et al., 2022; Hvedstrup et al., 2020; Klauser et al., 2022; Leong et al., 2016; Masaki et al., 2017; Masaki et al., 2019; Valera-Calero et al., 2021). In 21 studies (70.0%), the measurement boxes were much smaller than the visible muscle and of questionable representativeness. The exclusion of muscle fascia from the measurement box was clearly stated in eight articles (26.6%) (Aljinović et al., 2020; Friede et al., 2020; Gökşen et al., 2021; Klauser et al., 2022; Koppenhaver et al., 2020; Murillo et al., 2019; Takashima et al., 2017; Zhang et al., 2022). The example images of eight more studies (26.6%) suggest the avoidance of fascia in the measurements (Botanlioglu et al., 2013; Botanlıoğlu et al., 2019; Ertekin et al., 2021; Leong et al., 2016; Li et al., 2021; Lin et al., 2022; Sedlackova et al., 2021; Taş et al., 2018) (Figure S1).

4 | DISCUSSION

Current evidence on the stiffness of muscles in musculoskeletal pain syndromes is inconsistent. The very few studies of moderate-high quality suggest in part equal and in part higher stiffness of involved muscles in symptomatic individuals. The critical appraisal of the SWE methods identified several aspects that suggest a risk of bias. The reporting of the SWE methods was often incomplete.

Despite the perception of many patients, the evidence for objectively stiffer muscles in musculoskeletal pain is not clear. Most studies that examined more than one muscle report mixed results. The M. trapezius is clinically often described as tense and was the most-examined muscle in the included studies, with inconsistent findings both within the same and across different pain conditions. All four studies (13.3%) of moderate-high quality included trapezius measurements (Dieterich et al., 2020; Hvedstrup et al., 2020; Kolding et al., 2018; Leong et al., 2016). Only Leong et al. (2016) identified higher stiffness of the trapezius muscle in athletes with rotator cuff tendinopathy, but only in one of three examined shoulder positions. Dieterich et al. (2020) on women with chronic neck pain, Kolding et al. (2018) on tension-type headache and Hvedstrup et al. (2020) on migraine found equal trapezius stiffness. Five studies examined the lumbar multifidus muscle in individuals with low back pain. Three studies identified higher multifidus stiffness in the symptomatic individuals with low back pain (Koppenhaver et al., 2020; Masaki et al., 2017; Murillo et al., 2019). One study compared the stiffness of the multifidus muscle in individuals with ongoing low back pain to two different asymptomatic groups, one with a history of low back pain and one without. The results differed between the asymptomatic groups. The study found the lowest stiffness in the group without a history of back pain ($n = 19$), slightly higher stiffness in the symptomatic individuals ($n = 23$) and the highest stiffness of the multifidus muscle in asymptomatic individuals with a history of low back pain ($n = 16$) (Masaki et al., 2019). The last study reported no differences in multifidus stiffness between a symptomatic group with low back pain and asymptomatic individuals (Pinto et al., 2022).

Sample size may explain some of the inconsistent results. Two studies (6.6%) measured stiffness of the biceps femoris muscle in participants with knee osteoarthritis. The smaller study ($n = 40$) of Gökşen et al. (2021) found no statistic group difference while Li et al. (2021) ($n = 100$) demonstrated higher stiffness of the biceps femoris muscle. Similarly, Kolding et al. (2018) examined 17 participants with tension-type headache without demonstrating group differences, but Hvedstrup et al. (2020) included 48 participants with migraine and demonstrated increased stiffness in some but not all involved neck muscles. Apparently, large samples are required to demonstrate group differences in muscle stiffness.

Inconsistent findings of the mechanical muscle properties in pain conditions are supported by studies that examined the myoelectrical response of selective muscles to experimental pain. Individually different responses with more, equal and less activated muscles compared to the pain-free state have been documented using electromyography and SWE (Gizzi et al., 2015; Hug et al., 2014). A palpation study casts doubt on the clinical expectation that a painful body location can be detected by palpably increased muscle stiffness. Ninety-one patients with unilateral low back pain and 94 patients with unilateral neck pain were instructed to avoid any verbal or bodily response when they were palpated by two trained physicians to

identify the painful side. In low back pain, the painful side was correctly identified in 64.8%; and in neck pain in 58.8% (Maigne et al., 2012), which is only slightly better than chance. Possibly, stiffness and tenderness are often not clearly separated in clinical examinations.

Confidence in the results of the majority of the included studies is low due to their methodological quality. The reliability of the measurements was often unclear or referenced to studies performed under different conditions. The reliability of SWE measurements differs across muscles (Davis et al., 2019) and may depend on the size and positioning of the region of interest (Alfuraih et al., 2018). Inter-system reliability is not granted (Javed et al., 2022; Long et al., 2018; Mulabecirovic et al., 2016). The reliability of SWE measurements should be documented for the conditions of the study (Alfuraih et al., 2017).

Most authors declared measurements of the elastic shear modulus, but some reported very high values and/or no measures of recalculation, suggestive of Young's modulus. Following the recommendation to report shear wave speed (Davis et al., 2019) would prevent a wrong interpretation of stiffness measures in kPa.

The influence of the size of the measurement box has been discussed controversially (Alfuraih et al., 2017; Ateş et al., 2015; Gennisson et al., 2015; Kot et al., 2012; Wang et al., 2022c). Clearly, the more homogenous the stiffness of a muscle presents, the less problematic the size and the position of the measurement box. Many published example images support an inhomogeneous distribution of stiffness within muscles, in particular during activity (Davis et al., 2019; Ewertsen et al., 2018; Ferraioli et al., 2022; Gennisson et al., 2010; Gennisson et al., 2015; Lima et al., 2018). Representative measurements should therefore include a large area of the muscle. Third-party software for image analysis or custom-programmed software solutions enable standardized measures of representative muscle areas without the influence of the examiner (Dieterich et al., 2020; Doguet et al., 2020; Wolff et al., 2022).

Fascia and tendinous tissue are often stiffer than the respective muscle tissue. Therefore, the inclusion of fascial or tendinous tissue in the measurement box should be declared. Probably, structural boundaries, such as bones and fascia influence image quality and may produce artefacts within the elastogram (Davis et al., 2019; Lin et al., 2017; Săftoiu et al., 2019; Shiina et al., 2015). In current guidelines, the exclusion of fascia from the measurement box is not clearly recommended but likely enhances the comparability of studies that measure muscle stiffness.

Complementing the recommendations of Stiver et al. (2023), the following quality criteria are suggested. (i) Measurements of muscle stiffness should be performed blinded to the group status of the participant in a strictly standardized manner. Also, the analysis of the elastograms should be performed unaware of the group status or fully computed, that is, using exactly the same processing on all images. (ii) The scanning position and the position of the measurement box within the muscle should be standardized (Alfuraih et al., 2017, 2018). (iii) The measurement box should be as large as possible to include most of the visible muscle (Ewertsen et al., 2018; Lima et al., 2018; Ruby et al., 2019) and measurements should be repeated (Davis et al., 2019), at least thrice. (iv) Measures to control image quality and potential sources of artefacts must be pursued and documented to ascertain the maximal achievable quality for valid measurements. The inhomogeneous distribution of muscle stiffness may reflect true stiffness differences but may include also artefacts, especially at boundary conditions (Davis et al., 2019; Ferraioli et al., 2022; Lin et al., 2017).

The AXIS items that were most criticized refer to the target population and the sampling process. In several studies, the participants deemed not typical of the pathologic condition and sample sizes were likely not large enough to demonstrate significant differences between groups. With regard to published discussions of study limitations, authors rarely discussed the limitations of their measurement procedure. This may be interpreted as an insufficient awareness of biasing factors that may threaten the measurements' validity. Without the additional items specifying SWE methods, five studies would have been rated high quality according to the AXIS evaluation (Alfuraih et al., 2020; Dieterich et al., 2020; Hvedstrup et al., 2020; Klauser et al., 2022; Kolding et al., 2018).

5 | LIMITATIONS

The here presented work has several limitations. First, the systematic search was conducted in only two databases which carries some risk of missing eligible studies. However, we have supplemented our systematic by a hand search as well as additional searches using Google Scholar. Second, we included only studies that measured muscle stiffness using ultrasound SWE. The inclusion of other modalities, such as strain elastography, transient elastography (Shiina et al., 2015) or Myoton measurements (Lee et al., 2021) would have enabled a more complete map of the available evidence. We restricted our search to SWE, the currently most relevant method to quantify muscle stiffness, because each modality reflects a different aspect of the mechanical properties of a muscle, and the results of the different modalities may not be correlated (Lee et al., 2021; Shiina et al., 2015). Third, we included a variety of painful conditions, but muscles may react inconsistently across pathologies. We decided for this procedure to enable a first and general comparison of body regions and pathologies and to create a base for more specific questions (Miake-Lye et al., 2016; Schmucker et al., 2013; Snilstveit et al., 2016; White et al., 2020). Fourth, the identified studies include muscles in a variety of positions, in the active and the relaxed state. This variation influences the findings (Baumer et al., 2018; Dieterich et al., 2020). We do not assume that a consensus regarding examination positions is feasible; the variation reflects different clinical interests and muscle functions. Many studies stated to examine muscles in relaxation, but none confirmed muscle relaxation using electromyography. It is uncertain whether muscles were truly relaxed. Some studies described measures to foster muscle

relaxation (Li et al., 2021; Murillo et al., 2019), sometimes relaxation appeared questionable (Alfuraih et al., 2020). Fifth, we used the AXIS tool, which is designed for cross-sectional studies, on all studies regardless of the study design. We decided for this procedure because we used only the cross-sectional comparison of the baseline measurements of all included studies. Using a single checklist for quality control facilitated the comparison of the study quality.

6 | CONCLUSION

Existing evidence regarding the objective stiffness of muscles involved in musculoskeletal pain conditions is conflicting. The results of the four studies of moderate-high quality suggest that only few muscles stiffen with musculoskeletal pain. Methodological differences between studies and small sample sizes may explain many of the inconsistencies between findings. Methodological standards for SWE measurements on muscles are urgently required.

ACKNOWLEDGEMENTS

This study was supported by the Baden-Württemberg Ministry of Science, Research and Arts, Germany. The authors want to thank Maren Lickert for her support in the initial steps of this project. Open Access funding enabled and organized by Projekt DEAL.

CONFLICT OF INTEREST STATEMENT

The authors declare no conflict of interest. None of the authors was involved in the data extraction and quality assessment of studies to which they contributed or which examined a similar aim as their own study.

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

ORCID

Andreas Haueise http://orcid.org/0000-0002-5201-3968
Guillaume Le Sant http://orcid.org/0000-0001-7614-6389
Angelika Eisele-Metzger http://orcid.org/0000-0002-7405-4584
Angela V. Dieterich http://orcid.org/0000-0002-8522-8746

REFERENCES

Akagi, R. & Kusama, S. (2015) Comparison between neck and shoulder stiffness determined by shear wave ultrasound elastography and a muscle hardness meter. *Ultrasound in Medicine & Biology*, 41, 2266–2271.

Alfuraih, A.M., O'Connor, P., Hensor, E., Tan, A.L., Emery, P. & Wakefield, R.J. (2018) The effect of unit, depth, and probe load on the reliability of muscle shear wave elastography: Variables affecting reliability of SWE. *Journal of Clinical Ultrasound*, 46, 108–115.

Alfuraih, A.M., O'Connor, P., Tan, A.L., Hensor, E., Emery, P. & Wakefield, R.J. (2017) An investigation into the variability between different shear wave elastography systems in muscle. *Medical Ultrasonography*, 19, 392–400.

Alfuraih, A.M., Tan, A.L., O'Connor, P., Emery, P. & Wakefield, R.J. (2020) Muscle stiffness in rheumatoid arthritis is not altered or associated with muscle weakness: A shear wave elastography study. *Modern Rheumatology*, 30, 617–625.

Aljinović, J., Barišić, I., Poljičanin, A., Kuzmičić, S., Vukojević, K., Gugić Bokun, D. et al. (2020) Can measuring passive neck muscle stiffness in whiplash injury patients help detect false whiplash claims? *Wiener Klinische Wochenschrift*, 132, 506–514.

Ateş, F., Hug, F., Bouillard, K., Jubeau, M., Frappart, T., Couade, M. et al. (2015) Muscle shear elastic modulus is linearly related to muscle torque over the entire range of isometric contraction intensity. *Journal of Electromyography and Kinesiology*, 25, 703–708.

Barr, R.G., Nakashima, K., Amy, D., Cosgrove, D., Farrokh, A., Schafer, F. et al. (2015) WFUMB guidelines and recommendations for clinical use of ultrasound elastography: Part 2: Breast. *Ultrasound in Medicine & Biology*, 41, 1148–1160.

Baumer, T.G., Dischler, J., Davis, L., Labyed, Y., Siegal, D.S., van Holsbeeck, M. et al. (2018) Effects of age and pathology on shear wave speed of the human rotator cuff. *Journal of Orthopaedic Research*, 36, 282–288.

Bernabei, M., Lee, S.S.M., Perreault, E.J. & Sandercock, T.G. (2020) Shear wave velocity is sensitive to changes in muscle stiffness that occur independently from changes in force. *Journal of Applied Physiology*, 128, 8–16.

Bota, S., Sporea, I., Sirli, R., Popescu, A., Dănilă, M. & Sendroiu, M. (2011) Factors that influence the correlation of acoustic radiation force impulse (ARFI), elastography with liver fibrosis. *Medical ultrasonography*, 13, 135–140.

Botanlioglu, H., Kantarci, F., Kaynak, G., Unal, Y., Ertan, S., Aydingoz, O. et al. (2013) Shear wave elastography properties of vastus lateralis and vastus medialis obliquus muscles in normal subjects and female patients with patellofemoral pain syndrome. *Skeletal Radiology*, 42, 659–666.

Botanlıoğlu, H. (2019) The effect of shear-wave elastography on functional results and muscle stiffness in patients undergoing non-selective and selective open kinetic chain exercises. *Turkish Journal of Physical Medicine and Rehabilitation*, 65, 40–50.

Brandenburg, J.E., Eby, S.F., Song, P., Zhao, H., Brault, J.S., Chen, S. et al. (2014) Ultrasound elastography: The new frontier in direct measurement of muscle stiffness. *Archives of Physical Medicine and Rehabilitation*, 95, 2207–2219.

Davis, L.C., Baumer, T.G., Bey, M.J. & Holsbeeck, M.V. (2019) Clinical utilization of shear wave elastography in the musculoskeletal system. *Ultrasonography (Seoul, Korea)*, 38, 2–12.

Dieterich, A.V., Yavuz, U.Ş., Petzke, F., Nordez, A. & Falla, D. (2020) Neck muscle stiffness measured with shear wave elastography in women with chronic nonspecific neck pain. *Journal of Orthopaedic & Sports Physical Therapy*, 50, 179–188.

Doguet, V., Hug, F., Nordez, A. & Savery, D., Logiciel ElastoGUI (2020) *UR 4334—Laboratoire Motricité, Interactions, Performance*. Nantes: University of Nantes.

Dor, A. & Kalichman, L. (2017) A myofascial component of pain in knee osteoarthritis. *Journal of Bodywork and Movement Therapies*, 21, 642–647.

Downes, M.J., Brennan, M.L., Williams, H.C. & Dean, R.S. (2016) Development of a critical appraisal tool to assess the quality of cross-sectional studies (AXIS). *BMJ Open*, 6, e011458.

Eby, S.F., Song, P., Chen, S., Chen, Q., Greenleaf, J.F. & An, K.-N. (2013) Validation of shear wave elastography in skeletal muscle. *Journal of Biomechanics*, 46, 2381–2387.

Ertekin, E., Kasar, Z. & Turkdogan, F. (2021) Is early diagnosis of myofascial pain syndrome possible with the detection of latent trigger points by shear wave elastography? *Polish Journal of Radiology*, 86, 425–431.

Ewertsen, C., Carlsen, J., Perveez, M. & Schytz, H. (2018) Reference values for shear wave elastography of neck and shoulder muscles in healthy individuals. *Ultrasound International Open*, 4, E23–E29.

Ewertsen, C., Carlsen, J.F., Christiansen, I.R., Jensen, J.A. & Nielsen, M.B. (2016) Evaluation of healthy muscle tissue by strain and shear wave elastography—Dependency on depth and ROI position in relation to underlying bone. *Ultrasonics*, 71, 127–133.

Ferraioli, G., Barr, R.G., Farrokh, A., Radzina, M., Cui, X.W., Dong, Y. et al. (2022) How to perform shear wave elastography. Part II. *Medical ultrasonography*, 24, 196–210.

Friede, M.C., Klauser, A., Fink, C. & Csapo, R. (2020) Stiffness of the iliotibial band and associated muscles in runner's knee: Assessing the effects of physiotherapy through ultrasound shear wave elastography. *Physical Therapy in Sport*, 45, 126–134.

Gao, J., Caldwell, J., McLin, K., Zhang, M. & Park, D. (2020) Ultrasound shear wave elastography to assess osteopathic manipulative treatment on the Iliocostalis lumborum muscle: A feasibility study. *Journal of Ultrasound in Medicine*, 39, 157–164.

Gennisson, J.-L., Deffieux, T., Macé, E., Montaldo, G., Fink, M. & Tanter, M. (2010) Viscoelastic and anisotropic mechanical properties of in vivo muscle tissue assessed by supersonic shear imaging. *Ultrasound in Medicine & Biology*, 36, 789–801.

Gennisson, J.-L., Muller, M., Gabor, P., Frydman, R., Musset, D., Tanter, M. et al. (2015) Quantification of elasticity changes in the myometrium during labor using supersonic shear imaging: A feasibility study. *Ultrasonics*, 56, 183–188.

Gizzi, L., Muceli, S., Petzke, F. & Falla, D. (2015) Experimental muscle pain impairs the synergistic modular control of neck muscles. *PLoS One*, 10, e0137844.

Gökşen, A., Can, F. & Yilmaz, S. (2021) Does muscle stiffness predict early-onset knee osteoarthritis? *Journal of Surgery and Medicine*, 5, 1126–1130.

Haddaway, N.R., Page, M.J., Pritchard, C.C. & McGuinness, L.A. (2022) PRISMA2020: An R package and Shiny app for producing PRISMA 2020-compliant flow diagrams, with interactivity for optimised digital transparency and open synthesis. *Campbell Systematic Reviews*, 18, e1230.

Hashimoto, H., Shibahara, M., Mizuta, Y. & Mizushima, K. (2022) Measurement of piriformis muscle stiffness in athletes with lower back pain and piriformis syndrome: A shear-wave elastography study with consideration of the onset of piriformis syndrome. *Journal of Physical Education and Sport (JPES)*, 22, 1748–1760.

Hug, F., Hodges, P.W., Hoorn, W. & Tucker, K. (2014) Between-muscle differences in the adaptation to experimental pain. *Journal of Applied Physiology*, 117, 1132–1140.

Hug, F., Tucker, K., Gennisson, J.-L., Tanter, M. & Nordez, A. (2015) Elastography for muscle biomechanics: Toward the estimation of individual muscle force. *Exercise and Sport Sciences Reviews*, 43, 125–133.

Hvedstrup, J., Kolding, L.T., Ashina, M. & Schytz, H.W. (2020) Increased neck muscle stiffness in migraine patients with ictal neck pain: A shear wave elastography study. *Cephalalgia*, 40, 565–574.

Itoigawa, Y., Koga, A., Morikawa, D., Kubota, A., Uehara, H. & Maruyama, Y. et al. (2022) Posterior shoulder stiffness was associated with shoulder pain during throwing in college baseball players: Assessment of shear wave elastography. *European Journal of Orthopaedic Surgery & Traumatology: Orthopedie Traumatologie*, 33(4), 1237–1244.

Javed, H., Oyibo, S.O. & Alfuraih, A.M. (2022) Variability, validity and operator reliability of three ultrasound systems for measuring tissue stiffness: A phantom study. *Cureus*, 14, 31731.

Klauser, A.S., de-Koekkoek, F., Schwabl, C., Fink, C., Friede, M. & Csapo, R. (2022) Intraobserver assessment of shear wave elastography in tensor fasciae latae and gluteus maximus muscle: The importance of the hip abductor muscles in runners knee compared to healthy controls. *Journal of Clinical Medicine*, 11, 11.

Kolding, L.T., Do, T.P., Ewertsen, C. & Schytz, H.W. (2018) Muscle stiffness in tension-type headache patients with pericranial tenderness. *Cephalalgia Reports*, 1, 251581631876029.

Koppenhaver, S., Gaffney, E., Oates, A., Eberle, L., Young, B., Hebert, J. et al. (2020) Lumbar muscle stiffness is different in individuals with low back pain than asymptomatic controls and is associated with pain and disability, but not common physical examination findings. *Musculoskeletal Science & Practice*, 45, 102078.

Kot, B.C.W., Zhang, Z.J., Lee, A.W.C., Leung, V.Y.F. & Fu, S.N. (2012) Elastic modulus of muscle and tendon with shear wave ultrasound elastography: Variations with different technical settings. *PLoS One*, 7, e44348.

Lee, Y., Kim, M. & Lee, H. (2021) The measurement of stiffness for major muscles with shear wave elastography and myoton: A quantitative analysis study. *Diagnostics (Basel, Switzerland)*, 11(3), 524.

Leong, H.T., Hug, F. & Fu, S.N. (2016) Increased upper trapezius muscle stiffness in overhead athletes with rotator cuff tendinopathy. *PLoS One*, 11, e0155187.

Li, F., Wang, Z.-Y., Zhang, Z.-J., Shen, S.-H., Guo, J.-Y., Guo, Y.-X. et al. (2021) In hamstring muscles of patients with knee osteoarthritis an increased ultrasound shear modulus indicates a permanently elevated muscle tonus. *Frontiers in Physiology*, 12, 752455.

Lima, K.M.M.E., Costa Júnior, J.F.S., Pereira, W. & Oliveira, L.F. (2018) Assessment of the mechanical properties of the muscle-tendon unit by supersonic shear wave imaging elastography: A review. *Ultrasonography (Seoul, Korea)*, 37, 3–15.

Lin, C.-Y., Chen, P.-Y., Shau, Y.-W. & Wang, C.-L. (2017) An artifact in supersonic shear wave elastography. *Ultrasound in Medicine & Biology*, 43, 517–530.

Lin, L.-Z., Yu, Y.-N., Fan, J.-C., Guo, P.-W., Xia, C.-F., Geng, X. et al. (2022) Increased stiffness of the superficial cervical extensor muscles in patients with cervicogenic headache: A study using shear wave elastography. *Frontiers in Neurology*, 13, 874643.

Long, Z., Tradup, D.J., Song, P., Stekel, S.F., Chen, S., Glazebrook, K.N. et al. (2018) Clinical acceptance testing and scanner comparison of ultrasound shear wave elastography. *Journal of Applied Clinical Medical Physics*, 19, 336–342.

Maigne, J.Y., Cornelis, P. & Chatellier, G. (2012) Lower back pain and neck pain: Is it possible to identify the painful side by palpation only? *Annals of Physical and Rehabilitation Medicine*, 55, 103–111.

Masaki, M., Aoyama, T., Murakami, T., Yanase, K., Ji, X., Tateuchi, H. et al. (2017) Association of low back pain with muscle stiffness and muscle mass of the lumbar back muscles, and sagittal spinal alignment in young and middle-aged medical workers. *Clinical Biomechanics*, 49, 128–133.

Masaki, M., Ikezoe, T., Yanase, K., Ji, X., Umehara, J., Aoyama, J. et al. (2019) Association of pain history and current pain with sagittal spinal alignment and muscle stiffness and muscle mass of the back muscles in middle-aged and elderly women. *Clinical Spine Surgery: A Spine Publication*, 32, E346–E352.

Miake-Lye, I.M., Hempel, S., Shanman, R. & Shekelle, P.G. (2016) What is an evidence map? A systematic review of published evidence maps and their definitions, methods, and products. *Systematic Reviews*, 5, 28.

Mulabecirovic, A., Vesterhus, M., Gilja, O.H. & Havre, R.F. (2016) In vitro comparison of five different elastography systems for clinical applications, using strain and shear wave technology. *Ultrasound in Medicine & Biology*, 42, 2572–2588.

Murillo, C., Falla, D., Rushton, A., Sanderson, A. & Heneghan, N.R. (2019) Shear wave elastography investigation of multifidus stiffness in individuals with low back pain. *Journal of Electromyography and Kinesiology*, 47, 19–24.

Olchowy, A., Wieckiewicz, M., Winocur, E., Dominiak, M., Dekkers, I., Łasecki, M. et al. (2020) Great potential of ultrasound elastography for the assessment of the masseter muscle in patients with temporomandibular disorders. A systematic review. *Dento Maxillo Facial Radiology*, 49, 20200024.

Page, M.J., McKenzie, J.E., Bossuyt, P.M., Boutron, I., Hoffmann, T.C., Mulrow, C.D. et al. (2021) The PRISMA 2020 statement: An updated guideline for reporting systematic reviews. *BMJ (Clinical Research ed.)*, 372, n71.

Pinto, S.M., Cheung, J.P.Y., Samartzis, D., Karppinen, J., Zheng, Y., Pang, M.Y.C. et al. (2022) Are morphometric and biomechanical characteristics of lumbar multifidus related to pain intensity or disability in people with chronic low back pain after considering psychological factors or insomnia? *Frontiers in Psychiatry*, 13, 809891.

Proulx, L., Brizzolara, K., Thompson, M., Wang-Price, S., Rodriguez, P. & Koppenhaver, S. (2023) Women with chronic pelvic pain demonstrate increased lumbopelvic muscle stiffness compared to asymptomatic controls. *Journal of Women's Health*, 32, 239–247.

Raynaud, M., Zhang, H., Louis, K., Goutaudier, V., Wang, J., Dubourg, Q. et al. (2021) COVID-19-related medical research: A meta-research and critical appraisal. *BMC Medical Research Methodology*, 21, 1.

Rominger, M., Kälin, P., Mastalerz, M., Martini, K., Klingmüller, V., Sanabria, S. et al. (2018) Influencing factors of 2D shear wave elastography of the muscle—An ex vivo animal study. *Ultrasound International Open*, 4, E54–E60.

Ruby, L., Mutschler, T., Martini, K., Klingmüller, V., Frauenfelder, T., Rominger, M.B. et al. (2019) Which confounders have the largest impact in shear wave elastography of muscle and how can they be minimized? An elasticity phantom, ex vivo porcine muscle and volunteer study using a commercially available system. *Ultrasound in Medicine & Biology*, 45, 2591–2611.

Săftoiu, A., Gilja, O.H., Sidhu, P.S., Dietrich, C.F., Cantisani, V., Amy, D. et al. (2019) The EFSUMB guidelines and recommendations for the clinical practice of elastography in non-hepatic applications: Update 2018. *Ultraschall in der Medizin—European Journal of Ultrasound*, 40, 425–453.

Sawada, T., Okawara, H., Nakashima, D., Iwabuchi, S., Matsumoto, M., Nakamura, M. et al. (2020) Reliability of trapezius muscle hardness measurement: A comparison between portable muscle hardness meter and ultrasound strain elastography. *Sensors*, 20, 7200.

Schmucker, C., Motschall, E., Antes, G. & Meerpohl, J.J. (2013) Methoden des evidence mappings. eine systematische Ubersichtsarbeit. *Bundesgesundheitsblatt, Gesundheitsforschung, Gesundheitsschutz*, 56, 1390–1397.

Sedlackova, Z., Vita, M., Herman, J., Furst, T., Dornak, T. & Herman, M. (2021) Elasticity of neck muscles in cervicogenic headache. *Biomedical papers of the Medical Faculty of the University Palacky, Olomouc, Czechoslovakia*, 166(3), 328–333.

Shiina, T., Nightingale, K.R., Palmeri, M.L., Hall, T.J., Bamber, J.C., Barr, R.G. et al. (2015) WFUMB guidelines and recommendations for clinical use of ultrasound elastography: Part 1: Basic principles and terminology. *Ultrasound in Medicine & Biology*, 41, 1126–1147.

Snilstveit, B., Vojtkova, M., Bhavsar, A., Stevenson, J. & Gaarder, M. (2016) Evidence Gap Maps: A tool for promoting evidence informed policy and strategic research agendas. *Journal of Clinical Epidemiology*, 79, 120–129.

Stiver, M.L., Mirjalili, S.A. & Agur, A.M.R. (2023) Measuring shear wave velocity in adult skeletal muscle with ultrasound 2-D shear wave elastography: A scoping review. *Ultrasound in Medicine & Biology*, 49(6), 1353–1362.

Takasawa, E., Yamamoto, A., Kobayashi, T., Tajika, T., Shitara, H., Ichinose, T. et al. (2015) Characteristics of neck and shoulder pain in the Japanese general population. *Journal of Orthopaedic Science*, 20, 403–409.

Takashima, M., Arai, Y., Kawamura, A., Hayashi, T. & Takagi, R. (2017) Quantitative evaluation of masseter muscle stiffness in patients with temporomandibular disorders using shear wave elastography. *Journal of Prosthodontic Research*, 61, 432–438.

Taş, S., Korkusuz, F. & Erden, Z. (2018) Neck muscle stiffness in participants with and without chronic neck pain: A shear-wave elastography study. *Journal of Manipulative and Physiological Therapeutics*, 41, 580–588.

Valera-Calero, J.A., Sánchez-Jorge, S., Buffet-García, J., Varol, U., Gallego-Sendarrubias, G.M. & Álvarez-González, J. (2021) Is shear-wave elastography a clinical severity indicator of myofascial pain syndrome? An observational study. *Journal of Clinical Medicine*, 10, 2895.

Wang, C., Cui, L., Xu, Y., Xue, H., Sun, Y., Fu, Q. et al. (2022) Ultrasound and shear-wave elastography characteristics of the iliocapsularis muscle in pincer-type femoroacetabular impingement. *PM & R: The Journal of Injury, Function, and Rehabilitation*, 14, 1360–1367.

Wang, M., Liu, J., Meng, L., Fu, W., Gao, J., Ma, R. et al. (2022) Mapping of lumbar multifidus stiffness quantification in ankylosing spondylitis with shear-wave elastography. *BMC Musculoskeletal Disorders*, 23, 917.

Wang, X., Zhu, J., Gao, J., Hu, Y., Liu, Y., Li, W. et al. (2022c) Assessment of ultrasound shear wave elastography within muscles using different region of interest sizes, manufacturers, probes and acquisition angles: An ex vivo study. *Quantitative Imaging in Medicine and Surgery*, 12, 3227–3237.

White, H., Albers, B., Gaarder, M., Kornør, H., Littell, J., Marshall, Z. et al. (2020) Guidance for producing a Campbell evidence and gap map. *Campbell Systematic Reviews*, 16, e1125.

World Health Organization. (2022) *Musculoskeletal health*. Available at: https://www.who.int/news-room/fact-sheets/detail/musculoskeletal-conditions

Wolff, W.L., Heinemann, C.M. & Lipps, D.B. (2022) The influence of idiopathic chronic neck pain on upper trapezius and sternocleidomastoid muscle activity and elasticity during functional reaching: A cross-sectional study. *Journal of Biomechanics*, 141, 111223.

Yavuz, A., Bora, A., Bulut, M.D., Batur, A., Milanlioglu, A., Göya, C. et al. (2015) Acoustic radiation force impulse (ARFI) elastography quantification of muscle stiffness over a course of gradual isometric contractions: A preliminary study. *Medical Ultrasonography*, 17, 49–57.

Zhang, H., Peng, W., Qin, C., Miao, Y., Zhou, F., Ma, Y. et al. (2022) Lower leg muscle stiffness on two-dimensional shear wave elastography in subjects with medial tibial stress syndrome. *Journal of Ultrasound in Medicine*, 41, 1633–1642.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Haueise, A., Le Sant, G., Eisele-Metzger, A. & Dieterich, A. V. (2024) Is musculoskeletal pain associated with increased muscle stiffness? Evidence map and critical appraisal of muscle measurements using shear wave elastography. *Clinical Physiology and Functional Imaging*, 44, 187–204. https://doi.org/10.1111/cpf.12870