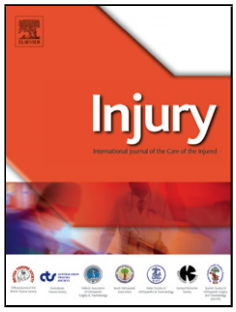
Accepted Manuscript



Title: Evidence for cervical muscle morphometric changes on magnetic resonance images after whiplash: a systematic review and meta-analysis

Authors: Daniel S. Owers, Diana M. Perriman, Paul N. Smith, Teresa Neeman, Alexandra L. Webb

PII: S0020-1383(17)30854-9
DOI: https://doi.org/10.1016/j.injury.2017.12.001
Reference: JINJ 7521

To appear in: *Injury, Int. J. Care Injured*

Accepted date: 3-12-2017

Please cite this article as: Owers Daniel S, Perriman Diana M, Smith Paul N, Neeman Teresa, Webb Alexandra L.Evidence for cervical muscle morphometric changes on magnetic resonance images after whiplash: a systematic review and meta-analysis.*Injury* https://doi.org/10.1016/j.injury.2017.12.001

Title: Evidence for cervical muscle morphometric changes on magnetic resonance images after whiplash: a systematic review and meta-analysis

Authors: Daniel S. Owers MBBS[1], Diana M. Perriman PhD[2], Paul N. Smith MBBS[2], Teresa Neeman PhD[3] and Alexandra L. Webb PhD[4*]

[1]Australian Capital Territory Health, Canberra Hospital, Woden, ACT, 2605

[2]Trauma and Orthopaedic Research Unit, Canberra Hospital, Woden, ACT, 2605

[3]Statistical Consulting Unit, Australian National University, Canberra, ACT, 2601

[4]Medical School, College of Medicine, Biology & Environment, Australian National University, Canberra, ACT, 2601

*Corresponding author: Dr Alexandra L. Webb, Medical School, Florey Building, 54 Mills Road, Australian National University, Canberra, ACT 2601, Australia. Phone: 02 6125 1970. E-mail: alexandra.webb@anu.edu.au

Abstract:

Introduction: Morphometric changes to cervical musculature in whiplash associated disorder have been reported in several studies with varying results. However, the evidence is not clear because only a limited number of cohorts have been studied and one cohort has been reported in multiple publications. The aim of this study was to assess the evidence for cervical muscle morphometric changes on magnetic resonance (MR) images after whiplash using a systematic review with meta-analysis.

Materials and Methods: PubMed, MEDLINE and Cochrane Library were searched without language restriction using combinations of the MeSH terms "muscles", "whiplash injuries", and "magnetic resonance imaging". Studies of acute and chronic whiplash were included if they compared whiplash and control cervical spine muscle morphometry measurements from MR images. The search identified 380 studies. After screening, eight studies describing five cohorts (one acute, three chronic, one both acute and chronic) met the inclusion criteria. Participant characteristics and outcome measures were extracted using a standard extraction format. Quality of eligible studies was assessed using the Newcastle-Ottawa Scale. Muscle cross-sectional area (CSA) and fat infiltrate (MFI) for acute and chronic whiplash cohorts were compared using mean difference and 95% confidence intervals. Meta-analysis models were created when data from more than two eligible cohorts was available, using inverse-variance random-effects models (RevMan5 version 5.3.5).

Results: Quality assessment was uniformly good but only two studies blinded the assessor. Analysis of the acute cohorts revealed no consensus with respect to CSA. MFI was not measured in the acute cohorts. Analysis of the chronic cohorts revealed CSA is probably increased in some muscles after whiplash but there is insufficient evidence to confirm whether MFI is also increased. Because the available data were limited, meta-analyses of only multifidus were performed. In chronic whiplash multifidus CSA was significantly increased at C5 ($Z=3.51$, $p<0.01$) and C6 ($Z=2.66$, $p<0.01$); and MFI was significantly increased at C7 only ($Z=2.52$, $p<0.01$) but the heterogeneity was unacceptably high ($I^2=83\%$).

Conclusions: The strength of the evidence for cervical muscle morphometric changes on MR images after whiplash is inconsistent for CSA and MFI. Future study

designs should be standardised with quantification of three-dimensional muscle morphometry.

Key Words:

Whiplash injuries, magnetic resonance imaging, cervical spine, muscle, muscle fat infiltrate, cross-sectional area

INTRODUCTION

Neck pain and disability from whiplash following motor vehicle trauma (MVT) is one of the most common debilitating injuries in the developed world [1, 2]. The reported incidence of whiplash associated disorders (WAD) after MVT is at least 300 per 100,000 in western countries with a consistent rising trend [3-6]. In the United States (US) alone, it is estimated that 3 million new cases of whiplash are reported every year [7]. Whilst the majority of cases recover, up to 50% develop chronic symptoms for which the efficacy of rehabilitation is variable [2, 8]. Whiplash imposes a significant economic burden on health-care systems with estimated annual costs totalling more than $29 billion spent on injuries and litigation in the US and €1.6 billion in the United Kingdom [9-11].

There has been continued debate about whether WAD is attributable to a defined pathoanatomical entity or to psychological or cultural factors [12-15]. It has been suggested that compensation seeking is associated with complaints of persistent pain after MVT and that some patients amplify their symptoms for financial gain [16]. The introduction of a no-fault insurance system in Canada which removed payments for pain and suffering resulted in a 43 percent reduction in claims by men and 15 percent for women [17]. In addition, a similar analysis in Australia found that outcome scores improved significantly when no-fault insurance was introduced [18]. However, more recent studies have proposed that adverse pain outcomes following MVT are not unique to litigious individuals and in fact are common among non-litigious individuals who are not engaged in compensation seeking [19-21]. Although no single pathognomonic entity has been identified in the cervical spine following

whiplash, advances in imaging technologies have led to reports of structural changes affecting the ligaments and muscles of the neck [22-24]. The evidence for signs of ligamentous damage on magnetic resonance (MR) imaging has been investigated in a systematic review with meta-analysis [25]. The results suggested that no significant differences exist between whiplash and control subjects for either alar or transverse ligaments. The evidence for changes in muscle morphology on MR images is not clear.

The MR measures that have been used to quantify the morphology of the cervical spine musculature after whiplash include cross-sectional area and muscle fat infiltrate [26-32]. An increase in cross-sectional area is thought to be due to an increase in injury-induced muscle fat infiltrate, making muscle fat infiltrate a potentially more robust marker for WAD than cross-sectional area [33, 34]. However, the cross-sectional area data is conflicting [28-32] and there is very little muscle fat infiltrate data [26, 27, 29, 30] at this point in time. Further, the cohorts are small but by combining them it is possible to ascertain whether the evidence supports the use of muscle fat infiltrate and/or cross-sectional area as a marker for WAD. If sensitive, these markers could potentially enable more precise rehabilitation strategies.

There has been one recent systematic review of muscle morphologic changes in chronic neck pain patients including WAD [35]. The authors concluded that there is some evidence for morphological changes in deep and higher cervical level muscles in chronic WAD with larger cross-sectional area measurements because of increased fatty infiltrate. In contrast, they concluded that idiopathic neck pain patients

have decreased cross-sectional area in most muscles because of disuse atrophy. However, this review had several limitations. First, rather than dissecting and comparing the actual study data, the review simply summarized the overall message from each of the studies. Second, both controlled and uncontrolled studies were included thereby allowing non-normalised results to be incorporated. Third, both MR and ultrasound imaging modalities were included. Finally, five studies by Elliott et al. [27-29, 34, 36] which reported data from the same cohort were presented as discrete studies thereby inflating their influence on the overall review. Therefore, the results of this review may not provide a clear picture of whether MR measurements of cervical muscle morphology are different in WAD compared to controls.

Therefore, the aim of this systematic review and meta-analysis was to rigorously assess the evidence for the presence of MR morphological changes (cross-sectional area and muscle fat infiltrate) in muscle after whiplash, and whether they represent a consistent marker which discriminates between WAD and control participants. The ability to be able to confidently identify WAD will have significant impact on diagnosis and the recognition of effective and non-effective management strategies. The study question was: in MR muscle measurement studies of acute and chronic WAD, does a systematic review and meta-analysis demonstrate evidence for increased cross-sectional area and/or muscle fat infiltrate in the cervical spine compared to non-WAD controls?

MATERIALS AND METHODS

The review protocol was specified in advance and registered on [Text hidden to ensure anonymity]. The review was conducted and reported according to the PRISMA statement [37].

Search Strategy and Study Selection

The PubMed (1974–December 2016), MEDLINE (1946– December 2016) and Cochrane Library (to December 2016) databases were searched without language restriction. Additional studies were identified through hand searching and reference lists. Our search term criteria consisted of MeSH headings linked by the Boolean term “AND” (("Muscles"[MeSH]) AND "Whiplash Injuries"[MeSH]) AND "Magnetic Resonance Imaging"[MeSH] (Appendix 1).

Studies of acute and chronic WAD were included if they compared WAD and control cervical spine muscle morphometry measurements from MR images. Exclusion criteria included: animal studies, participants <18 years, whiplash not caused by MVT and, patients with non-traumatic neck pain in the control cohort (Figure 1). The primary outcome of interest was measurement(s) of muscle morphometry.

Two authors [Text hidden to ensure anonymity] independently screened studies by title and abstract and selected studies that met the inclusion criteria. If the abstract was unavailable or further investigation was required, a full text examination was undertaken to ensure the study met the inclusion criteria. Data extraction was

undertaken by two authors [Text hidden to ensure anonymity] independently using a standard extraction format adapted from the Cochrane Haematological Malignancies Group [38]. This form facilitated systematic data extraction for accurate comparison between reviewers. Instruction and examples of data extraction forms can be found within the Cochrane Collaboration website [39]. Participant characteristics and outcome measures were recorded. All discrepancies in study selection and data extraction were resolved by discussion between three authors [Text hidden to ensure anonymity]. [Text hidden to ensure anonymity] all had previous experience conducting reviews and all negotiations were equitable.

Assessment of Study Quality

Two authors [Text hidden to ensure anonymity] independently assessed the quality of the included studies using the Newcastle-Ottawa scale (NOS) [40]. This scale was developed for quality assessment of non-randomised studies to be used in a systematic review. The NOS assigns up to a maximum of nine points for the least risk of bias in three domains: 1) selection of study groups (four points); 2) comparability of groups (two points); and 3) ascertainment of exposure and outcomes (three points). Inconsistency of scoring between reviewers was resolved by consensus. Quality assessment was undertaken to identify any studies for exclusion due to unacceptable bias. Unacceptable bias was defined as a zero score in any of the three domains.

Data Extraction and Analysis

Data Extraction

Data were reported separately for acute (<3 months since MVT) and chronic (>3 months since MVT) cohorts. These timeframes were chosen as they reflect current guidelines for the management of WAD [41]. Where there was insufficient data in the studies to enable extraction, the study authors were contacted to provide raw data.

Cross-sectional area and muscle fat infiltrate data for each muscle or muscle group were extracted for each cohort. Where left and right sided muscle data were provided the average was calculated. All cross-sectional area data were presented as millimetres squared for comparison. All muscle fat infiltrate data were presented as fat:muscle ratio except Elliott et al. [34] who reported estimated marginal means in mm^2 for cross-sectional area with and without fat, adjusted for BMI. We therefore calculated the muscle fat infiltrate by subtracting the mean muscle cross-sectional area (rmCSA) from the mean total cross-sectional area (rCSA) and recorded the resultant mean differences between the WAD and control muscle fat infiltrate with calculated 95% confidence intervals (CI) of the differences. A detailed explanation with formulae is shown in Appendix 2.

Cross-sectional area and muscle fat infiltrate data from Karlsson et al. [30] were transformed prior to extraction because mean and standard deviation (SD) for mild and severe WAD groups were presented separately. We therefore combined the mean and SD for the overall WAD group in order to calculate the mean differences (WAD minus control) and 95%CI of the differences. The formulae used are shown in Appendix 2.

Data Analysis

Differences between WAD and control data were calculated and presented as mean differences (WAD minus control) and 95%CI of the differences. The direction of the differences in WAD compared to control for cross-sectional area and muscle fat infiltrate were expressed as an increase, decrease or no change. These were based on the 95%CI i.e. if the 95%CI was positive and did not cross zero the difference was reported as an increase and *vice versa*; whereas if it crossed zero it was reported as no difference.

Meta-analysis models were created for multifidus in chronic WAD only because this was the only data to be reported in more than two cohorts. The eligible multifidus data was cross-sectional area at C4 to C6 levels and muscle fat infiltrate at C4 to C7 levels. Where the measurements were made between levels (e.g. C3/4) the inferior level was used i.e. C3/4 = C4 (Matsumoto et al. [31]) to correspond to measurements made at the most superior aspect of each vertebral level [27]. Studies were combined using inverse-variance random-effects models [42] with RevMan5 version 5.3.5 (Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2012). Both cross-sectional area and muscle fat infiltrate results were expressed as mean differences (MD) with 95% CI. Heterogeneity was analysed using a Chi^2 test with N-1 degrees of freedom (N = number of studies; alpha of 0.05) and also an I^2 test which quantifies heterogeneity in terms of percentage. I^2 values of 25%, 50% and 75% were considered to represent low, medium and high levels of heterogeneity [43]. Forest plots provided a visual representation of the model.

RESULTS

Study Selection

The search identified 380 studies. After screening, eight reports of five cohorts were included in the review [26-32, 34] (Figure 1; Appendix 3). One author was contacted for clarification of results and methods [31] and two authors for raw data [27, 29, 30], all but one author responded. The four studies published by Elliott et al. [27-29, 34] reported different muscles and morphometric parameters from the same cohort. Therefore, while the data was extracted from all four studies, the quality assessment and demographic reporting was derived from just one [27].

Assessment of Study Quality

All studies addressed the selection of WAD and controls adequately (Appendix 4). Exposure was established using appropriate methods such as accident and police reports, structured interviews or questionnaires. Comparability of WADs to controls included one to two matching factors (Appendix 4). There were some weaknesses. Only Ulbrich et al. [32] and Karlsson et al. [30] blinded their assessors of muscle morphometry. Matsumoto et al. [31] lost approximately 80% of their cohort over the ten-year follow-up period. However, this was investigated, and a selection bias was not demonstrated.

Study and Participant Characteristics

A comparison of raw data sets for 176 patients and 168 controls was undertaken. Six studies [26-30, 34] described chronic WAD, one study [32] described acute WAD

and one study described both acute and chronic WAD [31] (Table 1). Five studies reported cross-sectional area from four cohorts [28, 32] and five studies reported muscle fat infiltrate from three cohorts [26, 27, 29, 30, 34]. The WAD cohort size ranged from 5 to 79 and controls from 5 to 60 (Table 2). The majority of subjects were middle-aged with the exception of the Matsumoto et al. cohort [31] who was older because they were assessed 10 years post-MVT. Elliott et al. [27, 28],[29, 34] included female subjects only.

Acute WAD

The two studies assessing acute WAD patients reported cross-sectional area only [31, 32] (Table 1).

Ulbrich et al. [32] used axial Short Tau Inversion Recovery/Turbo Inversion Recovery Magnitude (STIR/TIRM) MR images to measure cross-sectional area of sternocleidomastoid and posterior muscle groups at C2, C4 and C5 levels (Table 1, Appendix 5). The images permitted measurement of groups of muscles only except for sternocleidomastoid. Their cohort (38 WAD; 38 controls) included equal numbers of males and females (Table 2). There were no differences between WAD and controls (Table 3).

Matsumoto et al. [31] used axial T2-weighted MR images to measure four individual muscles (multifidus, semispinalis cervicis, semispinalis capitis and splenius capitis) at C3/4, C4/5 and C5/6 levels (Table 1, Appendix 5). Their cohort (23 WAD; 60 controls) included both males and females (Table 2). Their data demonstrated an

increase in WAD cross-sectional area compared to control in multifidus at all levels and semispinalis cervicis at C3/4 (Table 3).

Chronic WAD

Four chronic WAD cohorts have been studied. In one cohort cross-sectional area only was reported in four posterior muscles [31]. In the second cohort, which was the largest, both cross-sectional area and muscle fat infiltrate in 10 anterior and posterior muscles were studied and published in four different studies [27-29, 34]. In a third, muscle fat infiltrate only was reported in two muscles from a very small cohort [26]. In the most recent cohort both cross-sectional area and muscle fat infiltrate were reported for multifidus exclusively [30].

Matsumoto et al. [31] rescanned their acute WAD and control groups after 10 years. T2-weighted MR sequences were used to measure the same four individual muscles (multifidus, semispinalis cervicis, semispinalis capitis and splenius capitis) at C3/4, C4/5 and C5/6 levels (Table 1, Appendix 5). Their cohort (23 WAD; 60 controls) included both males and females (Table 2). Their data demonstrated an increase in WAD cross-sectional area for multifidus at all levels, and semispinalis cervicis and capitis at C3/4 only (Table 4).

Elliott et al. [27-29, 34] used an axial T1-weighted spin echo sequence to measure cross-sectional area and muscle fat infiltrate of the multifidus, semispinalis cervicis, semispinalis capitis and splenius capitis, trapezius, rectus capitis posterior minor, rectus capitis posterior major, longus colli, longus capitis and sternocleidomastoid at various levels between C0 and C7 (Table 1, Appendix 5). Their cohort included

females only and the maximum size was 79 WAD and 34 controls (Table 2). The data extraction revealed an increase in cross-sectional area in the multifidus and longus colli/capitis at all levels, semispinalis capitis at C3 and C4, and splenius capitis at C3. However, cross-sectional area was decreased in semispinalis capitis at C6, semispinalis cervicis at C3, C5 and C6, and upper trapezius at C7. Muscle fat infiltrate was reported in two ways: two papers reported fat:muscle ratio [27, 29]) and one reported area [34]. Fat:muscle ratio measurements indicated a relative increase in fat content in all muscles, except sternocleidomastoid, at all levels (Table 5). Overall, this pattern was also seen in the extracted data when the fat content was expressed as an area and controlled for BMI apart from semispinalis cervicis at all levels, and the semispinalis and splenius capitis muscles at the C5-6 level [34].

Abbott et al. [26] used multi-echo gradient-echo MR to measure muscle fat infiltrate in multifidus and semispinalis cervicis combined, using a calculation of average fat content from axial slices at the C3, C4, C5, C6 and C7 levels (Table 1, Appendix 5). They reported the results for muscle fat infiltrate in an age and sex matched cohort of 5 WAD and 5 control participants. The data extracted revealed a minimal increase in WAD muscle fat infiltrate at C5 only (Table 5).

Karlsson et al. [30] used a gradient-echo sequence to measure cross-sectional area and muscle fat infiltrate of the multifidus at the C4, C5, C6 and C7 levels (Table 1, 4 & 5, Appendix 5). Their cohort consisted of 31 WAD and 31 control age and sex matched participants. They used manual segmentation to calculate both cross-sectional area and muscle fat infiltrate, with muscle fat infiltrate represented by a fat

fraction measured from fat- and water-separated images. When the data for WAD and control groups were extracted, there was no increase in cross-sectional area or muscle fat infiltrate in the WAD cohort (Table 4).

Multifidus Meta-analysis Models

Multifidus cross-sectional area was significantly increased at C5 and C6 (Z=3.51, p=0.0004; Z=2.66, p=0.008, respectively) but not C4 (Z=1.47, p=0.14). Heterogeneity was acceptable at C5 and C6 (I^2= 40% and 26%, respectively) but not at C4 (I^2= 72%) (Figure 2). Multifidus muscle fat infiltrate was only significantly increased at C7 (Z=2.52, p=0.01) but the heterogeneity between studies was unacceptably high (I^2= 83%) (Figure 3).

DISCUSSION

The aim of this review and meta-analysis was to assess the strength of the evidence for changes in cervical MR muscle morphometry in patients with WAD compared to control. The results suggest that, although cross-sectional area is probably increased in some muscles after whiplash, the evidence for an increase in muscle fat infiltrate is not consistently reported even though there appears to be an effect at C7. This is an important finding for radiologists, rehabilitation professionals, health policy makers, insurers and patients.

This is the first systematic review with meta-analysis to examine the evidence for cervical muscle morphometric changes based on MR images after whiplash injury in both acute and chronic cohorts. The strength of this review is that it has for the first time consolidated and compared sets of raw data in both acute and chronic cohorts for 176 patients and 168 controls. By extracting the raw data, we have been able to report mean differences between WAD and controls to determine whether there is any consistency between cohorts. In some cases, the data extraction was quite complex and would not have been easily accessible to the reader. A statistical meta-analysis with forest plots was also performed on the extracted multifidus data. Therefore, this review is the most rigorous and comprehensive to date.

A recently published systematic review of muscle morphological changes of chronic neck pain of idiopathic and traumatic origin arrived at different conclusions than those found in this analysis [35]. They suggested that the evidence overall supported an increased cross-sectional area in the deep and higher cervical muscles which the

authors attributed to an increase in muscle fat infiltrate. However, their conclusions were dependent on accepting the study findings as opposed to analysing the actual data. Further, consideration of the dominance of one cohort reported in multiple papers was not made. They also allowed any imaging modality, so both MR and ultrasound imaging studies were included. Our review and analysis utilised objective data extraction techniques and statistical modelling where sufficient data was available, and therefore the results are arguably more robust. The results of this review are also specific to whiplash and therefore more directly applicable.

Study Quality

All eligible studies were rated highly but only Ulbrich et al. [32] and Karlsson et al. [30] controlled for detection bias by blinding the assessor. Although this is reflected in the score, the influence of this factor is potentially underestimated in the Newcastle-Ottawa scoring system. In studies which involve manual muscle morphometry estimations, blinding is essential since the researcher is required to make subjective decisions about the boundaries of the muscle on each MR slice. There is therefore significant potential for detection bias where the researcher knows which group the participant belongs to.

Cross-sectional Area

An increase in muscle cross-sectional area after whiplash may indicate either inflammation or swelling in the acute phase or muscle fat infiltration in the chronic phase. A decrease in cross-sectional area may indicate atrophy. Atrophy has been reported in multifidus in both acute and chronic lumbar spine pain [44, 45]. In the

cervical spine increases, decreases and no change in cross-sectional area have all been reported in both acute and chronic WAD cohorts [28],[29, 31, 32, 34, 46]. Our analysis of the two acute whiplash studies [31, 32] revealed that there may be an increase in multifidus cross-sectional area but only Matsumoto et al. [31] measured this muscle in isolation thus this data needs corroboration. It is worth noting that Matsumoto et al. [31] did not discuss this finding in their paper preferring instead to emphasise the increased multifidus cross-sectional area apparent 10 years post-injury.

The results of our analysis suggest that there is good evidence for increased cross-sectional area in multifidus at C5 and C6 levels, but not C4, in chronic WAD. This finding is perplexing because there is no apparent reason for higher cervical spine levels to be spared from injury. On examination of the forest plots it is clear that the data from Karlsson et al. [30] drove the effect down generally, but particularly at C4. There were differences between these three WAD cohorts but they do not seem to explain the inconsistency in the results. Karlsson et al. [30] included both WAD II and III while Elliott et al. [27] included only WAD II. Therefore, morphological changes might be expected to be greater in the group with positive neurological signs (WAD III) but this was not the case [47]. Only 50% of the Matsumoto et al. [31] cohort reported head, neck or shoulder pain at 10 year follow-up. This indicates that half of the cohort were WAD 0 and yet our extracted data shows increased multifidus cross-sectional area at all levels in their whiplash group compared to controls (Table 4). This suggests that the presence of symptoms may not be relevant, although this does not comply with other published data which reports that muscle fat infiltrate is associated with the NDI score [48, 49]. However, it is important to note that

Matsumoto et al. [31] were only able to follow up 17% WAD and 26% control participants from their original cohort.

On examination of the other muscle data there is a lack of consistency. Of the two cohorts in whom semispinalis cervicis cross-sectional area was studied there are conflicting results [27, 31]. In contrast these two cohorts largely demonstrate no change in cross-sectional area for the semispinalis and splenius capitis muscles. The deep flexors (longus capitis/colli) were measured in the Elliott et al. cohort only [29]. That data suggests that cross-sectional area is consistently increased in these muscles to a similar magnitude and variance as those seen in multifidus in the same cohort. In a longitudinal study of acute whiplash patients at 48 hours, 3 months and 6 months, post-injury Ulbrich et al. [46] found no changes in the deep extensor muscle cross-sectional area. However, they grouped the muscles for measurement and therefore potentially lost the sensitivity required to evaluate individual muscles such as multifidus.

One hypothesis for explaining this apparent selectivity between muscles might be that propensity to injury is related to the fascicle length. The multifidus and longus capits/colli fascicle lengths (1.7-2.6 and 3.8 cm, respectively) are much shorter than semispinalis capitis and splenius (6.8 and 9.5 cm, respectively) [50, 51]. Therefore, they may be subject to a higher risk of overstretch injury [52]. This overstretching theory is enhanced by the fact that multifidus fascicles attach to the posterior aspect of the facet capsules [51, 53]. These deep attachments would therefore tether the

caudal end and reduce the dynamic fascicle length. This may be why multifidus is most consistently reported to be altered after whiplash.

Muscle Fat Infiltrate

Our results do not support the contention that increased muscle fat infiltrate in the cervical muscles can yet be considered a robust marker for chronic whiplash. Muscle fat infiltrate has been proposed as the likely cause for increased cross-sectional area in the cervical muscles after whiplash [29]. Theoretically, this makes sense since increased muscle fat infiltrate has been reported in other muscle groups following injury or disuse [54-56]. The mechanism for fatty infiltration after trauma is not known. However, it may involve fibro/adipogenic progenitor cells undergoing differentiation into adipocytes and fibrocytes after a muscle fails to regenerate following injury [57]. In this review, the evidence for an increase in muscle fat infiltrate after whiplash was almost exclusively limited to one cohort. Although the Elliott et al. [27-29, 34] cohort was the largest and most comprehensive in terms of muscles measured, our data extraction from the other two studies [26, 30] did not support their results. The data from Abbott et al. [26] demonstrated a very small increase at C5 only whilst the Karlsson et al. [30] data did not show any difference. Meta-analysis of the data for muscle fat infiltrate in multifidus revealed that there was no significant effect in the whiplash patients compared to control, except at C7. However, at C7, the high degree of heterogeneity between the studies invalidated this finding.

The inconsistency in the muscle fat infiltrate data may be driven by the measurement method rather than the marker. First, it has been shown that single image slices do

not represent 3D muscle measurements [58]. In all the included studies, muscle size and intra-muscular fat were quantified using cross-sectional area at selected vertebral levels. This may explain the discrepancies in the data which suggest differences in cross-sectional area and muscle fat infiltrate at some levels but not others. It is probable that, in these 2D methods, the data may have suffered from errors associated with partial volume effects and the alignment of the axial slice relative to the cervical vertebrae [59]. Advances in MR imaging techniques offer the possibility of obtaining accurate 3D measures of muscle morphometry with rapid acquisition times [33, 60, 61]. It is therefore clear that volumetric studies are required before the utility of muscle fat infiltrate as a reliable marker for chronic whiplash can be verified or refuted.

Another confounding factor with the use of muscle fat infiltrate as a marker for injury is that it increases with age [62, 63]. Increases in adipose tissue within skeletal muscle occur secondary to the denervation of muscle that accompanies aging [64]. In the Karlsson et al. [30] study the authors concluded that muscle fat infiltrate was significantly greater in severe WAD (Neck Disability Index >40%) compared to mild WAD and controls. However, this was a *post hoc* analysis which resulted in groups which were no longer balanced for age and possibly sex. The severe WAD group was older than both the controls and mild group. Therefore, the findings were confounded by a lack of equity with respect to age of approximately six years. An age effect on spinal muscle morphology has been established [63] but there is no precise data with respect to the cervical musculature. Therefore, age must be controlled in any further investigations of muscle fat infiltrate as a marker for whiplash.

The effect of sex on muscle morphometry has not been thoroughly studied. Elliott et al. [27] included only females while all other studies included both sexes [26, 30, 32]. Only Ulbrich et al. [32] analysed their data by sex. They found that the cross-sectional area in female WADs was greater than controls whereas in males the cross-sectional area for WADs was smaller. It is possible that female muscle may respond differently to injury and this needs further investigation given that females are reported to have a higher incidence of persistent pain following whiplash [27]. Body mass index (BMI) and body fat composition may be related to cross-sectional area and muscle fat infiltrate but the evidence is limited [27, 32, 65, 66]. However, until the relationship is clear, they should also be considered as potential confounding factors.

Study Limitations

This review should be considered in the light of its limitations. First, the number of eligible studies was very limited. We only included studies which compared WAD to controls and used MR to acquire the image data. The participant numbers in whiplash studies are typically small which probably reflects the difficulty of recruiting and the cost of MR imaging. Second, statistical meta-analysis was only possible for multifidus in the chronic cohort because the data for other muscles was not available. Further, we included data from Abbott et al. [26] in the meta-analysis even though their multifidus measurements incorporated semispinalis cervicis. Meta-analysis is an important method for obtaining robust evidence in this area. However, for this to happen, future studies need to be more homogenous with a control group matched for age, sex and BMI. Third, although a single cohort dominates the whiplash literature, our method mitigated this potential for bias by basing our analysis on

cohorts rather than reports. Fourth, there was considerable heterogeneity between the studies making the comparison of actual values problematic. Differences included time since MVT, WAD classification, MR sequences, patient age and sex, cervical spine level and muscles measured as well as the software used for morphometric analysis. We have therefore attempted to summarise the data 'effect' using the confidence intervals of the differences. Fifth, we compared whiplash to non-injured controls but there are two longitudinal studies which have examined the predictive power of muscle fat infiltrate with respect to recovery according to NDI score cut points [48, 49]. These studies reported early evidence for muscle fat infiltrate as a sensitive and specific metric for recovery after whiplash but were not within our review criteria and so were not included. Finally, the Newcastle-Ottawa scale has been criticised for its poor inter-observer reliability for some domains and vague decision rules [67]. However, we used two independent assessors and resolved any differences by consensus. The results are similar to those reported elsewhere [35].

Future Research

There are several important considerations for future studies which arise from this review. First, to develop useful comparative longitudinal data, meaningful and robust techniques which detect real changes in muscle morphology are required. Technological advances in MR have, in the past, made comparison difficult [31] but as measurement methods improve, particularly 3D imaging sequences, measurement should be more amenable to longitudinal comparison. Therefore, controlled longitudinal designs are required to establish whether cross-sectional area or muscle fat infiltrate are robust markers for WAD, and MR sequences should enable volumetric quantification of contractile muscle tissue compared to fat (muscle

fat infiltrate). Further, it is suggested that measurement of individual muscles should be undertaken because not all studies use the same groupings and intermuscular fat may potentially confound the MR measurements. Assessors should always be blinded because of the risk of detection bias. Finally, future studies should be more consistent with reporting and controlling for time since MVT, mechanism of MVT, WAD classification, optimized MR sequences, age and sex, cervical spine level and muscles measured, as well as the methods used for morphometric analysis.

CONCLUSIONS

In conclusion, this review and meta-analysis suggests that current reports of morphological changes in muscle after whiplash, using two-dimensional MR data, are not sufficiently consistent to support their use as a marker for chronic whiplash. Our data suggests that cross-sectional area may be increased in multifidus in chronic WAD, but there is insufficient evidence to confirm whether muscle fat infiltrate is also increased. It is possible that volumetric measurement of muscle fat infiltrate will clarify this issue. The most challenging feature of meta-analyses in this area is lack of data and study homogeneity. If we are to rely on MR muscle morphometry as a marker for whiplash it is crucial that it is robust in order to avoid incorrect diagnosis and false evidence in support of litigation and ineffective deployment of health resources. We recommend that adequately controlled designs and the inclusion of a standardized minimum dataset should be employed to facilitate meta-analysis in the future

Conflict of Interest

None of the authors has any conflict of interest to disclose.

\

Acknowledgements

Daniel Owers was paid a stipend of $AUD2100 as a Canberra Hospital Private Practice Fund Vacation Scholar while undertaking this review.

References:

[1] Carroll LJ, Holm LW, Hogg-Johnson S, Cote P, Cassidy JD, Haldeman S, et al. Course and prognostic factors for neck pain in whiplash-associated disorders (WAD): results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders. Spine (Phila Pa 1976). 2008;33:S83-92.

[2] Rodriquez AA, Barr KP, Burns SP. Whiplash: Pathophysiology, diagnosis, treatment, and prognosis. Muscle & Nerve. 2004;29:768-81.

[3] Casey PP, Feyer AM, Cameron ID. Identifying predictors of early non-recovery in a compensation setting: The Whiplash Outcome Study. Injury. 2011;42:25-32.

[4] Holm LW, Carroll LJ, Cassidy JD, Hogg-Johnson S, Cote P, Guzman J, et al. The burden and determinants of neck pain in whiplash-associated disorders after traffic collisions. Results of the Bone and Joint Decade 2000-2010 Task Force on Neck Pain and Its Associated Disorders (Reprinted from Spine, vol 33, pg S52-S59, 2008). Journal of manipulative and physiological therapeutics. 2009;32:S61-S9.

[5] Soderlund A. The role of educational and learning approaches in rehabilitation of whiplash-associated disorders in lessening the transition to chronicity. Spine (Phila Pa 1976). 2011;36:S280-5.

[6] Styrke J, Stalnacke BM, Bylund PO, Sojka P, Bjornstig U. A 10-year incidence of acute whiplash injuries after road traffic crashes in a defined population in northern Sweden. PM R. 2012;4:739-47.

[7] Freeman MD, Croft AC, Rossignol AM, Centeno CJ, Elkins WL. Chronic neck pain and whiplash: A case-control study of the relationship between acute whiplash injuries and chronic neck pain. Pain Research & Management : The Journal of the Canadian Pain Society. 2006;11:79-83.

[8] Sterling M, Hendrikz J, Kenardy J. Compensation claim lodgement and health outcome developmental trajectories following whiplash injury: A prospective study. Pain. 2010;150:22-8.

[9] Freeman MD. A Study of Chronic Neck Pain Following Whiplash Injury. Ann Arbor, MI Dissertation Services. 1998.

[10] Zaloshnja E, Miller T, Council F, Persaud B. Comprehensive and Human Capital Crash Costs by Maximum Police-Reported Injury Severity Within Selected Crash Types. Annual Proceedings / Association for the Advancement of Automotive Medicine. 2004;48:251-63.

[11] Bannister G, Amirfeyz R, Kelley S, Gargan M. Whiplash injury. J Bone Joint Surg Br. 2009;91:845-50.

[12] Joslin CC, Khan SN, Bannister GC. Long-term disability after neck injury. a comparative study. J Bone Joint Surg Br. 2004;86:1032-4.

[13] Sterling M, Jull G, Kenardy J. Physical and psychological factors maintain long-term predictive capacity post-whiplash injury. Pain. 2006;122:102-8.

[14] Sterling M, Jull G, Vicenzino B, Kenardy J, Darnell R. Physical and psychological factors predict outcome following whiplash injury. Pain. 2005;114:141-8.

[15] Sterling M, Kenardy J. The relationship between sensory and sympathetic nervous system changes and posttraumatic stress reaction following whiplash injury--a prospective study. J Psychosom Res. 2006;60:387-93.

[16] Ferrari R, Russell A. Epidemiology of whiplash: an international dilemma. Annals of the Rheumatic Diseases. 1999;58:1-5.

[17] Cassidy JD, Carroll LJ, Cote P, Lemstra M, Berglund A, Nygren A. Effect of eliminating compensation for pain and suffering on the outcome of insurance claims for whiplash injury. N Engl J Med. 2000;342:1179-86.

[18] Cameron ID, Rebbeck T, Sindhusake D, Rubin G, Feyer AM, Walsh J, et al. Legislative change is associated with improved health status in people with whiplash. Spine (Phila Pa 1976). 2008;33:250-4.

[19] Spearing NM, Connelly LB, Gargett S, Sterling M. Does injury compensation lead to worse health after whiplash? A systematic review. Pain. 2012;153:1274-82.

[20] Spearing NM, Gyrd-Hansen D, Pobereskin LH, Rowell DS, Connelly LB. Are people who claim compensation "cured by a verdict"? A longitudinal study of health outcomes after whiplash. J Law Med. 2012;20:82-92.

[21] McLean SA, Ulirsch JC, Slade GD, Soward AC, Swor RA, Peak DA, et al. Incidence and Predictors of Neck and Widespread Pain after Motor Vehicle Collision among US Litigants and Non-Litigants. Pain. 2014;155:309-21.

[22] Krakenes J, Kaale BR, Moen G, Nordli H, Gilhus NE, Rorvik J. MRI assessment of the alar ligaments in the late stage of whiplash injury--a study of structural abnormalities and observer agreement. Neuroradiology. 2002;44:617-24.

[23] Krakenes J, Kaale BR, Nordli H, Moen G, Rorvik J, Gilhus NE. MR analysis of the transverse ligament in the late stage of whiplash injury. Acta Radiol. 2003;44:637-44.

[24] Elliott JM. Are there implications for morphological changes in neck muscles after whiplash injury? Spine (Phila Pa 1976). 2011;36:S205-10.

[25] Li Q, Shen H, Li M. Magnetic resonance imaging signal changes of alar and transverse ligaments not correlated with whiplash-associated disorders. European Spine Journal. 2013;22:14-20.

[26] Abbott R, Pedler A, Sterling M, Hides J, Murphey T, Hoggarth M, et al. The geography of fatty infiltrates within the cervical multifidus and semispinalis cervicis in individuals with chronic whiplash-associated disorders. The Journal of orthopaedic and sports physical therapy. 2015;45:281-8.

[27] Elliott J, Jull G, Noteboom JT, Darnell R, Galloway G, Gibbon WW. Fatty infiltration in the cervical extensor muscles in persistent whiplash-associated disorders: a magnetic resonance imaging analysis. Spine (Phila Pa 1976). 2006;31:E847-55.

[28] Elliott J, Jull G, Noteboom JT, Galloway G. MRI study of the cross-sectional area for the cervical extensor musculature in patients with persistent whiplash associated disorders (WAD). Man Ther. 2008;13:258-65.

[29] Elliott JM, O'Leary S, Sterling M, Hendrikz J, Pedler A, Jull G. Magnetic resonance imaging findings of fatty infiltrate in the cervical flexors in chronic whiplash. Spine (Phila Pa 1976). 2010;35:948-54.

[30] Karlsson A, Leinhard OD, Aslund U, West J, Romu T, Smedby O, et al. An Investigation of Fat Infiltration of the Multifidus Muscle in Patients With Severe Neck

Symptoms Associated With Chronic Whiplash-Associated Disorder. The Journal of orthopaedic and sports physical therapy. 2016;46:886-93.

[31] Matsumoto M, Ichihara D, Okada E, Chiba K, Toyama Y, Fujiwara H, et al. Cross-sectional area of the posterior extensor muscles of the cervical spine in whiplash injury patients versus healthy volunteers--10 year follow-up MR study. Injury. 2012;43:912-6.

[32] Ulbrich EJ, Anderson SE, Busato A, Abderhalden S, Boesch C, Zimmermann H, et al. Cervical muscle area measurements in acute whiplash patients and controls. Journal of magnetic resonance imaging : JMRI. 2011;33:668-75.

[33] Elliott JM, Kerry R, Flynn T, Parrish TB. Content not quantity is a better measure of muscle degeneration in whiplash. Manual Therapy. 2013;18:578-82.

[34] Elliott JM, Pedler AR, Jull GA, Van Wyk L, Galloway GG, O'Leary SP. Differential changes in muscle composition exist in traumatic and nontraumatic neck pain. Spine (Phila Pa 1976). 2014;39:39-47.

[35] 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. 2016;22:42-9.

[36] Elliott J, Sterling M, Noteboom JT, Darnell R, Galloway G, Jull G. Fatty infiltrate in the cervical extensor muscles is not a feature of chronic, insidious-onset neck pain. Clin Radiol. 2008;63:681-7.

[37] Moher D, Liberati A, Tetzlaff J, Altman DG. Preferred reporting items for systematic reviews and meta-analyses: the PRISMA statement. J Clin Epidemiol. 2009;62:1006-12.

[38] Cochrane Haematological Malignancies Group. Data extraction form. http://chmgcochraneorg/sites/chmgcochraneorg/files/uploads/Template-Data%20Extraction-CHMGpdf Accessed: 22/03/2014. 2007.

[39] Cochrane Training. Data extraction forms. http://training.cochrane.org/ Accessed 10/10/2017.

[40] Wells GA SB, O'Connell D, Peterson J, Welch V, Losos M. The Newcastle-Ottawa Scale (NOS) for assessing the quality if nonrandomized studies in meta-analyses. Available from: URL: http://www.ohri.ca/programs/clinical_epidemiology/oxford.htm. [cited 2014 Jan 21]

[41] Sterling M. Physiotherapy management of whiplash-associated disorders (WAD). J Physiother. 2014;60:5-12.

[42] Higgins JPT GS. Cochrane Handbook for Systematic Reviews of Interventions [updated March 2011] Version 5.1.0. The Cochrane Collaboration Available from wwwcochrane-handbookorg. 2009.

[43] Higgins JP, Thompson SG, Deeks JJ, Altman DG. Measuring inconsistency in meta-analyses. BMJ. 2003;327:557-60.

[44] Hides JA, Stokes MJ, Saide M, Jull GA, Cooper DH. Evidence of lumbar multifidus muscle wasting ipsilateral to symptoms in patients with acute/subacute low back pain. Spine (Phila Pa 1976). 1994;19:165-72.

[45] 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;14:496-500.

[46] Ulbrich EJ, Aeberhard R, Wetli S, Busato A, Boesch C, Zimmermann H, et al. Cervical muscle area measurements in whiplash patients: Acute, 3, and 6 months of follow-up. Journal of magnetic resonance imaging : JMRI. 2012;36:1413-20.

[47] Higashino K, Matsuura T, Suganuma K, Yukata K, Nishisho T, Yasui N. Early changes in muscle atrophy and muscle fiber type conversion after spinal cord transection and peripheral nerve transection in rats. J Neuroeng Rehabil. 2013;10:46.

[48] Elliott J, Pedler A, Kenardy J, Galloway G, Jull G, Sterling M. The temporal development of fatty infiltrates in the neck muscles following whiplash injury: an association with pain and posttraumatic stress. PloS one. 2011;6:e21194.

[49] Elliott JM, Courtney DM, Rademaker A, Pinto D, Sterling MM, Parrish TB. The Rapid and Progressive Degeneration of the Cervical Multifidus in Whiplash: An MRI Study of Fatty Infiltration. Spine (Phila Pa 1976). 2015;40:E694-700.

[50] Kamibayashi LK, Richmond FJ. Morphometry of human neck muscles. Spine (Phila Pa 1976). 1998;23:1314-23.

[51] Anderson JS, Hsu AW, Vasavada AN. Morphology, architecture, and biomechanics of human cervical multifidus. Spine (Phila Pa 1976). 2005;30:E86-91.

[52] Azizi E, Roberts TJ. Geared up to stretch: pennate muscle behavior during active lengthening. J Exp Biol. 2014;217:376-81.

[53] Winkelstein BA, McLendon RE, Barbir A, Myers BS. An anatomical investigation of the human cervical facet capsule, quantifying muscle insertion area. J Anat. 2001;198:455-61.

[54] Kumar D, Karampinos DC, MacLeod TD, Lin W, Nardo L, Li X, et al. Quadriceps intramuscular fat fraction rather than muscle size is associated with knee osteoarthritis. Osteoarthritis Cartilage. 2014;22:226-34.

[55] Nozaki T, Tasaki A, Horiuchi S, Osakabe C, Ohde S, Saida Y, et al. Quantification of Fatty Degeneration Within the Supraspinatus Muscle by Using a 2-Point Dixon Method on 3-T MRI. AJR Am J Roentgenol. 2015;205:116-22.

[56] Wan Q, Lin C, Li X, Zeng W, Ma C. MRI assessment of paraspinal muscles in patients with acute and chronic unilateral low back pain. Br J Radiol. 2015;88:20140546.

[57] Joe AW, Yi L, Natarajan A, Le Grand F, So L, Wang J, et al. Muscle injury activates resident fibro/adipogenic progenitors that facilitate myogenesis. Nat Cell Biol. 2010;12:153-63.

[58] Vidt ME, Santago AC, 2nd, Tuohy CJ, Poehling GG, Freehill MT, Kraft RA, et al. Assessments of Fatty Infiltration and Muscle Atrophy From a Single Magnetic Resonance Image Slice Are Not Predictive of 3-Dimensional Measurements. Arthroscopy. 2016;32:128-39.

[59] McRobbie DW ME, Graves MJ, Prince MR. . MRI From proton to picture. 2nd ed: Cambridge; Cambridge University Press 200; 2007.

[60] Elliott JM, Walton DM, Rademaker A, Parrish TB. Quantification of cervical spine muscle fat: a comparison between T1-weighted and multi-echo gradient echo imaging using a variable projection algorithm (VARPRO). Bmc Medical Imaging. 2013;13.

[61] Hernando D, Haldar JP, Sutton BP, Ma J, Kellman P, Liang ZP. Joint estimation of water/fat images and field inhomogeneity map. Magn Reson Med. 2008;59:571-80.

[62] Delmonico MJ, Harris TB, Visser M, Park SW, Conroy MB, Velasquez-Mieyer P, et al. Longitudinal study of muscle strength, quality, and adipose tissue infiltration. Am J Clin Nutr. 2009;90:1579-85.

[63] Fortin M, Videman T, Gibbons LE, Battie MC. Paraspinal muscle morphology and composition: a 15-yr longitudinal magnetic resonance imaging study. Med Sci Sports Exerc. 2014;46:893-901.

[64] Sinanan AC, Buxton PG, Lewis MP. Muscling in on stem cells. Biol Cell. 2006;98:203-14.

[65] Elliott JM, Jull GA, T Noteboom J, Durbridge GL, Gibbon WW. Magnetic resonance imaging study of cross-sectional area of the cervical extensor musculature in an asymptomatic cohort. Clinical Anatomy. 2007;20:35-40.

[66] Elliott JM, Galloway GJ, Jull GA, Noteboom JT, Centeno CJ, Gibbon WW. Magnetic resonance imaging analysis of the upper cervical spine extensor musculature in an asymptomatic cohort: an index of fat within muscle. Clinical Radiology. 2005;60:355-63.

[67] Hartling L, Milne A, Hamm MP, Vandermeer B, Ansari M, Tsertsvadze A, et al. Testing the Newcastle Ottawa Scale showed low reliability between individual reviewers. J Clin Epidemiol. 2013;66:982-93.

Figure Legends

Figure 1

Flow diagram indicating search strategy

Figure 2

Forest plot of the mean differences in cross-sectional area (mm^2) of the multifidus muscle between whiplash associated disorder (WAD) and control by cervical spine level. Forest plot shows the mean difference and 95%CI for the studies included in the meta-analysis. For each study in the forest plot, the area of the square is proportional to study weight and the horizontal bar represents the 95% confidence interval. The center of the black diamond represents the overall estimate, and the

width or lateral points of the diamond indicate overall confidence intervals. The line of no effect is shown at 0. Z-score: the standardized expression of a value in terms of its relative position in the full distribution of values. The mean difference and 95%CI produced by RevMan5 version 5.3.5 (Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2012) are slightly different to those represented in Table 5.

SD = standard deviation; CI = confidence interval

Figure 3

Forest plot of the mean differences in muscle fat infiltrate (fat:muscle ratio) of the multifidus muscle between whiplash associated disorder (WAD) and control by cervical spine level. Forest plot shows the mean difference and 95%CI for the studies included in the meta-analysis. For each study in the forest plot, the area of the square is proportional to study weight and the horizontal bar represents the 95% confidence interval. The center of the black diamond represents the overall estimate, and the width or lateral points of the diamond indicate overall confidence intervals. The line of no effect is shown at 0. Z-score: the standardized expression of a value in terms of its relative position in the full distribution of values. The mean difference and 95%CI produced by RevMan5 version 5.3.5 (Copenhagen: The Nordic Cochrane Centre, The Cochrane Collaboration, 2012) are slightly different to those represented in Table 6.

Note. Abbott et al. [26] included both multifidus and semispinalis cervicis in their measurement of muscle fat infiltrate.

SD = standard deviation; CI = confidence interval

Table 1

Comparison of measurement methods and MR protocols between included studies.

Study	Outcomes measured	Muscles	Cervical levels[#]	Measurement software	Time between MVT and MR scan	MR scanner	MR sequence	MR slice thickness
Elliott et al. [28]*****	CSA	Mult, SemiSCerv, SemiSCap, SpCap, Trap, RCPm, RCPM	C0 – C7	MRIcro	Chronic > 3 months < 3 years	Australia: SONATA Siemens 1.5 T USA: Horizon LX General Electric 1.5 T	T1-W SE	4 mm
Elliott et al. [29]*****	CSA & MFI	LColli[a], LCap, SCM	C2/3, C5/6 & C0/1					
Elliott et al. [27]*****	MFI	Mult, SemiSCerv, SemiSCap, SpCap, Trap, RCPm, RCPM	C0 – C7					
Elliott et al. [34]*****	rCSA******& rmCSA	Mult, SemiSCerv, SemiSCap, SpCap, LColli[α], LCap, SCM	C2/3 & C5/6					
		RCPm, RCPM	C1/C2					
Matsumoto et al. [31]	CSA	Mult, SemiSCerv, SemiSCap, SpCap	C3 – C6	Image J (v1.42)	Acute < 2 weeks Chronic ≥ 10 years	Signa Excite HD General Electric 1.5 T	T2-W SE and GE	4 mm
Ulbrich et al. [32]	CSA	Deep extensors[†] Total extensors‡	C2 & C5	Philips PACS with SECTRA	Acute < 48hrs	SONATA Siemens 1.5 T	T1-W STIR/TIRM	3 mm

		SCM‖	C4					
Abbott et al. [26]	MFI	Mult & SemiSCerv¶	C3 – C7	MATLAB	Chronic > 3 months < 3 years	Not reported	3D multiecho GE	3 mm
Karlsson et al. [30]	CSA & MFI	Mult	C4 – C7	Analyze (VII)	Chronic > 6 months < 3 years	Phillips Ingenia 3.0 T	3D multiecho GE	0.75 mm

Note: All CSA and MFI measurements were made manually by tracing around the muscles/muscle groups of interest using the software identified in the table. Pixel intensity was used to calculate MFI.

RCPM - rectus capitis posterior major; RCPm – rectus capitis posterior minor; SemiSCap - semispinalis capitis; Trap - trapezius; LColli - longus colli; LCap - capitis; SpCap - splenius capitis; SemiSCerv - semispinalis cervicis; Mult - multifidus; SCM - sternocleidomastoid; IS – interspinales cervicis; ObCapInf - obliquus capitis inferior; PACS - Picture Archiving and Communication System software; MR – Magnetic Resonance; MFI – Muscle Fat Infiltrate; CSA – Cross-Sectional Area; 3D – three dimensional; SE – Spin Echo; GE – Gradient Echo; STIR/TIRM – Short TI Inversion Recovery/ Turbo Inversion Recovery Magnitude; T1-W=T1-weighted; T2-W=T2-weighted

*same cohort studied in Elliott et al. [27-29, 34]

[#] = refer to Appendix 5 for specific location of axial slice measurement sites

[a] = Longus colli and longus capitis were measured together at C2/3 and C5/6. Longus capitis only was measured at C0/1 in Elliott et al. [29] but not in Elliott et al. [34]

† = ObCapInf, RCPM at the C2 level; Mult, SemiSCerv, IS, spinalis at the C5 level

‡ = all extensor muscles at the C2 and C5 levels

‖ = SCM was a standalone measure [32]

¶ = Multifidus and semispinalis cervicis were measured together

**In Elliott et al. [34] CSA is characterised as relative CSA (rCSA) in order to account for errors due to slice angle; rmCSA represents the relative neck muscle CSA with fat removed

Table 2

Cohort demographics

Cohort	**Subject N (Female)**		**Age (years) Mean [SD]**		**BMI (kg/m^2) Mean [SD]**		**WAD grade†¶**	**NDI¶ score Mean [SD]**
	WAD	Control	WAD	Control	WAD	Control		
Elliott et al. [27]*****	79 (79)	34 (34)	29.7 [7.7]	27.0 [5.6]	25.1 [5.7]	23.0 [4.4]	II	45.5 [15.9]
Matsumoto et al. [31]	23 (10)	60 (24)	51.8 [14.3]	47.8 [12.3]	23.1 [3.2]	22.4 [3.2]	Not reported	Not reported
Ulbrich et al. [32]	38 (19)	38 (19)	35.2	35.1	23.2	23.0	I or II	Not reported
Abbott et al. [26]	5 (3)	5 (3)	30.6 [90]	35.0 [8.9]	29.6 [4.6]	25.8 [5.0]	Not reported	41.2 [13.6]
Karlsson et al. [30]	31 (17)	31 (17)	41.5 [10.9]	41.5 [10.6]	25.6 [3.8]	24.4 [3.2]	II or III	35.8 [14.1]

*same cohort studied in Elliott et al. [28, 29, 34], so Elliott et al. [27] selected for the reporting of cohort demographics

† = Quebec Taskforce Classification

¶ = only applicable to WAD subjects

NDI = Neck Disability Index

BMI = Body Mass Index

SD = Standard Deviation

WAD = Whiplash Associated Disorder

Table 3

Mean difference (95% confidence interval of the mean difference) in cross-sectional area (mm^2) for WAD minus control for all studies of acute whiplash injury (<3 months since motor vehicle trauma).

Spinal Level	C2		C3		C4		C5		C6
Matsumoto et al. [31]									
Multifidus				56 (29 − 83)		46 (9 − 84)		49 (12 − 86)	
Semispinalis cervicis				62 (25 − 98)		18 (-12− 47)		4 (-40 − 49)	
Semispinalis capitis				26 (-55 − 108)		30 (-42 − 102)		15 (-44 − 74)	
Splenius capitis				-1 (-70 − 67)		4 (-69 − 78)		7 (-61 − 76)	
Ulbrich et al. [32]									
Deep extensors	1.0 (-3 − 5)						-3 (-6 − 1)		
Total extensors	0 (-13 − 13)						4 (-15 − 23)		
Sternocleidomastoid					-0.2 (-4 − 4)				

Key: Increase in WAD compared to control | Decrease in WAD compared to control | No difference in WAD compared to control

WAD – whiplash associated disorder

Table 4

Mean difference (95% confidence interval of the mean difference) in cross-sectional area (mm^2) for WAD minus control for all studies of chronic whiplash injury (>3 months since motor vehicle trauma).

Spinal Level	Occ	C1	C2	C3	C4	C5	C6	C7
Matsumoto et al. [31]								
Mult				31 (3 − 59)	42 (3 − 81)	54 (15 − 94)		
SemiSCerv				51 (23 − 78)	30 (-2 − 62)	25 (-22 − 72)		
SemiSCap				124 (30 − 218)	45 (-27 − 117)	37 (-21 − 95)		
SpCap				-10 (-90 − 70)	21 (-55 − 97)	0.5 (-76 − 77)		
Elliott et al. [28]								
Mult				10 (3 − 17)	31 (22 − 41)	42 (28 − 55)	23 (7 − 39)	23 (9 − 37)
SemiSCerv				-16 (-27 − -4)	-8 (-19 − 3)	-22 (-34 − -10)	-21 (-33 − -9)	2 (-9 − 13)
SemiSCap				42 (22 − 61)	18 (3 − 34)	-9 (-23 − 6)	-17 (-28 − -6)	1 (-8 − 10)
SpCap				27 (12 − 41)	15 (-0.2 − 30)	-0.8 (-15 − 13)	-8 (-21 − 4)	5 (-7 − 17)
Trap				6 (-3 − 16)	-2 (-26 − 23)	-63 (-185 − 59)	-109 (-288 − 71)	-173 (-330 − -16)
RCPm		4 (-9 − 17)						
RCPM			8 (-7 − 24)					
Elliott et al. [29]								
LCap/Colli		18 (6 − 29)*	37 (28 − 45)			28 (18 − 38)		
SCM			26 (-25 − 77)			2 (-38 − 42)		
Karlsson et al. [30]								
Mult					-15 (-48 − 19)	9 (-25 − 43)	10 (-33 − 52)	19 (-30 − 67)

Key: Increase in WAD compared to control | Decrease in WAD compared to control | No difference in WAD compared to control

WAD – whiplash associated disorder; RCPM - rectus capitis posterior major; RCPm – rectus capitis posterior minor; SemiSCap - semispinalis capitis; Trap - trapezius; LColli - longus colli; LCap - capitis; SpCap - splenius capitis; SemiSCerv - semispinalis cervicis; Mult - multifidus; SCM - sternocleidomastoid; IS – interspinales cervicis; ObCapInf - obliquus capitis inferior

*Longus capitis only at this level

Table 5

Mean difference (95% confidence interval of the mean difference) in muscle fat infiltrate† (expressed as fat:muscle ratio or mm^2) for WAD minus control for all studies of chronic whiplash injury (>3 months since motor vehicle trauma).

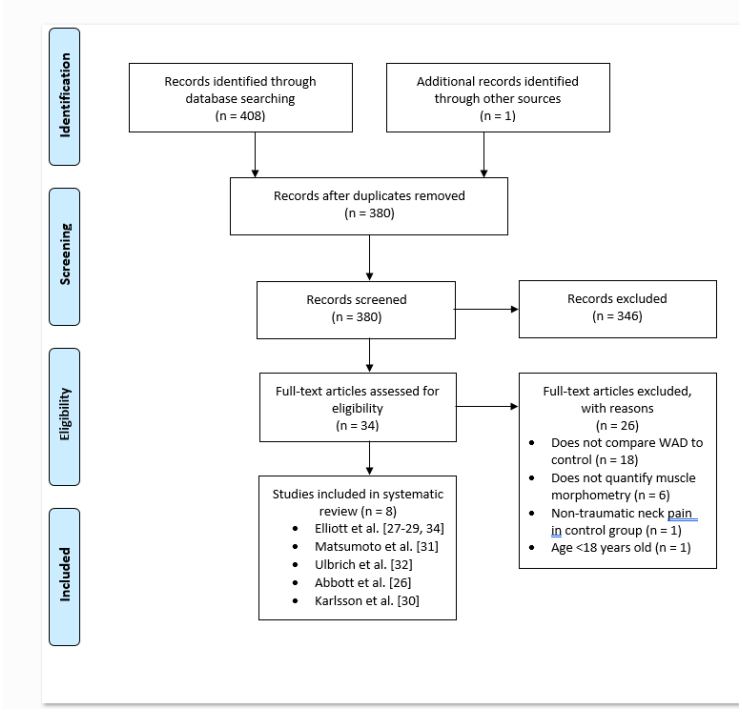
Spinal Level	Occ	C1	C2	C3	C4	C5	C6	C7
Elliott et al. [27]								
Mult				0.18 (0.15 − 0.21)	0.13 (0.10 − 0.15)	0.10 (0.08 − 0.12)	0.10 (0.08 − 0.12)	0.12 (0.10 − 0.14)
SemiSCerv				0.16 (0.13 − 0.18)	0.09 (0.07 − 0.11)	0.07 (0.05 − 0.09)	0.05 (0.03 − 0.07)	0.05 (0.03 − 0.06)
SemiSCap				0.11 (0.09 − 0.12)	0.08 (0.07 − 0.10)	0.06 (0.04 − 0.08)	0.05 (0.03 − 0.06)	0.04 (0.03 − 0.06)
SpCap				0.09 (0.07 − 0.11)	0.09 (0.07 − 0.10)	0.07 (0.05 − 0.09)	0.05 (0.03 − 0.07)	0.05 (0.03 − 0.06)
Trap				0.11 (0.09 − 0.14)	0.10 (0.08 − 0.12)	0.08 (0.06 − 0.10)	0.04 (0.02 − 0.07)	0.04 (0.01 − 0.06)
RCPm	0.20 (0.17 − 0.23)							
RCPM			0.15 (0.13 − 0.18)					
Elliott et al. [29]								
LCap/Colli	0.06 (0.04 − 0.08)*		0.09 (0.06 − 0.11)			0.10 (0.07 − 0.12)		
SCM			0.02 (-0.01 − 0.05)			0.01 (-0.03 − 0.05)		
Elliott et al. [34] – *MFI differences expressed in mm^2*								
Mult			*15.5 (3.5 − 27.5)*			*12.2 (0.2 − 24.2)*		
SemiSCerv			*-6.1 (-18.1 − 5.9)*			*4.1 (-7.9 − 16.1)*		
SemiSCap			*32.7 (20.7 − 44.7)*			*2.7 (-9.3 − 14.7)*		
SpCap			*23 (11.0 − 35.0)*			*7.8 (-4.2 − 19.8)*		
RCPm		*22.1 (10.4 − 34.4)*						
RCPM		*22.4 (10.1 − 34.1)*						
LCap/Colli			*14.6 (2.6 − 26.0)*			*14.0 (2.0 − 26.0)*		
SCM			*10.4 (-1.6 − 22.4)*			*8.4 (-3.6 − 20.4)*		
Abbott et al [26]								
Mult & SemiSCerv¶				0.09 (0 − 0.18)	0.08 (0 − 0.16)	0.08 (0.01 − 0.15)	0.06 (-0.02 − 0.15)	0.07 (-0.04 − 0.2)
Karlsson et al. [30]								

Mult	0.01 (-0.04 − 0.06)	0.003 (-0.03 − 0.04)	0.02 (-0.02 − 0.05)	0.04 (-0.01 − 0.09)

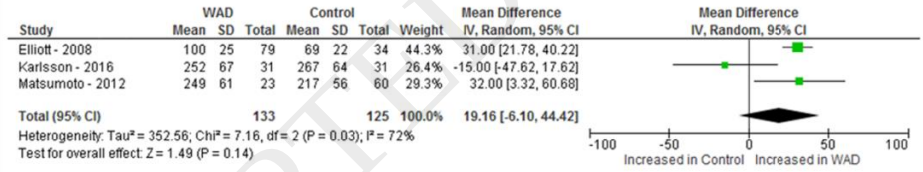
Key: Increase in WAD compared to control | Decrease in WAD compared to control | No difference in WAD compared to control

WAD – whiplash associated disorder; RCPM - rectus capitis posterior major; RCPm – rectus capitis posterior minor; SemiSCap - semispinalis capitis; Trap - trapezius; LColli - longus colli; LCap - capitis; SpCap - splenius capitis; SemiSCerv - semispinalis cervicis; Mult - multifidus; SCM - sternocleidomastoid; IS – interspinales cervicis; ObCapInf - obliquus capitis inferior

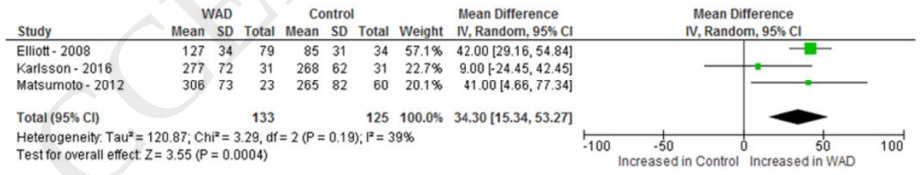
†Muscle fat infiltration was measured in three different ways but the data reported represents the differences within each study. Elliott et al. [27] [29] measured the fat:muscle ratio and Elliott et al. [34] measured the area of fat in the muscle in mm^2; Karlsson et al. [30] measured fat fraction (fat/[fat and water]); *Longus capitis only at this level; ¶ = Multifidus and semispinalis cervicis were measured together



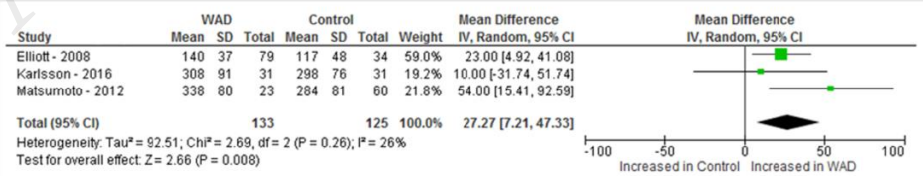
C4



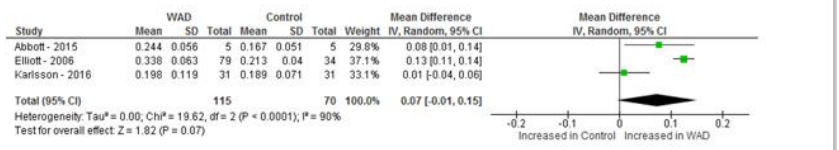
C5



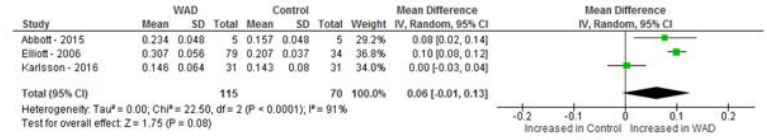
C6



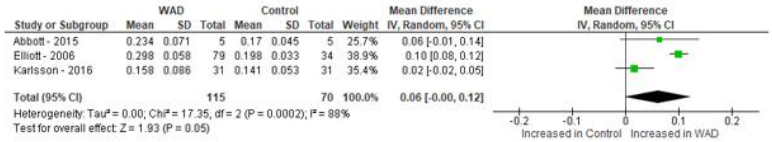
C4



C5



C6



C7