How is trunk muscle activity during forward bending in standing in chronic non-specific low back pain patients presenting with psychosocial characteristics such as fear avoidance beliefs, pain catastrophizing or kinesiophobia? A systematic review of the literature.

Brian Østergaard Sørensen

Dissertation as part of the requirements for the award of MSc Advanced Practice (Physiotherapeutic Practice)

University of Dundee

June 2016
## Contents

<table>
<thead>
<tr>
<th>Headings</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table and figures</td>
<td>III</td>
</tr>
<tr>
<td>Declaration and permission to copy</td>
<td>IV</td>
</tr>
<tr>
<td>Summary</td>
<td>V</td>
</tr>
<tr>
<td>Chapter 1. Introduction</td>
<td>1</td>
</tr>
<tr>
<td>Chapter 2. Background and Literature Review</td>
<td>3</td>
</tr>
<tr>
<td>Chapter 3. Method</td>
<td>23</td>
</tr>
<tr>
<td>Chapter 4. Presentation and Analysis of Findings</td>
<td>49</td>
</tr>
<tr>
<td>Chapter 5. Discussion</td>
<td>84</td>
</tr>
<tr>
<td>Chapter 6. Conclusion</td>
<td>93</td>
</tr>
<tr>
<td>Reference list</td>
<td>94</td>
</tr>
<tr>
<td>Appendix 1. EMBASE and CINAHL search</td>
<td>127</td>
</tr>
<tr>
<td>Appendix 2. Email correspondence</td>
<td>130</td>
</tr>
<tr>
<td>Appendix 3. Newcastle Ottawa Scales</td>
<td>131</td>
</tr>
<tr>
<td>Appendix 4. Email correspondence</td>
<td>135</td>
</tr>
<tr>
<td>Appendix 5. Data Extraction form for included studies</td>
<td>137</td>
</tr>
</tbody>
</table>
## Tables and Figures

<table>
<thead>
<tr>
<th>Headings</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>Table 1. Overview pain-related fear studies</td>
<td>9</td>
</tr>
<tr>
<td>Table 2. Overview musculoskeletal adaptation studies</td>
<td>12</td>
</tr>
<tr>
<td>Figure 1. Motor unit</td>
<td>20</td>
</tr>
<tr>
<td>Table 3. Initial search</td>
<td>33</td>
</tr>
<tr>
<td>Table 4. Search terms – second search</td>
<td>35</td>
</tr>
<tr>
<td>Table 5. Example of table used for retrieved articles</td>
<td>39</td>
</tr>
<tr>
<td>Table 6. Example of table for quality assessment of case-control studies</td>
<td>44</td>
</tr>
<tr>
<td>Table 7. Example of table for quality assessment of cross-sectional studies</td>
<td>45</td>
</tr>
<tr>
<td>Table 8. Example of data extraction form</td>
<td>47</td>
</tr>
<tr>
<td>Figure 2. PRISMA flowchart – study selection process</td>
<td>50</td>
</tr>
<tr>
<td>Table 9. Articles retrieved for eligibility</td>
<td>52</td>
</tr>
<tr>
<td>Figure 3. Study design distribution</td>
<td>57</td>
</tr>
<tr>
<td>Figure 4. Distribution of research subjects</td>
<td>60</td>
</tr>
<tr>
<td>Figure 5. Mean age of research subjects</td>
<td>61</td>
</tr>
<tr>
<td>Figure 6. Duration of symptoms</td>
<td>62</td>
</tr>
<tr>
<td>Figure 7. Placement of electrodes</td>
<td>64</td>
</tr>
<tr>
<td>Table 10. Quality assessment of cross-sectional studies</td>
<td>73</td>
</tr>
<tr>
<td>Table 11. Quality assessment of case-control studies</td>
<td>79</td>
</tr>
<tr>
<td>Table 12. Outcomes pain-related fear and muscle activity</td>
<td>81</td>
</tr>
<tr>
<td>Table 13. Outcomes disability and flexion relaxation ratio</td>
<td>82</td>
</tr>
</tbody>
</table>
Declaration

By submitting this dissertation electronically, I, (Brian Østergaard Sørensen, student id 130022765) am declaring that I am the sole author of this dissertation; that the work has not previously been accepted as part of any other degree submission; that all references cited have been consulted; that I have conducted all the work of which this is a record, and that the finished work contains (20513) words with allowable exclusions.

Permission to copy

By submitting this dissertation electronically, I (Brian Østergaard Sørensen, student id 130022765) give permission for my Dissertation to be copied and posted on My Dundee for future students.

Date: June 29, 2016
Summary

Background

Low back pain (LBP) is the most common single reason for addressing the health care system due to musculoskeletal pain. More than 80% will experience LBP at some point in life; 6 - 10% of those experiencing LBP will develop persistent and chronic pain with great variance in level of disability and pain.

Research indicates that LBP is multifactorial and calls for a biopsychosocial model when approaching the patient experiencing LBP. Despite this recognition LBP is poorly understood. A factor for debate has been the contribution of pain-related fear regarding maintenance of chronicity. Recently the importance of not treating all patients with LBP in the same way but trying to provide stratified care has been highlighted resulting in better outcomes.

Objective

To investigate if there is a correlation between increased muscle activity in a subgroup of patients with chronic non-specific LBP with pain-related fear during standing forward bending.

Method

A systematic review of available literature was conducted searching the databases Pubmed, EMBASE, CINAHL and The Cochrane Library databases from inception to March 2016. Searches used keywords related to LBP, electromyography, pain-related fear, motor control and the flexion relaxation phenomenon.
Quality assessment of included studies was conducted applying the Newcastle Ottawa Scale.

Results
The literature search revealed 211 studies out of which six were included but reporting were on five different studies in the systematic review. Included studies were all observational studies; two case-control and three cross-sectional studies. Included studies revealed measurements from two pairs of electrodes placed at the lumbar trunk muscles between L1 and L5 and one pair of electrodes placed at the abdominal wall.

Conclusion
Correlation between increased muscle activity during standing forward bending and the subgroup of interest was found. Based on the limited number of studies, results should be interpreted with caution.
1. Introduction

In 1987 Gordon Waddell received the Volvo award for his paper “A new clinical model for the treatment of low-back pain”. In this paper he suggested that a new framework for treatment of low back pain (LBP) was needed. He recommended that the traditional biomedical model should be replaced by a biopsychosocial model not considering LBP as something that exclusively needed treatment of the spine but also attended to disability, distress and illness behaviour of people with LBP. In general LBP is not well understood; this is supported by an increasing number of patients attending health care due to LBP in Denmark despite no signs of change in disease and an improved technology (Lind, 2011; MIDT, 2016).

LBP has been associated with risk factors such as previous LBP and genetic predisposition, where as the more physiological factors such as hard manual labour or flexion of the spine are less certain (Delitto et al., 2012; Villumsen et al., 2016). This systematic review will be attending one specific area of this musculoskeletal epidemic in low back pain. Namely trunk muscle activity in chronic non-specific LBP. The background chapter of this dissertation is presenting an overview into the complex field of LBP. An area which contribute greatly to the musculoskeletal are of the health sector both national and internationally.

First, the background will cover prevalence and epidemiology of LBP as well as the importance of identifying sub-groups and the rationale behind the biopsychosocial model and application of electromyography as a relevant measurement tool to investigate trunk muscle activity in patients with chronic non-specific LBP. Chapter 3
will detail the methods applied. Chapter 4 will cover the findings and results and chapter 5 will discuss the findings. Finally in chapter 6 a conclusion is presented.
2. Background and Literature Review

**Epidemiology and Prevalence**

Globally, low back pain (LBP) remains a leading cause of pain and disability and constitutes a considerable socioeconomic burden (Delitto *et al.*, 2012). The lifetime prevalence is approximately 80% (O’Sullivan, 2005), the recurrence rate within a six-month period is approximately 60% (Hestbaek *et al.*, 2003) and in 6-10% of LBP patients will become persistent and chronic (Pengel *et al.*, 2003). In Denmark, the population with LBP annually has 5.5 million more days absent from work compared with the asymptomatic population. Annually, 1,820 billions Danish kroner (DKK) are spent on treatment and 4,840 billion DKK are lost in production due to LBP (Sundhedsstyrelsen, 2015). Data from another group indicated annual costs around 13 billion DKK divided between treatment, short- and long-term absence from work and loss in productivity (MIDT, 2016). Norway has estimated the total costs of LBP to be an almost identical amount (13-15 billion Norwegian kroner) (FORMI, 2014). Deyo *et al.* (2009) presented data from the United States of America documenting a 629% increase in Medicare expenditures for epidural steroid injections, a 423% increase in expenditures for opioids for back pain, a 307% increase in the number of lumbar magnetic resonance images among Medicare beneficiaries and a 220% increase in spinal fusion surgery rates over a 12-year period. According to Lundquist *et al.* (2011), there was a 41% increase in the amount of discectomies over a period from 1999 to 2009 in Denmark and in the period from 2007 - 2010 there was an increase with 16.6% in back surgery procedures.

As illustrated above it is difficult to make exact measures in health care and one should be cautious to compare data from a western European country and the United
States of America where health care systems differ widely. However, it is very clear that the LBP epidemic is continuing and increasing.

**Pathology and Management**

Low back pain is defined as pain and discomfort, localised between the area from below Thoracic T12 and above the inferior gluteal folds, with or without leg pain. It is most of the time mentioned as non-specific LBP due to lack of ability to detect the pain generator, which could be any somatic structure either locally in the tissue or centrally driven from the central nervous system (van Tulder *et al.*, 2006). A specific structure is identified in less then 15% of people seeking care for LBP. Thus, despite advanced and improved diagnostics costs are still rising and the cause of LBP will remain unknown in approximately 85-90% of patients (Dagenais and Haldeman, 2012; Louw *et al.*, 2013).

The more specific causes of LBP are serious pathology such as cancer, infection, fracture, rheumatic disease i.e. ankylosing spondylitis or nerve root pain. Cancer and infection call for immediate specialist referral and treatment, whereas people diagnosed with rheumatic disease will need and receive rheumatologist prescribed medicine along with physiotherapy. Only a very little group of the 5-10% diagnosed with nerve root pain will need to undergo surgery (Lind, 2011). Those that undergo surgery will be at risk of complications such as nerve root injury, infection and recurrence and thus revision surgery at a rate between 1.4% - 11.4% (Epstein, 2016; Yoshihara *et al.*, 2016). Those that do not undergo surgery can improve spontaneously, often in combination with guided physiotherapy. In clinical practice a diagnostic triage suggested by Gordon Waddell (1998) is often used;
The group of patients with chronic non-specific LBP undoubtedly constitutes the biggest challenge. From a pathoanatomical perspective some would argue that this is due to the lack of ability to identify the specific source of a patient’s pain, that treatment is not working and that costs are rising. This argument would be based on the premises that the available diagnostic test cannot identify findings in a person’s spine that look abnormal or “damaged” (Hancock et al., 2011). However, based on recent advances in technology such as magnetic resonance imaging (MRI), a pathoanatomical source based on visual interpretation can easily be identified. Such findings are just as common in the asymptomatic population as in the symptomatic population and vice versa (Berg et al., 2013; Brinjikji et al., 2014). Therefore, the purely pathoanatomical approach is not able to both explain and treat and does thus not provide the sole solution.

As described by Pengel et al. (2003) and Hestbaek et al. (2003), recurrence and chronicity characterize the clinical course experienced by of a large number of patients with LBP. This cannot solely be explained by the pathoanatomical model either because structures often do heal spontaneously either with total or partial regression (Benson and Hartz, 2010; Chun-Chieh Chiu et al., 2015).

The Biopsychosocial Model
A Paradigm Shift

In 1987, Waddell advocated for a change in the approach to non-specific LBP. Waddell stated that the condition needed to be addressed with from a multimodal perspective including biological, psychological and social factors if we should fully understand it.

Over the last three decades, this change in paradigm from a bio/pathoanatomical model to a bio-psychosocial model has evolved and gradually changed the way non-specific LBP is considered by some health professionals (Hancock et al., 2011). Such biopsychosocial models are supported by guidelines; an example of it is the classification model of International Classification of Functioning, disability and Health (ICF) with its components of anatomy, activity, participation and environmental factors (WHO, 2001; Delitto et al., 2012).

Kendall et al. (1997) published a guide on how to assess psychosocial aspects in patients with non-specific LBP and used the term yellow flags. This does not only cover the patient’s status in regards to pain but also involving their relations to family, work, ongoing litigation circumstances and behaviour such as sleep pattern and coping strategies.

Kendall et al. (1997) also clearly stated that the psychosocial perspective should not be seen as a better or alternative solution but as a supplement to the pathoanatomical approach with the purpose of providing a holistic pattern that would help guide treatment. The trend within the health community including the profession of physiotherapy seems be be divided between either the pathoanatomical model or the psychosocial model when it comes to management of LBP (Ostelo et al., 2003; Bishop et al., 2007).

Pain-related Fear
With regards to patients’ behaviour it has been claimed that fear of or anxiety related to movement would be a central mechanism when it comes to the development of chronic non-specific LBP maintaining the patient in a vicious circle of pain and disability (Vlayen and Linton, 2000). The term fear avoidance is related to avoidance of movement and activities based on fear. This concepts has since expanded further and does now include learning, motivation and self-regulation theories. However, it is still the fear and cognitive behaviour processes that continue to be the major pillars of the model (Wideman et al., 2013). Another similar term is kinesiophobia which Kori, Miller and Tod (1990, p. 37) defined as “an irrational and debilitating fear of physical movement resulting from a feeling of vulnerability to painful injury or reinjury”. Such behaviours will then result in further withdrawal from normal activities which again will affect the patients physical and psychological state, which is associated with decreased pain tolerance meaning more pain and disability (Vlayen and Linton, 2000; Crombez et al., 2012; Gatchel et al., 2016). Similarly, a reduction in pain-related fear is linked to reduction in disability (Woby et al., 2004). The mechanisms of chronic pain including conditions such as chronic non-specific LBP are many and some better understood than others. However, it is beyond the scope of this systematic review to go further into this (Nijs et al., 2011).

Over time different researchers have studied the relationship between fear avoidance, pain-related fear, kinesiophobia and somatic reactions such as muscle and neural activity. Glombiewski et al. (2015) found that if they exposed people with higher fear avoidance scores to photos of daily activities it had an effect on increased skin conductance, interbeat interval (IBI) increase, and muscle tension increase in the lower back. Vlaeyen et al. (1999) found similar results by exposing people to a film showing daily activities with the spine in awkward positions. However, their
results where contradictory. Larivière et al. (2010) found poor back muscle endurance was related to catastrophizing. Finally Barke et al. (2012) performed a functional MRI (fMRI) study that revealed no difference between high fear avoidance group and controls with regard to neural activation when exposed to photos of back stressing movements. This leaves us with contradictory knowledge and results. However, it is important to be aware that outcomes in these studies where not the same and should therefore not be used to draw firm conclusions but merely show that fear to some extent is a very plausible aspect of the fear avoidance model.

Overview of Pain-related Fear Studies

To gain an understanding of the studies related to pain-related fear and their methodology an overview of the four studies (Vlaeyen et al., 1999; Larivière et al., 2010; Barke et al., 2012; Glombiewski et al., 2015) was conducted using the STROBE statement checklist for case-control and other observational studies (von Elm et al., 2008).

Table 1 presents the findings and which items included in the overview. It is important to be aware that STROBE does not assess quality but merely serves as reporting guideline for authors with regard to inclusion of items relevant for observational studies (Vandenbroucke et al., 2014).
<table>
<thead>
<tr>
<th>Study</th>
<th>Rational / objective</th>
<th>Study design</th>
<th>Setting</th>
<th>Number of participants</th>
<th>Sample size estimate</th>
<th>Control group</th>
<th>Recruitment</th>
<th>In- &amp; exclusion criteria stated</th>
<th>Exposure / intervention</th>
<th>Statistical methods</th>
<th>Outcome</th>
<th>Results</th>
<th>Bias and blinding</th>
</tr>
</thead>
<tbody>
<tr>
<td>Glombiewski 2015</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>71</td>
<td>Y</td>
<td>N</td>
<td>PHO DA &amp; Fear induction</td>
<td>Y</td>
<td>Y ANO-VA</td>
<td>Y</td>
<td>N</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Skin conductance</td>
<td>Muscle activity</td>
<td>Interbeat interval (IBI)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Larivière 2010</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>32</td>
<td>Y</td>
<td>Y</td>
<td>EMG endurance 4 pairs</td>
<td>Y</td>
<td>Y ANO-VA</td>
<td>EMG Endurance and strength</td>
<td>Y</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bark 2012</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>30 (clbp) (high &gt;35, 5 and low &lt;35, 5 fea)</td>
<td>Y</td>
<td>15 spider phobia</td>
<td>(Y) Som e criter ia</td>
<td>Y</td>
<td>Y ANO-VA</td>
<td>fMRI Y ANO-VA ANC O-VA</td>
<td>Y</td>
<td>N</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
By only performing an overview and not a thorough quality assessment, the risk of introducing bias is likely due to not detecting poor quality in these four studies. However, the main purpose of the brief overview was to make sure that the methodology was acceptable thus making their results trustworthy to justify a correlation between pain-related fear and muscle or neural activity.

A narrative analysis based on the four studies showed that all reported their objective, number of participants, application of statistical methods (commonly ANOVA), which is appropriate due to variance between groups (groups were divided into; symptomatic group, asymptomatic / control group and high / low fear avoidance group) (Jørgensen, Christensen and Kampmann, 2005). All stated very clearly how the intervention was carried out, which outcomes were used and finally presented results statistically (Tables and figures) and a cautious interpretation of their implications.

None of the studies stated which study design they had used, though they appeared to be cross-sectional or case-control studies (Habicht, 2011). Glombiewski et al. (2015) reported sample size estimation but none of the others attended to this. This is critical when it comes to interpretation of results and generalisability. Sample size calculations considering drop outs, confounders and missing data in these studies...
seem less problematic based on the construct of the studies since starting and ending data collection at the same test day. Only Larivière et al. (2010) and Glombiewski et al. (2015) clearly documented setting and recruitment procedure. Inclusion and exclusion criteria where at best only partially described and attention to confounders, bias and possible blinding of researchers regarding application of the intervention where not addressed in any of the studies. Symptomatic and control groups were as identical as possible regarding demographic data, thus reducing risk of confounding (Cao, Cox and Eslick, 2016).

Data from these studies presents contradictory results regarding the effect of pain-related fear on somatic processes. Part of the data are able to correlate pain-related fear to changes in muscle tension and skin conductance. However, despite investigating on a high fear avoidant group Barke et al. (2012) did not find a difference between the symptomatic and asymptomatic population on neural activity. The studies did explicitly present their objectives, their testing and used appropriate statistical analysis. Though replication and generalizability of the studies is difficult due to poor reporting of study design, setting and inclusion and exclusion criteria.

**Musculoskeletal Adaptations**

The chronic non-specific LBP population present with a different muscle activity pattern then healthy controls. Some of the following examples has been documented in the chronic non-specific LBP populations. Tsao and Hodges (2007; 2008) documented delayed or absence of feed-forwarded central nervous system (CNS) mediated pre-programmed reactions. Changes in anticipatory postural adjustments were shown by Jacobs et al. (2009). Abnormal trunk muscle responses have consistently been reported in chronic non-specific LBP patients during experimentally
stimulated postural and activities of daily living (ADL) (Hodges and Richardson, 1996; MacDonald, Moseley and Hodges, 2010; Neblett et al., 2013). The mentioned musculoskeletal dysfunctions or “abnormalities” are believed to be important postural control related biomechanical risk factors for occurrence or recurrence of back pain episodes. On the other hand, it would be just as relevant to look at these dysfunctions or abnormalities as new and advantageous “protection” strategies or just another way of doing what is the most efficient (Moseley 2003; 2004).

**Overview of Musculoskeletal Adaptation Studies**

Reporting such findings of musculoskeletal adaptations can have a considerable impact on clinical practice. However, based on an overview of the studies identified in the scoping review (Table 2), results should at best be interpreted with caution due to their observational nature.

**Table 2. Overview of musculoskeletal adaptation studies.**

<table>
<thead>
<tr>
<th>Author and date</th>
<th>Year (Y)</th>
<th>Number of participants</th>
<th>Sample size estimate</th>
<th>Control group</th>
<th>Recruitment</th>
<th>In- &amp; exclusion criteria</th>
<th>Exposure / intervention</th>
<th>Statistical methods</th>
<th>Outcomes</th>
<th>Results</th>
<th>Bias, blinding, limitation</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. Tsao &amp; Hodges, 2007</td>
<td>Y</td>
<td>22</td>
<td>N</td>
<td>N</td>
<td>TrA or ABS group</td>
<td>Y but some were odd</td>
<td>TrA or ABS</td>
<td>Y ANOVA &amp; Duncan T-test</td>
<td>Y FEM On set arm flex +</td>
<td>Y used a research assistant +</td>
<td></td>
</tr>
<tr>
<td>ITMS reported</td>
<td>Rational / objective</td>
<td>Study design</td>
<td>Setting</td>
<td>Number of participants</td>
<td>Sample size estimate</td>
<td>Control group</td>
<td>Recruitment</td>
<td>Exclusion / intervention criteria</td>
<td>Statistical methods</td>
<td>Outcome</td>
<td>Results</td>
</tr>
<tr>
<td>---------------</td>
<td>----------------------</td>
<td>--------------</td>
<td>---------</td>
<td>------------------------</td>
<td>---------------------</td>
<td>--------------</td>
<td>-------------</td>
<td>-------------------------------</td>
<td>-------------------</td>
<td>---------</td>
<td>---------</td>
</tr>
<tr>
<td>2. Tsao &amp; Hodg, 2008</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>9</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>TrA isolated workout</td>
<td>ANOVA &amp; Duncan</td>
</tr>
<tr>
<td>3. Jacobs, 2009</td>
<td>N</td>
<td>N</td>
<td>N</td>
<td>10</td>
<td>N</td>
<td>10</td>
<td>Y</td>
<td>Y</td>
<td>Arm flex sitt</td>
<td>Y</td>
<td>MANOVA</td>
</tr>
<tr>
<td>4. Hodges &amp; Richardson, 1996</td>
<td>Y</td>
<td>N</td>
<td>N</td>
<td>Y</td>
<td>15</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Arm flex x 10 on platform</td>
<td>Y</td>
</tr>
<tr>
<td>5. MacDonald, 2010</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>13</td>
<td>N</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>EM</td>
<td>G (S&amp;F)</td>
</tr>
</tbody>
</table>

Items such as sample size calculation and estimates such as power calculation, recruitment procedure and description of study design where lacking in most of the
studies. In Table 2 this relates to columns number 2, 5 and 7. Based on the opinion of the author of this review, two of six studies were cross-sectional (studies 5 and 6) and one of six was a cohort study design (study 1), three of six were case-control studies (studies 2, 3 and 4).

Furthermore, only one in six studies addressed these above mentioned issues. Where as five of the six were more concerned regarding equipment and if normalisation of EMG data should be executed.

All studies reported objective, exposure / intervention and especially results where presented in detail.

Studies from the scoping review clearly reveals an indisputable link between trunk muscle changes and the chronic non-specific LBP population however the purpose and the affect of it is not well understood.

Regarding the methodology of the studies in the overview they all stated their objective and applied statistical analysis as well as thorough presentation of their results. Items such as study design, setting, recruitment and sample size issues where with few exceptions such as Neblett, et al. (2013) inadequate reported.

An iatrogenic Problem

During the last two decades numerous approaches and interventions have been developed by the healthcare system and the fitness industry to target “normalisation” of these “musculoskeletal abnormalities”.

Especially the study by Hodges and Richardson (1996) that showed a delayed transversus abdominis (TrA) contraction during an arm flexion in the shoulder. When an individual without pain is moving the arm, a pre-contraction is initiated in the trunk muscles before the actual shoulder muscles contracts. The literature describes this as an anticipatory postural adjustment (Aruin and Latash, 1995; Shumway - Cook
and Woollacott, 2001). The phenomenon has also been well described in lower extremity movement (Shumway - Cook and Woollacott, 2001; Lehman et al., 2004).

Hodges and Richardson (1996) tested and compared this pre-contraction in a group of 15 patients with LBP with a group of 15 asymptomatic controls, based on the hypothesis that this pre-contraction was different in the LBP group (that TrA was dysfunctional).

The findings by Hodges and Richardson (1996) showed that in the symptomatic group, TrA was pre-contracted still but with a delay relative to the arm movement itself when compared to the asymptomatic control group. Similar findings of altered muscle activity in the trunk muscles was observed when analysing muscle activity during gait (Ghamkar and Kahlaee, 2015).

This contributed to the assumptions that specific and or single muscles where responsible for stabilizing the spine, especially TrA. These core muscles had some kind of unique ability to work independently of other muscles. Furthermore that strengthening and to some extent change in timing of these muscles would decrease pain and disability resulting in a more stable and thus less painful back. Thus despite this no high quality studies have shown a favorable effect of such approaches (Smith, Littlewood and May, 2014) and that Wong et al. (2013) concluded in their review that there were no indications that baseline TrA contraction ratio and anticipatory onset of trunk muscles morphometric change were related to the short- or long-term clinical outcomes of patients with chronic non specific low back pain after various exercise interventions.

This approach constitutes a large problem when dealing with chronic non-specific LBP since health personnel have a major influence on how patients think their problem should be solved. Health personnel sometimes explain to their patients that
their problem is of a pathoanatomical nature (i.e. the spine is unstable due to
dysfunctional motor control or reduced capacity to develop efficient endurance or
strength contraction of the trunk muscles); thus they need more core endurance or
strength training. However, when they do not improve neither regarding pain nor
disability this might reinforce their belief that they are still unstable and vulnerable
which is not necessarily the case. This makes health personnel liable of making them
fear-avoidant and maintaining them in the vicious circle (Domenech et al., 2011;
Darlow et al., 2013; Nijs et al., 2013).

**Measurement of Fear Avoidance and Kinesiophobia**

A very common and accepted way of measuring fear avoidance behaviour and
kinesiophobia is by using questionnaires such as the Fear Avoidance Belief
Questionnaire (FABQ) (Waddell et al., 1993) and the Tampa Scale of Kinesiophobia
(TSK) (Woby et al., 2005). The reliability and validity of the TSK scale has been rated
moderate to good (Swinkels-Meewisse et al., 2003; Goubert et al., 2004).

**Subgroup within the Chronic Non-specific Low Back Pain Population**

So far research related to management of this non-specific LBP has shown that the
condition is still not well understood and furthermore it has not been possible to
identify the most effective intervention resulting in research and guidelines providing
a more general approach (Delitto et al., 2012; Smith, Littlewood and May, 2014;
NICE, 2016). These guidelines are of course based on the available research.
However, the problem is that people with back pain are treated as if they were one of
a kind. This “one size fits all “ approach does not make sense since we already know
that certain factors and clinical findings promote and affect the course of LBP such as
genetics, level of education, coping strategies, depression and centralization (Werneke and Hart, 2001; Jarvick et al., 2005).

This knowledge calls for classification into subgroups and stratified care approaches. This was supported by Borkan et al. already in 1998 who emphasized that subgrouping was a top research priority. Some of the classification systems used in the clinical field are Mechanical Diagnosis and Therapy (MDT) also referred to as the McKenzie method (McKenzie and May, 2003) where subgroups are based on history taking and symptomatic response. MDT is supported by international guidelines and literature (DIHTA 1999; May and Donelson, 2008; Delitto et al., 2012; Lee et al., 2013). It has proven to be a reliable assessment tool when used by experienced therapists (Razmjou, Kramer and Yamada 2000; Kilpikoski et al. 2002). Petersen (2003) proposed a pathoanatomical classification system with 12 subgroups based on an algorithm which was effectuated by interpretation of history taking and symptomatic response. This system has been integrated in the Danish primary sector and has proven reliable (Petersen, 2003). Movement System Impairment (MSI) or Motor Control Impairment (MC) is another example of a commonly used classification system (Dankaerts et al. 2006; Henry et al. 2013). Fersum et al. (2009) demonstrated that the one suggested by Peter O´Sullivan demonstrated moderate reliability.

There are several examples showing that subgrouping seems promising (Long, May and Fung, 2008; Hill et al., 2011; Grotle et al., 2012). The subgroup defined by the criteria of chronic non-specific LBP and fear avoidance is relevant to study further due to the inclusion of factors that are strongly associated with a negative prognosis (Leeuw et al., 2007).
Differentiation Between the Symptomatic and the Asymptomatic Populations

Flexion Relaxation Phenomenon and Flexion Relaxation Ratio

Interestingly chronic non-specific LBP populations actually present with an increased muscle activity in the trunk muscles during ADL. Van Der Hulst et al. (2010) found that the chronic non-specific LBP group had an increased muscle activity during walking, basically like a “guardening” strategy which could be a sign of less ability to relax. This matches the findings of one of the most frequently used assessment tools used to differentiate the chronic symptomatic patient from the asymptomatic patient, namely surface electromyography (SEMG) measuring the flexion relaxation phenomenon. In the asymptomatic population the trunk muscles will show a decreased activity as the person tested in forward bending gets further down and the ligaments starts to take over. In the symptomatic population this decrease in muscle activity seems absent or reduced. This difference is not fully understood. However, it has been speculated that it could be due to reduced lumbar range of motion or reorganisation as a result of having pain (Geisser et al., 2005; Henchoz et al., 2013).

Such phenomena and explanations would be consistent with what is already mentioned regarding musculoskeletal dysfunctions in chronic non-specific LBP. In the different studies applying the flexion relaxation phenomenon it has been investigated and used in different ways but according to Alschuler et al. (2009) none has tested the validity of the different approaches. They also concluded that future studies where needed to address and examine which way(s) of measuring the flexion relaxation phenomenon would be best at distinguishing the symptomatic population from the asymptomatic population, thus accounting for its reliability and validity. This was also stated by Geisser et al. (2005) in their review.
Electromyography

Physiology

Electromyography (EMG) is useful to gain an insight in the muscles whether concerned with pathology, dysfunction or movements related to simple postural task or training.

An EMG signal is the electric signal picked up as a result of a muscle contraction (concentric, eccentric or isometric), which is produced by the motor unit action potential in the muscle membrane. This action potential can only be produced when the cell membrane has passed its threshold limit at about - 55 microvolts (mV). Normal resting potential at the muscle fiber membrane is between - 80 to - 90mV (Schibye and Klausen, 2003; Noraxon, 2005). For illustration, please see Figure 1.
Figure 1. Illustration of motor unit.

(Austin Community College (with permission), 2016).
The demand for an increased muscle activity level for a certain movement to take place is based on a recruitment process with an increasing number of motor units. Smaller motor units are recruited first and as the force requirements increase successively, larger units are recruited. Other parameters to take into account with increasing muscle activity demands are fiber type, speed of contraction, involvement of reflexes and most likely interpersonal differences (Sundhedsstyrelsen, 2011). During resting periods there should be no electrical activity in the muscles. However, there is some activity defined as “background noise”, which is explained to be caused by some of the chemical processes related to the cell membrane. This “background noise” should not exceed 10-15 mV (Shumway - Cook and Woollacott, 2001; Noraxon, 2005).

**Types of Electromyography**

There are two well described method of using EMG in the literature. Surface EMG (SEMG) or intramuscular EMG (needle- and fine-wire). The latter of these is invasive since inserted in the muscle bulk and has been evaluated to only providing a very local picture of muscle activity. SEMG on the other hand is mounted directly on the skin which makes application easier, less invasive providing a broader picture of muscle activity. However the downside of this method is that is has proven to be quite sensitive to the adjacent muscles and result in “Cross Talk” (Stokes *et al.*, 2003; Noraxon, 2005). The best solution to choose between either of these should be based on purpose. Use IEMG would be appropriate if the purpose was to measure activity in the multifidus muscle at a certain point such as L4. If the purpose is to examine general activity in the muscles, it could be a reasonable rationale to use SEMG. Another relevant more ethical consideration is related to using a less invasive procedure in a chronic population.
All the studies identified in the initial search using EMG applied SEMG. In general, the methodology used in different studies vary considerably regarding, number of electrodes, placing of electrodes and how the raw data material is analysed afterwards. This is critical and makes it difficult to draw conclusion and generalize. There is a need for a standardization as suggested by the Seniam group (Seniam, 2016).

**Background Conclusions and Research Aim**

Based on what is known about non-specific LBP it is the purpose of this systematic review to gain further insight into and establish a correlation between trunk muscle activity during forward bending and the subgroup of chronic non-specific LBP patients to inform and guide both clinicians and researchers dealing with this field of interest. The methodological strengths of a systematic review if conducted properly will be able to provide answers regarding the field of interest and on which foundation the answers are based.
3. Method

Systematic Review Study Design

In general, the purpose of a systematic review is to identify gaps in the literature and suggest where future scientific research should be targeted (Blaxter, Hughes and Tight, 2010; Dahlberg and McCaig 2010). The systematic review is considered the “gold standard” and should be conducted in a systematic, transparent and explicit way to ensure objectivity, replicability and appraisal of quality (Boland, Cherry and Dickson 2014). The need for a systematic review is based on the premise that the health care decision should be based on best available evidence. However, the amount of data in recent years makes it almost impossible for clinicians and researchers to keep up with the latest and best evidence (Aveyard, 2014).

Conclusions from single studies such as randomized controlled trials (RCT) or case-control studies often have conflicting results. Results are often attributed to methodological issues such as quality of studies or simply misinterpretation or
inability of the reader to critically appraise the studies or even worse by conflicts of interest including financial interests. The systematic review is important because the purpose is to summarise the evidence and evaluate the coherence of arguments including a critical analysis rather than just being descriptive and thereby assisting clinician and researcher (Bowling, 2002; CRD, 2009; Aveyard, 2014).

The term “gold standard” is related to the hierarchy of evidence which places systematic review conclusions based on randomized controlled trials (RCT) at the top (level 1) and expert statements at the bottom (Greenhalgh, 2014).

The aim of conducting this systematic review was to explore if the current literature could provide an answer to a clinically relevant subject such as how muscle activity is correlated to adult patients with chronic non-specific LBP presenting with psychosocial characteristics during ADL such as forward bending.

Registration

As this systematic review has the form of a dissertation it will not be sent for registration at an international database as PROSPERO. However, it would be good scientific practice to do so to avoid discrepancy between the intention of the review and what is actually reported during publication (National Institute for Health Research, 2016).

The Research Question

The method of a systematic review will be applied to analyse trunk muscle activity during forward bending in standing position in adult patients with chronic LBP presenting with psychosocial characteristics such as fear avoidance, pain catastrophizing or kinesiophobia. The research question has been guided by inspiration from the FINER criteria (Farrugia et al., 2009).
**Development of Research Question and Search Strategy**

In the development of the research question, which was guided by inspiration from the FINER criteria which is related to the subject of interest, its clinical relevance and ethical issues (Farrugia *et al.*, 2010) and the subsequent literature search, the PICO format was applied.

PICO is an abbreviation for Population (P), intervention (I), comparison (C) and outcome (O). Being very clear on the inclusion and exclusion criteria and the different components of the PICO will of course have a positive effect on the internal validity of this systematic reviews and reduce risk of bias but may affect external validity and thus the generalisability of the findings to clinical practice and vice versa (Farrugia *et al.*, 2010).

**PICO**

**Population:**

Patients with chronic non-specific LBP with either characteristics of fear avoidance, kinesiophobia or catastrophizing behavior.

**Intervention / Testprotocol:**

The test protocol involving standing flexion. The literature frequently describes the flexion relaxation phenomenon or flexion relaxation ratio as a relevant test for muscle activity in both asymptomatic people and symptomatic people with LBP.

**Comparison:**
An ideal control group would be a subgroup of patients with non-specific LBP without psychosocial characteristics such as fear avoidance behaviour. However, a control group is not necessarily relevant to include since the research question is to identify a correlation between a subgroup of patients with non-specific LBP with psychosocial characteristics and muscle activity. Thus, not to identify between subgroup differences, but to relate one subgroup related to a phenomenon.

**Outcome:**

Trunk muscle activity based on surface electromyography (SEMG) and validated questionnaires screening for psychosocial characteristics.

Trunk muscles of relevance are defined according to literature regarding SEMG including Noraxon (2005) and Seniam group recommendations (Seniam, 2016). The selected muscles are responsible for movement and postural adjustments of the trunk in standing positions. Abdominal muscles include transversus abdominis, rectus abdominis, internal obliques and external obliques. Back muscles include longissimus, iliocostalis, multifidus and erector spinae.

Questionnaires used to screen for psychosocial characteristics such as pain-related fear will be the Fear Avoidance Belief Questionnaire or the Tampa Scale of Kinesiophobia.

**Inclusion and Exclusion Criteria**

**Included Study Designs**

The included studies will not be from the top of the hierarchy of evidence since the RCT study design is not suited to answer the objective of this systematic review. The objective is investigation of correlation between variables at a particular point in time.
and not the development of the phenomenon or the effect of an intervention (Jørgensen, Christensen and Kampmann, 2005; Habicht, 2011).

However, if the purpose was to estimate if psychological parameters affected the flexion relation phenomenon and more specific the flexion relaxation ratio an observational study as for example a prospective cohort study would be ideal (Thise, 2014).

RCTs are suited to measure the effect of an intervention for example by comparing two different interventions or by comparing a group receiving an intervention and a control group of “usual care”/no intervention (Aveyard, 2014). RCTs are the best way of reducing bias, especially allocation bias. However, this should not be mistaken with a total elimination of bias (Hagen et al., 2008). It is unlikely that an RCT will contribute to answer the objective of the systematic review but it will be included if found relevant. A RCT could be relevant if the demographic data and outcome measures were correlated at baseline prior to an intervention.

Study designs included are observational studies similar to for example case-control, cross-sectional studies and case series (Bowling, 2002; CRD, 2009; Habicht, 2011).

Song and Chung (2010) and Chung et al. (2009) defined these as a subgroup of analytical studies with evidence level of 2 or 3 (Chung et al., 2009). Observational studies are criticized for being susceptible to bias due to a lower ability to control confounding factors thereby making it less obvious how variables affect each other. Though there is reason to believe that this can be adjusted for when knowing limitations and risks when assessing their quality (CRD, 2009; Song and Chung, 2010). Another criticism of observational study designs especially those without a control group has been problems related to consistently overestimating treatment
effect (Hagen et al., 2008). When comparing observational studies to RCTs, Benson and Hartz (2000) and Concato, Shah and Horwitz (2000) found no such systematical overestimation to the effects of treatment.

Based on these study design(s), the MOOSE guideline (Stroup et al., 2000) will be used as a guide and structure for the format of this systematic review.

**Language of Included Studies**

Only English language studies will be included. This might increase the risk of a language bias since publications in other languages could have an effect on the results. However, since non-specific LBP is an international epidemic it is fair to assume that studies will be submitted for publication in English language journals. This is supported by Galandi, Schawarzer and Antes (2006) who showed a decline in the number of German language publications. Cochrane suggested a case-to-case approach whether searches should be limited to English language studies only (Higgins and Green, 2011). Finally, the inclusion of English language studies was also a result of the limited resources of the reviewer.

**Measurement - Surface Electromyography**

Studies that used SEMG were included. This is the most common method of measuring the flexion relaxation phenomenon and the initial search of the literature revealed very limited studies using fine wire electromyography (Noraxon, 2005) or ultrasonography. Importantly, SEMG is suggested to be a useful tool to detect dysfunction associated with LBP (Alschuler et al., 2009). Another consideration could be that SEMG is a non-invasive procedure, which from an ethical and confounding point of view could be relevant. The argument is related to what is known about muscle activity in patients with chronic non-specific LBP, namely that just the
introduction to something fearful (such as movements or the application of a painful peripheral mechanical trigger such as a needle pierced into skin or muscle) will affect muscle activity (Glombiewski et al., 2015); this could give a less precise measurements of the activity intended to be investigated.

**Participants**

Adults with chronic non-specific LBP between 18 and 65 years are included. Low back pain is mainly reported in people from the early twenties to mid sixties with a peak prevalence from 35-55 years (van Tulder et al., 2006). There is, however, an increasing prevalence reported already from late adolescence (O’Sullivan and Lin, 2014). Restricting inclusion to the adult chronic population will increase internal validity. Exclusion of the acute population is based on data that suggesting changes in the psychosocial and the biological factors when comparing the acute to the chronic population. This suggests that the two groups have different presentations (Morsø et al., 2014) and that it is the transformation from acute to chronic and a better understanding of the chronic population that seems to be the cardinal points (Lind, 2011). Most acute patients seem to be able to spontaneously recover themselves within six to eight weeks (van Tulder et al., 2006). However, exposure to active treatment versus passive treatment modalities shows that those being active have a better outcome but only related to level of pain and not disability (DIHTA, 1999; Machado et al., 2006). It is important to keep in mind that “active” and “passive” treatment is very unspecific.

Background data supported the connection between behaviour and physical manifestation. The inclusion of a minimum of at least one validated questionnaire addressing psychosocial issues and characteristics such as pain related fear
behaviour and kinesiophobia (Waddell et al., 1993; Woby et al., 2005) is highly relevant. The literature supports the existence of validated questionnaire regarding these subjects in order to stratify treatment and predict prognosis (Hill et al., 2011; Morsø et al., 2013). Miles et al. (2011) also showed that the above factors played a strong role to predict outcome and a patient's ability to self manage.

**Types of publications**

Data published in peer-reviewed journals or presented as posters and conference presentations, abstracts and full text articles was searched. However, the relevance and quality of data was evaluated independently. The choices were made to ensure a sensitive search and to reduce publication bias (Bogduk and McGuirk, 2006).

**Exclusion Criteria**

**Measurement**

Studies that used other objective measurement tools to record muscle activity than SEMG were excluded. First of all because the flexion relaxation phenomenon is based on SEMG measurement (Nemblett, et al., 2013). Ultrasonography might be used clinically but it does not has the ability to quantify muscle contraction (activity), but merely to demonstrate some or none activity (lateral slide or thickness changes) based on the reflection of sound waves that form the image of the targeted structure. An establishment of the relationship between change in muscle size/thickness (ultrasonography) and level of muscle activity in (electromyography) is at best inconclusive (Hodges, 2005; Brown and McGill 2010). This would cause a problem regarding quantification in a possible meta-analysis. A pitfall in applying multiple measurement tools is that SEMG itself has limitations regarding reliability when it
comes to for example placement of electrodes (Noraxon, 2005). Adding another measurement tool would increase the possibility of a less trustworthy conclusion to the review. Finally, application of the ultrasound probe / transducer during a forward flexion from standing seems almost impossible. Especially on the abdominal muscles when the distance between the abdomen and the thighs is hardly there at full flexion.

**Unreported literature**

Unreported literature was not searched in this review since it was not feasible in practice to locate it. This matter could potentially result in some uncertainty and could constitute publication bias. An example of publication bias is when what is termed “negative” results meaning the expected outcome was not met and the author then does not attempt publication. Another possibility is that the “negative” results are less likely to be accepted for publication than “positive” results (Dickersin *et al*., 1987; Granquist, 2015). A reason for rejection for publication could be due to poor quality of the submitted study. If this is not attended to and included in the review, it could affect the final results and thus its validity of the review (Hagen *et al*., 2008). In addition, it would not have been feasible to contact all institutions and authors around the world for unpublished work and in languages that the reviewer is not familiar with (Maher *et al*., 2004).

**Participants**

This systematic review investigates chronic non-specific LBP which means a logical exclusion of the groups with back pain that can be diagnosed with confirmed structural cause of their back pain (X-ray or MRI confirmed fracture, ankylosing spondylitis etc.). The same applies to chronic and painful somatic conditions of a malignant nature such as cancer (Lind, 2011).
Pregnant women were excluded since it was reasonable to assume that they probably would have limited ability to reach full flexion (Gilleard, Crosbie and Smith, 2002).

**Search Strategy**

A three-step search strategy based on the research question with focus on population and outcome elements was applied in this systematic review, merely because the research question was not aiming at comparing interventions. Before initiating the actual three-step search strategy PROSPERO was searched to see if a review attending the subject had already been registered. This revealed 203 records but none regarding the subject of interest was registered (PROSPERO).

Databases searched in this systematic review was MEDLINE via Pubmed, EMBASE via ELSEVIER, CINAHL plus full text via EBSCOHOST and The Cochrane Library. These databases where selected based on their relevance to the subject of interest since these databases cover the disciplines of biomedicine, biotechnology and the health system in general (Hagen et al., 2008). They are also very extensive databases and this combination would result in references from journal articles and textbooks securing a certain level of sensitivity (Boland, Cherry and Dickson, 2014; Aveyard, 2014). Searching The Cochrane Library in particular also had the purpose of identifying if the review question had been investigated already. The exclusion of PEDro, which covers the discipline of physiotherapy, and other databases in the search might seem precarious but was based on the author’s decision that the included databases would fulfill the purpose of this review. Tsertsvadze *et al.* (2015) supported this in their study which showed that inclusion of other databases than a combination of EMBASE, MEDLINE and The Cochrane Library only identified 2.4%
more studies. Finally it was due to the given resources available with only one reviewer.

The search and retrieval of literature was conducted from either the University of Dundee Library resources or University College North Denmark library resources. An initial search conducted on 11.12. 2015 using the above mentioned databases was carried out with the purpose of covering the background of the subject of interest and identifying further keywords retrieved from either abstract or full text. Moreover to make sure that a systematic review had not already been conducted. This initial scoping of the literature was done with a time frame from 2005 to December 2015. The reason for this 10-year time frame was based two things. 1) The subject to be investigated in the systematic review has gained more interest during the last 10 years along with the paradigm shift within the profession of physiotherapy so it was reasonable to expect an increasing number of publications. 2) If a systematic review had been conducted prior to 2005, it seemed reasonable to conduct an updated search based on more recent publications. This initial search was done only by using thesaurus terms and in a combination with the boolean operators OR or AND. The search revealed 112 results and no systematic reviews were based on the research question.

Table 3. Initial search 11.12. 2015
The search (Table 3) revealed further keywords to be included in the development of the second search;

- Neuromuscular control
- Neuromuscular adaptation
- Flexion relaxation phenomenon
- Flexion relaxation ratio
- Electromyography
- Anxiety

Based on the initial scoping of the literature and the amount of data identified combined with the period of time this subject has been of interest in musculoskeletal physiotherapy, it was decided to search data from inception to present time in the second search in all the included databases. In Table 4 inclusion of the above mentioned keywords is applied to the search.

The use of keywords and index terms will most likely reduce the risk of selection bias. The literature seems rather inconsistent concerning terminology both with regards to its use of abbreviations and definitions, especially in relation to the column of “Intervention” and “Outcome” (Table 4). To avoid missing relevant data that could lead to a less reliable conclusion, a robust search in each included database with the purpose of not missing relevant data was made by a combination of searching

<table>
<thead>
<tr>
<th>Low back pain OR back pain OR chronic back pain</th>
</tr>
</thead>
<tbody>
<tr>
<td>AND</td>
</tr>
<tr>
<td>Motor control OR EMG OR muscle activity</td>
</tr>
<tr>
<td>AND</td>
</tr>
<tr>
<td>Fear avoidance OR kinesiophobia OR catastrophization</td>
</tr>
</tbody>
</table>
keywords, thesaurus terms and truncation. As shown in Table 4, the application of truncation on the abbreviation EMG was made to ensure that the inconsistency not affected sensitivity negatively. The second search was conducted between 17.03.2016 and 22.03.2016.

Table 4. Search terms - second search
The MEDLINE search was initially too specific as it only revealed 3 studies. To avoid missing relevant papers an approach to increase sensitivity by retrieving the citations

<table>
<thead>
<tr>
<th>Population</th>
<th>Intervention/Test</th>
<th>Comparison</th>
<th>Outcome</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. low back pain[Mesh]</td>
<td>11. motor control</td>
<td>30. electromyograph y</td>
<td></td>
</tr>
<tr>
<td>2. back pain[Mesh]</td>
<td>12. motor control*</td>
<td>y[Mesh]</td>
<td></td>
</tr>
<tr>
<td>3. chronic low</td>
<td>13. “motor control”</td>
<td>31. electromyograph y</td>
<td></td>
</tr>
<tr>
<td>Low back pain</td>
<td>14. muscle activity</td>
<td>y*</td>
<td></td>
</tr>
<tr>
<td>4. back pain*</td>
<td>15. muscle activity*</td>
<td>32. electromyograph y</td>
<td></td>
</tr>
<tr>
<td>5. low back pain*</td>
<td>16. “muscle activity”</td>
<td>y</td>
<td></td>
</tr>
<tr>
<td>6. chronic low</td>
<td>17. neuromuscular control</td>
<td>33. EMG</td>
<td></td>
</tr>
<tr>
<td>back pain*</td>
<td>18. neuromuscular control*</td>
<td>34. EMG*</td>
<td></td>
</tr>
<tr>
<td>8. back pain</td>
<td>20. neuromuscular adaptation</td>
<td>36. combine #30-35</td>
<td></td>
</tr>
<tr>
<td></td>
<td>21. neuromuscular adaptation*</td>
<td>using OR</td>
<td></td>
</tr>
<tr>
<td></td>
<td>22. “neuromuscular adaptation”</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>23. flexion relaxation phenomenon</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>24. flexion relaxation phenomenon*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>25. “flexion relaxation phenomenon”</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>26. flexion relaxation ratio</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>27. flexion relaxation ratio*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>28. “flexion relaxation ratio”</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>29. Combine #11-28 using OR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>30. electromyograph y</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>31. electromyograph y*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>32. electromyograph y</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>33. EMG</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>34. EMG*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>35. “EMG”</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>36. combine #30-35 using OR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>37. fear[Mesh]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>38. fear avoidance</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>39. fear avoidance*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>40. “fear avoidance”</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>41. kinesiophobia</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>42. kinesiophobia*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>43. “kinesiophobia”</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>44. catastrophization [Mesh]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>45. catastrophization</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>46. catastrophization*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>47. “catastrophization”</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>48. anxiety[Mesh]</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>49. anxiety</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>50. anxiety*</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>51. “anxiety”</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>52. combine #37-51 using OR</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>53. combine with AND</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>#9#29#36#52</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
at the second last step in the search was adopted which meant combining #9#29#52 (Table 4) (Bogduk and McGuirk, 2006).

The MEDLINE search is presented here.

((("Back Pain"[Mesh]) OR "Low Back Pain"[Mesh]) OR "chronic low back pain") OR chronic low back pain) OR back pain) OR back pain*)) AND (((((motor control) OR "motor control") OR motor control*) OR muscle activity) OR "muscle activity") OR "neuromuscular control") OR neuromuscular control*) OR "neuromuscular adaptation") OR neuromuscular adaptation*) AND ((((("Fear"[Mesh]) OR Fear avoidance) OR "fear avoidance") OR fear avoidance*) OR kinesiophobia) OR "kinesiophobia") OR kinesiophobia*) OR "Catastrophization"[Mesh]) OR Catastrophization) OR "Catastrophization") OR Catastrophization*) OR "Anxiety"[Mesh]) OR (anxiety)) OR "anxiety ") OR anxiety *)

An example from the EMBASE search and the CINAHL search can be found in appendix 1.

Finally, reference lists of included studies were searched for additional literature. Due to time constraints this process only took place in the included studies. Where full text or abstract or other data such as posters were found relevant to retrieve but not available the authors were contacted by email. There was one example of email-correspondence (Appendix 2) with an author of an abstract where data was not available to retrieve (Glombiewski et al., 2011). This revealed a poster presentation which was later excluded as it did not match the inclusion criteria of the systematic reviews search strategy.
In addition, grey literature was manually searched in the reviewer’s personal article database but this only resulted in more background literature. The search strategy applied to the grey literature search could affect selection bias since the author might have a selection of data that was not representative of the results related to the subject but instead based on what the author had retrieved based on personal perspective and interests (Juul, 2006).

**Selection Process of Included Studies**

A manual de-duplication procedure was then executed to avoid the same study being included more than once. Results of the process are demonstrated using a PRISMA flow diagram mapping out the process from identification, screening, eligibility, exclusion and final inclusion (Moher et al., 2009). The study selection process is detailed in Figure 2.

After this step another exclusion of studies was performed after screening through titles and abstracts. The last step in the selection process required retrieval of a number of full text studies, which were read for eligibility in relation to the systematic review. This was done either because they matched the criteria or because the abstract and title screening was not sufficient to clarify their relevance.

Table 5 is an example of the items that were included in this last step of the process. During the process, Table 5 helped the reviewer to maintain structure in the screening and retrieval of studies. Table 9 in the result section shows studies retrieved for eligibility. An explanation for exclusion of studies in combination with Table 9 contributes to increased transparency of the inclusion process. Included studies in Table 9 are marked with a Y and excluded studies with an X.
Access to these full text studies was provided at University of Dundee Library or University College North Denmark library resources.

All searched studies were independently assessed for inclusion by the author of the review. Ideally, this process should have included two independent reviewers reaching for consensus to avoid rejecting relevant studies as suggested by both Liberati et al. (2009) and Higgins and Green (2011). These recommendations were supported by Edwards et al. (2002). Edwards presented data demonstrating that single a reviewer missed on average 8% of eligible material.

Table 5. Example of table used when articles were retrieved for eligibility
<table>
<thead>
<tr>
<th>ITEMS</th>
<th>Population</th>
<th>Intervention</th>
<th>Comparison</th>
<th>Outcome</th>
<th>Study type</th>
<th>Language</th>
<th>Included</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study Title and author(s)</td>
<td>Low back pain, Chronic low back pain &amp; fear, fear avoidance, kinesiophobia, anxiety, catastrophization</td>
<td>Motor control, muscle activity, neuromuscular control or adaptation, flexion relaxation phenomenon or ration</td>
<td>Control-group, Cluster of high vs low fear avoidance</td>
<td>Electromyography, EMG &amp; fear, fear avoidance, kinesiophobia, anxiety, catastrophization</td>
<td>Meta analysis Systematic review, Randomized controlled trials, Cohort, Case control, Case series, Abstract or posters &amp; Textbooks</td>
<td>English</td>
<td>Y=YES N=NO</td>
</tr>
<tr>
<td></td>
<td>Y=YES N=NO</td>
<td>Y=YES N=NO</td>
<td>Y=YES N=NO</td>
<td>Y=YES N=NO</td>
<td></td>
<td>Y=YES N=NO</td>
<td>Y=YES N=NO</td>
</tr>
</tbody>
</table>

1.  
2.  
3.  
4.  
5.  
6.  
Ongoing

Quality Assessment

Quality assessment or critical appraisal is a part of the review process that has several important implications. From an ethical perspective, it should be considered that if the systematic review should justify the resources spent by both the author, librarians and supervisors, the review process should be as reliable as possible. The results of the systematic review are more likely to be reliable and have an impact on future research and clinical practice decisions if the studies included in the review
are of high enough quality. Moreover, quality should at least be presented in a
transparent way. Quality assessment could briefly be described as accessing the
weaknesses and strengths of included studies (Aveyard, 2014). If this part of the
systematic review is poorly executed, it will hamper the validity of the systematic
review. Quality assessment also serves as a final exclusion of papers either because
of low methodological quality or because thorough reading reveals that they can not
help answer the research question despite being included in the review process as
they were found relevant for the objective of the systematic review (Burls, 2009).
Quality assessment requires a reasonable amount of subjective judgement which of
course could constitute a bias concerning the author’s perspective. However, the
subjective judgement also help to avoid rejection of a study merely based on its
quality (Higgins and Green, 2011).

To overcome and minimize the risk of selection bias caused by the author’s
subjectivity, the quality assessment was carried out systematically using a quality
assessment tool matching the study designs. Selection of assessment tools for this
study was discussed with one of the supervisors at the University of Dundee.

The Newcastle Ottawa Quality assessment Scale was chosen for cohort and case-
control studies (Wells et al., 2014) and suggested by Higgins and Green (2011) and
Cook (2015). Furthermore, the scale has been applied in other systematic reviews
based on observational studies such as the one by Andrade et al. (2015), who
investigated the association between LBP, spondylolysis and spondylolisthesis. The
modified Newcastle-Ottawa Quality assessment Scale for cross-sectional studies
was used in cross-sectional studies, modified and applied by Herzog et al. (2013)
when chosen for this design.
Background reading revealed that most observational studies did not report their study design which is supported by Vandenbroucke et al. (2014). As a result of that the studies by Mann (2003), Song and Chung (2010) and Dekker et al. (2012) were used to categorize the study design if needed.

The Newcastle Ottawa Quality assessment Scale consists of three categories regardless of study designs however but differences in the third category.

- Selection
- Comparability
- Exposure (case control) - Outcome (cohort and cross-sectional )

Each of the categories are scored separately. A study can be awarded a maximum of one “star” for each numbered item within the Selection and Exposure / Outcome categories. A maximum of two stars can be given for Comparability (Wells et al., 2014). However, the scoring system using stars (from 0 - 9 stars) did not present with cut off scores with regard to level of study quality, “poor” or “good” quality. In the study by Park et al. (2015) a high quality score was defined as a score $\geq 7$ (stars). This cut off score was used in this systematic review

Examples of The Newcastle Ottawa Quality assessment Scales can be found in appendix 3.

Results from the quality assessment will be summarized and presented both narratively and in a table. The table format is based on the categories from The Newcastle Ottawa Quality assessment Scale. Table 6 and 7 presents templates of the quality assessment of case-control studies and of cross-sectional studies.
Running a pilot with The Newcastle Ottawa Quality assessment Scale (s) could have provided further familiarity with the strengths and limitations of each tool however this was not initiated.

Observational studies have traditionally been found to be accused less trustworthy and therefore ranked below randomized controlled trials in the hierarchy of evidence. To some extent this seems right due to their susceptibility of bias regarding confounders affecting the results in observational studies. This is mainly due to the randomization procedure. However, the design allows for investigation of certain questions that a randomized controlled trial would not be able to answer keeping ethical considerations in mind (Mann, 2003). It must also be kept in mind that randomized controlled trials or any other type of study design are no better than the rigor and transparency they are based on.

The quality assessment will be focused on how the population is representative of the objective of the study and how the population and control (if controls) are recruited and matched. This includes details related to demographics. This will form the premise of being able to generalise the study findings and is a way of accounting for confounders.

Regarding confounders the assessment also considered the application of statistical methods, (i.e. tests for variance such as ANOVA) within the studies to control for variance between variables and groups. The study designs also made it relevant to assess for reporting of odds ratio and risk ratio. The use of odds ratios included use of confidence intervals and p-values.
However, reporting of it will depend on several things such as objective of the studies which will dictate the relevance of it (Juul, 2006).

Sample size and estimation of power was assessed. It is, however, a difficult matter that depends on several aspects such as the variability of the population and the extent to which an error is accepted by the researchers conducting the study (Kamangar and Islami, 2013). Justification of sample size was as a minimum investigated in this review.

Finally, reporting of results and stated outcome(s) were assessed along with a conclusion of the overall study quality based on the rating of the Newcastle Ottawa Quality assessment Scale and the Newcastle Ottawa modified scale.
Table 6. Quality assessment of case-control studies

<table>
<thead>
<tr>
<th>ITEMS</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection</td>
<td>1. Case definition adequate</td>
</tr>
<tr>
<td></td>
<td>2. Representativeness of the cases</td>
</tr>
<tr>
<td></td>
<td>3. Selection of controls</td>
</tr>
<tr>
<td></td>
<td>4. Definition of controls</td>
</tr>
<tr>
<td>Comparability</td>
<td>1. Comparability of cases and controls on the basis of the design or analysis</td>
</tr>
<tr>
<td>Exposure</td>
<td>1. Ascertainment of exposure</td>
</tr>
<tr>
<td></td>
<td>2. Same method of ascertainment for cases and controls</td>
</tr>
<tr>
<td></td>
<td>3. Non-response rate</td>
</tr>
<tr>
<td>Number of *</td>
<td>Low - medium - high</td>
</tr>
</tbody>
</table>

* = satisfactory description
Table 7. Quality assessment of cross-sectional studies

<table>
<thead>
<tr>
<th>ITEMS</th>
<th>Study</th>
</tr>
</thead>
<tbody>
<tr>
<td>Selection</td>
<td>1. Representativeness of the sample</td>
</tr>
<tr>
<td></td>
<td>2. Sample size</td>
</tr>
<tr>
<td></td>
<td>3. Non-respondents</td>
</tr>
<tr>
<td></td>
<td>4. Ascertainment of the exposure</td>
</tr>
<tr>
<td>Comparability</td>
<td>1. The subjects in different outcome groups are comparable, based on</td>
</tr>
<tr>
<td></td>
<td>the study design or analysis. Confounding factors are controlled.</td>
</tr>
<tr>
<td>Outcome</td>
<td>1. Assessment of outcome</td>
</tr>
<tr>
<td></td>
<td>2. Statistical test</td>
</tr>
<tr>
<td>Number of *</td>
<td>Low - medium - high</td>
</tr>
<tr>
<td>Overall quality score</td>
<td>Low - medium - high</td>
</tr>
</tbody>
</table>
Data Extraction

The data extraction process was carried out by the author using a data extraction form suited for the purpose of collecting the dataset needed to answer the research question. The data was extracted on the 09.05.2016. The same form was used to ensure consistency in the process (CRD, 2009). It was decided to extract data from all the included into a single form instead of using one form per study. Table 8 shows a template of the final data extraction form.

The decision to use one form for data extraction was made despite the likelihood that more than one kind of study design would be included after searching the literature. Using one form to include data from multiple studies and study designs was found reasonable as long as the design was adequate. This was based on the premise that the data of interest was the same independently of study designs. The design of the form was based on criteria from Higgins and Green (2011) and inspired further based on items elicited in the background reading. Examples of such items were subgrouping of “low and high” pain-related fear and disability score. These are highly relevant regarding generalisability of findings. Another item was “full flexion” used as a synonym for “flexion relaxation phenomenon” or “flexion relaxation ratio”.

Importantly, it should be designed to ensure that the author could achieve information both regarding narrative and quantitative reporting later on in the review process. During the process of designing the form it was decided to add a column to the extraction form consisting of the item “limitation / discussion”. This was added to provide the author with some perspective on the data and the forthcoming thesis.
It was also decided that if the included studies examined multiple outcomes only the outcomes related to the research question of the systematic review were extracted. Data extraction is a very vulnerable process regarding consistency and it also carries a major risk of introducing bias. Data extraction is therefore advised to be performed by more than one person to minimize bias according to Higgins and Green (2011).

**Table 8.** Data extraction form for studies with cross-sectional and case-control designs

<table>
<thead>
<tr>
<th>Data extraction form</th>
<th>Data extracted by;</th>
<th>Date for data extraction;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study Author and year</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Item</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Study design</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subjects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Age (Mean)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Duration of symptoms</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Definition of chronic</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Sample size</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Setting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control subjects</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Age</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Gender</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Definition of control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Sample size</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Setting</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain related fear questionnaire</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low / high pain related fear</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disability / pain questionnaire</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Data extraction form</td>
<td>Data extracted by;</td>
<td>Date for data extraction;</td>
</tr>
<tr>
<td>----------------------</td>
<td>--------------------</td>
<td>---------------------------</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Low / high disability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMG back</td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMG abdominal</td>
<td></td>
<td></td>
</tr>
<tr>
<td>BP / BW</td>
<td></td>
<td></td>
</tr>
<tr>
<td>FRP / FRR / FF</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Statistical method</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Normalization</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ethics approval</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Results</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Limitation/ discussion</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

4. Presentation and Analysis of Findings

Study Selection

The result of the second search revealed 211 studies; one additional result was found by hand searching the reference lists of the obtained studies. After duplication procedure the number of studies were 164 and after abstract and title screening it were 23. Finally after full text screening it were down to six records reporting on five studies. The references from the second search were divided in the following way:

- The Cochrane Library; 43 references
- EMBASE; 34 references
- CINAHL plus with full text; 53 references
- MEDLINE; 81 references

The selection process is described in the PRISMA flowchart in Figure 2.
Figure 2. PRISMA flowchart
The systematic search for studies to answer the research question resulted in 23 articles assessed for eligibility (Table 9).

The studies highlighted in blue and a “yes” (Y) in the included column indicate which studies were included.

To avoid any misunderstanding when interpreting Table 9, it must be clarified that a yes (Y) in the first 6 columns was not an argument for inclusion but merely an indicator that the study included the items of interest based on the PICO format (Farrugia et al., 2009).
**Table 9. Articles retrieved for eligibility**

<table>
<thead>
<tr>
<th>ITEMS</th>
<th>Population</th>
<th>Intervention</th>
<th>Comparison</th>
<th>Outcome</th>
<th>Study type</th>
<th>Language</th>
<th>Included</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study Title and author(s) and year of publication</td>
<td>Low back pain, Chronic low back pain &amp; fear, fear avoidance, kinesiophobia, anxiety, catastrophization</td>
<td>Motor control, muscle activity, neuromuscular control or adaptation, flexion relaxation phenomenon or ration</td>
<td>Control-group, Cluster of high vs low fear avoidance</td>
<td>Electromyography, EMG &amp; fear, fear avoidance, kinesiophobia, anxiety, catastrophization</td>
<td>Meta analysis Systematic review, Randomized controlled trials, Cohort, Case control, Case series, Abstract or posters &amp; Textbooks</td>
<td>English</td>
<td>Y=YES N=NO</td>
</tr>
<tr>
<td>1. Thomas J.S. 2008</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Cross sectional</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>2. Abboud, J. 2014</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Case control</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>3. Vincent, H. 2013</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Cross sectional</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>4. Smeets, R. 2009</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Case control</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>5. Ledoux, E. 2012</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Case control</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>6. Astfalck, R.G. 2010</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Cohort</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>7. Mannion, A.F. 2011</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Cross sectional</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>8. Larivière, C. 2010</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Case control</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>9. Dubois, J.D. 2014</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Cross sectional</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>10. Henchoz, Y. 2013</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Cross sectional</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>11. Watson P.J. 1997</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Cross sectional</td>
<td>Y</td>
<td>Y</td>
</tr>
<tr>
<td>12. Glombiewsky. 2011</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Y= YES N=NO</td>
<td>Conference poster</td>
<td>Y</td>
<td>N</td>
</tr>
<tr>
<td>Study Title and author(s) and year of publication</td>
<td>Low back pain, Chronic low back pain &amp; fear, fear avoidance, kinesiophobia, anxiety, catastrophization</td>
<td>Motor control, muscle activity, neuromuscular control or adaptation, flexion relaxation phenomenon or ration</td>
<td>Control-group, Cluster of high vs low fear avoidance</td>
<td>Electromyography, EMG &amp; fear, fear avoidance, kinesiophobia, anxiety, catastrophization</td>
<td>Meta-analysis, Systematic review, Randomized controlled trials, Cohort, Case control, Case series, Abstract or posters &amp; Textbooks</td>
<td>English</td>
<td></td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td></td>
</tr>
<tr>
<td>14. Svendsen, J. 2013</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Case control</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>15. Alschuler, K. 2009</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Cross sectional</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>16. Vandamme, 2014</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Case control</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>17. Hedayati, R. 2014</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Case control</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>18. Larivière, C. 2013</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Cross sectional</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>19. Der Hulst, 2010</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Cross sectional</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>20. Thomas, J. 2008</td>
<td>Y</td>
<td>Y</td>
<td>Y (low vs high fear)</td>
<td>Y</td>
<td>Cross sectional</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>21. Masse-Alarie, H. 2016</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Case control</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>22. Unsgaard-Tøndel, 2013</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Y</td>
<td>Randomized Controlled trial</td>
<td>Y</td>
<td></td>
</tr>
<tr>
<td>23. Rampasad, M. 2011</td>
<td>Y</td>
<td>Y</td>
<td>N</td>
<td>Y</td>
<td>Cross sectional</td>
<td>Y</td>
<td></td>
</tr>
</tbody>
</table>
Reasons for Exclusion after Assessment of Eligibility

The exclusion of 16 out of 18 studies at this step was based on the inclusion criteria. (*= The same study)

Exclusion was mainly caused by:

• Not involving a test of full trunk flexion (i.e. flexion relaxation phenomenon was not measured). (n = 4)
• Exposure of the symptomatic population with films or photos which actually test expectations instead of the actual correlation between muscle activity and movement. (n = 1)
• Use of ultrasonography or intramuscular electromyography instead of surface electromyography. (n = 1*)
• Testing of paraspinal and abdominal muscle activity based on isometric or endurance / fatigue testing (usually flexion and extension directions) and not full flexion. (n = 8)
• Studies measuring effect of intervention such as randomized controlled trials where no pre intervention data were reported relevant for this systematic review. (n = *)
• Testing of paraspinal and abdominal muscle activity in activities of daily function such as walking (n = 2)

The study by Svendsen et al. (2013) (1 out of 18) was excluded as the population studied was subacute. Correspondence with one of the authors revealed considerable bias regarding flawed data due to change of study design and closing of the experiment prior to what was originally planned (appendix 4). The exclusion of
this study was finally made following discussion with one of the supervisors at University of Dundee.

Exclusion of the last study (1 out of 18) Alschuler et al. (2009) was made to avoid bringing bias into the review since two studies reported from the same dataset. It is reasonable to believe that these two studies could introduce bias and thus jeopardize the results of this review. The primary reference for this study is Geisser et al. (2004).

**Included Studies**

From the selection of the 23 full text studies, five studies were included in the systematic review. In Table 9 the 5 studies were number 9, 10, 11, 13 and 21. citations of the five studies is listed below with most recently published study first.


The five included studies had a cross-sectional or case-control design. A more detailed explanation and implications of these study designs is examined in this chapter.
Data Extraction

As identified earlier data extraction was plotted into one single form (Table 8). The form included an item, “low vs high disability sub-classification”, which was not needed.

Explanation of abbreviations are located under the data extraction form.

The data extraction form is found in appendix 5.

Description of Included Studies

Design and Confounding

The case-control design is described as ranking one step higher than the cross-sectional design on the hierarchy of evidence (Ho, Peterson and Masoudi, 2008). This would uncritically mean that case-control studies would contribute with a potentially more reliable conclusion in this review. However, based on the result of the quality assessment it was found that since all included studies with the exception of the high quality cross-sectional study by Geisser et al. (2004) were rated as being of medium quality, the risk of a large and low quality cross-sectional study skewing the results was less likely. Distribution of included studies regarding design is found in Figure 3.

Four of the five studies did not report which study design they used and Dubois et al. (2014) did not report this until the end of the article. This seems to be a general problem in observational designs though it is advised to describe the design either in the title, the abstract or at least early in the method section (von Elm et al., 2008).
All studies had a very clear description of their objective and compiled with these in relation to outcomes and reporting of it in the result section. The similarities regarding application of validated outcome measures and description of exposure had a satisfactory level and would in four of the five studies enable a replication of their procedure thus increasing internal validity. The study by Henchoz et al. (2013) was an exception to this.

All studies except Watson et al. (1997) adhered to good scientific practice regarding ethical approval and written informed consent.
Information on recruitment procedures was missing in three of five studies. The setting and being conducted in primary, secondary or tertiary sector were missing in three of five studies (Appendix 5). Information was sparse on inclusion and exclusions criteria and the different durations defined for chronic non-specific LBP were of great concern as it affect the generalisability of all studies and will affect the impact of the recommendations in this systematic review.

All studies included in this review had rather clear descriptions of demographic data and in the case-control studies it was possible to compare research subjects and controls.

None of the included studies reported sample size or power calculation or justified the missing information.

Despite blinding being difficult, none of the studies provided information on who conducted recruitment procedures, exposures and collection and final processing of the data. It would have been feasible to blind the assessors in several of the steps, which would have avoided decreased internal validity.

Confounding factors are likely to have affected the results of the included studies since it can be difficult to control for confounders, especially in the case-control studies that are likely to introduce confounders. With the exception of Watson *et al.* (1997) which lacked description of analyses conducted in the SPSS programme, the other four studies applied models of regression and multivariate analysis to account for relations between variables.

All studies applied p-value as an indicator of statistical significance; however, none of the studies reported odds ratios (OR) or confidence intervals in regarding full flexion or Flexion relaxation ratio or any other secondary outcomes. The lack of application
of OR or Risk Ratio (RR) could be due to the objective which was to measure SEMG activity and not causality concerning which exposures could predict risk of it occurring. From a methodological point of view, it would not make sense to calculate OR or RR.

The included studies had methodological limitations regarding reporting based on the STROBE recommendations (von Elm et al., 2008; Vandenbroucke et al., 2014). Included studies were assessed for methodological quality by the author of the review. It is a major limitation that only one reviewer took part in this assessment process (Boland, Cherry and Dickson, 2014), but this was the premise of this dissertation and caused by limited resources.

**Participants and Demographics**

Overall number of subject in the included studies was 198. Subjects were not equally divided between study designs with 34 subjects from the case-control studies and 164 subjects from the cross-sectional studies (Figure 4).
The studies by Henchoz et al. (2013) and Dubois et al. (2014) were from the same university in Quebec Canada and had an overlap of three researchers. Thus, their data applied to 74 of the research subjects, equivalent to more than one third of all the research subjects included in this review. It also means that their method would be likely to have a considerable effect on the overall results.

The mean age of the total number of research subjects ranged from 32.1 years (Henchoz et al., 2013) to 43.7 years (Watson et al., 1997) (Fig. 5) with a mean age of 38.12 years.
The age group in the included studies is representative of the younger part of the population with non-specific LBP. However, it is not representative of the older part of the population with non-specific LBP ranging from age forty to the mid-sixties (Freburger et al., 2009; Hoy et al., 2010). Thus, results are only applicable to the younger population with chronic non-specific LBP.

The five included studies were all conducted in western countries. It has been assumed that the LBP - epidemic is mainly known in the high-income countries. However, Louw, Morris and Grimmer-Somer (2007) found that the average lifetime
prevalence of LBP (62%) was on the rise in low / middle-income countries such as South Africa and Nigeria, though still presenting with only 75% of the prevalence reported in the high-income countries (O'Sullivan, 2005). The results from this review allow for application to western countries.

Included studies used different definitions of chronicity ranging from three months in Geisser et al. (2004) and Massè-Alarie et al. (2016) to twelve months in the study by Dubois et al. (2014). The different definitions of duration of chronicity among research subjects constituted a challenge. The differences between research subjects within studies such as in the study by Geisser et al. (2004) and between studies as well as very limited information of the research subjects in the study by Massè-Alarie et al. (2016) and Watson et al. (1997) added uncertainty concerning the homogeneity of the research subjects.

**Figure 6.** Duration of symptoms in the included studies

<table>
<thead>
<tr>
<th>Study</th>
<th>Mean duration (years)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Watson (1997)</td>
<td>3.75</td>
</tr>
<tr>
<td>Geisser (2004)</td>
<td>7.5</td>
</tr>
<tr>
<td>Dubois (2014)</td>
<td>11.25</td>
</tr>
<tr>
<td>Massè-Alarie (2016)</td>
<td>Not reported</td>
</tr>
<tr>
<td>Henchoz (2013)</td>
<td>15</td>
</tr>
</tbody>
</table>
The average duration of years with symptoms in the included studies was 8.25 years (Fig. 6) except from the study by Massè-Alarie et al. (2016) which did not report duration of symptoms. When comparing duration of years with symptoms to the outcome measures of disability, pain and pain-related fear, which in all studies were reported from low to moderate levels, it is interesting that there is no relation between them. Actually, the studies by Geisser et al. (2004) and Watson et al. (1997) reported the highest pain levels but were also the shortest average duration of symptoms and lowest pain-related fear scores. Verkerk et al. (2013) reported that a shorter duration of symptoms and a lower pain level at baseline were prognostic indicators of disability after five and 12 months. Comparison of the included studies with Verkerk et al. (2013) could indicate that level of pain or disability does not change as duration increases and a certain threshold has been passed. It is thus highly relevant to increase the understanding of the transition from acute to chronic non-specific LBP as suggested by Ramond-Roquin et al. (2015).

**Outcomes**

The included studies calculated the correlation between the flexion relaxation ratio (FRR) or full flexion (FF) and scores regarding pain-related fear. With exception of Dubois et al. (2014) who applied the term “chronic neuromuscular adaptation” and Massè-Alarie et al. (2016) who investigated muscle activity in TrA, the three other studies investigated the muscle activity in the lumbar erector spina (ES) with either one or two pairs of electrodes. Figure 7 presents the distribution between all the included studies regarding number and placement of electrodes.
This demonstrated that the dataset would be able to present results on lumbar muscle activity during FRR or FF (Fig. 7), though with a varying number of electrodes applied between studies. This applied to 94% of all investigated research subjects.

Regarding relation between pain-related fear and FRR or FF measuring on lumbar muscle activity, 60% of studies, equivalent to 68% of all investigated research subjects, were relevant.

Muscle activity in TrA during FRR or FF would be based on a single study of 12 research subjects and approximately the remaining 6% of research subjects. This study (Massè-Alarie et al., 2016) also contributed to investigate the relation between pain-related fear and FRR or FF measuring, TrA muscle activity.

Massè-Alarie et al. (2016) being the only study to report SEMG data from the abdominal muscles also made it inadequate to perform a meta-analysis.
It is important to note that the studies by Dubois et al. (2014) and Henchoz et al. (2013) included an intervention of experimental pain and expectation of fear. Considering the investigated population information regarding such protocols could have an effect on the SEMG signals.

SEMG data were normalised using three different methods in the included studies and the study by Watson et al. (1997) did not apply normalisation. Combining that with the large interpersonal differences in muscle activity in a symptomatic population and its affect on the raw SEMG signals between the different measures added to the concern regarding comparability between studies (Noraxon, 2005; Halaki and Ginn, 2012). Details of factors related to heterogeneity are provided in appendix 5.

An additional outcome of interest to this systematic review is the identification of disability as an independent variable with its possible correlation to FRR or FF. All included studies with the exception of Geisser et al. (2004) investigated level of disability and its relation to other variables.

Watson et al. (1997) and Massè-Alarie et al. (2016) reported the correlation to be a non-significant at both L1- L2, L4 - L5 and TrA. Dubois et al. (2014) did not report FRR or FF despite investigating the flexion relaxation phenomenon (FRP) but used the term “chronic neuromuscular adaptation” and found its correlation with disability to be significant (p = 0.006); they also reported that it was only one of several variables accounting for the level of disability. It is important to recognize that FRR or FF and “chronic neuromuscular adaptation” is not the same since the definition of “chronic neuromuscular adaptation” is the root mean square (RMS) SEMG measured during a full flexion used as a baseline of muscle activity.
Henchoz et al. (2013) found a significant correlation (p = 0.004) between disability and FRR / FF at L4 - L5 but not at L2 - L3. Thus, it is important to consider that this data was a result of introducing a task with high fear expectations. Furthermore, their results are not conclusive based on L2 - L3 being non-significant and L4 - L5 being significant.

Quality Assessment

The quality of the included studies was assessed using two different tools matching the study design (Sanderson, Tatt and Higgins, 2007). The Newcastle Ottawa Quality Assessment Scale was used for the case-control studies and the adapted Newcastle Ottawa Quality Assessment Scale was used for the cross-sectional studies. Tables 10 (p. 73) and 11 (p. 79) present the findings of the different study designs.

Quality Assessment of Cross-sectional Studies

Selection and Recruitment

None of the three cross-sectional studies explicitly included a description of their design. This was mentioned by Dubois et al. (2014) in the study limitations section. Categorized as cross-sectional was done by the author of this review based on measurement(s) on variables of interest conducted at one particular point in time, no control group and no follow up period (Mann, 2003; Song and Chung, 2010).

All of the three studies clearly stated their purpose in accordance with the objective of the systematic review. Inclusion criteria in Watson et al. (1997) and Dubois et al. (2014) were only mentioned very briefly and not at all in the study by Geisser et al. (2004), except for duration of symptoms.
The exclusion criteria in both Watson et al. (1997) and Dubois et al. (2014) were reported in more detail though not very detailed. According to the author of this systematic review, important information missing in the inclusion and exclusion criteria regarded referred or not referred symptoms in the lower limb(s), radiculopathy or psychological conditions such as depression; these are relevant and important prognostic factors (Turk, 2005; Dunn, Jordan and Croft, 2011; Ramond-Roquin et al., 2015; Verkerk et al., 2015). The lack of a detailed descriptions of inclusion and exclusion criteria also affects the internal validity of the studies (Jørgensen, Christensen and Kampmann, 2005). Moreover lack of detailed criteria also introduces an increased risk of selection bias (Meadows, 2003). On the other hand, if criteria were too strict it could mean that the generalisability of findings would be very limited (Jepsen et al., 2004).

The recruitment of research subjects was described in all of the studies. One study (Dubois et al., 2014) did not account for the location of the study. Watson et al. (1997) did not report the ethical approval of their study. It would be good scientific practice if a study investigating a population with chronic symptoms to obtain an ethical approval (DNVK, 2014; WMA, 2015; UoD, 2016).

There was no information on written informed consent in the study by Watson et al. (1997). However, both ethical approval and written consent may have been obtained since research subjects in the Watson et al. (1997) study were selected from a list of patients already included in another trial.

Geisser et al. (2004) also recruited research subjects from another trial; in this way both studies used a convenience sample (Meadows, 2003), which included subjects presenting with the “disease” of interest in their investigation. The problem could cause an overestimation of the findings compared to the general population and thus
affect generalisability (Song and Chung, 2010). However, the purpose and thereby the study design was made to examine for correlation between two variables at one particular point in time and not the relationship between cause and effect. It thus seems to be a fair procedure depending on presentation of demographic data in the studies.

Demographics and Chronicity

All three cross-sectional studies had an acceptable description of demographic data in relation to the item “representativeness of the sample” (Table 10). Subjects were rather similar concerning gender (percentage of males and females included) and age. Both gender and age matched the general population, which will affect the generalisability of the studies (Lind, 2011). However, caution is needed both concerning duration of symptoms and definition of duration, which is important to the classification as being chronic. These were different between studies. The study by Watson et al. (1997) described the shortest duration of symptoms (4.6 years) and Dubois et al. (2014) reported duration of symptoms for as long as 12.5 years. Definition of chronicity was the shortest in Geisser et al. (2004) with a cut-off point at >3 months and the longest in Dubois et al. (2014) applying a >12 months cut-off point.

These definitions are used in the literature although that there seems to be an agreement that duration >3 months is the cut-off point definition of chronic pain (Bogduk and McGuirk, 2006; Jensen et al., 2003).

The effect of duration of symptoms and reasons for chronicity cut-off point were not discussed in any of the studies.
To make a more accurate evaluation on how this affects external validity, more details should be included and discussed in all the cross-sectional studies included in this review. Geisser et al. (2004) is the only study presenting further relevant data regarding of prognosis of disease such as level of education, work status, receiving compensation or being involved in litigation (Hoy et al., 2010; Dunn, Jordan and Croft, 2011). To gain a better understanding and a more precise overview of the population, the researchers could have introduced an outcome measure on disability, psychological factors and level of pain and correlated these with duration of symptoms.

Contrary to both Watson et al. (1997) and Dubois et al. (2014), Geisser et al. (2004) did not include any outcome measure on disability. This could introduce a bias since such a confounder (disability) has been shown to affect prognosis (Mallen et al., 2007; Dunn, Jordan and Croft, 2011) and thus weaken generalisability.

**Sample Size**

None of the studies reported how sample size was reached or if a power calculation had been made (Table 10 “sample size”). Neither of the studies justified their sample size or mentioned it as a limitation.

Observational studies do not in general report sample size calculations which makes it difficult to assess if the researchers have considered the sample size needed for the purpose of their study (Martin et al., 2016). However, it is a general rule that the larger the sample size, the more likely it is to indicate a result that is close to the true mean of the population being investigated (Verhoeven, 2011). There are some principles that should be considered and which might justify their sample size but not their lack of reporting it. The studies could have reported how large an error was
found acceptable as less errors would inevitably call for a larger sample (Verhoeven, 2011; Kamangar and Islami, 2013). All research subjects in the studies had chronic LBP but they did not all have identical outcomes regarding disability, pain or psychosocial measures. This means that there is some variation in the population and it would thus not be sufficient to investigate just one subject (Kamangar and Islami, 2013). Since the included studies all had clear objective(s) of what they intended to measure combined with the likelihood that all subjects would present with the outcomes to some extent, a smaller sample size would be sufficient than if the outcome was not expected to be common (Ersbøll and Ersbøll, 2003).

Protocol

All studies included had a clear description of how the protocol was made on how patient information on outcome measures such as pain, disability (except from Geisser et al. 2004) and psychosocial factors / pain-related fear was obtained using questionnaires.

The included questionnaires regarding pain, disability and pain-related fear (appendix 5) had a moderate to high reliability and validity. Validation of pain measures (VAS and NRS) is hard to test since a gold standard for pain is not documented (Goubert et al., 2004; Grotle et al., 2006; Hawker et al., 2011). Use of questionnaires, however, strengthens internal and external validity (Habicht, 2011).

Application of SEMG was described by all studies, though less details were provided regarding skin preparation by Geisser et al. (2004). The procedure used during exposure included explanations before the actual movements, verbally guidance or other ways of using auditory cues. The details of this makes it possible to reproduce the procedure both in the actual study and in a possible replication of it.
Use of SEMG has many pitfalls; however, it measures microvolts based on muscle activity which matched the objective of the studies. Furthermore, it is commonly applied in other studies measuring muscle activity adaptation in a LBP population (Hodges and Richardson, 1997; Mannion et al., 1998).

Regarding placement of electrodes, all studies reported this in detail; however, none of the studies used a reference for recommendation of electrode placement sites such as the one proposed by the Seniam group. Placement of electrodes can be very challenging and variation may have a substantial effect to alter measures and give misleading results (Rainoldi, Melchiorri and Carruso, 2004).

**Confounding**

All three studies could have minimized the risk of introducing bias such as number of confounders if their study design had provided a more precise description of inclusion and exclusion criteria (Juul, 2006). Eliminating all confounders would not be realistic but those presented in the studies such as demographics should be considered if results should be reliable and residual confounding be minimized (Mann and Wood, 2012). None of the studies included as well as the majority of studies in non-specific chronic LBP populations reported level of physical activity or exercise though this has been proven to have a positive effect on recurrence of LBP (Steffens et al., 2016). Geisser et al. (2004) controlled for confounding factors and variables that could have an effect on their results. They accounted for each variable with a table showing the zero-order correlation between demographic data and all outcome measures and in between outcome measures expressed with a p-value, clarifying how variables were correlated and clearly describing their findings.
Watson et al. (1997) did not report any regulation of data in relation to confounding, which adds some uncertainty to the exact correlation between variables and credibility of results.

Dubois et al. (2014) applied a multiple regression analysis to determine how the different variables independently contributed to functional disability. Prior to that a principal component analysis was performed to make data easier to explore. This could be a way of controlling the number of independent factors for example the number of contributing confounders (Shlens, 2005).

Outcomes regarding use of statistical tests applied were reported similarly in Dubois et al. (2014) and Geisser et al. (2004). Both applied the Kolmogorov-Smirnov test to their data to see if data were normally distributed. Both applied multiple regression analysis to estimate relationship between variables and compare the relative effects of the variables. The use of a two-way ANOVA instead of a T-test by Dubois et al. (2014) makes sense when dealing with multiple levels and observing differences between groups (LAERD Statistics, 2013a).

Use of repeated T-tests in their example would have had an increased likelihood for presenting an overestimation of significance (Lund, 2004).

Watson et al. (1997) did not report a detailed description of what was done using the statistics programme SPSS. However, application of Pearson product moment correlation coefficients makes it possible to measure a linear relationship between variables. All studies used p-values in their result section to imply if results were statistically significant and the ones applying ANOVA reported their F-value as advised (LAERD Statistics, 2013b).

None of the studies described how data was stored, which from an ethical perspective seems relevant. It was not described who investigated the data. It would
have been feasible to apply a blind assessment of the dataset or to make two investigators go through the dataset (Vandenbroucke *et al.*, 2014).

More than one investigator investigating the data could have affected the data with more variability, which could affect the results both positively and negatively (Littlewood and May, 2013).

**Quality Score**

The quality scores for each of the studies based on the adapted Newcastle Ottawa Quality Assessment Scale for cross-sectional studies are shown in Table 10 (overall quality score). Geisser *et al.* (2004) was rated as a high quality study and the two others as medium quality studies.

**Table 10.** Quality assessment of cross-sectional studies

<table>
<thead>
<tr>
<th>ITEMS</th>
<th>Watson, PJ. 1997</th>
<th>Geisser, ME. 2004</th>
<th>Dubois, JD. 2014</th>
</tr>
</thead>
<tbody>
<tr>
<td>NR = Not reported</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>* = satisfactory description</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Selection</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Representativeness of the sample</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>2. Sample size</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>3. Non-respondents</td>
<td>NR</td>
<td>NR</td>
<td>NR</td>
</tr>
<tr>
<td>4. Ascertainment of the exposure</td>
<td>**</td>
<td>**</td>
<td>**</td>
</tr>
<tr>
<td><strong>Comparability</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. The subjects in different outcome groups are comparable, based on the study design or analysis. Confounding factors are controlled.</td>
<td>NR</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td><strong>Outcome</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Assessment of outcome</td>
<td>*</td>
<td>*</td>
<td>*</td>
</tr>
</tbody>
</table>
Quality Assessment for Case-control Studies

Classification of Study Design

The study by Henchoz et al. (2013) included a description of the cross-sectional, case-control study design. The author of this review did not find that the description wrong since they included an asymptomatic control group and a symptomatic population and measurements were made at a certain point in time.

Based on discussion with one of the supervisors at the University of Dundee and due to characteristics often related to case-control studies; 1) matching method being non randomized but frequency matching, 2) Cases were deliberately chosen based on presence of disease (Mann, 2003), it was decided to be quality assessed as a case-control study. The study by Massé-Alarie et al. (2016) did not report the design so as conducted with the cross-sectional studies the same procedure to categorize study design was used in regard to this matter.

Case-control studies are usually retrospective, which makes it harder to control for bias, but can be prospective and thus easier to control for bias (RCD, 2009; Habicht,
According to Vandenbroucke et al. (2014) such terminology of retrospective or prospective are poorly defined and used with such a variety that it should be abandoned and replaced with a clear description of what the researchers actually mean, including “how and when data collection took place” (Vandenbroucke et al., 2014, p. 1507).

In the included studies the objective of interest, meaning data collection of the exposure of interest, was first assessed after inclusion of groups resulting in neither a retrospective nor a classical prospective study design.

**Selection and Recruitment**

Both the included case-control studies had a clear description of their objective and complied with the objective of this systematic review.

Both studies had obtained ethical approval and used a written informed consent form in accordance with both national and international recommendations (WMA, 2015; UoD, 2016). However, neither of the studies described how data was stored, which from an ethical perspective seems relevant (UoD, 2016).

According to the NOS case-control quality assessment, all included research subjects should ideally be independently validated meaning that it should be traceable via for example health service or International Classification Codes (ICD) or objective measures such as an MRI. Delitto et al. (2012) for ICD codes related to chronic LBP. Such codes or objective measures were not reported in any of the studies; however, use of imaging as an objective measure to define a symptomatic
population is outdated and irrelevant due to a large prevalence of false positive results (Brinjikji et al., 2014; Fortin et al., 2014).

Neither Henchoz et al. (2013) nor Massè-Alarie et al. (2016) reported the recruitment process of research subjects (setting) or inclusion period (Table 11). Henchoz et al. (2013) defined adequate inclusion and exclusion criteria, which in a combination with their demographic dataset provided a clear description of research subjects and controls. Massè-Alarie et al. (2016) had an inadequate definition of research subjects and controls in their study based on limited information regarding inclusion and exclusion criteria and reference to such data from a previous study was not sufficient. The tables with detailed descriptions of demographic data were helpful though in their description. Massè-Alarie et al. (2016) did not report duration of symptoms either. In general, this study did not report details on subjects of interest in accordance with the STROBE recommendations (von Elm et al., 2008) and the generalisability of their results is threatened as it does not enable the reader to clearly identify research subjects and controls.

**Comparability and Confounding**

In a RCT the purpose of the randomization procedure is to reduce systematic errors and decrease the amount of confounders or make sure they are equally presented in all groups. However, the recruitment of research subjects (symptomatic population) and controls (asymptomatic population) in a case-control study is vulnerable to selection bias. Ideally this means that research subjects and controls have to be selected very carefully to decrease the differences between groups if the validity of
the study should be intact (Bowling, 2002; Littlewood and May 2013; Cao, Cox and Eslick, 2016).

When comparing cases to controls it is crucial that groups are comparable; however, none of the two studies reported if recruitment of groups was performed in the same or in multiple settings (i.e. hospitals, educational institutions, social classes, primary or secondary sector etc).

Information regarding dynamic or fixed population or period of time for selection is also missing. However, based on the demographic data, a very cautious interpretation could be that both studies selected research subjects and controls from dynamic populations according to different ages among research subjects (Knol et al., 2008). These aspects decreased the external validity (Bowling, 2002).

Matching is a way to adjust for confounders in case-control design; however, the effect of the matching factor (s) can not be used in the analysis (Juul, 2006; Cao, Cox and Eslick, 2016). Henchoz et al. (2013) strengthened the validity of their study using the matching procedure. They applied a frequency matching based on age or gender, which makes sense since these are not directly a part of the exposure (Juul, 2006). Massè-Alarie et al. (2016) had no description of matching which based on their relative small sample size would have made sense however depending on how they construct their analysis it might not be a problem (Vandenbroucke et al., 2014).

**Statistical Analysis**

Neither Massè-Alarie et al. (2016) nor Henchoz et al. (2013) applied a sample size calculation. This means that they did not account for the needed sample size to detect a significant difference and avoid a type 2 error (Ersbøll and Ersbøll, 2003;
Aveyard, 2014; Cao, Cox and Eslick, 2016). Neither of the two studies justified the absence of a sample size calculation.

Henchoz et al. (2013) assessed normal distribution of data with a Kolmogorov-Smirnov test. To test for group differences either a Mann-whitney U test or a Chi square test was applied to make sure introduction of bias regarding between group confounders was not the case (Lund, 2004).

As a result of multiple variables and interest in the independent relationship with other variables, a two-way ANOVA was introduced in both studies as well as Pearson’s correlation to test relations between variables of interest. Massè-Alarie et al. (2016) applied Bonferroni correction to correct family-wise error and trying to avoid a type 1 error after the application of ANOVA. Application of the Bonferroni correction has been criticized due to an increasing risk of introducing a type 2 error (Armstrong, 2014). The application of these tests in the statistical analysis increased the internal validity of their results (Juul, 2006).

In their result section both studies applied p-value based on the reported F-value to indicate if results were significant or non-significant (LAERD Statistics, 2013b). There were no use of neither odds ratio nor confidence intervals, which are commonly used in case control studies (Mann, 2003; Sistrom and Garwan, 2004). However, this could be explained by the objectives of included studies which were to identify a specific outcome related to their research subjects (Juul, 2006).

**Protocol**
With regard to ascertainment (Table 11 “Exposure”), both studies had applied the same procedure of ascertainment to both research subjects and control groups. Both studies had a detailed description of the electrode placement, though none of them referred to a particular protocol or for that matter to the Seniam recommendations (Seniam, 2016).

Henchoz et al. (2013) did not give a very accurate description of the procedure itself where as the study by Massè-Alarie et al. (2016) increased the possibility of a replication due to the description of a standardized procedure with a test trial and auditory cues. The included questionnaires regarding pain, disability and pain-related fear (presented in appendix 5) have a moderate to high reliability and validity with the exception of the State Trait Anxiety Inventory (STAI) (Goubert et al., 2004; Grotle et al., 2006; Hawker et al., 2011; Fernandes et al., 2012; Cleland et al., 2014). The author of this review could not find data that presented reliability or validity of the STAI. However Henchoz et al. (2013) also applied the Fear Avoidance Belief Questionnaire which measures similar factors. The application of these questionnaires and scales strengthens generalisability.

As the only study Massè-Alarie et al. (2016) provided details and rationale of possible exclusion or non-response rate of research- or control group subjects; this increases transparency.

Blinding can be a difficult process though at least a partial blinding as suggested by Mann (2003) of the assessor in both studies would have been possible. However
there were no reports on blinding procedures which is critical to credibility of results (Littlewood and May, 2013) and thus a limitation in the included studies.

**Quality Score**

The quality score for the two case-control studies based on the Newcastle Ottawa Quality Assessment Scale for case-control studies is shown in Table 11 (overall quality score). Both studies were rated to be of medium quality.

**Table 11. Quality assessment of case-control studies**

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>* = satisfactory description</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Selection</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Case definition adequate</td>
<td>*</td>
<td></td>
</tr>
<tr>
<td>2. Representativeness of the cases</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>3. Selection of controls</td>
<td></td>
<td></td>
</tr>
<tr>
<td>4. Definition of controls</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td><strong>Comparability</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Comparability of cases and controls on the basis of the design or analysis</td>
<td>**</td>
<td>*</td>
</tr>
<tr>
<td><strong>Exposure</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>1. Ascertainment of exposure</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2. Same method of ascertainment for cases and controls</td>
<td>*</td>
<td>*</td>
</tr>
<tr>
<td>3. Non-response rate</td>
<td></td>
<td></td>
</tr>
<tr>
<td>**Number of ***</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Overall quality score</strong></td>
<td>Low - medium - high</td>
<td>Medium</td>
</tr>
</tbody>
</table>

**Summary and Analysis**
Based on the quality assessment, certain similarities between case-control and cross-sectional studies became clear enabling the possibility to include them in a synthesis, though with some limitations.

According to the study protocol a meta-analysis of the results from the included studies was planned. Thus, included studies and their data did not meet an acceptable standard acquired concerning homogeneity; this precluded a meta-analysis.

A narrative summary of the data was thus chosen due to the similarities between studies despite factors related to heterogeneity which are summarized below.

Assessment of heterogeneity was based on Higgins and Green (2011).

**Results of Included Studies**

The overall purpose was to investigate if there is a correlation between muscle activity during forward bending from a standing position in a subgroup of chronic LBP patients presenting with pain-related fear. As a secondary objective findings between disability and muscle activity is presented.

**Pain – related Fear and Muscle Activity**

Study findings regarding FRR or FF and pain-related fear are summarized in Table 12.

Table 12. Correlation between pain-related fear and muscle activity in FRR /FF
<table>
<thead>
<tr>
<th>Electrode placement</th>
<th>Study</th>
<th>Correlation expressed with P-value</th>
<th>Between group differences in case-control</th>
<th>significant or non-significant</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1/L2 - L2/L3</td>
<td>Watson, 1997</td>
<td>p = 0.05 p = 0.04 NT</td>
<td></td>
<td>significant</td>
</tr>
<tr>
<td>L3/L4 - L4/L5</td>
<td>Geisser, 2004</td>
<td>p = 0.01* p = 0.01* NT</td>
<td></td>
<td>significant</td>
</tr>
<tr>
<td>TrA</td>
<td>Dubois, 2014</td>
<td>NT NT NT NT</td>
<td></td>
<td>NT</td>
</tr>
<tr>
<td></td>
<td>Henchoz, 2013</td>
<td>NR p = 0.012** NT Yes</td>
<td>significant at L4-L5</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Massè-Alarie, 2016</td>
<td>NT NT p = 0.03 Yes Yes</td>
<td>significant at TrA</td>
<td></td>
</tr>
</tbody>
</table>

* = Combined the 2 pairs of electrodes from the lumbar erector spina
** = when expecting pain
NR = Not reported, NT = Not tested
FRR = Flexion Relaxation Ratio
FF = Full Flexion
TrA = Tranversus Abdominus

Taken together the included studies by Watson et al. (1997), Geisser et al. (2004), Henchoz et al. (2013) and Massè-Alarie et al. (2016) regarding pain-related fear and FRR or FF suggest a significant correlation at L4-L5, L1/L2 - L2/L3 and the TrA correlation. L4-L5 significance was based on three studies, which makes the results more reliable than the L1/L2 - L2/L3 and TrA correlation significance, which were based on single studies.

Overall interpretation of these results should be done with caution due to the limited number of studies. Moreover, the study by Geisser et al. (2004) in which 57% of all L4-L5 research subjects contribute considerably to the conclusion. Thus, any possible bias from that study would affect overall conclusion of this review (Juul, 2006).
Disability and Muscle Activity

Results regarding FRR or FF or “chronic neuromuscular adaptation” and disability are summarized in Table 13.

Table 13. Correlation between disability and FRR or FF or “chronic neuromuscular adaptation”

<table>
<thead>
<tr>
<th>Electrode placement</th>
<th>Correlation expressed with P-value</th>
<th>Between group differences in case-control</th>
<th>significant or non-significant</th>
</tr>
</thead>
<tbody>
<tr>
<td>L1/L2 - L2/L3</td>
<td>L3/L4 - L4/L5</td>
<td>TrA</td>
<td></td>
</tr>
<tr>
<td>Study</td>
<td>p-value NR</td>
<td>p-value NR</td>
<td>NT</td>
</tr>
<tr>
<td>Watson, 1997</td>
<td>NT</td>
<td>NT</td>
<td>NT</td>
</tr>
<tr>
<td>Geisser, 2004</td>
<td>NT</td>
<td>NT</td>
<td>NT</td>
</tr>
<tr>
<td>Dubois, 2014</td>
<td>NT</td>
<td>p = 0.006***</td>
<td>NT</td>
</tr>
<tr>
<td>Henchoz, 2013</td>
<td>p-value NR</td>
<td>p = 0.004**</td>
<td>NT</td>
</tr>
<tr>
<td>Massè-Alarie, 2016</td>
<td>NT</td>
<td>NT</td>
<td>p-value NR</td>
</tr>
</tbody>
</table>

*= Combined the 2 pairs of electrodes from the lumbar erector spina
** = when expecting pain
*** = “chronic neuromuscular adaptation”
NR = Not reported, NT = Not tested
FRR = Flexion Relaxation Ratio
FF = Full Flexion
TrA = Transversus Abdominus

Regarding correlation between disability and FRR or FF results indicate that despite similar questionnaires (ODI and ODQ) (Appendix 5) it is not possible to draw firm conclusion regarding correlation due to conflicting data.
Conflicting data could be due to differences in protocols regarding both the introduction to pain and expectations of pain combined with the inconsistent use of FRR / FF and "chronic neuromuscular adaptation" between studies. The results presented in this chapter highlights the limited number of studies on the subject of interest and the quality of the included studies. Included studies had some methodological issues that to some extent made it difficult to compare them. Such as possible differences between them in terms of the inclusion and exclusion criterias and primary and secondary sector recruitment. Ignoring such differences and conducted a pooling of data would have invalidated the results. Details of factors related to heterogeneity are provided in appendix 5. However they were all from western countries and had almost similar age group of subjects. Despite differences the included studies can add some value to the synthesis.

5. Discussion

Increased Muscle Activity and Pain-related Fear

To the author’s knowledge, this systematic review is the first to summarize the evidence on trunk muscle activity during standing forward bending in a subgroup of
chronic non-specific LBP patients with pain-related fear. Despite increasing popularity regarding muscle activity and ADL in chronic non-specific LBP the number of studies is limited. However, based on the limited evidence a cautious conclusion can be drawn that increased muscle activity in the lumbar spine (ES) and the TrA during standing forward bending is correlated to a subgroup of chronic non-specific LBP patients.

This chapter will discuss the complexity of understanding the altered muscle activity and implications related to management based on the scientific evidence identified including the limitations of the fear avoidance model (FAM).

Theories regarding the need for individualized management in the acute phase of LBP and the implications of a model that makes subgrouping reliable will be suggested.

Finally thoughts regarding application of SEMG will be discussed along with the clinical and methodological strengths and limitations of this review.

An additional study not included in this review that investigated the subject of interest was Ghamkhar’s and Kahlaeé’s systematic review (2015) which showed increased trunk muscle activity during gait; this is similar to the results in this review, though in another function. Ghamkhar and Kahlaeé (2015) found an increased trunk muscle activity when LBP patients were compared to controls. In the studies by Henchoz et al. (2013) and Massè-Alarie et al. (2016), which were they only studies with a control group, a between group differences was also found. Both studies also found that the difference was most pronounced in those with chronic non-specific LBP with the highest catastrophizing or kinesiophobia scores. None of the other included studies reported a correlation between muscle activity and level of pain-related fear.
Ghamkhar and Kahlaee (2015) did not correlate disability, pain or pain-related fear levels to their results, which excludes further comparisons.

This opens the question if the change in muscle activity in the chronic non-specific LBP population is related to pain-related fear, pain, chronicity, disability or a combination of variables as suggested by Dubois et al. (2014). The identification of the most important variable contributing to increased muscle activity is only relevant if it is possible to determine that increased and changed muscle activity is directly linked to chronicity; this would support the hypothesis that poor and altered function in the trunk muscles was responsible for chronic non-specific LBP. If Henchoz et al. (2013) and Massè-Alarie et al. (2016) discovered that level of catastrophizing was correlated to muscle activity and that neither pain and disability level nor duration of symptoms seemed to play a role it would make sense to target patients’ pain-related fear as an intervention. Woby et al. (2004) supported that reduction in pain-related fear would decrease disability level. However, O’Keefe et al. (2016) did not find such interventions superior to for example a physically addressed intervention when outcome measures were pain and disability. It is of course related to some uncertainty to compare the results of O’Keefe et al. (2016) to this systematic review when not investigating the same outcome measure; however, it is within reason to conclude that if pain-related fear was addressed and proved to have an effect on muscle activity why would pain and disability not change as well?

O’keefe et al. (2016) noted that the non-significant differences despite very different interventions could be explained by therapist-patient relation, which has been shown to have a positive effect on outcomes (Hall et al., 2010).
Altered and increased muscle activity in chronic non-specific LBP might have been interpreted wrong merely because it is different from the asymptomatic population (Hodges and Richardson, 1996; MacDonald, Moseley and Hodges, 2010; Neblett et al., 2013). It is indisputable that changes are present and must somehow be related to an episode of LBP; however, it might not be an adaptation to maintain stability and avoid pain but might just as well be a reorganization (Moseley et al., 2003; Moseley, Nicolas and Hodges, 2004; Moseley and Flor, 2012;) that helps the patient during ADL.

Moseley and Flor (2012) implied that in the chronic non-specific LBP population reorganization was related to chronicity. However, it is important to note that their reference was the study by Flor (1996) who only studies a sample of 10 patients with chronic LBP. Thus, Wand et al. (2010) supported Flor et al. (1997) but stated that the changes were located in the chronic LBP patients with distress, which again was linked to the emotional and psychological aspects of chronic non-specific LBP and not to the pain itself. This fits very well with the results from the five included studies in this review, which were unsuccessful in showing correlation between either duration of symptoms or pain level.

Results are conflicting and data is sparse leaving the clinician with no clear direction of which of the variables in the biopsychosocial model that need to be addressed. The FAM with its cyclical relationship seems logical; however, the above results showed that involved variables and components are more complex and not fully understood. An example of the perspective from the FAM is that an acute episode of LBP is related to physiological processes and that long standing pain is merely driven by cognitive and behavioural components (Vlayen et al., 2000). This is not in accordance with the literature that shows physiological processes to be ongoing and
be directly linked to cognitive and psychological factors (Nijs and Houdenhove, 2009; Nijs, Houdenhove and Oostendorp, 2010; Jensen, Dahl and Arendt-Nielsen, 2003; Campbell and Edwards, 2009).

**Transition from Acute to Chronic LBP**

As a way of dealing with the uncertainty retained within this area, researchers have reinforced the importance of attention to the transition phase from acute to chronic LBP (Melloh et al., 2011; Ramond-Roquin et al., 2014).

The European guidelines on acute LBP (van Tulder et al., 2006) advocate that management should include reassurance, advice to stay active both at work and leisure activities, avoid bedrest, consider medication and if symptoms persist after 4 - 8 weeks a multidisciplinary programme should be considered.

Bunzli et al. (2015) investigated the beliefs underlying pain-related fear; the results indicated that when the pain level was high and pain was unpredictable it was threatening. Diagnostic uncertainty or diagnosis of underlying pathology that could not be fixed or did not respond to treatment increased the level of fear of causing further damage and left them with less ability to deal with their pain. Miles et al. (2011) discovered that self-efficacy was a predictor of outcome and suggested that it should be targeted in the early stage of pain. Both Bunzli et al. (2015) and Miles et al. (2011) present findings that if we follow guidelines on acute LBP we need to make sure that the reassurance is cognitive and not just affective (Pincus et al., 2013). This could most definitely mean that some individuals would need more than just a single consultation at their general practitioner as experiencing pain will result in some painful situations and events which should be discussed with a health personnel to avoid misinterpretation and causing more uncertainty. Another aspect is the misdiagnosis and the failure of getting better or reaching prior functional activity level,
which was found to contribute to uncertainty and lack of sensible decision-making. This could be directly correlated to some of the beliefs held by patients being told that their problem was purely a matter of core stability. However, lack of improvement reinforces beliefs negatively (Nijs et al., 2013; Bunzli et al., 2015). Bunzli et al. (2015) investigated subjects with high pain-related fear where as the research subject in this systematic review presented with low to moderate levels of pain-related fear. Thus, since the correlation between muscle activity and pain-related fear was established it would be interesting to discover the difference in muscle activity in a population with high pain-related fear to gain further understanding of how muscle activity develops.

Subgroups
Based on the results in the systematic review, the limitations of the FAM in combination with the results of Bunzli et al. (2015), Miles et al. (2011) and Wand et al. (2010), subgrouping becomes relevant as a first step to optimize management. Wideman et al. (2013) talks of subgroups based on cumulative interacting factors based on their ability to predict those developing prolonged pain and disability. Data supporting these thoughts are mentioned in the background section. Developing subgroups based on how risk factors predict certain outcome helps to classification into subgroups. It is important that classification can be done in a reliable way; this is what Spratt (2013) calls the Assessment - Diagnosis link (A-D). However, it does not necessarily direct the treatment that should improve outcomes meaning that when the A-D link is reliable, the Diagnosis - Treatment link (D-T) should be matched to finally evaluate the outcome link between Treatment - Outcome (T-O).
A very interesting finding in Geisser et al. (2004) that is related to subgrouping and predictors of prognosis is that they included research subjects that were waiting for litigation. Litigation and workers compensation are proven to have a negative effect on prognosis (Merrill, 1997; Turner et al., 2006; Gum, Gassman and Carreon, 2013). However, their results revealed a decreased muscle activity in the litigation group during standing forward bending, highlighting the complexity of chronic non-specific LBP.

**EMG**

**Comparison and Normalization of Data**

Small sample sizes is a very common finding in these observational studies to some extent influencing the generalisation of results (Aveyard, 2014). This could be solved by conducting a meta-analysis (Bowling, 2002). Thus, pooling data would not only demand homogeneity regarding population, intervention and outcomes in the studies but also a certain level of quality and standards for application of SEMG as well as normalisation of data. Comparison of SEMG between subjects and betweens studies should be cautiously interpreted, since SEMG measurements are vulnerable to many factors (Noraxon, 2005).

The ability to produce muscle activity in a symptomatic population is among others affected by de-conditioning, pain, not being willing to generate muscle contraction due to pain-related fear and time of day. Pain has been proven to change muscle activity due to excitatory and inhibitory mechanisms. This could mean that the overall muscle activity had just been distributed and not necessarily been either decreased on increased (Lund et al., 1991). Thus, muscle activity might be increased in ES but at the same time it might be decreased in the Gluteus Maximus muscle.
Tucker et al. (2012) showed that anticipation of pain and application of experimental pain caused similar alterations of motor unit recruitment strategies leaving uncertainty of overall muscle activity.

Combining this with other factors that can disturb accuracy of SEMG measurements such as placement of electrodes including skin preparation (Rainoldi, Melchiorri and Caruso, 2004), BMI, research subjects not being familiar with the exposure and cross talk between electrodes (Noraxon, 2005) result in SEMG being easily affected by confounders.

Due to these variations a comparison between studies demands normalisation of the raw SEMG signals. All studies with exception of Watson et al. (1997) normalised their data. Though applying different normalisations procedures do makes comparisons difficult. Dankaerts et al. (2004) investigated reliability of maximum voluntary contraction (MVC) and sub-maximal voluntary contraction (sub-MVC) in trunk muscles of patients with chronic LBP and concluded that sub-MVC was the preferable normalisation procedure when assessing EMG signal on multiple days. Included studies in this systematic review only measured on the same day. However, since measuring on a symptomatic population it would be reasonable to use sub-MVC since subjects would be more willing to perform such types of contraction compared to a maximum contraction based on their pain-related fear, which is supported by Noraxon (2005).

Normalization of SEMG signals is not always necessary and since the included studies measured at one point in time and did not compare muscle activity between muscles they did not need to normalize data. However, for comparison between studies normalization needs to be conducted (Halaki and Ginn, 2012).
Generally, it seems to be a limitation when comparing SEMG studies based on a lack of standardization procedures both regarding the normalisation of the raw SEMG signal, the exposure protocol and electrode placement protocols. The Seniam group suggested an anatomical landmark system with recommendations for application of electrodes, which has been adopted by Noraxon (2005). The author of this systematic review would recommend that future studies combined a strict standardization and description of what was conducted to minimize differences and increase possibility of comparison between studies. This is in line with recommendations from Vandenbroucke et al. (2014) regarding observational studies.

**Clinical Implications**

The clinical implications of these results are important since they highlight that targeting trunk muscles on the basis of an assumption that pain and disability is due to decreased strength is to simplistic, emphasizing the complexity of chronic non-specific LBP. Furthermore, the results of this systematic review cannot conclude anything about causation but merely that there is a correlation between muscle activity and chronic non-specific LBP with pain-related fear during standing forward bending, based on a limited number of muscles being measured during exposure. However, the findings are believed to add more knowledge regarding a specific subgroup of patients with LBP. Thus, generalisability of findings should be concluded with caution.

**Research Implications**
This systematic review found that the included studies differed substantially making comparison difficult. However, they were all of medium quality making their findings more reliable and thus allowing for comparison to some extent.

Another aspect were the issues of inadequate reporting in the studies. This is unfortunately not an issue isolated to the included studies in this systematic review but a general problem in observational studies (von Elm et al., 2008; Vandenbroucke et al., 2014). Future research should emphasize adherence to adequate reporting. Not least should correlation to important outcome measures and inclusion of prognostic indicators be addressed to clarify correlation between variables. This might also contribute to further subgroups based on for example low and high pain-related fear.

**Strengths and Limitations**

This review includes important methodological aspects that needs to be addressed. The use of only one reviewer in the process from searching the literature, identification of included studies, data extraction, quality assessment of included studies and summary of results is vulnerable to selection bias (Boland, Cherry and Dickson, 2014). However, the choices in the process is justified by the reported transparency of the systematic review.

The exclusion of articles in other languages than English leads to a risk of publication bias, which has been addressed in the method section (Jørgensen, Christensen and Kampmann, 2005; Boland, Cherry and Dickson, 2014).

6. Conclusion
The aim of this review was to investigate a possible correlation between trunk muscle activity in a subgroup of patients with chronic non-specific LBP with pain-related fear during standing forward bending. The results of this review indicate that a correlation is present. Furthermore, the results are based on a review that from a methodological perspective has been conducted in a systematic and transparent way ensuring objectivity and replicability.

The correlation between increased trunk muscle activity during standing forward bending and a subgroup of patients with chronic non-specific LBP with pain-related fear was found to be significant, especially at L4-L5. The review included a small number of studies (5) and lack of standardization between studies such as overall number of electrodes calls for a cautious conclusion. It must be emphasized that further studies with improved reporting and standardization are needed to substantiate the results of the present review.

This review has added to the research knowledge regarding trunk muscle activity in patients with chronic non-specific LBP emphasizing that decreased trunk muscle activity could not explain ongoing chronic non-specific LBP. Thus, a simplistic biomechanical approach to a complex problem should be avoided by the clinician.

This review would suggest that future studies should adhere to guidelines regarding reporting in observational studies and apply a standardized method when using SEMG to make comparison possible and in the future be able conduct a meta-analysis.

Reference list


centralization and directional preference: a useful tool for front line clinician?’,


model: a discussion of the relationship between chronic musculoskeletal pain and


Lundquist, C. B., Jacobsen, J. S., Nielsen, L. S., Jørgensen, P. B., Birch, S. and

Low Back Pain Repsond Differently to Trunk Loading Dispite Remission From

McKenzie method for low back pain: a systematic review of the literature with a meta-


To be published in The Journal of Pain [Preprint]. Available at: http://


**Appendix 1. EMBASE Search**
<table>
<thead>
<tr>
<th>#</th>
<th>Query</th>
<th>Database</th>
<th>Start Date</th>
<th>End Date</th>
<th>Count</th>
</tr>
</thead>
<tbody>
<tr>
<td>#22</td>
<td>'electromyography'/exp</td>
<td>All4 combined</td>
<td>2016-03-20</td>
<td>2016-03-20</td>
<td>69945</td>
</tr>
<tr>
<td>#21</td>
<td>#9 OR #10 OR #11 OR #12 OR #13 OR #14 OR #15 OR #16 OR #17 OR #18 OR #19 OR #20</td>
<td>All4 combined</td>
<td>2016-03-20</td>
<td>2016-03-20</td>
<td>433435</td>
</tr>
<tr>
<td>#20</td>
<td>flexion AND ('relaxation'/exp OR relaxation) AND ratio*</td>
<td>All4 combined</td>
<td>2016-03-20</td>
<td>2016-03-20</td>
<td>81</td>
</tr>
<tr>
<td>#19</td>
<td>'flexion relaxation ratio'</td>
<td>All4 combined</td>
<td>2016-03-20</td>
<td>2016-03-20</td>
<td>27</td>
</tr>
<tr>
<td>#18</td>
<td>flexion AND ('relaxation'/exp OR relaxation) AND phenomenon*</td>
<td>All4 combined</td>
<td>2016-03-20</td>
<td>2016-03-20</td>
<td>84</td>
</tr>
<tr>
<td>#17</td>
<td>'flexion relaxation phenomenon'</td>
<td>All4 combined</td>
<td>2016-03-20</td>
<td>2016-03-20</td>
<td>72</td>
</tr>
<tr>
<td>#16</td>
<td>neuromuscular AND adaptation*</td>
<td>All4 combined</td>
<td>2016-03-20</td>
<td>2016-03-20</td>
<td>1670</td>
</tr>
<tr>
<td>#15</td>
<td>'neuromuscular adaptation'</td>
<td>All4 combined</td>
<td>2016-03-20</td>
<td>2016-03-20</td>
<td>127</td>
</tr>
<tr>
<td>#14</td>
<td>'neuro muscular daptation'</td>
<td>All4 combined</td>
<td>2016-03-20</td>
<td>2016-03-20</td>
<td>0</td>
</tr>
<tr>
<td>#11</td>
<td>'neuromuscular function'/exp</td>
<td>All4 combined</td>
<td>2016-03-20</td>
<td>2016-03-20</td>
<td>180585</td>
</tr>
<tr>
<td>#10</td>
<td>'muscle contraction'/exp</td>
<td>All4 combined</td>
<td>2016-03-20</td>
<td>2016-03-20</td>
<td>111436</td>
</tr>
<tr>
<td>#9</td>
<td>'motor control'/exp</td>
<td>All4 combined</td>
<td>2016-03-20</td>
<td>2016-03-20</td>
<td>26424</td>
</tr>
<tr>
<td>#8</td>
<td>#1 OR #2 OR #3 OR #4 OR #5 OR #6 OR #7</td>
<td>All4 combined</td>
<td>2016-03-20</td>
<td>2016-03-20</td>
<td>119062</td>
</tr>
<tr>
<td>#7</td>
<td>chronic AND low AND ('back'/exp OR back) AND pain*</td>
<td>All4 combined</td>
<td>2016-03-20</td>
<td>2016-03-20</td>
<td>14026</td>
</tr>
<tr>
<td>#6</td>
<td>'back'/exp OR back AND pain*</td>
<td>All4 combined</td>
<td>2016-03-20</td>
<td>2016-03-20</td>
<td>98081</td>
</tr>
<tr>
<td>#5</td>
<td>'chronic low back pain'/exp OR 'chronic low back pain'</td>
<td>All4 combined</td>
<td>2016-03-20</td>
<td>2016-03-20</td>
<td>42173</td>
</tr>
<tr>
<td>#4</td>
<td>'back pain'/exp OR 'back pain'</td>
<td>All4 combined</td>
<td>2016-03-20</td>
<td>2016-03-20</td>
<td>88382</td>
</tr>
<tr>
<td>#3</td>
<td>'low back pain'/exp OR 'low back pain'</td>
<td>All4 combined</td>
<td>2016-03-20</td>
<td>2016-03-20</td>
<td>46737</td>
</tr>
<tr>
<td>#2</td>
<td>'backache'/exp</td>
<td>All4 combined</td>
<td>2016-03-20</td>
<td>2016-03-20</td>
<td>79986</td>
</tr>
<tr>
<td>#1</td>
<td>'low back pain'/exp</td>
<td>All4 combined</td>
<td>2016-03-20</td>
<td>2016-03-20</td>
<td>41675</td>
</tr>
</tbody>
</table>
CINAHL search URL

http://rss.ebscohost.com/AlertSyndicationService/Syndication.asmx/GetFeed?guid=4785118
Appendix 2.

E-mail correspondence

Re: Poster "Fear ...phobia in chronic lower back pain patients"

Julia Clombiewski
brian@fysioaalborg.dk

Dear Brian,

I will find the related paper. Good luck with you review!

Julia

Am 21.01.2016 um 13:19 schrieb brian@fysioaalborg.dk:

Dear Julia,

I’m currently doing a systematic review and in my search I retrieved the above mentioned poster. Is it related to an article? Have not been able to locate it – not even with the help of our library.

Best regards
Brian Sørensen, PT., Dip. MOT
Adjunct University college Nordjylland, PT department
MSc student University of Dundee, Scotland

Gutenbergstrasse 18
Appendix 3.

Examples of the Newcastle Ottawa Quality Assessment Scales

NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE for CASE CONTROL STUDIES
Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Exposure categories. A maximum of two stars can be given for Comparability.

Selection
1) Is the case definition adequate?
   a) yes, with independent validation
   b) yes, eg record linkage or based on self reports
   c) no description

2) Representativeness of the cases
   a) consecutive or obviously representative series of cases
   b) potential for selection biases or not stated

3) Selection of Controls
   a) community controls
   b) hospital controls
   c) no description

4) Definition of Controls
   a) no history of disease (endpoint)
   b) no description of source

Comparability
1) Comparability of cases and controls on the basis of the design or analysis
   a) study controls for _______________ (Select the most important factor.)
   b) study controls for any additional factor (This criteria could be modified to indicate specific control for a second important factor.)

Exposure
1) Ascertainment of exposure
   a) secure record (eg surgical records)
   b) structured interview where blind to case/control status
   c) interview not blinded to case/control status
   d) written self report or medical record only
   e) no description

2) Same method of ascertainment for cases and controls
   a) yes
   b) no

3) Non-Response rate
   a) same rate for both groups
   b) non respondents described
   c) rate different and no designation
NEWCASTLE - OTTAWA QUALITY ASSESSMENT SCALE for COHORT STUDIES

Note: A study can be awarded a maximum of one star for each numbered item within the Selection and Outcome categories. A maximum of two stars can be given for Comparability

Selection
1) Representativeness of the exposed cohort
   a) truly representative of the average _____________ (describe) in the community
   b) somewhat representative of the average _____________ in the community
   c) selected group of users eg nurses, volunteers
   d) no description of the derivation of the cohort

2) Selection of the non exposed cohort
   a) drawn from the same community as the exposed cohort
   b) drawn from a different source
   c) no description of the derivation of the non exposed cohort

3) Ascertainment of exposure
   a) secure record (eg surgical records)
   b) structured interview
   c) written self report
   d) no description

4) Demonstration that outcome of interest was not present at start of study
   a) yes
   b) no

Comparability
1) Comparability of cohorts on the basis of the design or analysis
   a) study controls for _____________ (select the most important factor)
   b) study controls for any additional factor (This criteria could be modified to indicate specific control for a second important factor.)

Outcome
1) Assessment of outcome
   a) independent blind assessment
   b) record linkage
   c) self report
   d) no description

2) Was follow-up long enough for outcomes to occur
   a) yes (select an adequate follow up period for outcome of interest)
   b) no

3) Adequacy of follow up of cohorts
   a) complete follow up - all subjects accounted for
   b) subjects lost to follow up unlikely to introduce bias - small number lost - > ____ % (select an adequate %) follow up, or description provided of those lost
c) follow up rate < ____% (select an adequate %) and no description of those lost
d) no statement

Newcastle-Ottawa Quality Assessment Scale adapted for cross-sectional studies

Selection: (Maximum 5 stars)

1) Representativeness of the sample:
   a) Truly representative of the average in the target population. * (all subjects or random sampling)
   b) Somewhat representative of the average in the target population. * (non-random sampling)
   c) Selected group of users.
   d) No description of the sampling strategy.

2) Sample size:
   a) Justified and satisfactory. *
   b) Not justified.

3) Non-respondents:
   a) Comparability between respondents and non-respondents characteristics is established, and the response rate is satisfactory. *
   b) The response rate is unsatisfactory, or the comparability between respondents and non-respondents is unsatisfactory.
   c) No description of the response rate or the characteristics of the responders and the non-responders.

4) Ascertainment of the exposure (risk factor):
   a) Validated measurement tool. **
   b) Non-validated measurement tool, but the tool is available or described.*
   c) No description of the measurement tool.

Comparability: (Maximum 2 stars)

1) The subjects in different outcome groups are comparable, based on the study design or analysis. Confounding factors are controlled.
   a) The study controls for the most important factor (select one). *
   b) The study control for any additional factor. *

Outcome: (Maximum 3 stars)

1) Assessment of the outcome:
   a) Independent blind assessment. **
   b) Record linkage. **
   c) Self report. *
   d) No description.

2) Statistical test:
   a) The statistical test used to analyze the data is clearly described and appropriate, and the measurement of the association is presented, including confidence intervals and the probability level (p value). *
This scale has been adapted from the Newcastle-Ottawa Quality Assessment Scale for cohort studies to perform a quality assessment of cross-sectional studies for the systematic review, “Are Healthcare Workers’ Intentions to Vaccinate Related to their Knowledge, Beliefs and Attitudes? A Systematic Review”.

We have not selected one factor that is the most important for comparability, because the variables are not the same in each study. Thus, the principal factor should be identified for each study.

In our scale, we have specifically assigned one star for self-reported outcomes, because our study measures the intention to vaccinate. Two stars are given to the studies that assess the outcome with independent blind observers or with vaccination records, because these methods measure the practice of vaccination, which is the result of true intention.
Appendix 4.

Email correspondence

Original email-correspondence (Danish)

Date of the first contact. April 26 2016

Hej [Name]

som du ved forsøger jeg jo at skabe det her master speciale.
i den forbindelse er jeg som tidligere nævnt landet på dit studie fra 2013 med Svendsen, Svarrer, Vollenbroek - Hutten og Madeleine. Standardized activities of daily living.....

Hvis jeg husker ret sagde du at det blev omdømt til et pilot studie pga hvad ? kan du huske hvad den oprindelige tanke var mht størrelsen på jeres sample og hvilken hypotese i forventede at få bekræftet.

mvh
Brian Sørensen

Response   May 11 2016

Hej Brian

Der var forskellige udfordringer med rekruttering, afprøvning af målemetoder og tidrammer i studiet, hvilket gav begrænsninger (det skulle jo gøres færdig indenfor rammerne af et PhD studie):

“Individual differences as well as the experimental protocol and the relative small subject populations contributed to the relative few significant differences among groups found in our study. Thus, the current case-control study has to be viewed as a pilot study as relatively small, mixed-gender populations were included. This limits any generalization of these findings beyond this sample population. Further studies with larger population size and longitudinal design are needed to infer cause-effect relationships between pain and motor pattern”.

Korrelationen giver vel god mening I forhold til at belyse om der sammenhænge mellem de foreslåede parametre. Gør det ikke? Jeg ved godt at det er et svagt mål,
Dear Brian

As you are aware of I’m currently writing my dissertation. In relation to that I’ve retrieved the study in which you were involved in 2013 together with Svendsen, Svarrer, Vollenbroek - Hutten and Madeleine. Standardized activities of daily living…..

If I remember it right You mentioned that it was changed to a pilot study during the process of collecting data but I can’t remember why? Can You remember what the original thoughts were regarding sample size and what the hypothesis was?

Best regards
Brian Sørensen

Dear Brian

There were various challenges with recruitment, testing of the different measuring methods and a time schedule which did set some limitations. (It had to be finished within the timeframe of a PhD study):

“Individual differences as well as the experimental protocol and the relative small subject populations contributed to the relative few significant differences among groups found in our study. Thus, the current case-control study has to be viewed as a pilot study as relatively small, mixed-gender populations were included. This limits any generalization of these findings beyond this sample population. Further studies with larger population size and longitudinal design are needed to infer cause-effect relationships between pain and motor pattern”.

Investigating the correlation between different variables does make sense. Doesn’t it?
It might be a weak outcome measure however in a pilot study it might reveal some indications of correlation between variables that need to be addressed further.

Good luck with your study

Regards

Uffe
# Appendix 5.

Data extraction form for studies with cross-sectional and case-control designs

<table>
<thead>
<tr>
<th>Data extraction table</th>
<th>Data extracted by: Brian Sørensen</th>
<th>Date for data extraction; 09/05-2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study Author and year</td>
<td>1. Watson, PJ. et al. 1997</td>
<td>2. Geisser, ME. et al. 2004</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Item</th>
<th>Cross-sectional</th>
<th>Cross-sectional</th>
<th>Case-control</th>
<th>Case-control</th>
</tr>
</thead>
<tbody>
<tr>
<td>Study design</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Subjects</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Age (Mean)</td>
<td>43.7 years of age</td>
<td>40.6 years of</td>
<td>39.8 years</td>
<td>32.1 years</td>
</tr>
<tr>
<td>-Gender</td>
<td>15 males / 21</td>
<td>34 males / 42</td>
<td>34 males / 18</td>
<td>11 males / 11</td>
</tr>
<tr>
<td>-Duration of symptoms</td>
<td>4.6 years</td>
<td>85.5 months /</td>
<td>12.5 years</td>
<td>8.7 years</td>
</tr>
<tr>
<td>-Definition of chronic</td>
<td>Definition &gt;6</td>
<td>7.2 years</td>
<td>Definition&gt;1</td>
<td>Definition&gt;6</td>
</tr>
<tr>
<td>-Sample size</td>
<td>36 subjects</td>
<td>76 subjects</td>
<td>2 months</td>
<td>6 months</td>
</tr>
<tr>
<td>-Setting</td>
<td>Back pain centre</td>
<td>Spine programme</td>
<td>(50% of days)</td>
<td>Setting ND</td>
</tr>
<tr>
<td>Control subjects</td>
<td>No control</td>
<td>No control</td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Age</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Gender</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Definition of control</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Sample size</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-Setting</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pain related fear</td>
<td>FABQ</td>
<td>TSK-2</td>
<td>FABQ, STAI-Y,PCS</td>
<td>FABQ, STAI-Y,PCS</td>
</tr>
<tr>
<td>questionnaire</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

| Study design              |                  |                  |              |              |
| Control subjects          | No control       | No control       | 31.3 years   | 37.6 years   |
| -Age                      |                  |                  | 12 males / 10| of age,      |
| -Gender                   |                  |                  | females      | 6 males / 7  |
| -Definition of control    |                  |                  | Definition: no| Definition No|
| -Sample size              |                  |                  | past or current| back pain   |
| -Setting                  |                  |                  | low back pain | with 12      |
| Pain related fear         | FABQ             | TSK-2            | FABQ, STAI-Y,PCS| Setting ND   |
| questionnaire             |                  |                  |              |              |

| Study design              |                  |                  |              |              |
| Control subjects          | No control       | No control       | 31.3 years   | 37.6 years   |
| -Age                      |                  |                  | 12 males / 10| of age,      |
| -Gender                   |                  |                  | females      | 6 males / 7  |
| -Definition of control    |                  |                  | Definition: no| Definition No|
| -Sample size              |                  |                  | past or current| back pain   |
| -Setting                  |                  |                  | low back pain | with 12      |
| Pain related fear         | FABQ             | TSK-2            | FABQ, STAI-Y,PCS| Setting ND   |
| questionnaire             |                  |                  |              |              |

139
<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Item</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Low / high pain related fear</td>
<td>One group</td>
<td>One group</td>
<td>One group</td>
<td>One group</td>
<td>One group</td>
</tr>
<tr>
<td>Disability / pain questionnaire</td>
<td>VAS ODQ</td>
<td>VAS MPQ</td>
<td>VAS ODQ</td>
<td>NRS ODI</td>
<td>VAS ODI and GPAQ</td>
</tr>
<tr>
<td>Low / high disability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EMG back</td>
<td>L1-2 + L4-5</td>
<td>L3 + L5</td>
<td>L3-4 (ES)</td>
<td>L2-3+L4-5 (ES)</td>
<td>L5 (MF-S) Onset purpose only</td>
</tr>
<tr>
<td>EMG abdominal</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>No</td>
<td>IO/TrA + EO</td>
</tr>
<tr>
<td>BP / BW</td>
<td>10-450 Hz, BW</td>
<td>30-1000 Hz, BP</td>
<td>10-450 Hz, BP</td>
<td>10-450 Hz, BP</td>
<td>10-450 Hz, BP</td>
</tr>
<tr>
<td>FRP / FRR / FF</td>
<td>FRR</td>
<td>FRR</td>
<td>FF</td>
<td>FRR / FF</td>
<td>FF</td>
</tr>
<tr>
<td>Statistical method</td>
<td>Linear regression, Kolmogorov-Smirnow Goodness of Fit Test, Kolmogorov-Smirnow Z. multiple regression</td>
<td>Kolmogorov-Smirnow ANOVA Principal Component Analysis Multiple Regression Analysis</td>
<td>Kolmogorov-Smirnow ANOVA Mauchly test of sphericity Geisser Greenhouse Pearson correlation</td>
<td>Two way ANOVA Bonferroni correction Pearson’s correlation Spearmans rank-order Grubb’s test</td>
<td></td>
</tr>
<tr>
<td>Normalization</td>
<td>Not normalized</td>
<td>Normalized</td>
<td>Normalized</td>
<td>Normalized</td>
<td>Normalized</td>
</tr>
<tr>
<td>Normalization Submax /Max voluntary isometric / third option (3)</td>
<td>ND (3)</td>
<td>ND (3)</td>
<td>ND (3)</td>
<td>ND (3)</td>
<td>ND (3)</td>
</tr>
<tr>
<td>Ethics approval</td>
<td>Not mentioned</td>
<td>Approved</td>
<td>Approved</td>
<td>Approved</td>
<td>Approved</td>
</tr>
</tbody>
</table>
### Study Author and year
2. Geisser, ME. et al. 2004
3. Dubois, JD. et al. 2014
4. Henchoz, Y. et al. 2013

### Data extraction

#### Results

<table>
<thead>
<tr>
<th>Item</th>
<th>Males showing greater Mm activity (0.33 P=0.01) and ROM (0.36 P=0.01). Pain duration not associated with EMG. Correlation between FRR and psychosocial (Pain related fear = PRF) (-0.45 P=0.01) Correlation between PRF and decreased flex. (-0.55) and max EMG during flex (-0.38 P=0.01) However zero order correlation between PRF and EMG was not statistically significant Used chronic neuromuscular adaptations as a baseline / control. Increased EMG activity during FF with noxious heat/pain applied. Innocuous heat no change in EMG during FF. both baseline and experimenta l trials Expectation of pain was more pronounced that noxious pain which made no difference in EMG activity. L4-5 had an increased Mm. activity in full flexion was associated with higher PCS (0.54 P=0.012). No difference between in low vs high expectations in chronic group but difference in control group</th>
<th>No effect of side of pain. A main effect between groups of almost significance P=0.057 Significance reached at 2 of the 5 phases of flexion-extension task P=0.047 Significance in correlation between TSK score and increased muscle activity in 3 of 5 phases P=0.001 P=0.04 P=0.0001 Also significant after Spearman’s rank-order All above in IO/TrA not significant in EO</th>
</tr>
</thead>
<tbody>
<tr>
<td>Data extraction table</td>
<td>Data extracted by: Brian Sørensen</td>
<td>Date for data extraction: 09/05-2016</td>
</tr>
<tr>
<td>Study Author and year</td>
<td>1.Watson, PJ. et al.1997</td>
<td>2. Geisser, ME. et al. 2004</td>
</tr>
</tbody>
</table>

---

Data extracted by: Brian Sørensen

Date for data extraction: 09/05-2016

<table>
<thead>
<tr>
<th>Data extraction</th>
<th>table</th>
<th>Data extracted by: Brian Sørensen</th>
<th>Date for data extraction: 09/05-2016</th>
</tr>
</thead>
<tbody>
<tr>
<td>Limitation/description</td>
<td>is data “normal” No cause effect due to design how does perception of mm-activity affect influence data Is limited ROM related to abnormal FRP? Pain, disability and psychological scores was low to moderate. No cause effect due to design Cutaneous thermal stimulation-can expectations modulate neuromuscular activity = yes Suggest subgroups of low vs high pain related fear score</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Data extraction table

Data extracted by; Brian Sørensen
Date for data extraction; 09/05-2016

2. Geisser, ME. et al. 2004
3. Dubois, JD. et al. 2014
4. Henchoz, Y. et al. 2013


142