Available online at www.sciencedirect.com
ScienceDirect
Manual Therapy 13 (2008) 349–356
MANUAL THERAPY
www.elsevier.com/locate/math

Original article

The reliability of paraspinal muscles composition measurements using routine spine MRI and their association with back function

Annina Ropponen[a,*], Tapio Videman[b], Michele C. Battié[b]

[a]Institute of Biomedicine, Physiology/Ergonomics, University of Kuopio, P.O. Box 1627, 70211 Kuopio, Finland
[b]Faculty of Rehabilitation Medicine, University of Alberta, Edmonton, Canada

Received 20 April 2006; received in revised form 24 January 2007; accepted 9 March 2007

Abstract

This study examines the reliability of quantitative and qualitative muscle composition measurements of paraspinal muscle cross-sectional areas (CSAs) from routine lumbar spine magnetic resonance images and their association with maximal isokinetic lifting performance. The extent of paraspinal muscle composition reflects back function is currently not known. Measurements were repeated 4–8 weeks apart and different measurements of related constructs were compared. Participants were a population-based sample of 169 males, 35–67 years old, without considering the presence or absence of a history of low back pain or related problems in the selection of subjects. The quantitative and qualitative muscle composition measurements for axial magnetic resonance (MR) images of paraspinal muscles at the L3–L4 lumbar spine level, isokinetic lifting force and work, and body fat percentage were the main outcome measures. Results showed that the reproducibility of different paraspinal muscle composition measurements at the L3–L4 level was excellent for CSAs (ICC = 0.95–0.99) and quantitative muscle composition measurements using cerebrospinal fluid adjusted signal intensity (ICC = 0.96–0.99), and moderate for qualitative muscle composition ratings (Kappa = 0.54–0.76). The correlations of the quantitative and qualitative muscle composition measurements with isokinetic lifting force and work were generally low ($r = 0.02$–0.41), and favoured the qualitative assessments. In conclusion, quantitative and qualitative muscle composition measurements of paraspinal muscles are highly reproducible tissue measures, have low associations with body fat and isokinetic lifting performance, and show that paraspinal muscle morphology using routine spine magnetic resonance imaging (MRI) is poorly related to back function.

Keywords: Reliability; Paraspinal muscles; Magnetic resonance imaging; Muscle composition measurements; Lumbar spine

1. Introduction

The wide availability of magnetic resonance imaging (MRI) devices and lumbar axial views of patients with severe or chronic spinal disorders provides as opportunity to examine morphological aspects of the paraspinal muscles in relation to back pain problems. Also the available magnetic resonance (MR) images adds to the epidemiologic research of the lower back, where the focus would be on methods suitable for large subject samples, such as low back pain (LBP) patients. Easily obtainable measurements such as measuring muscle quality from routinely available low back MRIs would provide new possibilities to assess low back, whereas diagnostic purposes require advantaged methods, such as tissue specific imaging sequences in MRI (see e.g. Schick et al., 2002). Paraspinal muscle cross-sectional area (CSA) measurements have been used in both cross-sectional and longitudinal follow-up studies of prognosis or treatment outcomes (Gibbons et al., 1997a; Peltonen et al., 1998; Marras et al., 2001; Kim et al., 2005). In addition, measurements of muscle morphology offer objective assessment of muscularity (Parkkola and Kormano, 1992; Ranson et al., 2005; Mengiardi et al., 2006), as compared to muscle function tests that may be

*Corresponding author. Tel.: +358 17 163111; fax: +358 17 163112.
E-mail address: annina.ropponen@uku.fi (A. Ropponen).

1356-689X/$ - see front matter
doi:10.1016/j.math.2007.03.004

influenced by such factors as motivation and pain (Lackner and Carosella, 1999; Keller et al., 1999). Paraspinal muscle CSA has been associated to some degree with the muscle's capacity to generate force and has been claimed to be an objective measurement for back function (Bruce et al., 1997; Keller et al., 1999; Guyton and Hall, 2000). However, the association between paraspinal muscle CSA and back function related factors, such as disability and back pain, have been controversial (Käser et al., 2001; Mannion et al., 2000). This lack of consistency in the relationship between CSA and back function may be because of failure to take into account paraspinal muscle composition, such as the degree of fatty infiltration, in addition to CSA. Isokinetic trunk extension performed with maximal force at constant speed has been found to be moderately associated with paraspinal muscle composition in one study ($r = 0.61$), and poorly in another ($R^2 = 4\%$) (Hultman et al., 1993; Keller et al., 1999). Back function tests of submaximal force levels and endurance had no clear association with paraspinal muscle composition (Hultman et al., 1993; Gibbons et al., 1997). Although the association between the paraspinal muscle CSA and the body fat percent has been shown to be moderate (Gibbons et al., 1998), we are not aware of studies based on the connection between paraspinal muscle composition and body fat percent.

The intra- and inter-rater reproducibility of paraspinal muscle CSA measurements were good to excellent, with intra-class correlation coefficients (ICC) ranging from 0.91 to 0.99 (Gibbons et al., 1997; Peltonen et al., 1998; Räty et al., 1999; Marras et al., 2001). However, the MR images of back muscles from healthy individuals and those with muscle atrophy look very different (Kader et al., 2000). Decreased general activity levels may influence the ratio of muscle, connective and fat tissues, reorganization of collagen fiber directions, atrophy in some sites, and fibrosis in others, without affecting the CSA of the whole muscle (Reid and Costigan, 1987; Mayer et al., 1989). Thus, the quantitative muscle composition measurements have been used in only a few studies investigating lumbar axial MR images (Parkkola and Kormano, 1992; Flicker et al., 1993; Gibbons et al., 1997; Ranson et al., 2005; Mengiardi et al., 2006). Of these, only Gibbons et al. (1997) used 20 healthy subjects and Ranson et al. (2005) used six healthy subjects, where they reported the reproducibility of the paraspinal muscle composition measurements, using muscle signal intensity as an indicator of composition (ICC$>$0.99). In comparison, paraspinal muscle density for 31 chronic LBP patients assessed from computed tomography images had a 2–4% measurement error due to the observer (Keller et al., 2003). Intra-rater agreement for the amount of intermuscular fat in upper cervical muscles from signal intensities of MR images was also high, ICC = 0.94–0.98 (Elliott et al., 2005). The single study we found based on the reliability of the qualitative ratings for paraspinal muscle fat content in LBP patients showed good inter-rater reliability (Kappa = 0.85) (Kader et al., 2000). We are aware of only three other studies that used qualitative ratings of muscle composition; Parkkola and Kormano (1992) studied 74 healthy subjects; Mooney et al. (1997) studied 8 patients with LBP, and Mengiardi et al. (2006) compared 25 chronic LBP patients with 25 matched asymptomatic volunteers but none reported the reproducibility of their measurements. A study combining both quantitative and qualitative muscle composition measurements (Parkkola and Kormano, 1992) has shown that the quantitatively analyzed non-muscle tissue categorized by a qualitative measurement of muscle composition corresponded well the quantitative measure of the intramuscular fat. Parkkola and Kormano (1992) concluded that the qualitative measurement would offer a reliable tool for muscle composition assessment whereas a recent study of Mengiardi et al. (2006) could not use the qualitative muscle composition measurement for differing LBP patients from volunteers.

The overall objectives of the present study were to investigate the intra-rater reliability of quantitative and qualitative measurements of total muscle CSA and composition, using the digital data from routine lumbar spine MRIs of subjects without current back disorders. In addition to repeating the measurements, we investigated the correlations of the quantitative and qualitative ratings of the muscle composition (e.g., fat infiltration) with body fat percentage and isokinetic lifting strength. The correlations of isokinetic lifting strength with total paraspinal muscle CSA and paraspinal muscle CSA after adjusting for measurements of muscle composition were compared. It is postulated that the latter correlation with adjusted muscle CSA will be higher. It is also expected that the reproducibility for quantitative muscle composition measurement will be higher in the well defined psoas major muscle than in the erector spinae muscle mass, which is a combination of three muscles (multifidus, longissimus and iliocostalis muscles) with overlapping muscle fascias, and deep or superficial cavitas between the muscles and fat deposits.

2. Methods

2.1. Subjects

Study subjects were a population-based sample of 169 men, ranging from 35 to 67 years old with a mean age of 49 years (standard deviation [SD] 7.8), a mean height of 176 cm (SD 7.3), and a mean weight 80 kg (SD 11.5). The subjects consisted of a random sample from a larger