ARTICLE IN PRESS

Journal of Biomechanics ∎ (∎∎∎∎) ∎∎∎–∎∎∎

Contents lists available at ScienceDirect

Journal of Biomechanics

journal homepage: www.elsevier.com/locate/jbiomech www.JBiomech.com

Identification of intrinsic and reflexive contributions to low-back stiffness: medium-term reliability and construct validity

Christian Larivière [a,f,*], Daniel Ludvig [b,c], Robert Kearney [d], Hakim Mecheri [a,f], Jean-Maxime Caron [e,f], Richard Preuss [e,f]

[a] Occupational Health and Safety Research Institute Robert-Sauvé, Montreal, Quebec, Canada
[b] Sensory Motor Performance Program, Rehabilitation Institute of Chicago, Chicago, IL, USA
[c] Department of Physical Medicine & Rehabilitation, Northwestern University, Chicago, IL, USA
[d] Biomedical Engineering Department, McGill University, Montreal, Canada
[e] School of Physiotherapy and Occupational Therapy, McGill University, Montreal, Canada
[f] Centre for Interdisciplinary Research in Rehabilitation of Greater Montreal (CRIR), Montreal Rehabilitation Institute, Montreal, Canada

ARTICLE INFO

Article history:
Accepted 25 November 2014

Keywords:
Low back pain
Rehabilitation
Reproducibility
Trunk stiffness
Reflex

ABSTRACT

This study aimed at testing the reliability and construct validity of a trunk perturbation protocol (TPP) that estimates the intrinsic and reflexive contributions to low-back stiffness. The TPP consists of a series of pseudorandom position-controlled trunk perturbations in an apparatus measuring forces and displacements at the harness surrounding the thorax. Intrinsic and reflexive contributions to low-back stiffness were estimated using a system identification procedure, leading to 12 parameters. *Study 1 methods (reliability):* 30 subjects performed five 75- s trials, on each of two separate days (eight weeks apart). Reliability was assessed using the generalizability theory, which allowed computing indexes of dependability (ϕ, analogous to intraclass correlation coefficient) and standard errors of measurement (SEM). *Study 2 methods (validity):* 20 healthy subjects performed three 75- s trials for each of five experimental conditions assumed to provide different lumbar stiffness; testing the construct validity of the TPP using four conditions with different lumbar belt designs and one control condition without. *Study 1 results (reliability):* Learning was seen between the first and following trials. Consequently, reliability analyses were performed without the first trial. Simulations showed that averaging the scores of three trials can lead to acceptable reliability results for some TPP parameters. *Study 2 results (validity):* All lumbar belt designs increased low-back intrinsic stiffness, while only some of them decreased reflex stiffness, which support the construct validity of the TPP. Overall, these findings support the use of the TPP to test the effect of rehabilitation or between-groups differences with regards to trunk stiffness.

1. Introduction

Understanding the central nervous system's (CNS) strategy in responding to trunk perturbations could help prevent back injuries and develop more effective rehabilitation and treatment programs for such injuries. To achieve this goal, several studies have used surface EMG measures, combined with trunk perturbations, to elicit and measure trunk muscle reflex responses (Lariviere et al., 2010b; Magnusson et al., 1996; Radebold et al., 2000; Reeves et al., 2005). However, EMG-based measures of back muscle reflex responses show, at best, modest reliability, even when data from several trials are combined (requiring lengthy experiment times) (Santos et al., 2011). Trunk muscles work in synergy to ensure an adequate overall reflex response; this allows different muscles to share the load a little differently from one disturbance to another. This may explain the low reliability of EMG measures in some conditions.

An appealing alternative to resolve these drawbacks is to collect the force and displacement signals during repeated random trunk perturbations, and to use a mathematical model to quantify the CNS's response. Previous studies have used this approach to separate and estimate the intrinsic and reflexive contributions to joint stiffness (Hendershot et al., 2012; Kearney et al., 1997; Moorhouse and Granata, 2007). For simplicity, the term "intrinsic stiffness" refers to all contributions to the restoring force following a perturbation, in the absence of a reflex or voluntary change in muscle recruitment. This includes the inertial, viscous and elastic properties (passive components) of the trunk (provided by joints, muscles,

* Corresponding author at: Occupational Health and Safety Research Institute Robert-Sauvé, Montreal, Quebec, Canada. Tel.: +514 288 1551x217; fax: +514 288 6097.
E-mail address: larchr@irsst.qc.ca (C. Larivière).

http://dx.doi.org/10.1016/j.jbiomech.2014.11.036
0021-9290/

connective tissues, etc.), as well as the steady-state muscle recruitment (active component). This approach has several advantages: (1) it requires much less time to collect the required data; (2) it provides a dynamic quantification of lumbar stiffness (3) it estimates the intrinsic mechanical contributions to lumbar stiffness, in addition to its reflexive contribution and (4) it measures the net response of all trunk muscles in synergy; (5) it helps separate the behavior of the system from the properties of the perturbation. We have adapted such a trunk perturbation protocol (TPP), based on previous work (Moorhouse and Granata, 2007).

We hypothesize that using a rich perturbation signal and looking at the net restoring force will lead to reliable estimates of intrinsic and reflexive contributions to lumbar stiffness, collectively termed trunk mechanical behaviors (TMB). Only one previous study has examined the reliability of TMB estimates; it found excellent within-day and moderate between-day reliability results when using a "more comfortable" plastic harness (Hendershot et al., 2012). Short-term reliability (between-day interval varying between 3 and 14 days approximately) was assessed in that study. Here we were interested to medium-term reliability (eight-weeks interval), which more likely mimics the duration of a rehabilitation program. In the present study, we were also interested in substantiating the validity of the TMB. There is no "gold standard" measure of lumbar joint stiffness to substantiate its criterion validity; rather we will test its construct validity by using lumbar belts that are known to increase lumbar stiffness (McGill et al., 1994). Consequently, the aims of this study were to assess the medium-term test-retest reliability and construct validity of a TPP that estimates the intrinsic and reflexive contributions to low-back stiffness.

2. Methods

2.1. Subjects and tasks

Study 1 (reliability). Fifteen healthy men (Age: 39 ± 14 yr; Height: 1.78 ± 0.08 m; Mass: 77 ± 11 kg) and 15 healthy women (Age: 40 ± 14 yr; Height: 1.64 ± 0.06 m; Mass: 63 ± 10 kg), aged between 18 and 65 (inclusion criterion), volunteered to participate. Exclusion criteria were back pain in the preceding year or any history of back pain lasting more than one week; surgery of the pelvis or spinal column; scoliosis; systemic or degenerative disease; body mass index (BMI) over 31.5 kg/m^2 (women) or 33 kg/m^2 (men); one positive response to the Physical Activity Readiness Questionnaire (Thomas et al., 1992); history of neurological disease or deficits not related to back pain (e.g. stroke, peripheral neuropathies, balance deficits); anticonvulsive, antidepressive and anxiolitic medication; pregnancy and claustrophobia. All subjects were informed about the experimental protocol and its potential risks, and gave written consent prior to their participation. The study and consent form were approved by the ethics committee of the Centre for Interdisciplinary Research in Rehabilitation of Greater Montreal.

The protocol followed these steps: (1) a 75- s trial was recorded without the subject in the apparatus; (2) the subject was positioned in the apparatus and instructions were provided; (3) two 30- s trials (the first without and the second with perturbations) were provided to familiarize the subject to the feedback and perturbations and (4) five 75- s trials were carried out, separated by 2- min. rest. This protocol was performed twice (same steps), with sessions separated by eight-weeks. The harness tightness, as measured by the distance between the front and back panels, was kept the same. The time of day (morning, afternoon, evening) the subjects were assessed was not controlled.

Study 2 (construct validity). Ten healthy men (Age: 26 ± 8 yr; Height: 1.80 ± 0.06 m; Mass: 80 ± 13 kg) and 10 healthy women (Age: 27 ± 11 yr; Height: 1.68 ± 0.07 m; Mass: 65 ± 10 kg) were recruited, with male and female subjects matched for age. The inclusion and exclusion criteria were the same as for study 1, but with different subjects recruited. The study and consent form were approved by the ethics committee of the Centre for Interdisciplinary Research in Rehabilitation of Greater Montreal.

The protocol was the same as for study 1, except that step 4 comprised three 75- s test trials for each of five experimental conditions. These conditions were assumed to provide different lumbar stiffness, thus testing the construct validity of the TPP. The conditions were as follows: (C1) control; (C2) extensible lumbar belt (LB); (C3) extensible LB with back panel; (C4) extensible LB with back and front panels; (C5) non-extensible LB. The C5 condition (non-extensible LB) was not intended to represent a given "level" of belt stiffness but may act on intra-abdominal pressure differently than the extensible belt and may consequently have an indirect effect on TMB estimates. The order of the conditions was randomly assigned, but all three trials for a given condition were performed before switching conditions.

The first LB model was extensible, allowing for insertion of dorsal and ventral panels (model LumboLux from Hope Orthopedic). The second was a non-extensible LB without panels (model 582 from MBrace). Both LBs had two layers of Velcro straps; the first layer allowing for initial adjustment and placement of the LB, and the second layer (elastic material for the extensible LB and two non-extensible nylon straps for the non-extensible LB) allowing the final tension adjustment. The optional ventral panel for the extensible LB was semi-rigid and covered with Velcro material, allowing it to be anchored with the first strap. The dorsal panel was a rigid Kydex insert, with a hole in the middle to allow the spine to flex without discomfort.

The tension of each LB was adjusted with the use of a thin (2 mm) 4.0 × 4.0 cm FSR sensor (Force Sensing Resistor; Interlink Electronics; model FSR400) fastened on the skin between the lateral aspect of the left iliac crest and the 12th rib. Using a feedback system, the subject adjusted the tension of the LB to reach a pressure of 70 mmHg or 9.3 KPa (Cholewicki et al., 2010), allowing a 5% error.

2.2. Trunk perturbation protocol (TPP)

The trunk perturbation assessment was comparable to a previously published protocol (Moorhouse and Granata, 2007). The subject stood upright in a position-controlled trunk perturbation apparatus (Fig. 1). The harness was tightened as much as possible according to the subject comfort (e.g. breathing). Each trial lasted 75 s, with the last 60 s retained for analysis (220 forward/backward perturbations). Throughout the trial, position perturbations consisting of a pseudorandom binary sequence with a switching rate of 150 ms and amplitude of 4 mm were applied at the T8 level. Sample position perturbation, measured force and estimated reflex and intrinsic components are shown in the frequency domain (Fig. 2) and time domain (Fig. 3). The profile of the perturbation (amplitude: 4 mm; velocity: 400 mm/s, acceleration and deceleration: 10 000 mm/s^2,) was designed so as to generate the 4-mm forward-backward displacements in less than 40 ms, the latter being less than back muscle reflex latencies (Granata et al., 2004; Lariviere et al., 2010b). Across the 36 subjects, the averaged rising time was 24 ± 3 ms (range: 14 to 41 ms). During the perturbation sequences, the subject was instructed to relax their abdominal muscles and to sustain a small preloading of the back muscles corresponding to a L5/S1 extension moment of 15 and 10 Nm in men and women, respectively (Fig. 3), taking into account the measured distance (lever arm) between the motor shaft and the L5/S1 joint. This preloading represents approximately 5% of back strength (L5/S1 extension moment), which was estimated at 300 Nm and 200 Nm in men and women, respectively, when the trunk is upright (Lariviere et al., 2002, 2006, 2009, 2010a). The load cell forces provided as feedback on a monitor were heavily filtered (low-pass zero-phase Butterworth filter; 4th order; cut-off frequency: 1 Hz) to emphasize the subject's voluntary torque by greatly reducing the torque produced in response to the perturbations. Subjects were asked to gradually adjust their back extension effort to maintain a 15-Nm or 10-Nm target preload through the perturbation sequence.

2.3. Data processing

The intrinsic (INT) and reflex (REF) contributions to trunk stiffness (Kearney et al., 1997; Moorhouse and Granata, 2007) were estimated from the force and displacement signals using the system identification methods described in the appendix and illustrated in Fig. 4. This produced the following performance statistics: (1) %VAF $F_{INT-TOTAL}$; (2) %VAF $F_{REF-TOTAL}$; (3) %VAF$_{TOTAL}$; (4) K (N/m): elastic component of intrinsic stiffness; (5) B (N · s/m): viscous component of intrinsic stiffness; (6) I (N · s^2/m): inertial component of intrinsic stiffness; (7) Gain (N · s^2/m): reflex gain. %VAF means percentage of variance accounted for. As detailed in the appendix, the %VAF values were first computed using the entire frequency spectrum and then, for three frequency range (low: 2–5 Hz; middle: 5–10 Hz; high: 10–20 Hz) using different bandwidths of the force signals. These variables will be hereafter collectively termed trunk mechanical behaviors (TMB).

2.4. Statistics

Study 1 (reliability). Reliability was assessed using the generalizability theory framework (Shavelson and Webb, 1991). Briefly, this theory consists of two parts: the first is the G-study that estimates the variance components (from the 2-way ANOVAs with repeated measures: 2 days × 5 TRIALS) and the second is the D-study, giving the estimated reliability for various designs aside from the one used in the G-study. These ANOVAs and the corresponding P values were also used to identify any systematic bias across trials and days, which would be ascribed to a learning effect. A random D-study model was used, allowing for all sources of variance to contribute to measurement error, which means that it will be possible to generalize the results to any day or trial. In the D-study, the sources of variances were used to calculate the standard error of measurement (SEM) and the dependability coefficients (ϕ) corresponding to the averaging of one, three or five trials within a day, which represent relatively acceptable measurement strategies. The coefficient ϕ is analogous