Review

Spinal muscle evaluation using the Sorensen test: a critical appraisal of the literature

Christophe Demoulin *, Marc Vanderthommen, Christophe Duysens, Jean-Michel Crielaard

Physical Medicine and Rehabilitation Unit, Liège University, Liège, Belgium

Received 3 May 2004; accepted 4 August 2004

Available online 05 October 2004

Abstract

The first test for evaluating the isometric endurance of trunk extensor muscles was described by Hansen in 1964. In 1984, following a study by Biering-Sorensen, this test became known as the “Sorensen test” and gained considerable popularity as a tool reported to predict low back pain within the next year in males. The test consists in measuring the amount of time a person can hold the unsupported upper body in a horizontal prone position with the lower body fixed to the examining table. This test has been used in many studies, either in its original version or as variants. Although its discriminative validity, reproducibility, and safety seem good, debate continues to surround its ability to predict low back pain; in addition, the gender-related difference in position-holding time remains unexplained and the influence of body weight unclear. A contribution of the hip extensor muscles to position holding has been established, but its magnitude remains unknown. The influence of personal factors such as motivation complicates the interpretation of the results. Despite these drawbacks, the Sorensen test has become the tool of reference for evaluating muscle performance in patients with low back pain, most notably before and after rehabilitation programs.

© 2005 Elsevier SAS. All rights reserved.

Keywords: Muscle endurance; Low back pain; Trunk extensor muscles; Physical evaluation; Spinal muscles

1. Introduction

Despite growing research efforts, nonspecific low back pain remains a major public health burden throughout the industrialized world. Epidemiological data indicate an annual prevalence of about 39–54% [1,2] and a lifetime prevalence of 60–65% [1–3]. Costs to society stem mainly from chronic forms, which account for only 5–10% of cases [4,5]. Chronic nonspecific low back pain results in both physical and psychological deconditioning that traps the patient in a vicious circle characterized by decreased physical performance, exacerbated nociceptive sensations, impaired social functioning, work disability, and depression. The physical component of deconditioning involves both stiffness of the lumbar spine-pelvic-femoral unit, decreased muscle strength and endurance, loss of cardiorespiratory adjustment to physical exertion, and neuromuscular inhibition (kinesiophobia) [6].

Several studies suggest that patients with chronic nonspecific low back pain may benefit from an active multidisciplinary approach involving an individually tailored reconditioning program [6–8]. Tools capable of quantitating individual deficiencies and of documenting the effects of reconditioning would be useful. Evaluating the endurance of trunk extensor muscles seems to have greater discriminative validity than evaluation of maximal voluntary contractile force [9–14].

In 1964, Hansen developed the first test for evaluating the isometric endurance of trunk extensor muscles and validated it in 168 healthy individuals and 90 patients who had had surgery for low back pain within the last 3–4 weeks [15]. In this test, the patient is in the prone position with the lower body fixed to the examining table and the upper body extending beyond the edge of the table. The test consists in holding the upper body horizontal as long as possible. In 1972, Troup et al. evaluated muscle fatigability by performing surface electromyography in patients during the test [16]. After a 1984 study by Biering-Sorensen [9] published in Spine, the test became known as the “Sorensen test”. Biering-Sorensen...
used the test together with several other evaluations in over 900 individuals and concluded that a shorter position-holding time during the Sorensen test predicted low back pain within the next year in males.

2. Methods

We conducted a Medline search for articles reporting the use of the Sorensen test to evaluate the back muscles. The keywords used for the search were “Sorensen test”, “back pain”, “muscle endurance”, “static endurance”, “muscle fatigue”, “function test”, “back extension”, and “trunk extensors”. In addition, the reference list of each article retrieved by the Medline search was examined for additional related articles.

3. Description of the Sorensen test

The patient lies on the examining table in the prone position with the upper edge of the iliac crests aligned with the edge of the table. The lower body is fixed to the table by three straps, located around the pelvis, knees, and ankles, respectively. With the arms folded across the chest, the patient is asked to isometrically maintain the upper body in a horizontal position (Fig. 1). The time during which the patient keeps the upper body straight and horizontal is recorded. In patients who experience no difficulty in holding the position, the test is stopped after 240 s.

The Sorensen test is the most widely used test in published studies evaluating the isometric endurance of trunk extensor muscles. It has been used either as described initially or with a number of modifications, as listed below.

- **Arm position**: the test has been used with the arms bent, the elbows held out, and the hands on the ears [17], forehead [18], or nape of the neck [19,20]; in another variant, the arms are held along the sides [13,21,22]. Because arm position influences the location of the center of gravity, these modifications affect the mass moment of the upper body and, therefore, test performance [23].
- **Location of the edge of the table**: in several studies, the anterior–superior iliac spines were placed at the edge of the table, instead of the upper edge of the iliac crests [24,25].
- **Number of straps**: two [19,26–28] to five straps [14] have been used to hold the lower body to the table; in the Roman chair variant (Fig. 2), the feet are fixed to the device and no straps are needed [29–34].
- **Starting position**: in several studies [31,34], the test was started with the upper body sloping downward toward the floor so that a concentric contraction of the trunk extensor muscles was needed initially to reach the horizontal position.
- **Hip flexion**: in theory, the hips remain fully extended throughout the Sorensen test. However, hip flexion was 6° in a study by Holmstrom et al. [10] and 40° in a study by Dedering et al. [29,30] (Fig. 3).
- **Method for documenting the horizontal position of the upper body**: whereas many authors (including Biering-Sorensen) simply trusted a visual evaluation [10,19,21,27,31,34–36], others used measurement devices (inclinometer [24,37,38], goniometer [18], or photoelectric cell [10,39]) or asked the patient to maintain contact between the back and a stadiometer or weight hanging from the ceiling [18,40,41].
- **Criteria for stopping the test**: in studies that used measurement devices or contact with an object to define the horizontal position, specific test-stopping criteria were used, such as trunk downsloping by more than 5–10° [24,37,42] or loss of contact with the object for more than 10 s [43].
• Test duration: some individuals can hold the position for longer than 240 s [9,15,46], and Jørgensen and Nicolai suggested that the test should be continued beyond this time [46]. The populations also varied across studies, from patients with prior low back pain [10,24] or current chronic low back pain [20,31,39,42] to patients with no history of low back pain [10,24,34,35] or 15-year-old students [47,48]. In addition, several studies failed to separate data from males and females [22,24].

As expected, these numerous methodological variations translate into considerable discrepancies in study findings (Table 1). Biering-Sørensen [9] and Holmstrom et al. [10] found high mean position-holding times, not only in healthy individuals (198 and 171.5 seconds, respectively), but also in patients with low back pain (163 and 137.5 s, respectively). Most other studies found lower values in healthy individuals [21,22,24,35,44].

4. Predictive validity of the Sorensen test

Biering-Sørensen [9] reported that a position-holding time less than 176 s predicted low back pain during the next year in males, whereas a time greater than 198 s predicted absence of low back pain. Importantly, the test had no predictive validity in females. In a study by Luoto et al. [13], separating the participants into three groups based on position-holding times showed that a time less than 58 s was associated with a threefold increase in the risk of low back pain, as compared to a time greater than 104 s.

Sjolie and Ljunggren et al. [49] and Adams et al. [50] found that the Sorensen test predicted low back pain in both males and females, whereas others found no predictive value [51–54]. Mannion et al. [44] reported that the risk of low back pain was independent from the position-holding time but was correlated with muscle fatigability as measured by surface electromyography during the test. In miners, Stewart et al. [55] found no significant differences in position-holding time between individuals with and without a history of low back pain.

5. Discriminative validity of the Sorensen test

In many studies, the position-holding time was significantly decreased in patients with chronic low back pain [9,11,15,22,24,39,56,57]. This finding suggests that chronic low back pain may be associated with decreased isometric endurance of the trunk extensor muscles.

6. Biometric characteristics

In several studies, neither body weight [42,58] nor mass moment of the trunk [10] influenced the position-holding time. Other studies, however, found a negative correlation between body weight and position-holding time [9,19,21,27,40]. The potential influence of age and stature remains debated [9,10,21,27]. In contrast, there is general agreement that differences exist between males and females. With a few exceptions [21,42,59], studies found significantly longer position-holding times in females [9,35,40,44,60]. Similarly, spectral analysis of electromyography signals recorded during the test indicated greater muscle fatigability in males [14,29,40,44,58]. Several hypotheses have been put forward to explain this gender-related difference. In females, the weight of the upper body is less and the center of gravity of the trunk lowers, as compared to males [9,11]. However, the position-holding time remained longer in females wearing weights attached to the upper body [61] or performing isometric trunk muscle endurance tests in the standing position [11]. The greater degree of lumbar lordosis in females may afford a mechanical advantage by lengthening the lever arm of the spinal erector muscles [62,63]. An influence of sex hormones has been suggested also [11,35]. Nevertheless, the most compelling hypothesis involves differences in muscle composition. Mannion et al. [64] suggested that the spinal muscles may show better adaptation to aerobic exercise in females as a result of a larger proportion of slow Type I fibers in the cross-sectional muscle area.

Table 1

<table>
<thead>
<tr>
<th>Endurance (s)</th>
<th>Males</th>
<th></th>
<th></th>
<th>Females</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Healthy</td>
<td>Prior LBP</td>
<td>LBP</td>
<td>Healthy</td>
<td>Prior LBP</td>
<td>Current LBP</td>
</tr>
<tr>
<td>Biering-Sørensen [9] n &gt; 900</td>
<td>198 a</td>
<td>176 b</td>
<td>163</td>
<td>197</td>
<td>210</td>
<td>177</td>
</tr>
<tr>
<td>Holmstrom et al. [10] n = 203</td>
<td>171.5 a</td>
<td>166.7 a</td>
<td>137.5 b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mannion and Dolan [35] n = 229</td>
<td>116 a</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hultman et al. [39] n = 148</td>
<td>150 a</td>
<td>136 b</td>
<td>85 b</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mannion et al. [44] n = 200</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

The position-holding times are given in seconds; n is the number of individuals in the study; and LBP means low back pain. a and b indicate populations of LBP patients with and without an impact of their disease on their daily activities. Different letters in superscript indicate significant differences (P < 0.05).
7. Reproducibility

The reproducibility of the Sorensen test has been evaluated, but the studies either included small numbers of individuals [10,11,30,35,60] or used the correlation coefficient \( r \), which is not optimal for assessing test reproducibility [10,21,34,36,38,57]. Investigations that relied on the intraclass coefficient of correlation (ICC) usually found that reproducibility was satisfactory (ICC > 0.75) [65] both in healthy individuals and in patients with low back pain. Simmonds et al. [22] reported that ICC values were 0.73, 0.68, and 0.99 for within-session, session-to-session, and intraobserver reproducibility, respectively, in healthy individuals; corresponding values in patients with low back pain were 0.91, 0.88, and 0.99. In a study of interobserver reproducibility, Latimer et al. found ICC values of 0.77, 0.83, and 0.88 in individuals with prior low back pain, no symptoms, and current low back pain, respectively [24].

Reproducibility was poorer when a Roman chair was used for the test. In a study of 12 healthy individuals conducted by Mayer et al. [34], the correlation coefficient was 0.2. Keller et al. [31] reported that the coefficient of variation was 20–21% in 31 patients with low back pain and 31 healthy individuals.

8. Validity

Although Hansen [15] described the test as a tool for evaluating back strength, studies subsequently established that it assesses isometric muscle endurance. Furthermore, the muscle contractions elicited by the test were found to be no greater than 40–52% of the maximal voluntary contractile force [10,25,35,45,66,67]. Similarly, the electromyographic activity of the spinal erector muscles rarely exceeded 40% of its maximal value [67,68].

The Sorensen test has been misinterpreted as a specific tool for evaluating the back muscles [13]. Published studies demonstrate that the test assesses the endurance of all the muscles involved in extension of the trunk, which include not only the paraspinal muscles, most notably the multifidus muscle [25,69], but also the hip extensor muscles. The contribution of the gluteus maximus and hamstring muscles remains controversial [42,70,71]. Sparto et al. [72] and Arzokusky et al. [69] stated that these muscles played a minor role, whereas others reported a correlation between the position-holding time and the time-course of hip extensor fatigability as assessed by surface electromyography, suggesting a significant role for the hip extensor muscles [40,42].

Another challenge to the validity of the Sorensen test comes from evidence that the position-holding time is unrelated to the cross-sectional area of the paraspinal muscles [39,73]. An influence of individual factors such as motivation, pain tolerance, and competitiveness has been suggested [66,74]. In individuals who stop the test because of pain or breathing problems [9,24,27,28,35], the result may not reflect muscle performance. In this situation, a submaximal Sorensen test coupled with electromyography or a Borg scale evaluation might be sufficient for documenting muscle fatigability while limiting the role for individual factors such as motivation [29,35].

9. Sensitivity to change

The position-holding time has been reported to increase significantly after active rehabilitation therapy [38,75]. A training program involving dynamic exercises performed on a Roman chair regularly over several weeks was followed by a significant increase in the position-holding time [71], although measurements obtained using a dedicated dynamometer showed no increase in back extensor muscle strength [32,71].

10. Spinal loads induced by the Sorensen test

Callaghan et al. [76] estimated that the compression load imposed on the spine during the brief Sorensen test was 4000 N, which is slightly above the value recommended by the National Institute of Occupational Security and Health in 1981 [77].

Pain, including spinal pain, may cause the patient to discontinue the test [9,24,27,28,35]. However, no persistent adverse effects such as pain exacerbation have been reported. Simmonds et al. [22] found high within-session reproducibility and recorded no instances of lasting pain induced by the test. Measurements of the lumbar curvature with the trunk in a horizontal position showed no increase in the normal lumbar lordosis [76]. Nevertheless, the suggestion by Tsuboi et al. [14] that the test be performed with the trunk extended 5° above the horizontal seems inappropriate.

11. Variants of the Sorensen test

11.1. Dynamometric measurements

Tests for evaluating isometric spinal muscle endurance using a dynamometer have been developed [11,44,78]. Jorgensen et al. [11] reported that dynamometric measurements obtained at a fixed percentage of the voluntary maximal force (usually 50–60%) were superior over the Sorensen test with a 4-min maximum in several ways: the influence of anthropometric factors was smaller, reproducibility was better (with a correlation coefficient of 0.89 as compared to 0.82), discriminative validity was greater, and the time needed for the test was shorter, resulting in a smaller influence of motivation. However, whereas the Sorensen test requires only submaximal muscle contraction, dynamometric tests require determination of the maximal voluntary force, which may be inappropriate in some patients. Furthermore, pain may result
in underestimation of the maximal voluntary force and therefore in spurious endurance test results.

11.2. The Ito test (Fig. 4)

Ito et al. [79] developed a test for evaluating isometric endurance of the trunk extensor muscles. The patient lies in the prone position with a pad under the abdomen and the arms along the sides. When a signal is given, the individual lifts the upper body while flexing the neck as much as possible and contracting the gluteus maximus muscles to stabilize the pelvis. The test consists in holding this position as long as possible (without exceeding 5 min) while breathing normally. This simple test was described by Ito et al. as inducing less lumbar lordosis than the Sorensen test. Shirado et al. [80] reported that maximum neck flexion together with pelvic stabilization resulted in maximum activity of the erector spinae muscles. In a study of 190 individuals, Ito et al. [79] found that their test discriminated between healthy controls and patients with low back pain. Evaluation of session-to-session reproducibility showed that the ICC was 0.97 in healthy controls and 0.93 in low-back-pain patients [79].

The Ito test does not seem to induce pain and may result in less spinal loading than the Sorensen test. Furthermore, as the lower body is not fixed, the contribution of the hip extensor muscles may be smaller [68]. Factors that limit international recognition of the Ito test include the absence of studies, the absence of a standardized test procedure (type of pad, extent of upper-body lifting, etc.), and the theoretical risk of exaggerating the degree of lumbar lordosis.

11.3. McIntosh tests

Moreau et al. [81] noted that two other tests have been described in the literature, for evaluating the isometric endurance of the upper and lower spinal extensor muscles, respectively. However, the only published study of these tests [82] fails to provide information on validity, safety, and potential use in low-back-pain patients.

11.4. Dynamic evaluations

Repetitive arch-up tests provide a dynamic evaluation of the trunk extensor muscles without requiring the use of a dynamometer [21,59,83–86]. The position is derived from the Sorensen test. The test consists in flexing the trunk at 45° then returning to the horizontal position as many times as possible, at a rate of one arch-up every 2–3 s.

Moreland et al. [28] used a test in which the lower limbs were fixed to a triangular pad and the patient was asked to flex the trunk so as to touch the table with the nose then to return to the horizontal position at a rate of 25 arch-ups per minute (Fig. 5). In a study by Uderman et al. [87], the test was done on a Roman chair and the patients asked to arch up repeatedly over a 90° angle.

Whereas the static version of the Sorensen test has been widely used in published studies, the dynamic variant has received less attention. Available data suggest that results may be similar in males and females [61] and that the discriminative validity may be decreased as compared to the static test [21,59]. As with the static test, the occurrence of pain or cramps may limit the validity of the test [28,83]. In addition, the contribution of the hip extensor muscles increases gradually from one repetition to the next, further limiting the validity of the test [88,89]. Thus, in-depth studies are needed to evaluate the usefulness and limitations of these dynamic tests.

12. Conclusions

The Sorensen test allows for a rapid, simple, and reproducible evaluation of the isometric endurance of the trunk extensor muscles. It discriminates between healthy individuals and patients with low back pain and may predict the occurrence of low back pain in the near future. Although the Sorensen test has been extensively studied, the better performance among females remains partly unexplained and the contribution of the hip extensor muscles is unknown. The absence of a single standardized test procedure is an impediment to comparative studies. The role for motivation, the fact that pain can lead patients to stop the test, and the impossibility of quantifying the relative muscle strength developed by the individual constitute the major shortcomings of the Sorensen test. Nevertheless, data in the literature argue in favor of the Sorensen test for evaluating the isometric endurance of the trunk extensor muscles. The Ito test and the dynamic variants of the Sorensen test need to be investigated further.
References


