

# Test-retest repeatability of strength capacity, aerobic power and pericranial tenderness of neck and shoulder muscles in children - relevant for tension-type headache

Birte Tornøe<sup>1,2,5,6</sup>

Lars L Andersen<sup>3</sup>

Jørgen H Skotte<sup>3</sup>

Rigmor Jensen<sup>4</sup>

Gunvor Gard<sup>1</sup>

Liselotte Skov<sup>2</sup>

Inger Hallström<sup>1</sup>

<sup>1</sup>Department of Health Sciences, Lund University, Scania, Sweden; <sup>2</sup>Children's Headache Clinic, Department of Pediatrics, University of Copenhagen, Herlev Hospital, Herlev, Denmark; <sup>3</sup>National Research Centre for the Working Environment, Copenhagen, Denmark; <sup>4</sup>Danish Headache Center, Department of Neurology, University of Copenhagen, Glostrup Hospital, Glostrup, Denmark; <sup>5</sup>Department of Physiotherapy and Occupational Therapy, University of Copenhagen, Glostrup Hospital, Glostrup, Denmark; <sup>6</sup>Department of Physiotherapy, Medical Department, University of Copenhagen, Herlev Hospital, Herlev, Denmark

**Background:** Frequent or chronic tension-type headache in children is a prevalent and debilitating condition for the child, often leading to medication overuse. To explore the relationship between physical factors and tension-type headache in children, the quality of repeated measures was examined. The aim of the present study was to determine the test-retest repeatability of parameters determining isometric neck and shoulder strength and stability, aerobic power, and pericranial tenderness in children.

**Methods:** Twenty-five healthy children, 9 to 18 years of age, participated in test-retest procedures within a 1-week interval. A computerized padded force transducer was used for testing. The tests included the isometric maximal voluntary contraction and force steadiness of neck flexion and extension, and the isometric maximal voluntary contraction and rate of force of the dominant shoulder. Pericranial tenderness was recorded by means of standardized manual palpation, and a submaximal cycle ergometer test predicted maximal oxygen uptake ( $VO_2$  max). The measurements were evaluated in steps, using the intraclass correlation coefficient (ICC); changes in the mean between the two test occasions; the levels of agreement, visualized in Bland-Altman Plots; and by quantifying the variability.

**Results:** The results showed an acceptable test-retest repeatability of isometric maximal voluntary contraction (ICC 0.90–0.97). The force steadiness measurements revealed a trend of systematic changes in the direction of neck flexion and need further examination in both healthy and ill children. The rate of force development, Total Tenderness Score, and prediction of  $VO_2$  max showed repeatability, with ICC 0.80–0.87.

**Conclusion:** The measurements of strength capacity, aerobic power, and tenderness provide acceptable repeatability, suitable for research in children.

**Keywords:** muscle strength, intratester reliability, cardiovascular fitness

## Introduction

Frequent or chronic tension-type headache (TTH) in children is a prevalent and debilitating condition for the child, often leading to medication overuse.<sup>1,2</sup> In its infancy, research in children with TTH is limited, though findings from adults highlight the areas that need to be examined. The pathophysiological factors underlying TTH are of high interest, as new knowledge will aid in updating prevention and treatment activities. Several parameters have attracted attention. For example, research in adolescents links impaired function of the neck and shoulder muscles with TTH.<sup>3,4</sup> The rate of force development (RFD), as a component of

Correspondence: Birte Tornøe  
Department of Pediatrics E, Herlev  
Hospital, Arkaden, Herlev Ringvej 75,  
2730 Herlev, Denmark  
Tel +45 3868 1067  
Email birte.tornoe@regionh.dk

strength activation in muscles, was likewise found impaired in painful muscles, in adult chronic pain research<sup>5</sup> and could be hypothesized to be affected in children with headache. Together with muscular deficits, increased pericranial tenderness is a symptom of interest pointing to hypersensitivity of the peripheral or central nervous system linked with chronic headache.<sup>6,7</sup> Interesting research results from the longitudinal Nord-Trøndelag Health Study (HUNT) I and II also show that low physical activity is a risk factor for developing non-migraine headache.<sup>8</sup> The first step involved in exploring related parameters and their association with chronic headache in children is to examine the repeatability of current test methods. Repeatability represents fundamental knowledge essential to interpreting clinical trial results and refers to examining changes and the variability of the mean between two identical repeated test occasions.

The repeatability of isometric strength testing has been found acceptable in research with adolescents 17 years of age with and without TTH,<sup>9</sup> and the repeatability showed high values with intraclass correlation coefficients (ICC) (95% confidence interval [CI]) for neck flexion/extension in healthy subjects was 0.99/0.98 (0.97–0.99) and in subjects with TTH was 0.99/0.99 (0.98–0.99). Likewise was the repeatability of isometric strength measurements in adults with neck pain compared with healthy controls,<sup>10–12</sup> and in studies of healthy adults found acceptable.<sup>13–15</sup> RFD was also examined and found reliable in adults.<sup>5,16</sup> Earlier studies by Bendtsen et al validated use of the Total Tenderness Scoring system (TTS) to examine pericranial muscle tenderness.<sup>17,18</sup> The first to examine the repeatability of the TTS in children was the study by Soe et al, and it found the correlation coefficient between the two test days was acceptable ( $r = 0.81$ ).<sup>19</sup> Confirmation of the repeatability of these measurements in both healthy and children ill with TTH is warranted.

The aim of the present study was to determine the test-retest repeatability of isometric maximal voluntary contraction (MVC) and force steadiness (FS) of neck flexion and extension from a neutral position, and of isometric MVC and RFD of dominant unilateral shoulder abduction, in healthy children 9 to 18 years of age. Another objective of this study was to determine the test-retest repeatability of the TTS and of a submaximal cycle ergometer test for predicting maximal oxygen uptake ( $\text{VO}_2 \text{ max}$ ). To the best of our knowledge, the present study is the first to examine these combined measurements in children 9 to 18 years of age.

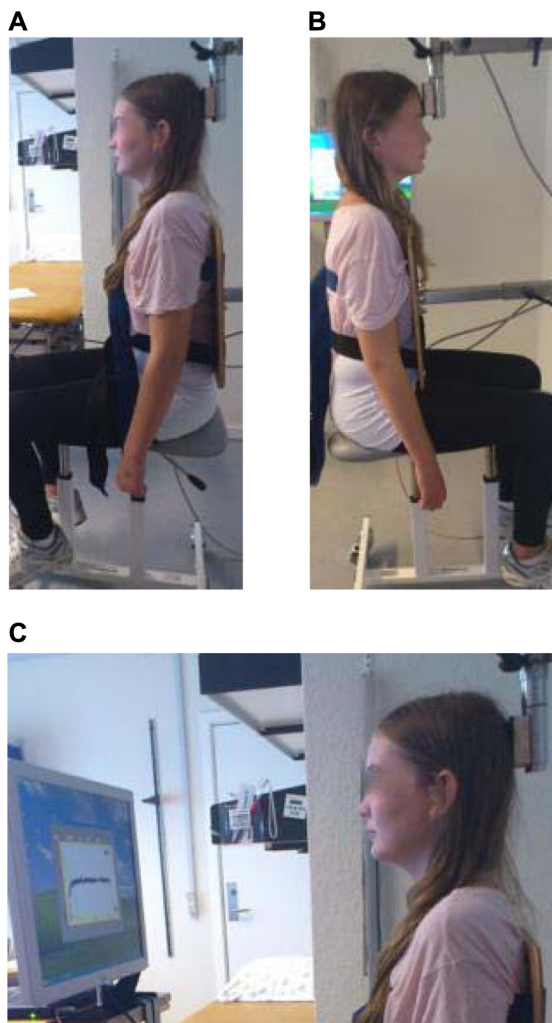
## Methods

### Participants and design

This study is a part of a larger case-control study examining the associations between headache and physical capacity in children with primary headache compared with healthy controls. Forty-one healthy children accepted participation in the case-control study and out of these, a convenience sample of 28 children consented to participate in two rounds of testing. Two girls failed to follow the schedule and one boy did not show up for the second test date. Thus, only 25 of the children (ten boys and 15 girls) participated in the test-retest design within a 1-week interval. None of the children suffered from frequent episodic or chronic TTH, migraine, or any other diseases, and all of them attended mainstream schools. The children were recruited after May 1, 2010 from primary or secondary schools in Copenhagen and its metropolitan area, based on the guidelines and with approval of the Danish National Committee on Biomedical Research Ethics. Sealed envelopes containing detailed information were distributed to the children by their teachers at school, notifying them about the conditions involved in participating in a medical study, the right to confidentiality, and the right to freely withdraw from the study at any time. In order for the children to participate, their parents had to actively contact the researcher and provide written informed consent and the children had to provide oral informed consent. Children over the age of 15 also received special age-appropriate written information. None of the test procedures involved putting the children at risk of harm. A relative or parent of the child was always present at testing. The study followed the ethical principles of the World Medical Association 2008 Declaration of Helsinki<sup>20</sup> and the United Nations 1989 Convention on the Rights of the Child.<sup>21</sup> The study was also approved by the Danish Data Protection Agency and registered with ClinicalTrials.gov (identifier: H-3-2009-081).

### Instrumentation

The strength measurement device was a Vishay Nobel, type KIS-2, max. 2 kN, computerized force transducer (Vishay Precision Group, Malvern, PA, USA) wall mounted on a custom-built adjustable stand (Figure 1). The system was calibrated with standard 20 kg weights. The signals from the force transducer were generally sampled with a frequency of 100 Hz and low pass filtered with a cutoff frequency of 10 Hz. For the FS measurement, the signal was low pass filtered with a cutoff frequency of 5 Hz and the standard deviation (SD) was calculated. The 30% MVC level was calculated



**Figure 1** Test positions and instrumentation. Photos by permission from child and family.

**Notes:** (A) Left side view of neck extension test. (B) Right side view of neck flexion test. (C) Left side view of neck extension force steadiness test.

for a period of 25 seconds. The RFD was calculated as the maximum slope ( $N \times 10^{-1} \times s^{-1}$ ) for a 0.1-second linear fit to the force signal. The bicycle, recalibrated on each test day, was a calibrated Monark Ergonomic 939E PC (Monark Exercise AB, Stockholm, Sweden).

## Overall test procedures

The measurements were conducted at the Danish Headache Center, Glostrup Hospital, Copenhagen, Denmark. An experienced research lab technician did all the testing and was blinded to the history of the child. All test procedures followed a standardized protocol. The participants were given detailed information to familiarize them with the procedure. The tests were conducted in a warm, comfortable room. A coassistant, also blinded, assisted the child during test procedures. The test-retest

was carried out on the same weekday and time within a 1-week interval. Age and anthropometric measures were noted.

## General joint mobility

General joint mobility was screened using the revised Beighton score,<sup>22</sup> which is a bilateral examination in a standing position, of five joint movements that includes the fifth metacarpophalangeal joint, the thumb, elbow, knee, and hands flat on the floor with stretched knees. The maximum score is 9 points. The cutoff level for hypermobility was  $\geq 5/9$ . A goniometer (Smith & Nephew Rolyan Inc, London, UK, A441-1) was used to measure the passive bilateral dorsiflexion of the fifth metacarpophalangeal joint and the passive bilateral hyperextension of the elbow and knee.

## Palpation of tenderness

Tenderness by palpation was scored in a sitting position, in a chair with an adjustable neck and leg support, using the validated TTS,<sup>18</sup> with a four-point scale scored by the lab technician as follows: 0 = denial of tenderness, no visible reaction; 1 = verbal report of discomfort or mild pain, no visible reaction; 2 = verbal report of moderate pain, with or without visible reaction; and 3 = verbal report of marked pain and visible expression of discomfort. The palpation was conducted with small rotational movements, and pressure was maintained for 4 to 5 seconds. The palpation pressure was initially controlled by a palpometer<sup>17</sup> on an arbitrary scale 80–200 arbitrary units (AU), reaching a standardized pressure of 120 AU.<sup>18</sup> The palpometer was the original instrument prepared at the Danish Headache Center. A total score was calculated from palpation of seven bilateral sites (masseter, frontalis, temporalis, processus mastoideus, occipital insertion, trapezius, and sternocleidomastoideus). The maximum possible score was 42 points. The addition of scores is recognized worldwide as a ratio scale and an important tenderness measure used in headache and muscle-related pain research.<sup>23,24</sup>

## Muscle tests

The child was positioned on a chair with the upper and lower trunk fixed with belts, in order to avoid associative work of the trunk muscles,<sup>14</sup> while the arms hung relaxed at the child's sides. Legs were positioned on the chair to prevent push off from the floor. The child's position was corrected until an upright and symmetrical posture was obtained, with the neck in a neutral position in relation to the sagittal, frontal, and rotational axis. For neck extension measures, the padded force

transducer was positioned with the lower edge corresponding with the protuberantia occipitalis. For neck flexion measures, the lower edge of the pad corresponded with the line between the child's eyebrows (Figure 1A and B).

For isometric MVC of neck extension and neck flexion, the child was instructed to build up pressure for a few seconds to reach a maximum amount of tension, to hold the tension briefly, and then to release it. If the force of the third trial exceeded 5%, the computer automatically added a fourth trial.

For FS in neck extension and flexion, the child was instructed, by means of a computer screen, to build up tension (10 seconds) according to a line (30% of MVC) and then to hold that tension at the line for 25 seconds (Figure 1C).

For shoulder measurements, the child was asked to lie on a mattress on his or her back in a supine position. The dominant shoulder was then positioned in external rotation and abduction, while the elbow was stretched, to position the back of the wrist at the pad. The child was instructed to press the wrist of the dominant arm against the pad as quickly (RFD) and as forcefully (MVC) as possible.

## Prediction of $\text{VO}_2$ max

For cycle ergometer testing, the seat height was adjusted to allow the child's knee to be slightly flexed when the plantar region of the foot was in its lowest position. The armrests were also adjusted. Heart rate was monitored using an adjustable Polar pulse belt (ProTerapi A/S, Brøndby, Denmark) and recorded continuously on a computer. The child cycled 6 to 7 minutes, with a cadence of 60 repetitions per minute and at a resistance starting at 50 watts. If the child was over the age of 14 or appeared to be in the later stages of puberty, the start resistance was estimated based on the status of the child's general level of fitness and raised accordingly to 75 watts. Power was added until the child's pulse exceeded 120 beats/min. In order to obtain a submaximal steady-state pulse, time was added (maximum, 7 minutes) according to original test procedures.<sup>25</sup>  $\text{VO}_2$  max was predicted from a steady-state submaximal pulse, using the Åstrand nomogram.<sup>26</sup>

## Statistical analyses

Descriptive statistics of age and anthropometric measures of the study group and test values were presented as means and SDs. Three trials of strength measurements were chosen for data analysis because the results from the third and fourth trial showed a trend of learning and building up more power. For the MVC and RFD tests, the largest force of three trials was chosen as the peak value. For FS, the

minimum of the measured SD was calculated as the peak value. Additionally, the mean of three trials was calculated. The test-retest repeatability was examined using Lexell and Downham's approach<sup>27</sup> and divided into four steps. In step one, the ICC was calculated using a one-way random analysis of variance (ANOVA) for simple replication. ICC values  $>0.75$  represent "excellent reliability," and values between 0.4 and 0.75 represent "fair to good reliability."<sup>27</sup> Step two involved calculating changes in the mean for peak data by studying the indices, in order to detect systematic changes. The mean differences (test 2 – test 1) were calculated together with the SD of the differences, the standard error of the mean, and the 95% CI. If the 95% CI did not include zero, this indicated a significant systematic change in the mean between the two test occasions, thus suggesting either a learning effect or a fatigue effect. The distributional assumption of the differences was examined by plots and the Kolmogorov–Smirnov test. In step three, a visual plot and calculation were used according to Bland and Altman<sup>28,29</sup> in order to visualize the level of agreement between the measurements. According to Bland and Altman, the plot can be used for measurements taken with different devices or for repeated measures. The mean difference (test 2 – test 1) was organized as a Y-axis reference line, with two additional lines at  $\pm 2$  times the SD as the limits of agreement. The "n" differences of the measurements were plotted against their corresponding means. A random scatter of points above and below the reference line was considered acceptable as a measure of repeatability. If scatter points were clustered or repeated identically, systematic errors, lack of agreement or distribution transformation should be considered. Data that was not significantly normally distributed was transformed using a lognormal transformation for further examination. The last step,<sup>27</sup> step four, involved quantifying the variability of the measurements by (A) the method error (ME) defined by  $\text{ME} = \text{SD}_{\text{diff}} / \sqrt{2}$ , where the  $\text{SD}_{\text{diff}}$  refers to the standard deviation of the mean difference between test 2 and test 1. (B) the coefficient of variation (CV%) defined by  $\text{CV}\% = (\text{ME}/\text{mean}) \times 100$ , where the mean is the mean of all data from test 2 + test 1.<sup>27</sup> The ME and CV% make it possible to determine the typical variation between two test occasions and means that an improvement after an intervention smaller than the variation does not indicate a clinically important improvement.

Differences between the two test occasions in tenderness for individual muscles were analyzed using a Wilcoxon rank test. A significance level was set at



**Table 1** Age and anthropometric measures of study group, N = 25

Variables	Mean (SD)
Age (years)	13.7 (1.8)
Height (cm)	164.4 (7.6)
Weight (kg)	51.3 (9.0)
BMI (w/h <sup>2</sup> )	18.8 (2.0)

**Abbreviations:** BMI, body mass index; SD, standard deviation.

$P \leq 0.05$ , while SPSS 17.0 (SPSS Inc, Chicago, IL, USA) and 19.0 (IBM, Armonk, NY, USA) were used for statistical analysis.

## Results

Twenty-five children with a mean age 13.7 years (SD 1.8) completed the study and cooperated well during the test procedures. Table 1 presents the ages and anthropometric measures of the group studied. None of the children were obese, and two children reached the hypermobility cutoff of  $\geq 5/9$ . Table 2 provides descriptive statistics of the test measures.

The ANOVA analysis revealed a test-retest repeatability of the peak maximal voluntary strength measurements for neck extension with an ICC of 0.91 (95% CI: 0.81–0.96); for neck flexion, with ICC 0.91 (95% CI: 0.82–0.96); and for shoulder abduction, with ICC 0.97 (95% CI: 0.92–0.98). The RFD showed an ICC of 0.79 (95% CI: 0.58–0.90), while the TTS likewise, showed repeatability, with an ICC of 0.87 (95% CI: 0.74–0.94). The prediction of  $VO_2$  max showed repeatability with the ICC 0.80 (95% CI: 0.60–0.91). The FS measurements, however, showed only fair-to-good repeatability, with the lowest ICC in the direction of neck flexion – the mean ICC was 0.39 (95% CI: 0.01–0.67), while the peak ICC was 0.67 (95% CI: 0.39–0.84). Table 3 presents the ICCs.

The indices of change in the mean between the two test occasions (Table 4) show a systematic change in the FS measurements, especially in the direction of flexion.

The Bland–Altman plots reveal a similar picture in that the plots of MVC and RFD show acceptable agreement. The plots of TTS and the prediction of  $VO_2$  max show a random scatter of points above and below the mean difference line, thus showing good agreement. The FS measurements, however, show, in keeping with data from Tables 3 and 4, an unequal spread, especially regarding the direction of flexion, thus illustrating a trend of systematic error. Figure 2 provides TTS,  $VO_2$  max  $l \cdot \text{min}^{-1}$ , and peak plots. No significant changes were found in TTS for individual muscles between the two test occasions.

Quantifying the size of the variability by ME and CV%, as shown in Table 5, reveals a larger variability in the measurements of RFD and TTS.

## Discussion

In this study, 25 healthy children 9 to 18 years of age were tested twice, by a single trained and experienced, blinded research lab technician, for isometric strength measurements in the neck and shoulder, pericranial tenderness, and prediction of  $VO_2$  max using a cycle ergometer test and the Åstrand nomogram. The maximal voluntary shoulder test revealed an ICC exceeding 0.90. This agreed with earlier findings in adults by Andersen et al.<sup>30</sup> The mean and peak MVC isometric strength of the neck also showed repeatability without systematic changes and an acceptable agreement between measurements. A Finnish study using a comparable device system has previously shown repeatable mean and peak force measurements of adolescents with and without headache,<sup>9</sup> in both extension and flexion directions. In our study, the peak forces seemed slightly more stable, thus indicating that this should be an area of focus in future

**Table 2** Descriptive statistics of mean and peak values of three trials, N = 25

Variables	Test 1	Test 2	Test 1	Test 2
	Mean variables Mean (SD)	Mean variables Mean (SD)	Peak variables Mean (SD)	Peak variables Mean (SD)
Shoulder MVC ( $N \times 10^{-1}$ )	5.32 (2.58)	5.39 (2.80)	5.72 (2.63)	5.88 (2.96)
Shoulder RFD ( $N \times 10^{-1} \times s^{-1}$ )	19.93 (13.04)	20.89 (15.49)	24.24 (16.28)	24.68 (17.17)
Neck extension MVC ( $N \times 10^{-1}$ )	9.62 (4.87)	9.65 (5.45)	10.38 (4.90)	10.30 (5.86)
Neck flexion MVC ( $N \times 10^{-1}$ )	5.47 (2.64)	5.34 (3.02)	6.00 (2.99)	5.84 (3.29)
Neck extension FS ( $N \times 10^{-1}$ )	0.08 (0.04)	0.08 (0.03)	0.07 (0.03)	0.07 (0.03)
Neck flexion FS ( $N \times 10^{-1}$ )	0.08 (0.04)	0.07 (0.02)	0.07 (0.03)	0.06 (0.02)
TTS (0–42)	11.48 (8.25)	10.00 (7.64)	–	–
$VO_2$ max ( $l \cdot \text{min}^{-1}$ )	2.49 (0.71)	2.50 (0.67)	–	–

**Abbreviations:** FS, force steadiness; MVC, maximal voluntary contraction; RFD, rate of force development; SD, standard deviation; TTS, total tenderness score;  $VO_2$  max, predicted maximal oxygen uptake.

**Table 3** Intraclass correlation coefficients with 95% confidence interval of mean and peak data, N = 25

Variables/ICC	Mean ICC (95% CI)	Peak ICC (95% CI)
Shoulder MVC ( $N \times 10^{-1}$ )	0.97 (0.94–0.99)	0.97 (0.92–0.98)
Shoulder RFD ( $N \times 10^{-1} \times s^{-1}$ )	0.80 (0.61–0.91)	0.79 (0.58–0.90)
Neck extension MVC ( $N \times 10^{-1}$ )	0.93 (0.85–0.97)	0.91 (0.81–0.96)
Neck flexion MVC ( $N \times 10^{-1}$ )	0.90 (0.79–0.95)	0.91 (0.82–0.96)
Neck extension FS ( $N \times 10^{-1}$ )	0.63 (0.32–0.82)	0.76 (0.53–0.89)
Neck flexion FS ( $N \times 10^{-1}$ )	0.39 (0.01–0.67)	0.67 (0.39–0.84)
TTS (0–42)	0.87 (0.74–0.94)	–
VO <sub>2</sub> max ( $l \cdot \text{min}^{-1}$ )	0.80 (0.60–0.91)	–

**Abbreviations:** CI, confidence interval; FS, force steadiness; ICC, intraclass correlation coefficient; MVC, maximal voluntary contraction; RFD, rate of force development; TTS, total tenderness score; VO<sub>2</sub> max, predicted maximal oxygen uptake.

research. The test of maximal voluntary contraction is an unusual situation for the participant and means a learning effect might have existed during the trial. As a result, considering the highest value as the maximal performance supported is relevant in the present study. Our study comprised a 1-week interval between two measurements, and in contrast to comparable studies where the time between measurements was considerably shorter, issues involving fatigue or increased results after repetitive learning are less obvious.<sup>9,10</sup> Our results reveal that the FS measurements, especially regarding neck flexion, showed systematic changes between the two test occasions. A study by Ylinen et al<sup>10</sup> showed that in adults, the extension measures are more reliable than flexion measures. The researchers concluded that this phenomenon might be due to the fact that the extension muscle force is stronger than that of flexion and thus more stable. Another factor that might influence the FS measurements is the level of resistance chosen.

The present measures were conducted at 30% of MVC in a neutral upright position. O’Leary et al,<sup>31</sup> who examined the steadiness of cervical flexion at 20% and 50% of MVC in 20 adults in a supine position, found that at 20% the measures had a perfect reliability by repeat, while at 50% in a small group of eight healthy controls, the measures demonstrated very poor reliability. As children are reportedly less stable than adults,<sup>32</sup> a trend of systematic changes likely due to fatigue is, to some degree, understandable with a FS measurement at 30% of MVC. As a result, we recommend that further study in children in this area use resistance at 20% and 30% of MVC.

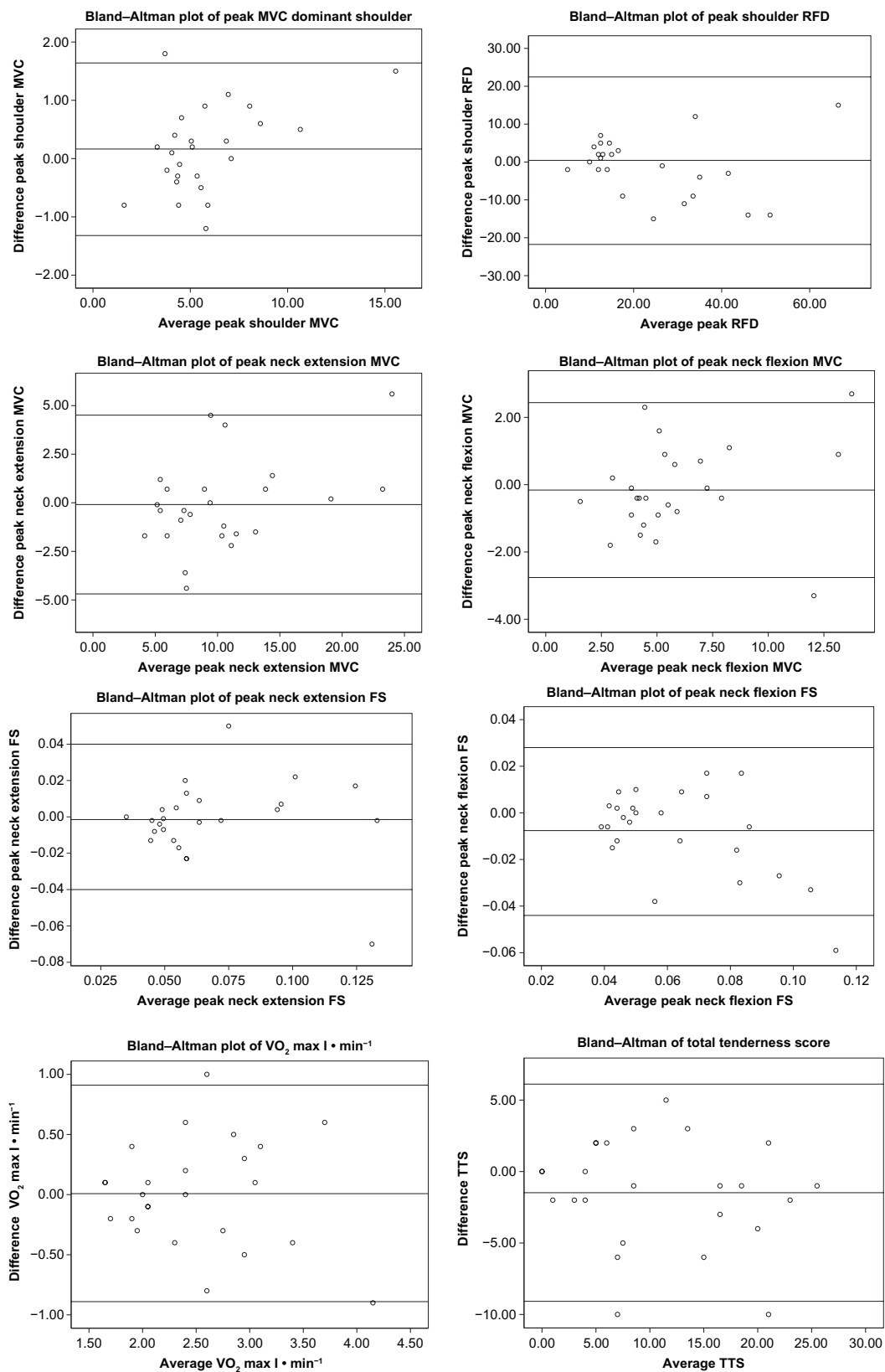
The tenderness and the aerobic power tests show acceptable repeatability and agreement between measurements. In accordance with the TTS and following procedures validated by Bendtsen et al,<sup>18</sup> we used the same calibrated palpometer, the same experienced lab technician, and the same lab for both sets of tests. In the early original study, the intraobserver reliability was particularly consistent, and the pressure of palpation was stable if the pressure was initially controlled by the palpometer. The degree of tenderness is usually reported as an aggregate of the scores. In keeping with the recommendations of Bendtsen et al,<sup>17,18</sup> our studies are based on two assumptions: first, the observer is assumed to have the ability to evaluate pain during palpation using both verbal and behavioral scoring; second, the assumption is made that an ordinal scale score, summed to a total score (TTS), is considered as a ratio scale. Bendtsen et al found a positive linear relationship between TTS and pain intensity scores during palpation of the trapezius muscle and concluded that the four-point tenderness scale, and thus the aggregate sum, can be considered a ratio scale.<sup>17,18</sup> As a result, it can be

**Table 4** Indices of changes in the mean between two test occasions (test 2 – test 1) for peak data, the total tenderness score, and VO<sub>2</sub> max  $l \cdot \text{min}^{-1}$ , N = 25

Variables/indices	Mean diff test 2 – test 1	SD*	SE*	95% CI*
Shoulder MVC ( $N \times 10^{-1}$ )	0.16	0.74	0.15	–0.14–0.47
Shoulder RFD ( $N \times 10^{-1} \times s^{-1}$ )	0.44	11.11	2.22	–4.14–5.02
Neck extension MVC ( $N \times 10^{-1}$ )	–0.09	2.29	0.46	–1.04–0.86
Neck flexion MVC ( $N \times 10^{-1}$ )	–0.16	1.33	0.27	–0.71–0.39
Neck extension FS ( $N \times 10^{-1}$ )	–0.00	0.02	0.00	–0.01–0.01
Neck flexion FS ( $N \times 10^{-1}$ )	–0.01	0.02	0.00	–0.02–0.00
TTS (0–42)	–1.48	3.80	0.76	–3.05–0.09
VO <sub>2</sub> max ( $l \cdot \text{min}^{-1}$ )	0.01	0.45	0.09	–0.18–0.19

**Note:** \*refers to the mean difference between test 2 and test 1.

**Abbreviations:** CI, confidence interval; FS, force steadiness; MVC, maximal voluntary contraction; RFD, rate of force development; TTS, total tenderness score; VO<sub>2</sub> max, predicted maximal oxygen uptake; SE, standard error of the mean; SD, standard deviation; diff, difference.



**Figure 2** Bland-Altman plots of peak variables,  $VO_2$  max  $l \cdot \text{min}^{-1}$  and total tenderness score.

**Notes:** The middle reference line shows the mean difference (test 2 – test 1) and the two additional lines  $\pm 2$  SD.

**Abbreviations:** FS, force-steadiness; MVC, maximal voluntary contraction; RFD, rate of force development; SD, standard deviation; TTS, total tenderness score;  $VO_2$  max, predicted maximal oxygen uptake.

**Table 5** Indices of measurement variability between two test occasions (test 2 – test 1) for peak data, total tenderness score and  $\text{VO}_2 \text{ max l} \cdot \text{min}^{-1}$ ,  $N = 25$

Variables/indices	Method error (ME)	Coefficient of variation (CV%)
Shoulder MVC ( $N \times 10^{-1}$ )	0.52	8.97
Shoulder RFD ( $N \times 10^{-1} \times \text{s}^{-1}$ )	7.86	32.13
Neck extension MVC ( $N \times 10^{-1}$ )	1.62	15.67
Neck flexion MVC ( $N \times 10^{-1}$ )	0.94	15.88
Neck extension FS ( $N \times 10^{-1}$ )	0.01	14.29
Neck flexion FS ( $N \times 10^{-1}$ )	0.01	16.67
TTS (0–42)	2.69	25.05
$\text{VO}_2 \text{ max (l} \cdot \text{min}^{-1})$	0.32	12.80

**Abbreviations:** FS, force steadiness; MVC, maximal voluntary contraction; RFD, rate of force development; TTS, total tenderness score;  $\text{VO}_2 \text{ max}$ , predicted maximal oxygen uptake.

concluded that repeated recording of TTS is of good quality for examining pericranial tenderness in children. However, upon close examination of the quantified variability, the TTS revealed a large ME and corresponding CV%, which is an issue that must be taken into consideration when interpreting clinical improvements in trials. This was also the case for RFD measurements.

## Limitations and comments

One main limitation of this study is that it did not involve children suffering from frequent or chronic TTH. Future testing on children with headache, based on the method presented here, would increase our knowledge even further. As a result, families with ill children must be encouraged to seek treatment at the headache clinic, and the burden of test-retest procedures on families must be taken into consideration. Despite current drawbacks, the present study provides fundamental support for future studies examining underlying pain mechanisms in children with TTH.

## Conclusion and implications for practice

In conclusion, the measurements of strength capacity, aerobic power and pericranial tenderness show an acceptable repeatability, suitable for research in children. Our study demonstrates that two of the examined test procedures, the TTS and the submaximal aerobic power test, both of which only require a limited amount of equipment, can have a positive impact on clinical practice. Moreover, the procedures are easily learned, and the test situation is comfortable for the participant. Results from future research in children will hopefully provide additional knowledge about children's muscle capacities and their

possible relation to headache and other pain disorders. This knowledge, in turn, can have profound implications for the choice of treatment and the interpretation of intervention results.

## Acknowledgments

This study was financed with support from the Master Carpenter Jørgen Holm and Wife Elisa F Hansen Memorial Trust, the Beckett-Foundation, the Lundbeck Foundation, the Axel Muusfeldt Foundation, and the Dagmar Marshalls Foundation. We would like to thank the children and parents who took the time to participate in this study and also to extend our appreciation to the school boards for allowing the schools to participate. We would like to thank the study research lab technician, Hanne Andresen, and the Danish Headache Center, Glostrup Hospital, for conducting all the testing. We are also grateful to the nursing and physiotherapy students, Christian Tilsted, Pernille Hansen, May Bang, and Marianne Villadsen, who assisted in the testing procedures. We thank statistician Claus Jensen, MSc, the Danish Ministry of Education and Children, and statistician Jørgen Vinsløv Hansen of the National Research Centre, for the working environment for supervising the statistical analyses. Finally, we would also like to thank Beth Kanijo, MD, and Nancy Aaen, MA, for participating in the editing process.

## Disclosure

The authors report no conflicts of interest in this work.

## References

- Ozge A, Sasmaz T, Cakmak SE, Kaleagasi H, Siva A. Epidemiological-based childhood headache natural history study: after an interval of six years. *Cephalalgia*. 2010;30(6):703–712.
- Grazzi L, D'Amico D, Usai S, Solari A, Bussone G. Disability in young patients suffering from primary headaches. *Neurol Sci*. 2004; 25 Suppl 3:S111–S112.
- Oksanen A, Metsähonkala L, Viander S, et al. Strength and mobility of the neck-shoulder region in adolescent headache. *Physiother Theory and Pract*. 2006;22(4):163–174.
- Oksanen A, Pöyhönen T, Ylinen JJ, et al. Force production and EMG activity of neck muscles in adolescent headache. *Disabil Rehabil*. 2008;30(3):231–239.
- Andersen LL, Holtermann A, Jørgensen MB, Sjøgaard G. Rapid muscle activation and force capacity in conditions of chronic musculoskeletal pain. *Clin Biomech (Bristol, Avon)*. 2008;23(10): 1237–1242.
- Anttila P, Metsähonkala L, Aromaa M, et al. Determinants of tension-type headache in children. *Cephalalgia*. 2002;22(5):401–408.
- Ashina S, Jensen R, Bendtsen L. Pain sensitivity in pericranial and extracranial regions. *Cephalalgia*. 2003;23(6):456–462.
- Varkey E, Hagen K, Zwart JA, Linde M. Physical activity and headache: results from the Nord-Trøndelag Health Study (HUNT). *Cephalalgia*. 2008;28(12):1292–1297.
- Oksanen A, Ylinen JJ, Pöyhönen T, et al. Repeatability of electromyography and force measurements of the neck muscles in adolescents with and without headache. *J Electromyogr Kinesiol*. 2007;17(4):493–503.



10. Ylinen J, Salo P, Nykänen M, Kautiainen H, Häkkinen A. Decreased isometric neck strength in women with chronic neck pain and the repeatability of neck strength measurements. *Arch Phys Med Rehabil.* 2004;85(8):1303–1308.
11. Ylinen JJ, Rezasoltani A, Julin MV, Virtapohja HA, Mälkiä EA. Reproducibility of isometric strength: measurement of neck muscles. *Clin Biomech (Bristol, Avon).* 1999;14(3):217–219.
12. Chiu TT, Sing KL. Evaluation of cervical range of motion and isometric neck muscle strength: reliability and validity. *Clin Rehabil.* 2002;16(8):851–858.
13. Peolsson A, Oberg B, Hedlund R. Intra- and inter-tester reliability and reference values for isometric neck strength. *Physiother Res Int.* 2001;6(1):15–26.
14. Rezasoltani A, Ylinen J, Bakhtiary AH, Norozi M, Montazeri M. Cervical muscle strength measurement is dependent on the location of thoracic support. *Br J Sports Med.* 2008;42(5):379–382.
15. Garcés GL, Medina D, Milutinovic L, Garavote P, Guerado E. Normative database of isometric cervical strength in a healthy population. *Med Sci Sports Exerc.* 2002;34(3):464–470.
16. Holtermann A, Roeleveld K, Vereijken B, Ettema G. The effect of rate of force development on maximal force production: acute and training-related aspects. *Eur J Appl Physiol.* 2007;99(6):605–613.
17. Bendtsen L, Jensen R, Jensen NK, Olesen J. Muscle palpation with controlled finger pressure: new equipment for the study of tender myofascial tissues. *Pain.* 1994;59(2):235–239.
18. Bendtsen L, Jensen R, Jensen NK, Olesen J. Pressure-controlled palpation: a new technique which increases the reliability of manual palpation. *Cephalalgia.* 1995;15(3):205–210.
19. Soee AB, Thomsen LL, Tornøe B, Skov L. Reliability of four experimental mechanical pain tests in children. *J Pain Res.* 2013;6:103–110.
20. World Medical Association. *Declaration of Helsinki; Principles for Medical Research Involving Human Subjects.* Ferney-Voltaire: World Medical Association; 2008. Available from: <http://www.wma.net/en/30publications/10policies/b3/index.html>. Accessed June 8, 2013.
21. United Nations. *Convention on the Rights of the Child.* Geneva: United Nations; 1989. Available from: <http://www.ohchr.org/EN/Professional-Interest/Pages/CRC.aspx>. Accessed June 8, 2013.
22. Smits-Engelsman B, Klerks M, Kirby A. Beighton score: a valid measure for generalized hypermobility in children. *J Pediatr.* 2011;158(1):119–123.
23. Fernández-de-Las-Peñas C, Cuadrado ML, Arendt-Nielsen L, Pareja JA. Side-to-side differences in pressure pain thresholds and pericranial muscle tenderness in strictly unilateral migraine. *Eur J Neurol.* 2008;15(2):162–168.
24. Sundstrup E, Jakobsen MD, Andersen CH, et al. Participatory ergonomic intervention versus strength training on chronic pain and work disability in slaughterhouse workers: study protocol for a single-blind, randomized controlled trial. *BMC Musculoskelet Disord.* 2013;14:67.
25. Cink RE, Thomas TR. Validity of the astrand-ryhming nomogram for predicting maximal oxygen intake. *Br J Sports Med.* 1981;15(3):182–185.
26. Åstrand PO, Rodahl K, Dahl HA, Strømme SB. *Textbook of Work Physiology; Physiological Bases of Exercise*, 4th ed. Champaign, IL: Human Kinetics; 2003.
27. Lexell JE, Downham DY. How to assess the reliability of measurements in rehabilitation. *Am J Phys Med Rehabil.* 2005;84(9):719–723.
28. Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet.* 1986;1(8476):307–310.
29. Bland JM, Altman DG. Comparing methods of measurement: why plotting difference against standard method is misleading. *Lancet.* 1995;346(8982):1085–1087.
30. Andersen LL, Saervoll CA, Mortensen OS, Poulsen OM, Hannerz H, Zebis MK. Effectiveness of small daily amounts of progressive resistance training for frequent neck/shoulder pain: randomised controlled trial. *Pain.* 2011;152(2):440–446.
31. O’Leary SP, Vicenzino BT, Jull GA. A new method of isometric dynamometry for the craniocervical flexor muscles. *Phys Ther.* 2005;85(6):556–564.
32. Maslen B, Straker L. A comparison of posture and muscle activity means and variation amongst young children, older children and young adults whilst working with computers. *Work.* 2009;32(3):311–320.

## Journal of Pain Research

### Publish your work in this journal

The Journal of Pain Research is an international, peer-reviewed, open access, online journal that welcomes laboratory and clinical findings in the fields of pain research and the prevention and management of pain. Original research, reviews, symposium reports, hypothesis formation and commentaries are all considered for publication.

Submit your manuscript here: <http://www.dovepress.com/journal-of-pain-research-journal>

Dovepress

The manuscript management system is completely online and includes a very quick and fair peer-review system, which is all easy to use. Visit <http://www.dovepress.com/testimonials.php> to read real quotes from published authors.