Original Contribution

REPRODUCIBILITY OF ULTRASONOGRAPHIC MEASUREMENTS OF THE ULNAR NERVE AT THE CUBITAL TUNNEL

ALEXANDRA FINK,∗†, MARLIN TEGGELER,∗†, MARC SCHMITZ,∥, JAAP JANSSEN,∥ and MARTIJN PISTERS*‡∥#

*Physical Therapy Science; †Program in Clinical Health Sciences; ‡Department of Rehabilitation; ∥Nursing Science and Sport; Brain Center Rudolf Magnus; University Medical Center Utrecht, Utrecht, The Netherlands; †Hand Therapy Center Enschede (Handtherapie Nederland), Enschede, The Netherlands; ‡Physical Therapy Practice Emmastraat, Enschede, The Netherlands; ∥Stichting Ondernemingen Medische Technologie (SOMT), Amersfoort, The Netherlands; †Institute of Movement Studies, University of Applied Sciences Utrecht, Utrecht, The Netherlands; ‡Department of Health Innovation and Technology, Fontys University of Applied Sciences, Eindhoven, The Netherlands; and ∥Center for Physical Therapy Research and Innovation in Primary Care, Leidsche Rijn Julius Health Care Centers, Utrecht, The Netherlands

(Received 26 February 2016; revised 18 September 2016; in final form 27 September 2016)

Abstract—The cross-sectional area (CSA) of the ulnar nerve is thought to be indicative of ulnar nerve entrapment. The purpose of the study was to determine reproducibility of ultrasonographic measurements of CSA of the ulnar nerve at the cubital tunnel in healthy adults. Two sonographers tested 69 participants using a standardized protocol. The inter-rater reliability intra-class correlation coefficient was 0.63, and the intra-rater reliability intra-class correlation coefficient was 0.85 for sonographer 1 and 0.88 for sonographer 2. The smallest detectable changes were 2.47 and 2.63 mm² (25% and 26% of the mean CSA). The mean difference and 95% limits of agreement for sonographers 1 and 2 were −0.13 (−2.25 to 2.29) and −0.38 (−2.93 to 2.18). Based on the fair to good inter-rater reliability, the excellent intra-rater reliability and the clinical applicable intra-rater agreement, ultrasonography seems to be a valuable tool with which to assess the CSA of the ulnar nerve for diagnostic and evaluative purposes. (E-mail: a.fink@xpertclinic.nl) © 2016 World Federation for Ultrasound in Medicine & Biology.

Key Words: Ultrasonography, Reproducibility, Elbow, Ulnar nerve.

INTRODUCTION

Ulnar nerve entrapment (UNE) usually occurs at the cubital tunnel underneath Osborne's ligament, manifesting as a neuropathy caused by nerve compression (Elhassan and Steinmann 2007; Latinovic et al. 2006). Common symptoms of UNE are pain, numbness, paresthesia, muscle weakness and atrophy of the muscles innervated by the ulnar nerve (Elhassan and Steinmann 2007; Latinovic et al. 2006; Mondelli et al. 2005; Stewart 1987). The incidence of UNE was 24.7 per 100,000 persons in Western society in 2005 (Mondelli et al. 2005).

Ultrasound (US) is a quickly available, pain-free, non-invasive and low-cost imaging modality (Heinemeyer and Reimers 1999; Pompe and Beekman 2013). Beekman et al. (2004) stated that US is valid for diagnosing UNE. Entrapments at the cubital tunnel can be detected using US measurements of increased cross-sectional areas (CSA) reflecting thickening or swelling (Thoirs et al. 2008; Volpe et al. 2009; Wiesler et al. 2006; Yoon et al. 2008). Furthermore, CSA is more appropriate than thickness measurements because of the varying appearances of the ulnar nerve (Wiesler et al. 2006).

Considerable research has been performed on US of peripheral nerves of the upper extremities (Boehm et al. 2014; Sugimoto et al. 2013; Tagliafico and Martinoli 2013; Thoirs and Williams 2001; Thoirs et al. 2005), with fair to excellent inter-rater reliability. However, no attention has been paid to intra-rater agreement parameters by means of the smallest detectable change (SDC).

Previous studies have calculated intra-rater reliability outcomes based on a small sample (Boehm et al. 2014; Thoirs and Williams 2001). Inter- and intra-rater reliability was determined for different nerves together (Boehm et al. 2014; Tagliafico and Martinoli 2013;...
Thoirs et al. 2005). Furthermore, different statistical analyses were performed, a Pearson’s correlation coefficient or coefficient of variation, instead of the preferred ICC (Sugimoto et al. 2013; Thoirs et al. 2005).

Reliability and agreement are important concerns for categorization and quality assertion of an US measurement. Reliability and agreement outcomes offer evidence of measurement error in CSA measurement, the amount of error determining the applicability of the results (Dunn 2004). In general, US has been described as a highly sonographer-dependent diagnostic tool (Martinoli et al. 2002). Earlier research recommended the need for a more standardized and detailed measurement protocol, which requires training of the sonographers for reliability of US in other structures (O’Connor et al. 2005; Rutten et al. 2006). Therefore, the use of a standardized, reproducible measurement protocol seems very important.

Given these facts, additional research is needed to determine the reproducibility of the ulnar nerve at the cubital tunnel separately. Therefore, the research question of the present study is: What is the reproducibility of ultrasonicographic measurements of the cross-sectional area of the ulnar nerve at the cubital tunnel in healthy adult patients?

METHODS

Design and participants

The approval for the test–retest study was obtained from the Research Ethics Committee of the University Medical Center Utrecht. Additionally, all participants gave written informed consent before the measurements. A convenience sample at Fontys University of Applied Sciences in Eindhoven, The Netherlands, was recruited from February 2013 up to April 2013. Participants had to be 18 y or older to be included. Exclusion criteria were (i) complaints about one or more parts of the dominant arm; (ii) physical therapy or surgery because of complaints over the dominant arm within the last 3 mo; and (iii) subluxation of the ulnar nerve in the test position, tested during the US measurement. In line with Walter et al. (1998), a minimal sample size was determined at n = 39. Because of the COSMIN (consensus-based standards for the selection of health status measurement instrument) (Mokkink et al. 2010) rating criteria, establishing the methodological quality of a study, the minimal sample size was determined to be n > 50.

Demography and anthropometry

Before the US measurement, demographic characteristics, including age, gender, anthropometric measures such as height (cm) and weight (kg) were collected, and the dominant arm of each participant was noted. Additionally, the body mass index (kg/m²) was calculated.

Raters and US protocol

Two physical therapists, with 3 and 7 y of experience in clinical US, carried out the US measurements. Measurements were performed with two Mylab One (Esaote Benelux, Maastricht, Netherlands) scanners with a 6- to 13-MHz linear array probe using the elbow preset for superficial tissue.

Two-dimensional measurements were made, in line with Wiesler et al. (2006) who reported that 2-D measurements are more accurate than 1-D measurements. To make the US measurements reproducible, the position of the participant was specified. Imaging was performed on the subject’s dominant elbow exclusively, with the participant sitting parallel to the examination table. The flattened hand was placed on the table in maximal internal rotation of the upper arm and pronation of the forearm. A position of 70–80° of elbow flexion was chosen in this protocol because of possible physiologic changes of the ulnar nerve above 90° of elbow flexion. The olecranon had to be facing the sonographer during the measurement. The specified posture of the participants and the position of the probe are illustrated in Figure 1. To become familiar with and to improve the measurement protocol, the sonographers practiced the technique for about 1 h on four participants and discussed possible difficulties.

Fig. 1. Specified posture of the participant during ultrasonographic measurement of the ulnar nerve in the cubital tunnel.
As published in previous articles, the medial epicondyle and the olecranon were defined as bony landmarks to determine the exact position of the probe at the cubital tunnel (Cartwright et al. 2007; De Maeseneer et al. 2012; Martinoli 2010; Tagliafico and Martinoli 2013; Wiesler et al. 2006). Similar to real-time measurements in daily practice, the sonographers used the ellipse tool of the US scanner on the fixed picture to measure the CSA (in mm²) of the ulnar nerve at the cubital tunnel. Measurements were performed on the hyper-echoic rim of the nerve. Care was taken to position the transducer perpendicular to the nerve. Elliptical US measurement of the ulnar nerve in the cubital tunnel is illustrated in Figure 2. Data were registered by research assistants to reduce recall bias of the sonographers to each other’s as well as to their own previous measurements, and all images were saved digitally. The test procedure lasted approximately 30 min for each participant. Both sonographers started simultaneously, measuring different participants (measurements 1.1 and 2.1), and participants switched after two consecutive measurements (measurement 1.2 and 2.2). The averages of these consecutive measurements were used for statistical analysis. After an interval of 15 min this process was repeated. An interval of 15 min was used to diminish recall bias. In this interval participants were expected to remain stable and were not subject to change (Beekman et al. 2004). The procedure is illustrated in Figure 3.

**Statistical analysis**

Statistical analysis was performed using the Statistical Package for the Social Sciences (Version 19.0, IBM, Armonk, NY, USA). At first, data were graphically checked for normal distribution using quantile–quantile plots. Descriptive statistics were presented as means ± standard deviation (SD) and range for continuous data or numbers (n) and percentages (%). The threshold of statistical significance was established for all measurements at \( p < 0.05 \).

**RESULTS**

In total, 73 patients were recruited. Four patients were excluded because of subluxation of the ulnar nerve. This resulted in the inclusion of 69 patients (28 men and 41 women) with a mean (±SD) age of 36 (±15) y. No missing values were registered, and data were normally distributed. Additional demographic and anthropometric data are summarized in Table 1. Means ± SD and 95% confidence intervals of the CSA of the ulnar nerve of both measurements are reported in Table 2. No adverse events were stated. Inter-rater reliability was fair to good with an ICC value of 0.63 (\( p < 0.001 \)). Intra-rater reliability was excellent, with an ICC of 0.85 for sonographer 1 and 0.88 for sonographer 2. P values < 0.001. With respect to intra-rater agreement, the SDCs of sonographers 1 and 2 were 2.47 mm² (25% of the mean CSA) and 2.63 mm² (26% of the mean CSA), respectively. The mean difference and 95% LoA for sonographers 1 and 2 were −0.13 (−2.56, 2.29) and −0.38 (−2.93, 2.18), respectively. All reproducibility outcomes are provided in Table 3.

![Fig. 2. Elliptical ultrasonographic measurement of the ulnar nerve in the cubital tunnel. CSA = cross-sectional area.](image_url)
DISCUSSION

The purpose of the present study was to determine the inter- and intra-rater reliability, as well as the intra-rater agreement, of US measurements of the CSA of the ulnar nerve at the cubital tunnel. Excellent intra-rater reliability outcomes are obtained by both raters in this study. The ICC values of both raters are quite similar, suggesting that the different levels of experience between the raters had little influence on intra-rater reliability in this study. Furthermore, a fair to good ICC value is presented for inter-rater reliability. The SDCs of the present study indicate that exceeding 25%–26% of the mean CSA reflects a real change.

To our knowledge, this is the first study determining all reproducibility outcomes by computing the ICC, SDC and LoA. The reliability results of the present study confirm the results of previous research (Tagliafico and Martinoli 2013; Thoirs and Williams 2001; Thoirs et al. 2005). Thoirs and colleagues (Thoirs and Williams 2001; Thoirs et al. 2005) reported excellent intra-rater reliability (ICC of 0.88), and fair to good inter-rater reliability (Pearson’s correlation coefficient of 0.81) for US measurements of the ulnar nerve in healthy individuals. Tagliafico and Martinoli (2013) and Boehm et al. (2014) also reported fair to excellent intra-rater (ICCs of 0.63–0.96 and 0.93, respectively) and inter-rater (ICCs of 0.59–0.93 and 0.98, respectively) reliability for combined US measurements of different nerves in healthy patients. As a result, caution should be exercised when comparing these results. Furthermore, the results of this study cannot confirm the outcome of previous research (Martinoli et al. 2002), which states that US is sonographer dependent. The ICC values of both

Table 1. Demographic and anthropometric characteristics of participants (N = 69)

<table>
<thead>
<tr>
<th>Characteristics</th>
<th>Mean ± SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>Age (y)</td>
<td>36 ± 15</td>
</tr>
<tr>
<td>Gender (men/women)</td>
<td>28/41</td>
</tr>
<tr>
<td>Height (m)</td>
<td>1.74 ± 0.10</td>
</tr>
<tr>
<td>Weight (kg)</td>
<td>72.1 ± 13.8</td>
</tr>
<tr>
<td>Body mass index (kg/m²)</td>
<td>23.82 ± 3.65</td>
</tr>
<tr>
<td>Dominant arm (right/left)</td>
<td>60/9</td>
</tr>
</tbody>
</table>

SD = standard deviation.

Table 2. Mean, SD and 95% CI of the CSA of the ulnar nerve of both measurements

<table>
<thead>
<tr>
<th>Measurement (mm²)</th>
<th>Mean ± SD</th>
<th>95% CI</th>
</tr>
</thead>
<tbody>
<tr>
<td>Measurement 1</td>
<td>9.83 ± 2.54</td>
<td>4.85–14.80</td>
</tr>
<tr>
<td>Measurement 2</td>
<td>10.08 ± 2.46</td>
<td>4.90–14.90</td>
</tr>
</tbody>
</table>

SD = standard deviation; CI = confidence interval; CSA = cross-sectional area.
sonographers differed only slightly in the present study. A possible explanation for the small difference might be the well-defined and detailed measurement protocol.

This is the first study establishing the SDC of US measurements of the ulnar nerve. A change of 25%–26% of the mean CSA is needed before a real change besides measurement error can be detected. The SDC has to be interpreted with reference to changes in CSA outcomes of the ulnar nerve of individuals diagnosed with UNE compared with healthy individuals or with the asymptomatic arm. Previous studies determined the CSA values of healthy individuals and individuals diagnosed with UNE (Ayromlou et al. 2012; Bayrak et al. 2010; Wiesler et al. 2006). Statistical significant differences of the outcomes of the CSA of the ulnar nerve in both groups were found in all three studies (Ayromlou et al. 2012; Bayrak et al. 2010; Wiesler et al. 2006). In these studies, the CSA of the ulnar nerve increased by 56% (Bayrak et al. 2010), 75% (Ayromlou et al. 2012) and 192% (Wiesler et al. 2006) in individuals diagnosed with UNE compared with healthy individuals. Considering these detected increases in CSA, it can be presumed that an entrapped nerve will transcend the SDC found in the present study. On this basis, it seems that the SDC of 25%–26% found in this study will be clinically applicable to detect changes other than measurement error. The SDC reported in the present study cannot directly be compared with the statistically significant changes reported in the three aforementioned studies, because measurement protocols differed. For example, elbow flexion ranged from 15° (Bayrak et al. 2010; Wiesler et al. 2006) to 70°–80° in the present study and 90° in the study of Ayromlou et al. (2012). Further research is needed to determine whether different test positions influence the SDC.

Standardization of the measurement protocol was employed to increase reproducibility. Elbow flexion was chosen to be 70°–80°. This position is based on physiologic changes and varying representations of the ulnar nerve from approximately 90° to full elbow flexion, as well as the higher chance of subluxation of the nerve because of discontinuity of Osborne’s ligament from about 90° of flexion (James et al. 2011). In previous studies, the measurements were performed at different angles of elbow flexion, resulting in varying means for the CSA of the ulnar nerve (Ayromlou et al. 2012; Jacob et al. 2004; Tagliafico and Martinoli 2013; Thoirs and Williams 2001, Thoirs et al. 2005). The mean CSA values of the ulnar nerve in the present study are higher. One explanation could be that the measurements in the present study were performed on the hyper-echoic rim, instead of inside the rim. Many protocols chose to exclude the rim for practical reasons and comparability with outcomes of other studies. The rim was included in this study, because excluding the rim means excluding a part of the nerve, the epineurium. Even the epineurium can change pathologically during entrapment (Thoirs et al. 2005; Wiesler et al. 2006). Therefore changes in the epineurium might signify differences in the mean CSA (Thoirs et al. 2005). Further research is needed to determine the influence of the pathologic change of the hyper-echoic rim on the increase in CSA and on reproducibility in individuals with UNE, to decide whether the rim should consequently be included or excluded during the measurements. Furthermore, future research should focus on the influence of the test positions on the mean CSA outcomes and on the reproducibility of the US measurements.

A limitation of the present study might be the 15-min interval between measurements. This short interval could have caused the sonographers to remember the outcomes of previous measurements. This interval was chosen because of organizational considerations and because of the expectation that participants would remain stable between measurements. In the present study, outcomes were not expected to be influenced by recollection because recall bias was reduced by using research assistants to blind the sonographers to the outcomes. Another limitation might be the generalization of the outcomes. In the present study, a convenience sample of young healthy adults was measured before various individuals with clinical presentation of UNE. In addition, obese individuals or those with increased muscle echogenicity as a result of denervation and aging may possibly be more difficult to image. Future research needs to determine and compare the reliability and agreement values of US measurements.

<table>
<thead>
<tr>
<th>Reliability</th>
<th>ICC (95% CI)</th>
<th>SDC, mm² (% of mean CSA)</th>
<th>Mean difference (95% LoA)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Inter-rater reliability</td>
<td>0.63* (0.46–0.76)</td>
<td>2.63 (25)</td>
<td>−0.13 (−2.56 to 2.29)</td>
</tr>
<tr>
<td>Sonographer 1</td>
<td>0.85* (0.76–0.90)</td>
<td>2.76 (25)</td>
<td>−0.10 (−2.50 to 2.29)</td>
</tr>
<tr>
<td>Sonographer 2</td>
<td>0.88* (0.80–0.92)</td>
<td>2.63 (26)</td>
<td>−0.38 (−2.93 to 2.18)</td>
</tr>
</tbody>
</table>

ICC = intra-class correlation coefficient; CI = confidence interval; SDC = smallest detectable change; LoA = limits of agreement; CSA = cross-sectional area.

* p < 0.001.
measures of the ulnar nerve at the cubital tunnel in both groups. The raters performing the measurements were experienced in using the protocol. Future research should determine if the protocol is utilisable for less experienced raters.

Based on the inter- and intra-rater reliability outcomes, the present study indicates that US is a usable modality for physical therapists and radiologists. In daily practice, comparison of measurements is mostly made by one sonographer. Based on the excellent intra-rater reliability of both the less and the more experienced sonographers, it seems that US can contribute to the quality of the diagnostic process and to the evaluation of changes over time. Furthermore, sonographers with different levels of experience seem to be able to obtain highly reproducible US measurements of the ulnar nerve. According to the outcome of the SDC compared with changes in the CSA in individuals with UNE, it seems that tissue changes can be detected after treatment.

CONCLUSIONS

On the basis of the fair to good inter-rater reliability and excellent intra-rater reliability for both raters, US seems to be a valuable tool for determining CSA outcomes of the ulnar nerve at the cubital tunnel for diagnostic purposes and repetitive measurements in clinical practice. Because of the clinically applicable intra-rater agreement values, US seems to be useful in determining treatment effects such as tissue changes over time.

REFERENCES
