Review

ASSESSING THE RELIABILITY OF ULTRASOUND IMAGING TO EXAMINE PERIPHERAL NERVE EXCURSION: A SYSTEMATIC LITERATURE REVIEW

Ben Kasehagen,* Richard Ellis,† Rodney Pope,* Nicholas Russell,* and Wayne Hing*  
* Bond Institute of Health and Sport, Faculty of Health Sciences and Medicine, Bond University, Robina, Gold Coast, Queensland, Australia; and † Health and Rehabilitation Research Institute, School of Clinical Sciences, Auckland University of Technology, Auckland, New Zealand

(Received 29 March 2017; revised 22 August 2017; in final form 28 August 2017)

Abstract—Ultrasound imaging (USI) is gaining popularity as a tool for assessing nerve excursion and is becoming an important tool for the assessment and management of entrapment neuropathies. This systematic review aimed to identify current methods and report on the reliability of using USI to examine nerve excursion and identify the level of evidence supporting the reliability of this technique. A systematic search of five electronic databases identified studies assessing the reliability of using USI to examine nerve excursion. Two independent reviewers critically appraised and assessed the methodological quality of the identified articles. Eighteen studies met the eligibility criteria. The majority of studies were of “moderate” or “high” methodological quality. The overall analysis indicated a “strong” level of evidence of moderate to high reliability of using USI to assess nerve excursion. Further reliability studies with consistency of reporting are required to further strengthen the level of evidence.

Key Words: Reliability, Nerve excursion, Ultrasound imaging.

INTRODUCTION

The peripheral nervous system is constantly exposed to mechanical loads imposed upon it by movements and postures of the body. To cope with and respond to these external forces (tensile, shear and compressive forces), a peripheral nerve must be able to withstand compression, elongate and stretch, glide and slide relative to its interfacing tissues (Topp and Boyd 2012). Compromise of any one (or combination) of these neural biomechanical responses is believed to be part of the multifactorial etiology of many peripheral neuropathies (Dilley et al. 2008; Greening et al. 2005). For example, one of the most commonly reported features of carpal tunnel syndrome is impaired median nerve movement through the carpal tunnel or forearm (Filiius et al. 2013; Hough et al. 2007; Nakamichi and Tachibana 1995). A recent systematic review concluded that reduced median nerve excursion, observed with ultrasound imaging (USI), was a commonly reported and significant feature of carpal tunnel syndrome (Ellis et al. 2017).

With this in mind, the assessment of nerve biomechanics is rapidly evolving, with in vivo methods superseding traditional cadaveric methods. For example, contemporary methods for evaluating peripheral nerve excursion in vivo utilize real-time USI. A range of research reports have detailed the use of USI to assess in vivo nerve excursion for the median (Coppieters et al. 2009; Dilley et al. 2003; Filiius et al. 2013; Hough et al. 2007), ulnar (Dilley et al. 2007), radial (Kasehagen et al. 2016), sciatic (Coppieters et al. 2015; Ellis et al. 2012; Ridehalgh et al. 2012), tibial (Carroll et al. 2012; Ellis et al. 2008) and common fibular (Boyd et al. 2012) nerves. Newer technologies, such as sonoelastography, have been reported as methods to examine nerve excursion and also parameters such as shear strain and passive stiffness (Andrade et al. 2016; Greening and Dilley 2017; Yoshii et al. 2017). Furthermore, USI can also be used to assess peripheral nerve morphology and structure (e.g., thickness and cross-sectional area) (Alshami et al. 2009; Fink et al. 2017).

On this basis, there exists a significant opportunity to utilize USI to assess peripheral nerve structure and excursion in clinical populations, particularly those with...
peripheral neuropathies. There is growing support for di-
agnostic use of USI to evaluate nerve biomechanics and
structure in clinical conditions such as carpal tunnel syn-
drome (Ellis et al. 2017; McDonagh et al. 2015). However,
the expansion of these in vivo techniques for use in clinical
populations is still in its infancy.

For researchers and clinicians to have confidence in
using USI to assess in vivo nerve excursion, the avail-
able methods need to have been assessed to confirm ade-
quately levels of measurement reliability. Furthermore,
the growing popularity, cost effectiveness and accessibil-
ity of USI highlights the importance of determining its
reliability when used for such measurements. On this basis,
the purpose of this systematic review was to identify, crit-
ically appraise and synthesize key findings from studies
that have assessed the reliability of USI measurements of
peripheral nerve excursion in vivo. The specific aims of
the review were to identify those peripheral nerves that
have been examined in this way, to document the ap-
proaches used for making such measurements, to investigate
the reported levels of reliability for these measurements
while considering the methodological quality of the iden-
tified reliability studies and to establish the current level
of evidence that supports the reliability of using USI to
quantify nerve excursion. The scope of the review was
limited to examining nerve excursion and does not include
examination of any other biomechanical properties of the
nerve.

METHODS

The design and reporting of this systematic review
with critical narrative synthesis have been guided by the
Centre for Reviews and Dissemination’s Guidance for Un-
dertaking Reviews in Health Care (Centre for Reviews and
Dissemination 2009) and the Preferred Reporting Items
for Systematic Reviews and Meta-analyses (PRISMA)
statement (Moher et al. 2009).

Search strategy

A systematic literature search was conducted to iden-
tify all relevant articles for inclusion in this review. The
search sought articles published up until May 2016, with
no restriction on publication date. The search was con-
ducted across the relevant health and science electronic
databases (PubMed, Embase, CINAHL, SportDiscus and
Scopus), with additional hand searching of reference lists
of eligible articles also performed. The search was con-
ducted using a consistent search strategy across all databases
and included key words from three main concepts: USI
(ultrasound, ultrasonography, sonography), reliability (re-
liability, repeatability) and nerve excursion (nerve,
peripheral nerve, nervous system, neural system, nerve
movement, nerve excursion, neurodynamic, neural mo-
bilization, neural mobilization, neural glide, nerve glide,
nerual slide, nerve slide). The Boolean operators “Or” and
“And” were used to link the key words from each concept
and to link the concepts themselves, respectively. After
article selection, a final hand search was performed of the
reference lists of the included articles to identify any other
potentially eligible articles.

Screening and selection

One reviewer (B.K.) screened all titles and ab-
stracts of the 1592 articles identified in the literature search
to assess potential eligibility. Duplicates (151) and ar-
ticles that were clearly ineligible (1412) were excluded
during this initial screening process. Full text was ob-
tained of the remaining 29 potentially eligible studies. One
of these 29 studies was reported in two separate articles
(Ridehalgh et al. 2012, 2014); these articles utilized the
same participant group and were therefore considered
within the remainder of this systematic review in refer-
tence to the original work (Ridehalgh et al. 2012).

Two reviewers (B.K. and N.R.) independently ap-
praised all identified studies against the following inclusion
and exclusion criteria to determine final eligibility. Dif-
fences in judgments were discussed with an additional
reviewer (R.E.), who acted as an arbiter to determine the
final judgment of eligibility.

Inclusion criteria

Ultrasound imaging was used to quantify in vivo nerve
excursion. Reliability of assessing nerve excursion was
quantified, and the method of statistical analysis used to
assess reliability was specified.

The sample included human participants; no restric-
tions were made regarding type of study cohort (e.g.,
healthy or clinical populations).

Informed consent was provided for all study partici-
pants, as was protocol approval by an ethics committee
or institutional review board.

Exclusion criteria

Articles not available in English or full text not
available. Those studies deemed eligible formed the final
set of included studies for the review. The results of the
search, screening and selection processes were docu-
mented in a PRISMA diagram (Fig. 1) (Moher et al. 2009).

Critical appraisal of methodological quality. The
methodological quality of all included studies was criti-
cally appraised using a standardized Critical Appraisal Tool
(CAT) developed by Brink and Louw (2012). The CAT
was designed specifically to critically appraise the meth-
odological quality of reliability and validity studies that
have assessed clinical outcome measures and objective tests
(Brink and Louw 2012). The CAT consists of 13 items in
total, some of which are pertinent to assessing reliability methods and some of which are pertinent to assessing clinical validity methods. As previously reported (Rabelo et al. 2015), four items of the CAT (items 3, 7, 9 and 11) are relevant only to methods used to assess validity; therefore, these items were excluded from use in the current review. This left a total of nine items to be used in the modified CAT (MCAT) for appraising the methodological quality of the included studies. From these items, both items 4 and 5 relate to blinding of the raters—to either the findings of other raters (item 4, inter-rater reliability) or to their own findings (item 5, intra-rater reliability) (Brink and Louw 2012). With this in mind, for this review, items 4 and 5 of the MCAT were combined to form one item to be assessed, with an overall score of “yes” or “no” being awarded depending on whether blinding of raters was reported in a way that was appropriate to the particular study design (i.e., based on whether the study assessed inter-rater reliability, intra-rater reliability or both).

The CAT developed by Brink and Louw (2012) does not yield an overall quality score or rating, and therefore, a scoring and rating system previously developed by Prowse et al. (2016) for use with this MCAT was employed. One point was awarded for meeting each item of the MCAT, with a maximum score of eight points possible, as items 4 and 5 were combined to form one item. Methodological quality ratings were then assigned as follows, based on total quality score: 0–2 = poor; 3–4 = fair; 5–6 = moderate; and 7–8 = high (Prowse et al. 2016). Two reviewers (B.K. and N.R.) independently appraised all identified studies using the MCAT. Differences in judgments for individual items were discussed with a third reviewer (R.P.), who acted as an arbiter to determine a final judgement for the item.

Data extraction and analysis

Methodological details and key findings of relevance to the aim of this review were extracted from each included study using a systematic approach and then tabulated to allow for comparisons across the included studies. Meta-analysis was not performed because of the heterogeneity of the different methods employed across the included studies (e.g., differences in the nerves examined, the body sites selected for USI and the body
movements utilized to induce nerve movement). Instead, descriptive analysis and a critical narrative synthesis of key findings were performed.

After data extraction and tabulation, a descriptive analysis was conducted to determine the levels of evidence provided by the included studies regarding the reliability of USI measurements of nerve excursion (van Tulder et al. 2003). This approach to assessing levels of evidence has been used in other systematic reviews (Barrett et al. 2014; Prowse et al. 2016). During this analysis, for each nerve imaged in the included studies, a rating was assigned to reflect the level of evidence provided by the included studies that related to that particular nerve. The level of evidence in each instance was rated as strong, moderate, limited or conflicting based on the number of relevant studies, along with the methodological quality, and using the rating system reported by van Tulder et al. (2003) (Table 1).

A critical narrative synthesis was then conducted, first to elucidate commonalities and variations in the protocols used in the different studies and for different nerves to measure nerve excursion using USI. This was important, because the reliability of measurements can clearly be affected by the measurement protocol. Second, the critical narrative synthesis considered the key findings from included studies regarding the reliability of USI measurements of nerve excursion. The synthesis considered reliability of these measurements for each nerve specifically and overall.

RESULTS

Literature search and selection

Results of the literature search, screening and selection processes are summarized in the PRISMA diagram in Figure 1. Eighteen studies, including four studies identified from manual searching of the reference lists of eligible studies, met the eligibility criteria and were included in the final review. The study conducted by Ridehalgh et al. (2014) reported reliability data that were previously presented by the same group in 2012 (Ridehalgh et al. 2012). For this reason, the data reported in Ridehalgh et al. (2012) were reported in the current review without further inclusion of the 2014 study.

Peripheral nerves examined

Of the 18 studies included in this review, 9 examined the median nerve (Coppieters et al. 2009; Farooq 2012; Filius et al. 2013; Gonzalez-Suarez et al. 2015; Hough et al. 2000, 2007; Martínez-Payá et al. 2015; Paquette et al. 2015; Wang et al. 2014), 1 examined the radial nerve (Kasehagen et al. 2016), 4 examined the sciatic nerve (Coppieters et al. 2015; Ellis et al. 2008, 2012; Ridehalgh et al. 2012), 5 examined the tibial nerve (Boyd and Dilley 2014; Boyd et al. 2012; Carroll et al. 2012; Ellis et al. 2008; Shum et al. 2013) and 1 examined the common fibular nerve (Boyd et al. 2012).

Participant characteristics

A total of 392 participants were involved across the 18 studies included in the current review, with 59% of these participants being female. The sample sizes in individual studies ranged from 6 participants (Gonzalez-Suarez et al. 2015) to 56 participants (Hough et al. 2007). The average age across all groups was 36 y (range: 18–86 y).

Thirteen studies recruited solely healthy participants (Carroll et al. 2012; Coppieters et al. 2009, 2015; Ellis et al. 2008, 2012; Filius et al. 2013; Gonzalez-Suarez et al. 2015; Hough et al. 2000; Kasehagen et al. 2016; Martínez-Payá et al. 2015; Ridehalgh et al. 2012; Shum et al. 2013; Wang et al. 2014). Four studies compared healthy participants with participants with known clinical conditions, which included type II diabetes mellitus (Boyd et al. 2012), type I or type II diabetes mellitus (Boyd and Dilley 2014), whiplash-associated disorder (Farooq 2012) and carpal tunnel syndrome (Hough et al. 2007). The remaining study pooled the data from healthy and carpal tunnel syndrome participants and did not compare between groups (Paquette et al. 2015). See Table 2 for further participant information.

Protocols for USI measurements of nerve excursion

A number of different protocols and techniques have been reported for using USI to quantify nerve excursion. The most commonly reported is frame-by-frame digital analysis of grey speckle features from within the ultrasound image (also known as speckle tracking). Several different software packages have been reported that use cross-correlation algorithms to compare the movement of gray-scale, speckle features (within specified regions of interest) between individual ultrasound frames (Dilley et al. 2001; Nicoud et al. 2011). Thirteen of the 18 included studies (Boyd and Dilley 2014; Boyd et al. 2012; Carroll et al. 2012; Coppieters et al. 2009, 2015; Ellis et al. 2008, 2012; Farooq 2012; Gonzalez-Suarez et al. 2015; Kasehagen et al. 2016; Paquette et al. 2015; Ridehalgh et al. 2014).
<table>
<thead>
<tr>
<th>Study</th>
<th>Study population</th>
<th>Item 1 Demographic characteristics</th>
<th>Item 2 Adequate description of observer and competence</th>
<th>Testing circumstances and execution</th>
<th>Item 10 Sufficient description of test procedure</th>
<th>Data analysis</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Median nerve</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Martínez-Payá et al. (2015)</td>
<td>n = 22 11 M:11 F</td>
<td>Y Healthy: 22 y</td>
<td>Y Sonographer 12 y of USI experience</td>
<td>N Not stated Intra-observer</td>
<td>N Not stated</td>
<td>Y Y No withdrawals</td>
<td>Y κ coefficient (95% CI)</td>
</tr>
<tr>
<td>Gonzalez-Suarez et al. (2015)</td>
<td>n = 6 2 M:4 F</td>
<td>Y Healthy: 24 y</td>
<td>Y Physician 6 y of MSK USI experience</td>
<td>N Not stated Intra- and inter-observer</td>
<td>N Not stated</td>
<td>Y Y No withdrawals</td>
<td>Y ICC (95% CI), SEM, MDC</td>
</tr>
<tr>
<td>Paquette et al. (2015)</td>
<td>n = 7 2 M:5 F 6 M:5 F</td>
<td>Y Uni/bilateral CTS: 56 y Healthy: 37 y</td>
<td>Y Trained physical therapist</td>
<td>N Not stated Intra-observer</td>
<td>N Not stated</td>
<td>Y 30-min interval</td>
<td>Y Y No withdrawals</td>
</tr>
<tr>
<td>Wang et al. (2014)</td>
<td>n = 10 4 M:6 F (bi-lateral wrists)</td>
<td>Y Healthy: 39 y</td>
<td>Y Radiologist 5 y MSK USI experience</td>
<td>N Not stated Intra- and inter-observer</td>
<td>N Not stated</td>
<td>Y ≥3 mo between analyses</td>
<td>Y Y No withdrawals</td>
</tr>
<tr>
<td>Filius et al. (2013)</td>
<td>n = 20 10 M:10 F</td>
<td>Y Healthy: 28 (21–72) y</td>
<td>Y Physician MSK USI experience</td>
<td>Y Blinded Intra- and inter-observer</td>
<td>Y Randomized</td>
<td>2-d interval</td>
<td>Y Y No withdrawals</td>
</tr>
<tr>
<td>Farooq (2012)</td>
<td>n = 7 2 M:5 F 5 M:5 F</td>
<td>Y WAD: 35 Healthy: 25 y</td>
<td>Y Trained MSK physiotherapist</td>
<td>N Not stated Inter-observer</td>
<td>N One test condition</td>
<td>N Not stated</td>
<td>Y Y No withdrawals</td>
</tr>
<tr>
<td>Coppeters et al. (2009)</td>
<td>n = 15 7 M:8 F</td>
<td>Y Healthy: 30</td>
<td>N Not stated</td>
<td>N Not stated Inter-observer</td>
<td>Y Randomized</td>
<td>N Not stated</td>
<td>Y Y No withdrawals</td>
</tr>
<tr>
<td>Hough et al. (2007)</td>
<td>n = 19 8 M:11 F 6 M:5 F</td>
<td>Y Idiopathic CTS: 57 (35–86) y Healthy: 48 (21–64)</td>
<td>N Not stated</td>
<td>Blinded Intra-observer</td>
<td>Y Reversed order</td>
<td>N Not stated</td>
<td>Y Y No withdrawals</td>
</tr>
<tr>
<td>Hough et al. (2000)</td>
<td>n = 16 7 M:9 F (bilateral arms)</td>
<td>Y Healthy: 38 (26–61) y</td>
<td>Y Not stated</td>
<td>Blinded Intra-observer</td>
<td>NA One test condition</td>
<td>N Not stated</td>
<td>Y Y No withdrawals</td>
</tr>
<tr>
<td><strong>Radial nerve</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kasehagen et al. (2016)</td>
<td>n = 30 12 M:18 F</td>
<td>Y Healthy: 30 (19–49) y</td>
<td>Y Sonographer 10 y of USI experience</td>
<td>Y Blinded Intra-observer</td>
<td>Y Randomized</td>
<td>N Not stated</td>
<td>Y Y No withdrawals</td>
</tr>
</tbody>
</table>

(continued on next page)
<table>
<thead>
<tr>
<th>Study</th>
<th>Study population</th>
<th>Item 1 Demographic characteristics</th>
<th>Item 2 Adequate description of observer and competence</th>
<th>Testing circumstances and execution</th>
<th>Data analysis</th>
<th>Quality</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>Cohort: Mean age (range)</td>
<td>Intra- and/or inter-observer blinding</td>
<td>Stability of variable and suitability of time interval</td>
<td>Sufficient description of test procedure</td>
<td>Description results with explanation of withdrawals</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sciatic nerve</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coppieters et al. (2015)</td>
<td>n = 15 6 M:9 F</td>
<td>Y Healthy: 28</td>
<td>N Not stated</td>
<td>Y Blinded Intra-observer</td>
<td>N Randomized</td>
<td>N Not stated</td>
</tr>
<tr>
<td>Ridehalgh et al. (2012, 2014)</td>
<td>n = 18 9 M:9 F</td>
<td>Y Healthy: 29 (19–68) y</td>
<td>N Not stated</td>
<td>Y Blinded Intra-observer</td>
<td>N Same order</td>
<td>Y 48 h—one wk interval</td>
</tr>
<tr>
<td>Ellis et al. (2012)</td>
<td>n = 31 9 M:22 F</td>
<td>Y Healthy: 29 (21–61) y</td>
<td>Y Sonographer 5 y experience</td>
<td>Y Blinded Intra-observer</td>
<td>Y Randomized</td>
<td>Y 1-min interval</td>
</tr>
<tr>
<td>Sciatic and tibial nerves</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Ellis et al. (2008)</td>
<td>n = 27 13 M:14 F</td>
<td>Y Healthy: 23 (18–38)</td>
<td>Y Experienced sonographer</td>
<td>Y Blinded Intra-observer</td>
<td>N Not stated</td>
<td>Y 1-min interval</td>
</tr>
<tr>
<td>Tibial nerve</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boyd and Dilley (2014)</td>
<td>n = 20 6 M:14 F</td>
<td>Y Type I and/or II DM: 51 (25–66) y Healthy: 46 (23–66)</td>
<td>N Not stated</td>
<td>Y Blinded Intra-observer</td>
<td>N Not stated</td>
<td>N Not stated</td>
</tr>
<tr>
<td></td>
<td>n = 20 10 M:10 F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Shum et al. (2013)</td>
<td>n = 25 11 M:14 F</td>
<td>Y Healthy: 29 y</td>
<td>N Not stated</td>
<td>N Not stated</td>
<td>N Not stated</td>
<td>Y 2-min interval</td>
</tr>
<tr>
<td>Carroll et al. (2012)</td>
<td>n = 16 6 M:10 F</td>
<td>Y Healthy: 35 y</td>
<td>Y Examiner 3 mo of USI training</td>
<td>N Not stated</td>
<td>N Not stated</td>
<td>Y 1-min interval between repeated scans. 5-min interval between sessions 1 and 2</td>
</tr>
<tr>
<td>Tibial and common fibular nerves</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Boyd et al. (2012)</td>
<td>n = 5 4 M:1 F</td>
<td>Y Type II DM: 57 y Healthy: 40 y</td>
<td>N Not stated</td>
<td>N Not stated</td>
<td>N Not stated</td>
<td>N Not stated</td>
</tr>
<tr>
<td></td>
<td>n = 5 1 M:4 F</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Y = met MCAT criteria; N = did not meet MCAT criteria; NA = not accessible; M = male; F = female; DM = diabetes mellitus; WAD = whiplash-associated disorder; CTS = carpal tunnel syndrome; ICC = intraclass correlation; CI = confidence interval; MDC = minimal detectable change; SRD = smallest real difference; SEM = standard error measurement; USI = ultrasound imaging; MSK = musculoskeletal.

Two studies utilized real-time spectral Doppler USI to quantify nerve excursion (Hough et al. 2000, 2007). By measuring the deflections of the Doppler/B-mode signals, in several planes, velocity of nerve movement was calculated with excursion quantified from measuring the area under the deflection traces (Hough et al. 2000).

For the measurement of transverse nerve movement, several studies compared the position of the relevant nerve (often the centroid point or the nerve outline) and digitally measured the change in nerve position (representing the amount of excursion) between the first and final frames of an ultrasound video (Ellis et al. 2008; Filius et al. 2013; Martínez-Payá et al. 2015; Wang et al. 2014).

Critical appraisal of methodological quality

The methodological quality of the studies included in this review was rated as high for 4 studies (Ellis et al. 2008, 2012; Filius et al. 2013; Kasehagen et al. 2016), moderate for 13 studies (Boyd and Dilley 2014; Carroll et al. 2012; Coppieters et al. 2009, 2015; Furooq 2012; Gonzalez-Suarez et al. 2015; Hough et al. 2000, 2007; Martínez-Payá et al. 2015; Paquette et al. 2015; Ridehalgh et al. 2012; Shum et al. 2013; Wang et al. 2014) and fair for 1 study (Boyd et al. 2012). Across all of the included studies, the items regarding description of participant characteristics (item 1) and test procedure (item 10), explanation of withdrawals (item 12) and description of appropriate statistical analyses (item 13) were all satisfied.

There was some variability across the studies regarding the description of the level of experience or competence of the sonographer (item 2). Seven studies did not describe the person performing USI (Boyd and Dilley 2014; Boyd et al. 2012; Coppieters et al. 2009, 2015; Hough et al. 2000, 2007; Shum et al. 2013), whilst a further three studies (Furooq 2012; Paquette et al. 2015; Ridehalgh et al. 2012) stated a “trained” physiotherapist conducted the imaging, with the implication that they were trained in USI.

There was also variability in the reporting of observer (i.e., rater) blinding, irrespective of whether the studies examined intra-rater reliability or inter-rater reliability (items 4 and 5 combined). Ten studies (Boyd et al. 2012; Carroll et al. 2012; Coppieters et al. 2009; Furooq 2012; Gonzalez-Suarez et al. 2015; Martínez-Payá et al. 2015; Paquette et al. 2015; Ridehalgh et al. 2012; Shum et al. 2013; Wang et al. 2014) did not report blinding of the observers during the data analysis.

Item 6 concerned the order in which participants were assessed on each measurement occasion. Randomization of participant order can reduce the possibility of confounding through recall of ordered participants and results. There was variability in the scores for this item across studies.

The stability of a measured variable (in this case, nerve excursion) over time (item 8) is a feature of reliability studies that should be taken into account. It is possible that calculations of the reliability of measurement may be affected not only by lack of repeatability in the measurement methods or technique, but also because the measured variable itself alters over time. In situations where the measured variable is quite unstable and rapidly changing, the interval between measurement occasions must be kept short to minimize the instability effect on reliability scores. This review did not seek to determine whether the intervals reported were appropriate, but rather documented the extent to which study authors reported they considered this issue or reported the time interval. Nine studies (Boyd and Dilley 2014; Boyd et al. 2012; Coppieters et al. 2009, 2015; Furooq 2012; Hough et al. 2000, 2007; Kasehagen et al. 2016; Martínez-Payá et al. 2015) did not report the specific intervals between measurement occasions.

Reliability of ultrasound imaging measurements of nerve excursion

Several types of statistics and plots have been recommended as the optimal indicators of the reliability of measurements performed using USI. These include intraclass correlation coefficients (ICCs), standard errors of measurement (SEMs), minimum detectable changes (MDCs) and Bland–Altman plots (Whittaker and Stokes 2011; Whittaker et al. 2007). Most of the studies included in this review used an ICC as the main reliability statistic, with the exceptions being use of a κ coefficient in one study (Martínez-Payá et al. 2015) and a dependability coefficient in another (Paquette et al. 2015), as noted in Table 3. Fourteen of the 18 studies reported an ICC coefficient, whilst 4 of the 18 studies reported a MDC. A majority of studies, 15 of 18, did not provide a Bland–Altman plot.

Table 3 provides details of the levels of reliability determined for the USI measurements of peripheral nerve excursion in each of the 18 included studies and for each nerve considered in the studies. Different types of reliability were assessed with 10 studies (Boyd and Dilley 2014; Boyd et al. 2012; Ellis et al. 2008, 2012; Filius et al. 2013; Hough et al. 2000, 2007; Kasehagen et al. 2016; Shum et al. 2013; Wang et al. 2014) examining intra-rater within-session reliability, 3 studies (Carroll et al. 2012; Paquette et al. 2015; Ridehalgh et al. 2012) examining intra-rater between-session reliability, 7 studies (Coppieters et al. 2009, 2015; Furooq 2012; Filius et al. 2013; Gonzalez-Suarez et al. 2015; Martínez-Payá et al. 2015; Wang et al. 2014).
Table 3. Reliability of ultrasound imaging measurements of peripheral nerve excursion

<table>
<thead>
<tr>
<th>Nerve Level</th>
<th>USI protocol</th>
<th>Study quality</th>
<th>Reported statistic</th>
<th>Intra-rater reliability</th>
<th>Inter-rate reliability</th>
<th>Bland–Altman plot</th>
<th>Nerve Level of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Inter-rater reliability</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>USI protocol</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Median nerve: Longitudinal movement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Gonzalez-Suarez et al. (2015)</td>
<td>Speckle tracking</td>
<td>Moderate</td>
<td>ICC</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Paquette et al. (2015)</td>
<td>Speckle tracking</td>
<td>Moderate</td>
<td>Dependability coefficient</td>
<td></td>
<td>0.49–0.91 (ICC range); no 95% CI; 0.41–1.21 (SEM range); 0.95–2.82 (MDC range)</td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Farooq (2012)</td>
<td>Speckle tracking</td>
<td>Moderate</td>
<td>ICC</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Coppieters et al. (2009)</td>
<td>Speckle tracking</td>
<td>Moderate</td>
<td>ICC</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Hough et al. (2007)</td>
<td>Spectral Doppler</td>
<td>Moderate</td>
<td>ICC</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Martinez-Payá et al. (2015)</td>
<td>Nerve location comparison, start and end frames</td>
<td>Moderate</td>
<td>κ coefficient</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Wang et al. (2014)</td>
<td>Nerve location comparison, start and end frames</td>
<td>Moderate</td>
<td>ICC</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Filius et al. (2013)</td>
<td>Nerve location comparison, start and end frames</td>
<td>High</td>
<td>ICC</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Farooq (2012)</td>
<td>Speckle tracking</td>
<td>Moderate</td>
<td>ICC</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Median nerve: Transverse movement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Martinez-Payá et al. (2015)</td>
<td>Nerve location comparison, start and end frames</td>
<td>Moderate</td>
<td>κ coefficient</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Radial nerve: Longitudinal movement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Kasehagen et al. (2016)</td>
<td>Speckle tracking</td>
<td>High</td>
<td>ICC</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
</tr>
<tr>
<td>Sciatic nerve: Longitudinal movement</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Coppieters et al. (2015)</td>
<td>Speckle tracking</td>
<td>Moderate</td>
<td>ICC</td>
<td></td>
<td></td>
<td></td>
<td>No</td>
</tr>
<tr>
<td>Ridehalgh et al. (2012)</td>
<td>Speckle tracking</td>
<td>Moderate</td>
<td>ICC</td>
<td></td>
<td></td>
<td></td>
<td>Yes</td>
</tr>
</tbody>
</table>

(continued on next page)
Table 3. (continued)

<table>
<thead>
<tr>
<th>Study Quality</th>
<th>USI protocol</th>
<th>Nerve Level of evidence</th>
<th>Within session</th>
<th>Between sessions</th>
<th>Reliability (ICC/κ coefficient/dependability coefficient; 95% CI; SEM; MDC/SRD [mm])</th>
<th>Bland–Altman plot</th>
<th>Nerve Level of evidence</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td></td>
<td>Intra-rater reliability</td>
<td>Inter-rater reliability</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>Within session</td>
<td>Between sessions</td>
<td>Within session</td>
<td>Between sessions</td>
<td></td>
</tr>
<tr>
<td>High ICC</td>
<td>Speckle tracking</td>
<td>Ellis et al. (2012)</td>
<td>0.95; 0.92–0.96; 0.20; no MDC</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High ICC</td>
<td>Speckle tracking</td>
<td>Ellis et al. (2008)</td>
<td>0.75; 0.59–0.87; 0.75; no MDC</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High ICC</td>
<td>Nerve location comparison, start and end frames</td>
<td>Ellis et al. (2008)</td>
<td>0.76; 0.60–0.87; 1.01; no MDC</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High ICC</td>
<td>Nerve location comparison, start and end frames</td>
<td>Ellis et al. (2008)</td>
<td>0.39; 0.15–0.63; 0.62; no MDC</td>
<td>Yes</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate ICC</td>
<td>Speckle tracking</td>
<td>Boyd and Dilley (2014)</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fair ICC</td>
<td>Speckle tracking</td>
<td>Boyd et al. (2012)</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High ICC</td>
<td>Speckle tracking</td>
<td>Carroll et al. (2012)</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High ICC</td>
<td>Speckle tracking</td>
<td>Ellis et al. (2008)*</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate ICC</td>
<td>Speckle tracking</td>
<td>Boyd and Dilley (2014)</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fair ICC</td>
<td>Speckle tracking</td>
<td>Ellis et al. (2008)</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High ICC</td>
<td>Nerve location comparison, start and end frames</td>
<td>Ellis et al. (2008)</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate ICC</td>
<td>Speckle tracking</td>
<td>Boyd and Dilley (2014)</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fair ICC</td>
<td>Speckle tracking</td>
<td>Boyd et al. (2012)</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High ICC</td>
<td>Nerve location comparison, start and end frames</td>
<td>Ellis et al. (2008)</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Moderate ICC</td>
<td>Speckle tracking</td>
<td>Boyd and Dilley (2014)</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fair ICC</td>
<td>Speckle tracking</td>
<td>Boyd et al. (2012)</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>High ICC</td>
<td>Nerve location comparison, start and end frames</td>
<td>Ellis et al. (2008)</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Common fibular nerve: Transverse movement</td>
<td>Speckle tracking</td>
<td>Boyd et al. (2012)</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Common fibular nerve: Superficial-deep movement</td>
<td>Speckle tracking</td>
<td>Boyd et al. (2012)</td>
<td>No</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

USI = ultrasound imaging; PC = popliteal crease; PMT = posterior midthigh; ICC = intra-class correlation coefficient; CI = confidence interval; SEM = standard error of measurement; MDC = minimal detectable change; SRD = smallest real difference.

* This statistic was reported, but based on n = 3 participants.
examining inter-rater within-session reliability and 1 study (Gonzalez-Suarez et al. 2015) examining inter-rater between-session reliability. Comparisons between these different types of reliability was difficult as a majority of studies looked primarily at one form of reliability testing. One study compared within-session versus between-session inter-rater reliability (Gonzalez-Suarez et al. 2015), with reliability being less desirable for the between-session measures. Two studies (Filius et al. 2013; Wang et al. 2014) examined both intra-rater and inter-rater within-session reliability for the assessment of transverse movement of the median nerve, with comparable results.

Irrespective of the specific nerve imaged and the direction of excursion, the reliability of measurements of nerve excursion performed with USI was typically moderate to high. The exceptions to this finding were few. Paquette et al. (2015) reported a dependability coefficient of 0.49 (low reliability) for assessing median nerve excursion during a tensioner technique with the arm at 45° shoulder abduction. Ellis et al. (2008) reported an ICC of 0.39 (low reliability) for assessing superficial–deep (superficial = toward the skin, deep = away from the skin) sciatic nerve excursion at the posterior mid-thigh during a sliding technique performed at the ankle joint with the participants sitting.

Level of evidence for the reliability of USI measurements of nerve excursion

Consideration of the numbers of available studies and the methodological quality of each of those studies together (as per Table 1), along with the reliability findings described, revealed several findings regarding the levels of evidence that supported the key findings of this review. First, the pooled data, including all studies, nerves and directions of nerve excursion, provided a strong level of evidence to support the finding that the reliability of USI measurements of peripheral nerve excursion is typically moderate to high.

With respect to specific nerves assessed in included studies, the level of evidence to support the findings regarding reliability of USI measurements of excursion was found to be moderate for the median, sciatic and tibial nerves, and limited for the common fibular and radial nerves. Table 3 provides further details of the levels of evidence and USI reliability findings for each nerve.

DISCUSSION

The current systematic review identified 18 published studies that have assessed the reliability of measurements of peripheral nerve excursion derived from USI. These 18 studies investigated five peripheral nerves, including the median, radial, sciatic, tibial and common fibular nerves, across a variety of participant demographic characteristics.

One of the key aims of this review was to assess the methodological quality of those studies that have used USI as a method of quantifying nerve excursion. Across the 18 included studies, the methodological quality assessed via the MCAT was varied, with 4 studies of high, 13 studies of moderate and 1 study of low quality.

It should be noted that methodological limitations and specific measurement protocols may have directly influenced the levels of measurement reliability that were reported in the studies included in this review, and may also affect the reliability of such measurements in practice contexts. For example, it is widely accepted that USI is an operator-dependent tool. Therefore, the level of experience of the person who performs the USI may potentially influence the level of measurement reliability. The experience level of the sonographer has been reported to be positively associated with more consistent findings compared with less experienced examiners (Cartwright et al. 2013). Although not formally assessed from studies examining peripheral nerves with USI, differences in USI measurement reliability have been reported between experienced and novice sonographers when examining muscle morphology and function (Hides et al. 2007; Iwan et al. 2014). Eight of the included studies specifically described the experience level of the sonographer, whilst another three studies inferred that the sonographer had USI training. The remaining studies did not describe the sonographer.

The clear reporting of the experience level of the sonographer in future studies will be a key methodological feature to further increase confidence in findings regarding USI measurement reliability. Similarly, in practice contexts, the experience level of the sonographer and the extent to which the measurement protocol is both standardized and optimized will influence the precision of the measurements and the confidence we can have as to their accuracy.

Further inconsistencies in the reporting of methods included the description of observers/raters, particularly with respect to blinding, where 10 of the included studies failed to comment on this. Furthermore, the time between measurements was inconsistently reported. This may become a significant factor if the variable of interest (nerve excursion) has the capacity to change over time.

To facilitate the interpretation of reliability studies for USI measurements, experts in the field have advocated for optimization and consistency in both the statistical analyses used and the methods of reporting (Whittaker and Stokes 2011; Whittaker et al. 2007). Across the 18 studies included in this review, there was evidence of such consistency developing. For example, the majority of the included studies used statistics such as the ICC to report
their main reliability findings. Those that did not use the ICC instead used appropriate alternatives. However, there was variability in the use of supporting statistics, such as SEM, MDC, Bland–Altman plots and 95% confidence intervals.

It was clear from the review that few measurement protocols have been used to quantify nerve excursion from ultrasound data. By far, the majority of studies utilized speckle tracking via digital processing of specific greyscale, speckle features between individual frames of an ultrasound video or cine loop. Although several different analysis methods were reported that utilized speckle tracking, the vast majority (12 of the 18 included studies) utilized the method reported by Dilley et al. (2001), and one study utilized the method reported by Nicoud et al. (2011). Two studies (Hough et al. 2000, 2007) reported the use of spectral Doppler USI to assess longitudinal median nerve movement. However, it was not apparent that this method has been utilized since 2007, with more contemporary studies preferring speckle-tracking protocols. It must be noted that more contemporary ultrasound methods, such as sonoelastography, are emerging as tools to examine nerve excursion (among other biomechanical parameters). Although these relevant studies did not meet inclusion for this review, it will be of interest to follow these emerging technologies.

Pooled data from the 18 studies yielded “strong” evidence that the reliability of USI measurements of peripheral nerve excursion is typically moderate to high, with few exceptions. This finding provides substantial support for the use of USI in clinical and healthy populations to measure peripheral nerve excursion. However, on a specific nerve basis, the levels of evidence vary, with a “moderate” level of evidence supporting the finding that the reliability of USI measurements of median, sciatic and tibial nerve excursion is, on balance, high to very high, but only “limited” evidence to support findings of very high reliability for USI assessment of the common fibular nerve and moderate to high reliability for the radial nerve. No published evidence was identified to elucidate the reliability of such measurements for other peripheral nerves. Differences in nerve anatomy, course through the body, depth and surrounding structures may all influence the ease (or lack of) for imaging certain nerves, which may influence the reliability of respective measurements. This highlights the need for further research to comprehensively examine the reliability of this measurement approach for those nerves for which limited or no evidence exists. Future research should consider the methodological weaknesses identified in the studies included in this review, which are evident in Table 2 and highlighted in the synthesis of key findings from the critical appraisal of included studies.

A number of possible benefits exist for quantifying nerve excursion in clinical practice. The technique may be used as a diagnostic tool, for example, for the assessment of entrapment neuropathies (such as carpal tunnel syndrome) in which impaired nerve excursion is believed to be a key aetiological factor (Ellis et al. 2017). Furthermore, the selection of therapeutic techniques such as neural mobilization exercises, which aim to promote optimal peripheral nerve mechanics (Basson et al. 2017), may be better targeted to conditions where a known impairment of nerve movement against the interfacing tissues is identified with USI.

Strengths and limitations

This review represents the first review we have been able to identify on this topic and, thus, valuably informs practice and future research. The strong levels of evidence for some key findings, particularly those indicating USI measurements of peripheral nerve excursion are typically reliable, ensure the review is valuable to practitioners and researchers alike. Researchers will also be usefully informed by the areas of the findings where the levels of evidence were limited or non-existent, as these gaps in the literature can inform research planning. Methodological deficiencies identified in the included studies can also usefully inform future research design.

The review has a number of limitations that must be acknowledged. First, a single reviewer conducted the initial literature search across all databases. Two reviewers appraised the identified articles for eligibility before critically appraising the selected articles, thus reducing human error and selection bias. Second, because the data from included studies were not amenable to meta-analysis because of the heterogeneity in, for example, methods, nerves examined and populations from which participants were drawn, the findings were synthesized using a critical narrative approach supported by descriptive quantitative analyses. Finally, it was accepted that there are many different methodological variables that can influence reliability. A pragmatic approach was taken for this review to present results in a manner in which an overall impression of the evidence was considered.

CONCLUSIONS

The current systematic review identified a strong level of evidence to support the main finding that, typically, measurements of peripheral nerve excursion are moderately to highly reliable, with few exceptions. Although further research is needed to extend findings to nerves other than those examined in the included studies, the results of this review indicate that measurement of peripheral nerve excursion through USI is a promising technique that can be used with increasing confidence. Nevertheless, attention should be given to ensuring sonographer competence in the technique and ensuring the measurement protocols are...
standardized and optimized. Further research is needed to build our knowledge of what optimization of these protocols should entail, to continue to explore the use of USI for such measurements in nerves other than those reported in this review and to further elucidate the clinical implications of different measurement values in varying clinical contexts and populations.

REFERENCES

Alshami AM, Cairns CW, Wylie BK, Souvlis T, Coppieters MW. Reliability and size of the measurement error when determining the cross-sectional area of the tibial nerve at the tarsal tunnel with ultrasonography. Ultrasound Med Biol 2009;35:1098–1102.


