Objective/Background: The figure-of-eight technique is a measure for hand volume that has been validated among experienced American clinicians and physical therapy students, but not among Middle Eastern occupational therapy students. The purpose of this study was to assess the intrarater and inter-rater reliability as well as concurrent validity of the figure-of-eight technique of measuring hand volume by 4th year (of a 5-year curriculum) occupational therapy students.

Methods: This study used a cross-sectional design of a single group with three-level repeated measures of five raters. Twenty-three healthy students participated in this study. Five raters (4th year occupational therapy students) performed three separate blinded figure-of-eight measurements of hand volume for each hand. Two independent examiners performed one volumetric measurement for each hand using a water volumeter. Intraclass correlation coefficient (ICC) were used to examine the intrarater and inter-rater reliability as well as concurrent validity of the figure-of-eight technique of measuring hand volume by 4th year (of a 5-year curriculum) occupational therapy students.

Results: Intrarater reliability (ICC, 3k) ranged from .98 to .99 and the inter-rater reliability (ICC, 2k) was .99. The Pearson correlation coefficient for the concurrent validity was $r = .929$ ($p < .001$).

Conclusion: This study demonstrated that five occupational therapy students in the 4th year of a 5-year curriculum were reliable raters for hand volume using the figure-of-eight technique,
Introduction

Because occupational therapy students are required to conduct standardized evaluations and use the evaluation data to plan and implement treatment during their clinical fieldwork affiliations, it is important to know whether their evaluation skills match the standards of experienced clinicians (Hinojosa, Kramer, & Crist, 2010; Verma, Paterson, & Medves, 2006). Furthermore, occupational therapy students are often hired to assist underclass colleagues in the orthopaedic assessment lab or as data collectors in orthopaedic research projects. Furthermore, it is not always possible to have access to patients, or impose the burden of multiple assessments on patients, when a faculty needs to assess intrarater and inter-rater reliability of students’ evaluation skills. However, it is important for students to extend the knowledge obtained in the classroom to clinical practice and learn how to conduct assessments in a uniform manner, even if it means practicing their skills on their peers. Although multiple studies of inter-rater reliability of the figure-of-eight technique for measuring hand volume have been published, our study focuses on occupational therapy students in a Middle Eastern context. This is different from other studies that were performed by experienced occupational therapists and physical therapists on burn patients (Dewey, Hedman, Chapman, Wolf, & Holcomb, 2007) and by physical therapy students (Leard et al., 2004; Maihafer et al., 2003; Pellecchia, 2003).

In the Kuwait University Department of Occupational Therapy, measurement of hand volume to determine oedema is taught in two ways: the volumeter procedure and the figure-of-eight technique. In clinics, measurement of oedema is important, because it is an inflammatory tissue response to trauma that is characterized by accumulation of excessive amount of interstitial fluid in the limb. Oedema can be a serious problem if not resolved in timely manner as it can cause functional limitations, slow the recovery process, cause pain, joint stiffness, immobility, and may lead to the formation of adhesions and excessive fibrosis (Post, Visser-Meily, Boomkamp-Koppen, & Prevo, 2003; Rebeiro, Lima, Carreira, Masiero, & Chamlian, 2012).

Oedema measurements are also frequently used as an essential outcome measure in clinical practice. There are several methods to measure hand oedema, including girth circumference measurements (Lewis, 2010; Pellecchia, 2003), water displacement (Rebeiro et al., 2012), electronic balance measurements (Hughes, 2005; Hughes & Lau, 2008), bioelectric impedance, and computer modelling (Karges, Mark, Stikeleather, & Worrell, 2003). The two most commonly used tools to measure oedema in a clinical setting are volume measurements (with a water volumeter) and girth measurements (with a tape measure).

The volumeter was first introduced into medicine by Glisson in 1622 (Rebeiro et al., 2012) and it utilizes the same principle of water displacement first discovered by the ancient Greek mathematician, Archimedes, which states that the water volume displaced is equal to the volume of the object immersed in the water (Karges et al., 2003). The volumeter procedure involves immersing the oedematous hand into a water-filled translucent tank, which will displace excess water into a receiver cup. The displaced water is then measured by a graduated cylinder to quantify the hand volume (volumeter). The psychometric properties of upper extremity volumeter measurements are well documented (Dodds et al., 2004; Farrell, Johnson, Duncan, Offenbacker, & Curry, 2003; Rebeiro et al., 2012) and the method has been shown to be reproducible, with less than 1% error (Karges et al., 2003).

The volumeter is considered the “gold standard” method for measuring hand volume of oedematous hands and is regarded as one of the most useful standardized tools for measuring hand oedema because it can be quantified and used to assess the effectiveness of treatment (Rebeiro et al., 2012). Despite its wide popularity, however, the volumeter has several disadvantages when used in clinical settings. The volumeter has to be set up several minutes ahead of time (filling the tank with water), it is difficult to move once filled with water, and it is difficult to transport it between locations. Furthermore, volumetric measurements are messy as they require the patients to immerse their hands in water, and it is therefore unsuitable for certain patient populations.

The figure-of-eight technique is a girth measurement method in which the hand is measured with a standard tape measure that is circumferentially wrapped around the hand. The figure-of-eight technique is also a valid and reliable method of measuring hand oedema (Dewey et al., 2007; Leard et al., 2004; Maihafer et al., 2003; Pellecchia, 2003). The figure-of-eight technique is clinically more feasible than water volumetry to measure hand oedema because the technique is easy to administer and is cost and time efficient (Leard et al., 2004; Pellecchia, 2003). The time required to set up and perform the figure-of-eight technique is approximately 1 minute, whereas it takes several minutes to set up and perform measurements with a volumeter. The tool required to perform the figure-of-eight method is readily available in most clinical settings and is less expensive than the volumeter, with an average cost of $13 for a standard tape measure in comparison with $275 for a hand volumeter (as of December, 2012). Another advantage of using the figure-of-eight technique for measuring oedema over the volumeter is that it provides an alternative measuring method for patients with skin ulcers (Pani, Vanamail, & Yuvaraj, 1995), burns (Dewey et al., 2007), sutures, open wounds (Leard et al., 2004), externally fixated devises, and plastered splints (Karges et al., 2003), who cannot be tested using the volumeter.

Rehabilitation therapists frequently assess oedema in the hand to assess the effectiveness of treatment methods designed to decrease oedema for patients with upper
extremity pathology. It is important that the therapists perform these measurements in a reliable manner. When establishing therapists’ measurement consistency, the main forms of reliability that need to be ensured when measuring oedema are intrarater and inter-rater reliability (Pellecchia, 2003; Portney & Watkins, 2009). Intrarater reliability is established by the same data collector administering an assessment repeatedly. However, inter-rater reliability is established with more than one data collector administering the same assessment to the same individuals within the same time frame. Intrarater reliability measures the consistency of a single assessor, and inter-rater reliability measures consistency among raters, and whether raters are interchangeable (one assesses the patient at admission and another at discharge). Both forms of reliability are essential because differences in scores by the same therapists, or between different therapists can produce unacceptable measurement errors (Lindstrom-Hazel, Kratt, & Bix, 2009).

Most reliability studies in rehabilitation rely on professional therapists for data collection (Chen, Kasven, Karpatkin, & Sylvester, 2007). The psychometric properties of the figure-of-eight technique have been well validated when used by experienced occupational and physical therapists in clinical settings (Dewey et al., 2007; Leard et al., 2004; Maihafer et al., 2003). However, reliability studies of the technique have not been reported for occupational therapy students, or students in curricula outside of the United States, and these students are also required to carry out standardized evaluations and implement treatment plans during their clinical fieldwork training.

The purpose of this study was to assess the intrarater and inter-rater reliability as well as concurrent validity of the figure-of-eight technique of measuring hand volume by 4th year occupational therapy students.

Methods

Study design

The study used a cross-sectional design of a single group with repeated measures of 23 hands by five raters. Intrarater reliability refers to the consistency of scores of the same rater at three time intervals and without any change in the condition being measured; therefore, healthy participants with no conditions affecting the hand were included in this study. Inter-rater reliability refers to the degree of consistency of scores among the five raters. Concurrent validity refers to the relationship between scores from the target measure (i.e., the figure-of-eight technique) and the "gold standard" validated tool (i.e., volumeter). The study was reviewed and approved by the Institutional Committee for Research Involving Human Participants of the Kuwait University.

Participants

Twenty-four healthy individuals between 18 and 23 years of age were recruited by an advertisement in the university campus to participate in the study. One volunteer was dropped from the study because she had a schedule conflict and had to leave in the middle of the study. Therefore, a total of 23 participants (7 males and 16 females) completed this study. The mean age of the volunteers was 20.19 with a standard deviation (SD) of 1.35 years. All of the participants were right-hand dominant. After the participants were briefed on the study protocol, they signed an informed consent before any data were collected. The participants completed a brief survey that included questions about their age, gender, hand dominance, and past medical conditions. Inclusion criteria required that there be no injury to the upper extremity in the past 6 months.

Five occupational therapy students were randomly selected from the 4th year (of a 5-year curriculum) to be the "raters." The raters mean (SD) age was 21.8 (0.9) years and grade point average (GPA) was 3.61 (0.23). The five raters had already completed an orthopaedics course during their 3rd year of school in which they learned how to perform oedema measurements using both techniques and were required to apply the methods in their clinical fieldwork placements.

Instrumentation

A standard 0.635-cm (0.25 in.) wide retractable tape was used to measure the figure-of-eight hand volume, and a standard commercial 500-mL graduated cylinder volumeter was used to measure the amount of water displaced.

Procedures

Before starting the data collection, the researcher reviewed the figure-of-eight method of measuring hand volume with the five raters to ensure consistent measurement procedures. The figure-of-eight measurements were performed according to the method described by Maihafer et al. (2003), as follows: The tip of the tape measure was positioned on the pisiform bone (start point) and was then drawn in an ulnar direction along the wrist passing over the tendon of flexor carpi ulnaris. The tape measure was then directed distally and obliquely across the dorsum of the hand passing over the midpoint of the second metacarpal and moving towards the proximal palmar crease, where it was then directed in an ulnar direction across the palmar surface with the tape measure aligned with the distal palmar crease of the fifth digit. The tape measure was then drawn back across the dorsum of the hand towards the radial wrist crease (creating the "eight shape" on the dorsum of the hand) where the distal end of the tape was re-aligned with the wrist crease and directed to the starting point (the pisiform bone). All the measurements were performed without any markings on the participants’ hands to guarantee rater independence in relocating the anatomical landmarks for each measurement.

For testing, the 23 participants were seated in a large circle and were labelled from 1 to 23. Each of the five raters performed a single figure-of-eight measurement on the 23 participants starting in order from participant #1 to #23. The same procedure was repeated two more times, and therefore each participant was measured three times by each of the five raters. The time required to complete one measurement on all the 23 participants was approximately
20 minutes. That is, each rater had a 20-minute gap between measurements of the same participant. The data were recorded on different forms to make sure that the raters were masked to their previous measurements.

Two of the raters (randomly chosen from the five initial raters) jointly performed a single volumetric measurement on each participant using the standardized procedure. All the volumetric measurements were performed using the same volumeter.

Statistical analysis

All data analysis was done using the Statistical Package for the Social Sciences version 19.0 (IBM SPSS Statistics). Statistical significance level was set at \( p < .05 \). Intrarater reliability was calculated using the intraclass correlation coefficient (ICC, 3k), and inter-rater reliability was calculated using ICC, 2k. To establish concurrent validity, Pearson product-moment correlations were used to calculate the relationship between the mean figure-of-eight value for each rater and the volumeter values. Standard error of the mean (SEM), which determines absolute reliability, was calculated to describe the precision of the measurement.

Results

Descriptive statistics for figure-of-eight measurement trials are presented in Table 1. The intrarater reliability (ICC, 3k) for figure-of-eight measurements of 23 hands over the three trials ranged from .98 to .99, with SEM ranging from .17 to .32 cm.

The inter-rater reliability (ICC, 2k) among the five raters for all three measurement trials was .99 (SEM, .31 cm), with Trial 1 = .99 (SEM, .39 cm), Trial 2 = .99 (SEM, .37 cm), and Trial 3 = .99 (SEM, .18 cm).

Pearson correlation coefficients between the mean of each rater’s three figure-of-eight measurements and the volumetric measurements were .95 (\( p < .001 \); Rater 1); .93 (\( p < .001 \); Rater 2); .90 (\( p < .001 \); Rater 3); .92 (\( p < .001 \); Rater 4); and .96 (\( p < .001 \); Rater 5).

Discussion

The purpose of this study was to assess the intrarater and inter-rater reliability as well as concurrent validity of the figure-of-eight technique of measuring hand volume by 4th year occupational therapy students. The results of this study demonstrated that the students had excellent intrarater and inter-rater reliability using the figure-of-eight method, as well as excellent concurrent validity with the "gold standard" volumetric method of measuring hand volume.

The results of the student raters in our study yielded results comparable with therapist raters in the Pellecchia (2003) and Maihafer et al. (2003) studies, which also used similar figure-of-eight methods to measure hand volume in individuals with no hand injuries. Pellecchia (2003) reported high intratester and intertester reliability, with ICCs of .99 and .97, respectively, while Maihafer et al. (2003) reported ICCs of .99 for both intratester and intertester reliability. Both studies reported Pearson correlation coefficients between figure-of-eight and volumetric measurements to be at least .94, and considered these findings as indicative of high concurrent validity.

The excellent intrarater reliability for the figure-of-eight technique means that a student can be confident that changes from initial to re-evaluation measurements indicate actual changes in hand volume (oedema) size, and not measurement errors, given that the measurements were performed using the standardized procedure described in the literature. The figure-of-eight technique was also shown to have high inter-rater reliability. This means that the initial and follow-up measurements can be conducted by different therapists and still produce comparable results. The high concurrent validity for the figure-of-eight technique means that the measurements obtained by the

| Table 1 Descriptive Statistics (cm) and Intrarater Reliability for Figure-of-eight Repeated Trials with 23 Hands. |
|---------------------------------|-----------------|-----------------|-----------------|-----------------|
| Raters | Figure-of-eight trials | Range for 23 participants | Mean | SD | Intrarater reliability ICC, 3k (SEM) |
|--------|-----------------|-----------------|-----------------|-----------------|
| Rater 1 | Trial 1 | 33.5–43.5 | 38.45 | 2.78 | .99 (.27) |
| Rater 1 | Trial 2 | 33.5–44.0 | 38.46 | 2.89 | .99 (.25) |
| Rater 1 | Trial 3 | 34.0–44.0 | 38.52 | 2.77 | .99 (.18) |
| Rater 2 | Trial 1 | 34.0–43.4 | 38.80 | 2.81 | .99 (.31) |
| Rater 2 | Trial 2 | 34.0–42.5 | 38.33 | 2.33 | .99 (.37) |
| Rater 2 | Trial 3 | 34.0–43.0 | 38.28 | 2.43 | .99 (.25) |
| Rater 3 | Trial 1 | 33.5–43.5 | 38.69 | 2.68 | .99 (.31) |
| Rater 3 | Trial 2 | 33.5–43.0 | 37.79 | 2.74 | .99 (.25) |
| Rater 3 | Trial 3 | 33.5–43.0 | 37.95 | 2.55 | .99 (.17) |
| Rater 4 | Trial 1 | 33.5–43.0 | 38.06 | 2.62 | .99 (.32) |
| Rater 4 | Trial 2 | 34.0–43.5 | 38.19 | 2.45 | .99 (.25) |
| Rater 4 | Trial 3 | 34.0–43.0 | 38.04 | 2.45 | .99 (.17) |
| Rater 5 | Trial 1 | 34.5–42.5 | 38.29 | 2.02 | .99 (.32) |
| Rater 5 | Trial 2 | 34.0–43.8 | 38.56 | 2.58 | .99 (.25) |
| Rater 5 | Trial 3 | 34.0–43.5 | 38.85 | 2.56 | .99 (.25) |

SD = standard deviation; SEM = standard error of the mean.
The accuracy of measuring hand volume may not be hand pathology or oedema for the past 6 months. Although the raters in this study had no history of future studies by stratifying students according to GPA. In addition, the participants in this study had no history of hand pathology or oedema for the past 6 months. Although the student raters in this study were randomly selected from the whole cohort of students in the 4th year class, their GPAs were relatively high (mean = 3.61, SD = 0.23) in comparison with other students. These students may not be representative of all occupational therapy students and students with lower GPAs may not perform as well as the raters in this study. This could be avoided in future studies by stratifying students according to GPA. In addition, the participants in this study had no history of hand pathology or oedema for the past 6 months. Although the accuracy of measuring hand volume may not be affected by lack of pathology, the implications of this study may be more clinically relevant if the participants included actual patients undergoing hand rehabilitation.

Implications for occupational therapy practice

Based on the excellent intrarater and inter-rater reliability and concurrent validity (relative to the volumeter), the students’ skills met or exceeded the skills needed to validly assess oedema using the figure-of-eight technique. Moreover, their individual performance was stable over time (intrarater reliability), and the excellent inter-rater reliability indicates that their skills are interchangeable (i.e., if on fieldwork Student 1 assesses a patient at admission and Student 2 re-assesses the same patient at discharge, any change would not be due to variance among assessors). Our findings also indicate that 4th year students, who assist 3rd year students in measurement labs, with a short refresher demonstration can be valid and reliable assistants. In addition, 4th year students can achieve clinical levels of reliability to be data collectors on research projects requiring use of either the figure-of-eight or the volumeter.

Compared with volumetry, the figure-of-eight method is more practical for clinical use, and students entering clinical fieldwork can cite the results of this study as evidence in those clinics that only use volumetry. Because the figure-of-eight procedure is simple, takes less time to perform, requires equipment that is cost effective and readily available in most clinical settings, students and clinicians can use it with more patient populations. Moreover, clinical fieldwork students can become as reliable as experienced therapists with minimal training.

Conclusion

The 4th year students in a 5-year baccalaureate program in Kuwait were able to establish excellent intrarater and inter-rater reliability using the figure-of-eight method to measure hand volume. Concurrent validity of the figure-of-eight measures with the gold standard Volumeter measures was also excellent. The simple, cost-effective figure-of-eight method of measuring hand volume can be validly and reliably mastered by students at a level that matches experienced therapists.

References


Student rater reliability with figure-of-eight technique