

## Reliability and Feasibility of Methods to Quantitatively Assess Peripheral Edema

Kimberly G. Brodovicz, DrPH; Kristin McNaughton, MHS; Naoto Uemura, MD;  
Gary Meiningner, MD; Cynthia J. Girman, DrPH; and Steven H. Yale, MD

**Objective:** To evaluate methods to assess peripheral edema for reliability, feasibility and correlation with the classic clinical assessment of pitting edema.

**Design:** Cross-sectional observational study.

**Setting:** Large primary care clinic in Marshfield, Wisconsin, USA.

**Participants:** Convenience sample of 20 patients with type 2 diabetes and a range of edema severity, including patients without edema.

**Methods:** Eight methods of edema assessment were evaluated: (1) clinical assessment of pit depth and recovery at three locations, (2) patient questionnaire, (3) ankle circumference, (4) figure-of-eight (ankle circumference using eight ankle/foot landmarks), (5) edema tester (plastic card with holes of varying size pressed to the ankle with a blood pressure cuff), (6) modified edema tester (edema tester with bumps), (7) indirect leg volume (by series of ankle/leg circumferences), and (8) foot/ankle volumetry by water displacement. Patients were evaluated independently by three nurse examiners.

**Results:** Water displacement and ankle circumference had high inter-examiner agreement (intraclass correlation coefficient 0.93, 0.96 right; 0.97, 0.97 left). Agreement was inconsistent for figure-of-eight (0.64, 0.86), moderate for indirect leg volume (0.53, 0.66), and low for clinical assessments at all locations. Agreement was low for the edema testers but varied by the pressure administered. Correlation with the classic, subjective clinical assessment was good for the nurse-performed assessments and patient questionnaire. Ankle circumference and patient questionnaires each took 1 minute to complete. Other tools took >5 minutes to complete.

**Conclusions:** Water displacement and ankle circumference showed excellent reliability; however, water displacement is a time-consuming measure and may pose implementation challenges in the clinical and clinical trial environments. Patient-reported level and frequency of edema, based on an unvalidated questionnaire, was generally well correlated with the physician assessment of edema severity and may prove to be another reliable and accurate method of assessing edema. Additional study is needed to evaluate the validity and responsiveness of these methods.

**Keywords:** Measurement; Peripheral edema; Reliability

Peripheral edema is a common clinical sign in a wide variety of conditions affecting the cardiovascular, renal, and hepatic systems, as well as in inflammatory and metabolic diseases, malnutrition, pregnancy, and as a post-surgical complication.<sup>1,2</sup> Additionally, edema has been reported as an adverse reaction to drugs.<sup>3-9</sup> Drug-induced edema typically is confined to the lower extremities, occurs within weeks after drug initiation, and is generally dose dependent.<sup>1,7-9</sup> Patients with type 2 diabetes mellitus (T2DM), for example, not only

have a higher prevalence of peripheral edema than healthy subjects, but use of certain anti-hyperglycemic medications, especially in combination with insulin, is associated with an increased incidence of edema.<sup>7,10,11</sup> Edema is often an early sign of significant fluid retention, which could eventually result in significant cardiac overload and conditions such as heart failure. Although many physicians are aware of the importance of monitoring edema, accuracy and consistency in assessing edema has been a challenge.

**Reprint Requests:** Kimberly G. Brodovicz, DrPH, Department of Epidemiology, Merck Research Laboratories, PO Box 1000, UG1D-60, North Wales, PA 19454, Tel: 267-305-7490, Fax: 215-616-2042, Email: kimberly\_brodovicz@merck.com

**Research Support:** Fully supported by Merck & Co., Inc. SH Yale has received grant support from Merck & Co., Inc. and has been a consultant to and a clinical investigator for the company.

Received: August 27, 2008  
Revised: December 2, 2008  
Accepted: December 17, 2008

doi: 10.3121/cmr.2009.819

A variety of methods to quantitatively measure peripheral edema have been proposed (e.g., volumetry and ankle circumference) but their use has generally been limited to physical therapy and sports medicine settings.<sup>12-16</sup> The most widely-used technique is a subjective clinical assessment where an examiner applies pressure with his/her index finger to a single location on the patient's ankle.<sup>17,18</sup> This technique captures pit depth and the time needed for the skin to return to its original appearance (recovery time) as a single edema score ranging from 0 to 4.<sup>19</sup> While this classic method is commonly used, it has not been proven to be a sufficiently objective, reliable, or sensitive assessment of edema. In clinical studies, peripheral edema, as a pharmacologically-mediated adverse event, has typically been captured through water-displacement volumetry, patient self-report, or other, often non-validated, investigator assessments.<sup>20-25</sup> Water-displacement volumetry is widely considered to be the reference method of edema assessment;<sup>26</sup> however, it is also considered to be time-consuming, difficult to perform, and inappropriate in some clinical situations (e.g., postoperative), limiting the use of this method.<sup>12,27</sup> Documentation of the accuracy and reliability of other methods of measuring peripheral edema vary. While the classic, subjective clinical assessment of edema described above is often sufficient in routine clinical practice, more rigorous objective assessment of peripheral edema is required in research settings and when careful evaluation of edema is needed as part of clinical care.

The purpose of this study was to examine the reliability, feasibility, and correlation with the classic clinical assessment of eight methods of evaluating peripheral edema: (1) separate clinical assessments of pit depth and recovery at three locations (lower calf above the medial malleolus, behind the medial malleolus, dorsum of the foot), (2) patient questionnaire, (3) ankle circumference, (4) figure-of-eight (ankle circumference using eight ankle/foot landmarks), (5) edema tester (plastic card with holes of varying sizes pressed to the ankle while a blood pressure cuff applies varying pressures), (6) modified edema tester (edema tester with bumps instead of holes), (7) indirect leg volume (by series of ankle/leg circumferences), and (8) water-displacement volumetry.

## METHODS

### Patient Population

Ambulatory, non-hospitalized patients, aged 21 years or older, with a diagnosis of T2DM were recruited. This study was limited to patients with T2DM given the increased incidence of peripheral edema in this population. Patients with a range of edema severity, including patients without edema, were targeted for inclusion to determine if the methods could distinguish both the presence or absence of edema and between levels of edema severity. Presence and severity of edema were initially determined by physician assessment based upon the classic, subjective clinical assessment.

Patients were excluded if any of the following criteria were met: current hospitalization, presence of superficial skin

ulcers, open sores, wounds, or other skin conditions on the lower extremity, or history of an ankle injury or lower extremity surgery within the past 30 days.

This study was approved by the Marshfield Clinic Research Foundation Institutional Review Board (IRB). Written, informed consent was obtained from all study participants after full explanation of the study and before enrollment.

### Study Design

The first 20 patients who met eligibility criteria were enrolled in the study. Three senior research nurses with extensive clinical trial and research experience served as examiners in this cross-sectional, observational study. Examiners were trained on all methods prior to study start. Each method was explained and demonstrated to the examiners followed by a short practice session where the examiners were able to gain hands-on experience with the methods. Each patient's demographic information and list of current and prior (within the last 7 days) medications were recorded along with a brief medical history, including conditions that increase the risk of edema and any procedures performed to evaluate the etiology of the edema. Height, weight, and blood pressure were obtained for all patients and a simplified physical examination was performed including assessment for ankle/leg injuries or skin conditions that may exclude the patient from participation. The study physician classified the patients into edema categories (none, mild, moderate, severe) based upon the classic subjective clinical assessment and any edema present as lipedema, lymphedema, or venous stasis.<sup>1</sup> Lipedema is an entity distinguished clinically from lymphedema by the presence of soft symmetrical tissue enlargement of the lower extremities with sparing of the foot, ankle, and upper extremities.<sup>1</sup>

Each patient was evaluated by all three nurse examiners. All evaluations for a given patient were completed on the same day. Two to three methods were performed together to ensure any temporal changes in edema severity would not affect inter-examiner agreement. All examiners would complete one group of assessments before the next group was performed. Examiners were blinded to the other examiners' evaluations and results.

Evaluations were performed on both the left and right ankles/legs. Measures for ankle circumference and figure-of-eight were obtained in duplicate to assess intra-examiner reliability. One patient was evaluated with all eight methods. To reduce the duration of each visit, subsequent patients were evaluated with the clinical assessments of pit depth and recovery, patient questionnaire, ankle circumference, water displacement, and the 6 mm modified edema tester at 150 mmHg and randomized to either figure-of-eight and indirect volume or to the original and modified edema testers.

The feasibility of these methods was assessed by analyzing the amount of time needed to complete each assessment, the

magnitude of measurement error between examiners for each method, and feedback from patients and examiners. Examiners recorded the start and stop times of each evaluation. The magnitude of measurement error between examiners was estimated using the average absolute difference between the examiners. Patient and examiner feedback was obtained through standardized, self-administered questionnaires. Examiners were asked to provide their feedback on each method regarding perceived clinical relevance, ease of use, and potential for standardization.

### Assessment Tools

With the exception of the clinical assessment and the patient questionnaire, the tools evaluated are designed to quantify edema through serial measurements (i.e., change from baseline).

#### Clinical Assessment

The classic clinical assessment described by Seidel et al<sup>19</sup> is an unvalidated, subjective measure of edema. The clinical assessment evaluated in this study, an adaptation of this classic technique, was repeated at three anatomical locations (the lower calf at 7 cm proximal to the midpoint of the medial malleolus, behind the medial malleolus, and the dorsum of the foot). The scoring definitions were modified to capture pit depth and recovery separately. Pit depth was estimated visually and scored as follows: 0=no clinical edema, 1=slight pitting (2 mm depth) with no visible distortion, 2=somewhat deeper pit (4 mm) with no readily detectable distortion, 3=noticeably deep pit (6 mm) with the dependent extremity full and swollen, and 4=very deep pit (8 mm) with the dependent extremity grossly distorted. Recovery time was recorded in seconds.

#### Patient Questionnaire

Patient-reported edema was collected using a standardized, unvalidated questionnaire. The questionnaire was examiner-administered and included five questions to assess the presence and severity of self-reported edema over the past week and one question to assess the ease of completing the five questions. Only those patients responding they had experienced edema in the last week were asked about edema severity and frequency.

#### Ankle Circumference

Ankle circumference was measured in centimeters at a single location as described by Mora et al.<sup>15</sup> For consistent ankle circumference measurements, each ankle was marked with a semi-permanent marker at approximately 7 cm proximal to the midpoint of the medial malleolus. Unlike the method outlined by Mora et al,<sup>15</sup> a tension-controlled measuring tape (Gulick II, Lafayette Instrument Company, Lafayette, IN), rather than a standard measuring tape, was used to minimize measurement error due to differences in the amount of tension applied.

#### Figure-of-Eight

The figure-of-eight method uses eight landmarks on the ankle and foot to measure ankle circumference in centimeters:

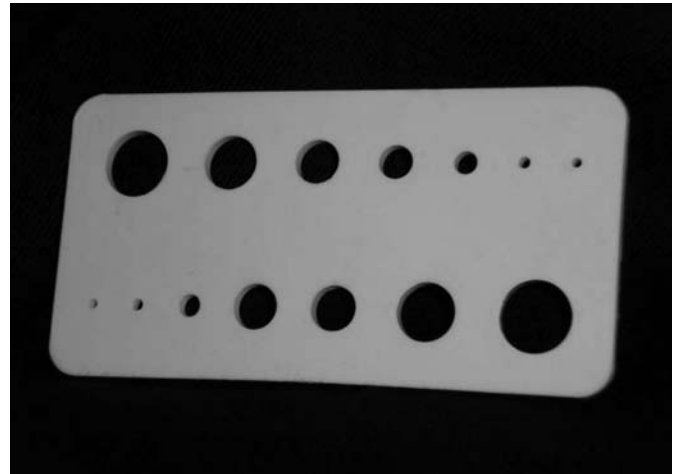


Figure 1. Edema tester.

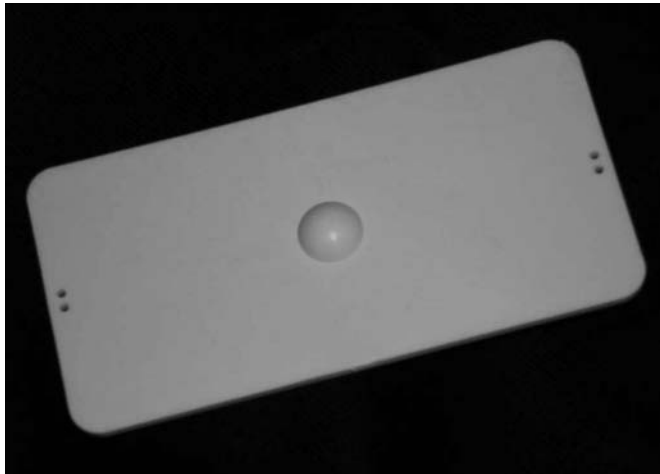
(1) midway between the tibialis anterior tendon and the lateral malleolus, (2) distal to the tuberosity of the navicular, (3) proximal to the base of the 5<sup>th</sup> metatarsal, (4) tibialis anterior tendon, (5) distal to the distal tip of the medial malleolus, (6) Achilles tendon, (7) distal to the distal tip of the lateral malleolus, and (8) back to origin.<sup>14,16,26,28</sup> Unlike the previously published methods, a tension-controlled measuring tape, rather than a standard measuring tape, was wrapped around the ankle/foot following the eight landmarks.

#### Edema Tester

The edema tester is a 5 cm x 10 cm flexible plastic plate with two parallel rows of seven holes ranging from 2 mm to 12 mm in diameter arranged by size, with one row of holes increasing in size and the other line decreasing in size (figure 1) (ACI Medical, San Marcos CA, USA).<sup>29</sup> The edema tester was placed with the long axis in the vertical plane and placed superior to the medial malleolus. A standard blood pressure cuff was placed over the tester and inflated to a standard pressure. No standard guidelines exist for the optimal amount or duration of pressure applied. In the original article, a pressure of 50 mmHg was held for 2 minutes.<sup>29</sup> Several thresholds of pressure and duration were evaluated for use in this study. Applied pressures of 100 mmHg and 150 mmHg, held for 3 seconds, were used. The cuff and edema tester were then removed. The numbers of impressions left by the holes were counted and the time for the last impression to disappear was measured in seconds. A large number of impressions and a long recovery time indicate more severe edema.

#### Modified Edema Tester

The unvalidated modified edema tester is similar to the edema tester except the holes were replaced by bumps in an attempt to mimic the standard clinical assessment (figure 2) (ACI Medical). Several modified edema testers were evaluated for inclusion in this study. The versions included in this study had a single bump of 4 mm or 6 mm in height. As with the original edema tester, the modified edema testers were used with a blood pressure cuff. Pressures of 100 mmHg and



**Figure 2.** Modified edema tester.

150 mmHg were held for 3 seconds, and the recovery time in seconds measured.

#### *Indirect Measure of Leg Volume*

Leg volume was calculated indirectly using the disk model.<sup>12,13</sup> A series of ankle and leg circumferences was used to calculate the volume of each cross-section in millimeters. The sum of the disk volumes provides an estimate of total leg volume. For consistent measurements, each extremity was marked with a semi-permanent marker at 3 cm intervals beginning 3 cm proximal to the midpoint of the medial malleolus to the tibial medial condyle. Unlike the method outlined by Latchford and Casley-Smith<sup>13</sup> and Kaulesar Sukul et al,<sup>12</sup> a tension-controlled measuring tape, instead of a standard measuring tape, was used to minimize measurement error.

#### *Water Displacement*

Foot/ankle volume was measured through water displacement using a commercially available foot volumeter (Baseline, Fabrication Enterprises Inc., White Plains, NY). The volumeter, a clear acrylic rectangular box (13"x5"x9") with a spout at the top of one of the short sides was filled with water until water rushed out of the spout (figure 3). Once the water level was stable, the patient placed one foot in the volumeter and the displaced water collected and measured in a graduated cylinder. The amount of water displaced in milliliters equals the volume of the foot/ankle.

#### **Statistical Methods**

Inter-examiner reliability was assessed with graphical techniques (e.g., scatterplots of examiner 1 versus examiner 2), Kappa statistics (categorical), and intraclass correlation coefficients (ICC) with 95% confidence intervals (CI) for continuous measures. ICC was calculated based on three examiners assessing each patient and represents the expected reliability of a single examiner's rating (referred to as case 2.1 in Shrout and Fleiss 1979).<sup>30</sup> Intra-examiner test-retest reliability was assessed by ICC (case 1.1).<sup>30</sup> The first of the replicate measurements is presented and was used in inter-examiner

reliability calculations. ICCs of 1 represent perfect agreement, ICCs of at least 0.75 are considered favorable.<sup>31</sup> Although precision was limited with only 20 patients evaluated, it was estimated that if good agreement exists between examiners, 20 patients would be adequate to allow estimation of ICC within  $\pm 0.2$  (half-width for 90% CI). For measures assessed in fewer patients (figure-of-eight and indirect volume  $n=12$ ; original and modified edema testers  $n=9$ ), estimation of ICC was within  $\pm 0.3$  and  $\pm 0.4$ . Correlations between continuous metrics were calculated using Pearson correlation coefficients. Cross-tabulations were used to describe categorical endpoints. Two elements of validity were assessed, convergent validity (i.e., tools expected to be related to each other are related) and measurement sensitivity (i.e., ability to discriminate between different levels of edema severity). Correlations between the initial physician classification of edema, nurse examiners' clinical assessment, and patient self-assessments (via the patient questionnaire) were assessed. Time to complete each method was calculated from the start and stop times recorded by the examiners for each evaluation.

## **RESULTS**

### **Patient Characteristics**

The mean age was 70 years (range 59 to 85 years) and 14 patients (70%) were women (table 1). Mean body mass index was 33.1 kg/m<sup>2</sup> (range 20.8 to 45.8 kg/m<sup>2</sup>). The initial physician assessment, based upon the classic clinical assessment, classified 10 patients as having mild edema, 3 as having moderate edema, 3 as having severe edema, and 4 as having no edema (table 1). Edema secondary to venous stasis was the most common edema type, occurring in 50% of all patients with edema. Severity of edema was positively related to the length of time since initial T2DM diagnosis. The median time since initial T2DM diagnosis for patients with no edema was 6 years, compared to 12 and 21 years for patients with moderate and severe edema, respectively.



**Figure 3.** Baseline volumeter measuring device, foot set, 5x13x6 inches (reproduced with permission from WisdomKing.com).

**Table 1.** Demographic and clinical characteristics (n=20).

	Number of patients (%)
Severity of edema*	
None	4 (20) <sup>†</sup>
Mild	10 (50)
Moderate	3 (15)
Severe	3 (15)
Type of edema	
Venous stasis	10 (50)
Lipedema <sup>‡</sup>	4 (20)
Lymphedema	1 (5)
Two or more	2 (10)
None	3 (15)
Mean age, years (range)	70.2 (59 to 85)
Gender, women	14 (70)
Race, white	20 (100)
Mean weight, kg (range)	91.8 (59.4 to 133.4)
Mean body mass index (kg/m <sup>2</sup> )	33.1 (20.8 to 45.8)
Concurrent conditions	
Hypertension	17 (85)
Dyslipidemia/hyperlipidemia	17 (85)
Coronary artery disease	6 (30)
Chronic heart failure	4 (20)
Dyspnea	2 (10)
Venous insufficiency	1 (5)
Renal insufficiency	4 (20)
Arthritis	6 (30)
Concomitant medications <sup>§</sup>	
Anti-hypertensive medication, excluding diuretics	18 (90)
Diuretics	14 (70)
Lipid-altering medication	9 (45)
Non-steroidal anti-inflammatory agents	2 (10)
Thiazolidinediones	1 (5)
Insulin	7 (35)

\* Initial classification of edema was determined by a physician and based upon the classic clinical assessment of edema. Efforts were made to recruit patients with a range of edema severity and patients with no edema.

<sup>†</sup> One patient was classified as having a medical history of edema consistent with lipedema but no edema was present at the time of the study.

<sup>‡</sup> Lipedema is a entity distinguished clinically from lymphedema by the presence of soft symmetrical tissue enlargement of the lower extremities with sparing of the foot, ankle and upper extremities.[1]

<sup>§</sup> Medications taken on the day of the visit.

Hypertension and lipid disorders were the most common co-morbid conditions with 17 patients (85%) having each condition (table 1). Other common conditions included coronary artery disease (30%), heart failure (20%), and renal insufficiency/failure (20%). The most commonly used medications were anti-hypertensives not including diuretics (90%), diuretics (70%), and lipid-altering medications (45%) (table 1).

The clinical assessment, patient questionnaire, ankle circumference, 6 mm modified edema tester at 150 mmHg, and water displacement were completed in all 20 patients; figure-of-eight and indirect leg volume were completed in 12 patients; and the original edema tester and the 4 mm

(100 mmHg and 150 mmHg) and 6 mm (100 mmHg) modified edema testers were completed in 9 patients. Examiner 3 administered the questionnaire to 19 patients; however, all 20 patients completed the questionnaire at least once.

#### Inter-Examiner Reliability

Water displacement and ankle circumference were the only tools that showed high inter-examiner agreement, with ICCs over 0.90 and tight 95% CIs for ICC (table 2). Agreement for figure-of-eight was inconsistent, with the agreement for the left ankle (ICC=0.86) much better than for the right ankle (ICC=0.64). Agreement for indirect leg volume was moderate (ICC=0.53 right; 0.66 left) and agreement for the recovery times of the clinical assessment was low for all locations.

Agreement in the pit depth score of the clinical assessment was variable with weighted kappa statistics ranging from low to moderate: lower calf: range 0.34 to 0.61, behind the medial malleolus: -0.01 to 0.68, and dorsum of the foot: 0.17 to 0.46. Agreement in recovery times was also generally low for both the original and modified edema testers but varied by the pressure administered. The mean absolute difference between the examiners for the number of impressions left by the original edema tester ranged from 0.5 to 2.7 at 100 mmHg and 0.1 to 1.4 at 150 mmHg.

#### Magnitude of Measurement Error Between Examiners

The range in average absolute differences between examiners was generally wide and not necessarily related to the method's reliability (table 2). Water displacement was a highly reliable method, however, the range of average absolute differences was 5.1 mL to 71.9 mL. The range of average absolute differences for both figure-of-eight, a less reliable method, and ankle circumference, a highly reliable method, was <1 cm.

#### Patient Questionnaire

Half the patients (n=10) stated they did not have any swelling in their ankles over the prior week (table 3). Nine of the remaining 10 patients stated they experienced swelling in their ankles every day over the prior week; one patient experienced ankle swelling on only 1 day. About 40% of the patients reported mild or moderate swelling but were not at all bothered by this swelling. Over 90% of the patients stated that the edema did not limit their normal activity. The level and bothersomeness of the swelling, as reported by patients, differed slightly between examiners. Even though each examiner asked these questions within 15 minutes of each other, patients answered these two questions differently. However, three patients did comment that they had difficulty choosing between the "moderate" and "extreme" response categories.

#### Intra-Examiner Reliability

Measures for ankle circumference and figure-of-eight were obtained in duplicate. The intra-examiner reliability for ankle circumference was very high with individual ICCs for the

**Table 2.** Inter-examiner agreement by interclass correlation coefficient (ICC).

Edema assessment tool	ICC* (95% confidence interval)		Median recovery time (maximum)	Average absolute difference (range)
	Right	Left		
Water displacement (n=20)	0.93 (0.88, 0.97)	0.96 (0.93, 0.98)	-	5.1-71.9 mL†
Ankle circumference (n=20)	0.97 (0.94, 0.98)	0.97 (0.95, 0.99)	-	0.06-0.2 cm
Figure of eight (n=12)	0.64 (0.41, 0.85)	0.86 (0.75, 0.95)	-	0.08-0.63 cm
Clinical assessment – recovery (n=20)				
Lower calf	0.53 (0.34, 0.73)	0.03 (-0.12, 0.29)	1 sec (3 min)	1-21.6 sec
Behind medial malleolus	0.28 (0.07, 0.54)	-0.01 (-0.16, 0.25)	1 sec (4.5 min)	2.7-26.1 sec
Dorsum of foot	0.05 (-0.13, 0.33)	0.11 (-0.08, 0.39)	1 sec (61 sec)	0.3-2.2 sec
Indirect leg volume (n=12)	0.53 (0.28, 0.80)	0.66 (0.44, 0.86)	-	9.4-238.8 mL
Original edema tester – recovery				
100 mmHg (n=9)	0.48 (0.18, 0.81)	0.12 (-0.17, 0.60)	15.5 sec (4.3 min)	0.25-34.2 sec
150 mmHg (n=9)	0.75 (0.55, 0.92)	0.43 (0.15, 0.78)	57 sec (8.6 min)	9.5-83.2 sec
Modified edema tester – recovery				
4 mm, 100 mmHg (n=9)	0.65 (0.40, 0.88)	0.24 (-0.06, 0.68)	1.2 min (10 min)	16 sec-1 min 5 sec
4 mm, 150 mmHg (n=9)	0.37 (0.14, 0.72)	0.68 (0.44, 0.89)	1.6 min (20 min)	23 sec-3 min 43 sec
6 mm, 100 mmHg (n=9)	0.38 (0.15, 0.74)	0.72 (0.48, 0.91)	1.8 min (18 min)	27 sec-2 min 30 sec
6 mm, 150 mmHg (n=20)	0.40 (0.21, 0.63)	0.75 (0.62, 0.87)	3.8 min (31 min)	1 min 37 sec-5 min 28 sec

\* Interclass correlation coefficients of at least 0.75 are considered favorable.

† Left 14.7 to 71.9 mL; Right 5.1 to 52.9 mL.

**Table 3.** Patient reported edema by questionnaire.

		Average response (%)	Examiner (%)		
			1	2	3*
1. Over the last week, how often did you have swelling in your ankles?	Every day	44.0	9 (45)	9 (45)	8 (42.1)
	6 days	--	--	--	--
	5 days	--	--	--	--
	4 days	--	--	--	--
	3 days	--	--	--	--
	2 days	--	--	--	--
	1 day	5.1	1 (5)	1 (5)	1 (5.3)
	Not at all	50.9	10 (50)	10 (50)	10 (52.6)
2. What level of swelling did you experience?	Extreme swelling	14.1	--	2 (20)	2 (22.2)
	Quite a bit of swelling	20.4	3 (30)	2 (20)	1 (11.1)
	Moderate swelling	10.0	2 (20)	1 (10)	--
	Mild swelling	34.4	4 (40)	3 (30)	3 (33.3)
	Very little swelling	21.1	1 (10)	2 (20)	3 (33.3)
3. How much were you bothered by swelling in your ankles?	Extremely	7.0	1 (10)	1 (11.1)	--
	Quite a bit	3.7	--	--	1 (11.1)
	Moderately	13.7	3 (30)	--	1 (11.1)
	A little	32.2	3 (30)	3 (33.3)	3 (33.3)
	Not at all	43.3	3 (30)	5 (55.6)	4 (44.4)
4. During what times of day did you have swelling in your ankles?	Morning only	--	--	--	--
	Afternoon only	14.1	--	2 (20)	2 (22.2)
	Evening only	48.5	4 (40)	5 (50)	5 (55.6)
	Morning & afternoon	--	--	--	--
	Morning & evening	--	--	--	--
	Afternoon & evening	6.7	2 (20)	--	--
	Morning, afternoon, & evening	30.7	4 (40)	3 (30)	2 (22.2)
5. Did swelling in your ankles limit your normal activity?	Extremely	--	--	--	--
	Quite a bit	--	--	--	--
	Moderately	--	--	--	--
	A little	7.0	1 (10)	--	1 (11.1)
	Not at all	93.0	9 (90)	10 (100)	8 (88.9)
6. Did you have any difficulty answering these questions?	No	90	7 (70)	10 (100)	9 (100)
	Yes	10	3 (30)	--	--

\* Only 19 patients evaluated.

© Merck & Co., Inc. Whitehouse Station, NJ, USA

right ankle ranging from 0.99 to 1.0 and 0.98 to 1.0 for the left ankle. Generally, the intra-examiner reliability for figure-of-eight was good and was not influenced by the severity of edema. Individual ICCs for the figure-of-eight assessment were not calculated due to the low overall ICCs and the small number of observations for this randomized measure.

### Correlation Between Assessments

#### *Clinical Assessment: Depth Versus Recovery Time*

In general, pit depth was moderately correlated with recovery time, with mean correlations ranging from 0.42 to 0.71 (table 4). However, there were outliers in recovery time for all examiners and locations (table 2). In addition, the upper end of the range of the depth score (4=very deep pit, 8 mm) was rarely used.

*Clinical Assessment Versus Physician Classification of Edema*  
Pit depth was moderately associated with edema severity (mean correlations ranging from 0.36 to 0.59, table 4). There were a few instances when pit depth was recorded for a patient classified as not having edema; however, these depths were all equal to 1 (slight pitting, no visible distortion). There were also several instances when no pit depth was recorded for patients classified as having severe edema.

There was a weak correlation between recovery time and edema severity (mean correlations ranging from 0.23 to 0.50, table 4). There were outliers for recovery time for all examiners and locations (table 2). While the majority of these outliers were seen in patients classified as having moderate or

**Table 4.** Mean Pearson correlation coefficients between assessments of lower extremity edema.

	Recovery time*	Physician classification†
Pit depth, lower calf	0.53 (right) 0.67 (left)	0.59 (r) 0.54 (l)
Pit depth, medial malleolus	0.42 (r) 0.66 (l)	0.53 (r) 0.46 (l)
Pit depth, dorsum of foot	0.71 (r) 0.51 (l)	0.36 (r) 0.51 (l)
Recovery time, lower calf		0.31 (r) 0.50 (l)
Recovery time, medial malleolus		0.23 (r) 0.35 (l)
Recovery time, dorsum of foot		0.23 (r) 0.33 (l)
Patient-reported level of swelling‡		0.72
Patient-reported frequency of swelling‡		0.61

\* Pit depth and recovery time from clinical assessment.

† Physician classification of edema (none, mild, moderate, severe) by standard clinical assessment.

‡ By questionnaire

severe edema, there were a few instances where patients with mild edema had longer than average recovery times.

#### *Patient Questionnaire Versus Physician Classification of Edema*

The initial physician classification of edema (none, mild, moderate, severe by the classic clinical assessment) was compared to responses from the patient questionnaire related to the frequency and severity of edema. There was general agreement between the initial classification of edema severity and patient-reported level of swelling ( $r=0.72$ , table 4). All three patients identified as having mild edema by the initial physician assessment reported their level of swelling as either “very little swelling” or “mild swelling.” Those with moderate edema ( $n=3$ ) reported their level of swelling as either “mild” or “moderate.” Those with severe edema ( $n=3$ ) reported their level of swelling as either “moderate” or “quite a bit.”

Initial classifications of moderate or severe edema correlated reasonably well with patient-reported frequency of ankle swelling ( $r=0.61$ , table 4). However, the patient-reported frequency of swelling did not agree with edema severity for patients classified as having no or mild edema (e.g., 70% of the patients identified with mild edema reported they never have swelling in their ankles). While three of the four patients identified as having no edema reported they never have swelling, one of these four patients reported having ankle swelling every day.

#### **Measures of Feasibility**

##### *Time to Complete Assessments*

The average time required to complete each method ranged from 1 minute (patient questionnaire and ankle circumference) to nearly 10 minutes (6 mm modified edema tester at

150 mmHg). The clinical assessment and figure-of-eight were the only other tools that took <3 minutes to perform (2.3 and 2.7 minutes, respectively). Water displacement took the second longest time to perform at 8 minutes, excluding set-up time. With the exception of the modified edema tester, all three examiners took roughly the same amount of time to complete each assessment (data not shown). Examiner 2 took 4 to 5 minutes longer than the other examiners to complete the modified edema tester assessment.

##### *Patient Feedback*

The patients’ feedback was favorable with all methods rated as causing little or no discomfort. Three patients stated they felt a little discomfort during the assessment with the modified edema tester, and two patients experienced a little discomfort during the clinical assessment. There was generally no difficulty on the part of the patient in completing any of the assessments. Overall, 90% to 95% of patients stated they would be willing to have the assessments performed again. Only 78% of the patients randomized to receive the edema tester assessments stated they would be willing to have those particular assessments repeated.

##### *Examiner Feedback*

The examiners feedback varied. There was disagreement between the examiners in terms of the tools’ clinical relevance, with ratings from “not at all clinically meaningful” to “very clinically meaningful.” While there was more agreement between the examiners regarding the difficulty of using the methods, responses remained variable. In the overall ranking of preference, accounting for perceived ease of use, clinical relevance, and potential for standardization, the top three tools for examiner 1 were the modified edema tester,



clinical assessment, and ankle circumference. For examiner 2, water displacement was the top tool followed by the clinical assessment and patient assessment. The top tools for examiner 3 were ankle circumference followed by the edema tester and, in a four-way tie for third, clinical assessment, patient-reported assessment, modified edema tester, and indirect leg volume technique. All examiners reported that figure-of-eight, the edema tester, and the modified edema tester were specifically challenging to perform.

## DISCUSSION

This study assessed the inter-examiner and intra-examiner reliability and feasibility of various methods used to quantify peripheral edema. Water-displacement volumetry and ankle circumference performed better than the other methods assessed, showing excellent inter-examiner agreement.

The water displacement method has been shown to be a sensitive and reliable measure of lower extremity edema in patients with ankle injuries and with ICCs of 0.99.<sup>16,26,32</sup> The high reliability of water-displacement volumetry observed in the current study is similar to that seen previously.<sup>26</sup> However, measurements in the current study were more variable with a range of average absolute differences between examiners of 5.1 mL to 71.9 mL. Petersen reported average differences of 11 mL to 18 mL. The variability observed in the current study may be due to the position of the patients – all patients were standing for this assessment. Patients' ability to stabilize themselves on one foot while placing the other foot carefully into the volumeter may have had a large impact upon the amount of excess water lost, since patients' stability and strength may vary by leg, contributing to the differences observed between the legs.

Ankle circumference had nearly perfect inter-examiner and intra-examiner agreement in the current study; however, these high levels of agreement may be overestimates, as the mark made by examiner 1 at the location of the measurement was typically still faintly visible for examiners 2 and 3 despite attempts to remove the mark with alcohol swabs. While ankle circumference generally shows good reliability in both healthy volunteers and in patients,<sup>16,33</sup> the use of a tension-controlled tape and markings to standardize the location of the measurement in the current study likely improved the reliability of this method.

While the figure-of-eight method has been shown to be a reliable measure in both healthy volunteers and volunteers with recent ankle injury,<sup>14,16,26</sup> reliability was inconsistent in the current study, with the agreement for the left ankle (ICC=0.86) much better than for the right ankle (ICC=0.64). In studies of healthy volunteers<sup>14</sup> and in patients with recent ankle or lower leg injuries,<sup>16,26</sup> reliability, as measured by the ICC, was 0.99. Variations in these measures were <1 cm in these studies. While the measurement variation observed in the current study was similar to these previous results, reliability was much lower (0.64 on the right leg, 0.86 on the

left). The lower reliability observed in this study may be due to the experience of the examiners and/or the ease of performing the measure on one side versus another. The examiners in Tatro-Adams,<sup>14</sup> Petersen<sup>26</sup> and Mawdsley<sup>16</sup> were masters-level physical therapy students, who are likely more familiar with these types of methods than clinical research nurses. Examiners in the current study also consistently had more trouble performing the measure on the left side than on the right, likely related to the dominant handedness of the examiners. This phenomenon was not reported in Tatro-Adams, Petersen or Mawdsley. Measurements were limited to the left ankle in the Tatro-Adams study. In the Mawdsley study, measures were obtained on the injured ankle and averaged across patients (6 of the 15 patients had left side injuries); side of the injury was not reported in Petersen. While not a fully validated method, figure-of-eight has correlated well with volumetric measurement ( $r>0.90$ ).<sup>16,26</sup>

In addition, there are definite implementation challenges with both water displacement and figure-of-eight. Water displacement is a time-consuming method, can be awkward, and is unlikely to translate well to the clinic or clinical trial environment. Kaulesar Sukul et al<sup>12</sup> also found this measure to be time-consuming and impractical in certain clinical situations (e.g., immediate postoperative period). However, given the excellent inter-examiner agreement, water displacement may be beneficial in select clinical research settings. Examiners in the current study found figure-of-eight to be cumbersome and inconsistent, with inter-examiner agreement better for right than left ankles.

The patient questionnaire was easy to administer and there was general agreement between the patients' assessments and the initial physician classification of edema, even though there was some slight variation in the answers given to each examiner. However, this patient questionnaire is not a fully validated instrument. This study represents the first time patients were exposed to this questionnaire. Further development of this patient questionnaire is underway.

The clinical assessment, indirect leg volume, and the original and modified edema testers did not perform well, showing low to average inter-examiner agreement. In addition, both components of the clinical assessment, pit depth and recovery time, were only moderately associated with the initial physician classification of edema and with each other. Previous research, however, has shown the disk model method of calculating leg volume to be highly correlated with water-displacement volumetry in healthy volunteers,<sup>12</sup> and the original edema tester has been shown to have discriminant validity.<sup>29</sup>

While the sample size for this study was relatively small, especially for some methods, and the population limited to T2DM patients, this study was able to identify reliable methods to quantify edema as well as challenges to implementation of the methods reviewed. In addition to the implementation issues for water displacement and figure-of-eight described

above, there were also difficulties with other methods. For the indirect leg volume assessment, examiners often found it difficult to keep the tape measure level around the leg. During the edema tester assessments, examiners found it difficult to maintain a constant pressure for 3 seconds. Clearly visualizing the impressions left by these devices was also difficult. All the methods were also subject to learning bias. As the examiners gained experience with the methods, the performance of the method may have improved over the course of the study resulting in overestimation of the method's reliability. To minimize this learning bias, especially when using these methods in a clinical trial setting, extensive pre-study examiner training workshops, with certification, are essential. Also, obtaining measurements during a run-in or screening period would allow the examiners to become comfortable with the method(s) and would further minimize learning bias. Examiner training and experience with the methods is critical to the reliability of the method; therefore, the results observed in this study are limited to settings with trained examiners.

Due to the cross-sectional nature of this study, the responsiveness of the tools (i.e., their ability to detect change over time) could not be examined. Assessment of the tools' validity (i.e., whether they capture the true nature of patients' edema) was also limited. Since there is no reference standard for a cross-sectional, objective measurement of edema, assessing the validity of each method in this study was difficult. In addition, with the exception of the clinical assessment and the patient questionnaire, the methods evaluated are designed to quantify edema through serial measurements. However, two elements of validity were assessed, convergent validity (i.e., tools expected to be related to each other are related) and measurement sensitivity (i.e., ability to discriminate between different levels of edema severity).

## CONCLUSION

Water displacement and ankle circumference showed excellent inter-examiner reliability, but water displacement may pose implementation challenges in the clinical and clinical trial environment. Patient-reported level and frequency of edema, based on an unvalidated questionnaire, were generally well correlated with the physician assessment of edema severity and may prove to be another reliable and accurate method of assessing edema. The low inter-examiner agreement and high variability for many of these methods limit their usefulness in both clinical research and practice. High variability within and between examiners could lead to misleading results when used to assess changes in edema presence or severity over time. Even in large trials where less precision can be balanced by a large sample size, highly variable measures may lead to uninterpretable or misleading results, and the inability to detect true changes that may have occurred. Additional study is needed to evaluate the validity and responsiveness of these methods.

## ACKNOWLEDGMENTS

Presented in part at the Society of General Internal Medicine 29<sup>th</sup> Annual Meeting, April 2006, Los Angeles, CA. The authors thank Jane Carl and Debra Krieg for their help in conducting this study and Marshfield Clinic Research Foundation for its support through the service of Alice Stargardt in the preparation of this manuscript.

## REFERENCES

1. Yale SH, Mazza JJ. Approach to diagnosing lower extremity edema. *Compr Ther* 2001;27:242-252.
2. Ely JW, Osheroff JA, Chambliss ML, Ebell MH. Approach to leg edema of unclear etiology. *J Am Board Fam Med* 2006;19:148-160.
3. Bell DS. Unilateral edema due to a thiazolidinedione. *Diabetes Care* 2003;26:2700.
4. Rozzini L, Ghianda D, Vicini Chilovi B, Padovani A, Trabucchi M. Peripheral oedema related to quetiapine therapy: a case report. *Drugs Aging* 2005;22:183-184.
5. Azzoli CG, Miller VA, Ng KK, Krug LM, Hensley ML, O'Reilly EM, Muller RJ, Kris MG. Gemcitabine-induced peripheral edema: report on 15 cases and review of the literature. *Am J Clin Oncol* 2003;26:247-251.
6. Talbert RL, Bussey HI. Update on calcium-channel blocking agents. *Clin Pharm* 1983;2:403-416.
7. Mudaliar S, Chang AR, Henry RR. Thiazolidinediones, peripheral edema, and type 2 diabetes: incidence, pathophysiology, and clinical implications. *Endocr Pract* 2003;9:406-416.
8. Messerli FH. Vasodilatory edema: a common side effect of antihypertensive therapy. *Am J Hypertens* 2001;14:978-979.
9. Messerli FH. Calcium antagonists in hypertension: from hemodynamics to outcomes. *Am J Hypertens* 2002;15:94S-97S.
10. Uemura N, Abbi S, Denker AE, Wright H, Mistry G, Moreau A, Bowden C, Kipnes MS, Ahmann A, Lunde NN, Meininger G, Thach C, Brodovicz K, Gottesdiener K, Wagner J. Time course of body composition changes following pioglitazone in patients with type 2 diabetes: increase in body fluid is detected earlier (at 2-4 weeks) than fat mass (at 4-8 weeks). 66<sup>th</sup> ADA Annual Meeting; 2006 [Abstract 573-P].
11. Kalambokis GN, Tsatsoulis AA, Tsianos EV. The edematogenic properties of insulin. *Am J Kidney Dis* 2004;44:575-590.
12. Kaulsar Sukul DM, den Hoed PT, Johannes EJ, van Dolder R, Benda E. Direct and indirect methods for the quantification of leg volume: comparison between water displacement volumetry, the disk model method and the frustum sign model method, using the correlation coefficient and the limits of agreement. *J Biomed Eng* 1993;15:477-480.
13. Latchford S, Casley-Smith JR. Estimating limb volumes and alterations in peripheral edema from circumferences measured at different intervals. *Lymphology* 1997;30:161-164.
14. Tatro-Adams D, McGann SF, Carbone W. Reliability of the figure-of-eight method of ankle measurement. *J Orthop Sports Phys Ther* 1995;22:161-163.
15. Mora S, Zalavras CG, Wang L, Thordarson DB. The role of pulsatile cold compression in edema resolution following ankle fractures: a randomized clinical trial. *Foot Ankle Int* 2002;23:999-1002.
16. Mawdsley RH, Hoy DK, Erwin PM. Criterion-related validity of the figure-of-eight method of measuring ankle edema. *J Orthop Sports Phys Ther* 2000;30:149-153.
17. Raju S, Owen S Jr, Neglen P. The clinical impact of iliac venous stents in the management of chronic venous insufficiency. *J Vasc Surg* 2002;35:8-15.

18. Malacco E, Vari N, Capua no V, Spagnuolo V, Borgnino C, Palatini P; Val-Syst Study. A randomized, double-blind, active-controlled, parallel-group comparison of valsartan and amlodipine in the treatment of isolated systolic hypertension in elderly patients: the Val-Syst study. *Clin Ther* 2003;25:2765-2780.
19. Seidel HM, Ball JW, Dains JE, Benedict GW. Heart and blood vessels. In: Schreffer S, ed. *Mosby's Guide to Physical Examination*, 3rd ed. St. Louis, MO: Mosby; 1995:419.
20. Guardia JA, Ortiz-Butcher C, Bourgoignie JJ. Oncotic pressure and edema formation in hypoalbuminemic HIV-infected patients with proteinuria. *Am J Kidney Dis* 1997;30:822-828.
21. Whyte KF, Douglas NJ. Peripheral edema in the sleep apnea/hypopnea syndrome. *Sleep* 1991;14:354-356.
22. Meler JD, Solomon MA, Steele JR, Yancy CW Jr, Parkey RW, Fleckenstein JL. The MR appearance of volume overload in the lower extremities. *J Comput Assist Tomogr* 1997;21:969-973.
23. Blankfield RP, Finkelhor RS, Alexander JJ, Flocke SA, Maiocco J, Goodwin M, Zyzanski SJ. Etiology and diagnosis of bilateral leg edema in primary care. *Am J Med* 1998;105:192-197.
24. Lund-Johansen P, Strandén E, Helberg S, Wessel-Aas T, Risberg K, Rønnevik PK, Istad H, Madsbu S. Quantification of leg oedema in postmenopausal hypertensive patients treated with lercanidipine or amlodipine. *J Hypertens* 2003;21:1003-1010.
25. Fogari R, Malamani G, Zoppi A, Mugellini A, Rinaldi A, Fogari E, Perrone T. Effect on the development of ankle edema of adding delapril to manidipine in patients with mild to moderate essential hypertension: a three-way crossover study. *Clin Ther* 2007;29:413-418.
26. Petersen EJ, Irish SM, Lyons CL, Miklaski SF, Bryan JM, Henderson NE, Masullo LN. Reliability of water volumetry and the figure of eight method on subjects with ankle joint swelling. *J Orthop Sports Phys Ther* 1999;29:609-615.
27. Strandén E. A comparison between surface measurements and water displacement volumetry for the quantification of leg edema. *J Oslo City Hosp* 1981;31:153-155.
28. Esterson PS. Measurement of ankle joint swelling using a figure of 8. *J Orthop Sports Phys Ther* 1979;1:51-52.
29. Cesarone MR, Belcaro G, Nicolaidis AN, Arkans E, Laurora G, De Sanctis MT, Incandela L. The edema tester in the evaluation of swollen limbs in venous and lymphatic disease. *Panminerva Med* 1999;41:10-14.
30. Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bulletin* 1979;86:420-427.
31. Deyo RA, Diehr P, Patrick DL. Reproducibility and responsiveness of health status measures. Statistics and strategies for evaluation. *Control Clin Trials* 1991;12:142S-158S.
32. Nilsson S, Haugen GB. Volumetry in the evaluation of swelling in the ankle and foot. *J Oslo City Hosp* 1981;31:11-15.
33. Liehr P, Todd B, Rossi M, Culligan M. Effect of venous support on edema and leg pain in patients after coronary artery bypass graft surgery. *Heart Lung* 1992;21:6-11.

## **AUTHOR AFFILIATIONS**

*Kimberly G. Brodovicz, DrPH*  
*Department of Epidemiology*  
*Merck Research Laboratories*  
*North Wales, Pennsylvania*

*Kristin McNaughton, MHS*  
*Department of Epidemiology*  
*Merck Research Laboratories*  
*North Wales, Pennsylvania*

*Naoto Uemura, MD*  
*Department of Clinical Research—Clinical Pharmacology*  
*Merck Research Laboratories*  
*Rahway, New Jersey*

*Gary Meininger, MD*  
*Department of Clinical Research—Metabolism*  
*Merck Research Laboratories*  
*Rahway, New Jersey*

*Cynthia J. Girman, DrPH*  
*Department of Epidemiology*  
*Merck Research Laboratories*  
*North Wales, Pennsylvania*

*Steven H. Yale, MD*  
*Department of Internal Medicine, Marshfield Clinic and*  
*Clinical Research Center, Marshfield Clinic Research*  
*Foundation*  
*Marshfield, Wisconsin*