Research Report

Reliability and Validity of Arm Volume Measurements for Assessment of Lymphedema

Background and Purpose. Arm lymphedema following breast cancer surgery is a continuing problem. In this study, we assessed the reliability and validity of circumferential measurements and water displacement for measuring upper-limb volume. **Subjects.** Participants included subjects who had had breast cancer surgery, including axillary dissection—19 with and 22 without a diagnosis of arm lymphedema—and 25 control subjects. **Methods.** Two raters measured each subject by using circumferential tape measurements at specified distances from the fingertips and in relation to anatomic landmarks and by using water displacement. Interrater reliability was calculated by analysis of variance and multilevel modeling. Volumes from circumferential measurements were compared with those from water displacement by use of means and correlation coefficients, respectively. The standard error of measurement, minimum detectable change (MDC), and limits of agreement (LOA) for volumes also were calculated. **Results.** Arm volumes obtained with these methods had high reliability. Compared with volumes from water displacement, volumes from circumferential measurements had high validity, although these volumes were slightly larger. Expected differences between subjects with and without clinical lymphedema following breast cancer were found. The MDC of volumes or the error associated with a single measure for data based on anatomic landmarks was lower than that based on distance from fingertips. The mean LOA with water displacement were lower for data based on anatomic landmarks than for data based on distance from fingertips. **Discussion and Conclusion.** Volumes calculated from anatomic landmarks are reliable, valid, and more accurate than those obtained from circumferential measurements based on distance from fingertips. [Taylor R, Jayasinghe UW, Koelmeyer L, et al. Reliability and validity of arm volume measurements for assessment of lymphedema. *Phys Ther.* 2006;86:205–214.]

Key Words: *Arm circumference, Arm volume, Axillary surgery, Breast cancer, Lymphedema, Reliability, Validity, Volume measurement.*

Richard Taylor, Upali W Jayasinghe, Louise Koelmeyer, Owen Ung, John Boyages

Lymphedema is the result of a functional overload of the lymphatic system in which lymph volume exceeds transport capabilities.[1] Arm lymphedema has long been recognized as a complication of treatment by surgery or radiotherapy, or both, of the axilla for breast cancer.[2,3] Approximately 15% to 20% of patients with breast cancer develop lymphedema following breast cancer treatment.[4] In the United States, approximately 400,000 people with breast cancer are living with lymphedema, and their quality of life may be affected by disfigurement, discomfort, and disability associated with arm and hand swelling.[5] Lymphedema may cause limb swelling, heaviness, pain, pitting of skin, tightness or hardness in the limb, inflammation, and reduced mobility in the shoulder. For many women, lymphedema may be one of several arm symptoms that adversely affect quality of life and functional status; other impairments may result from muscle, tendon, or ligamentous damage as a consequence of treatment. Delaying intervention in reducing lymphedema may result in poor quality of life and greater emotional distress.[4] Lymphedema is an important issue for women who have had breast cancer and is of concern to consumer groups.

Arm lymphedema needs to be measured quantitatively to aid in the assessment of severity at the time of diagnosis and remeasured to assess response to treatments that may be administered. Furthermore, accurate measurement of arm volume is needed in observational studies of arm lymphedema as a complication of local treatment for breast cancer and in research trials of prevention or treatment. Evaluation of the effectiveness of treatments for lymphedema requires an accurate, easy-to-use method for the calculation of arm volume. Circumferential measurements usually are made (at different distances from the fingertips) and compared with those for the other arm as a sum or average or as a computed volume of an arm segment.

The most widely accepted measure of lymphedema is limb volume compared with that of the unaffected limb or compared with that of the same limb before the interventions or events that led to lymphedema. Volumes are most accurately measured by water displacement,[4,6,7] although with limbs there are difficulties in defining and implementing the upper level for immersion, and water displacement is not convenient for routine clinical use. Megens et al[4] and Sander et al[7] found an intraclass correlation coefficient (ICC) of .99 for interrater reliability of water displacement volumes. However, many researchers choose not to use the water displacement method because it is time-consuming, is

R Taylor, MB, BS, DTMH, PhD, FRCP, FAFPHM, is Professor, School of Population Health, Faculty of Health Sciences, University of Queensland, Herston, Queensland, Australia, and Consultant Epidemiologist, New South Wales Breast Cancer Institute, University of Sydney, Westmead Hospital, Westmead, New South Wales, Australia.

UW Jayasinghe, MSc, GradDipStat, GradDipApplStat, PhD, is Consultant Statistician, New South Wales Breast Cancer Institute, University of Sydney, Westmead Hospital, Westmead, New South Wales, Australia.

L Koelmeyer, BApplSc(OT), is Occupational Therapist, Occupational Therapy Department, Westmead Hospital, Westmead, New South Wales, Australia.

O Ung, MB, BS, FRACS, is Clinical Associate Professor and Clinical Services Director, New South Wales Breast Cancer Institute, University of Sydney, Westmead Hospital, Westmead New South Wales, Australia.

J Boyages, MB, BS, PhD, FRACR, is Associate Professor and Executive Director, New South Wales Breast Cancer Institute, University of Sydney, Westmead Hospital, PO Box 143, Westmead, New South Wales, Australia 2145 (johnb@bci.org.au). Address all correspondence to Dr Boyages.

All authors provided concept/idea/research design. Dr Taylor, Dr Jayasinghe, and Ms Koelmeyer provided writing. Ms Koelmeyer provided data collection, and Dr Taylor and Dr Jayasinghe provided data analysis. Ms Koelmeyer and Dr Boyages provided subjects. Dr Boyages provided fund procurement and clerical support. Dr Jayasinghe, Ms Koelmeyer, and Dr Boyages provided consultation (including review of manuscript before submission). The authors acknowledge Gregory Heard, PhD, Research Manager, New South Wales Breast Cancer Institute, University of Sydney, Westmead Hospital, PO Box 143, Westmead, New South Wales, Australia, for contributing to project management and consultation. They also acknowledge the occupational therapists from the Occupational Therapy Department, Westmead Hospital, Westmead, New South Wales, Australia, who participated in the measurement study on which this article is based.

This study was approved by the Western Sydney Area Health Service Human Research Ethics Committee.

The New South Wales Breast Cancer Institute receives funding from the New South Wales Department of Health. Jobst Beiersdorf Australia Limited funded the participation of a staff occupational therapist in the measurement study on which this article is based.

This research, in part, was presented at the Australian Lymphoedema Association Conference, March 26–28, 2004, Brisbane, Queensland, Australia, and at the Leura V International Breast Cancer Conference, November 10–14, 2004, Sydney, New South Wales, Australia.

This article was received January 3, 2005, and was accepted August 1, 2005.

not portable, and can be nonhygienic.[4] Considering the difficulties related to the water displacement method, searching for alternative methods of volume determination remains a worthwhile pursuit. Volumes can be calculated from multiple circumferential measurements of the limb on the basis of the assumption of truncated conical segments (frustum). Intervals for circumferential tape measurements can be based on the distance from the middle fingertip or on the distance from or between anatomic landmarks.

Sander et al[7] compared volumetric measurements of an upper extremity obtained by water displacement with geometric measurements. They used geometric volume formulas for a cylinder, frustum (truncated cone), rectangular solid, and trapezoidal solid to compute volumes of the arm and hand at different measurement intervals and found that the frustum assumption produced the smallest standard error of measurement (SEM). Arm circumferences also can be measured with a Perometer,[8,*] which involves automated readings obtained at 0.4-cm intervals along the arm while the arm is positioned in a frame with a mobile source of infrared light; volumes are computed from the summation of elliptical segment volumes by use of specially designed software. Perometer measurements are reliable, convenient, and highly reproducible, with each measurement taking only a few seconds,[9] but the equipment is costly and not used widely. Limb volume also can be estimated from the anterior and lateral silhouettes with a device that uses beams of infrared light; however, this method is expensive, and its accuracy is uncertain.[10]

The purpose of this study was to determine whether volumes calculated from anatomic landmarks are reliable, valid, and more accurate than those obtained from circumferential measurements based on distance from fingertips. The accuracy of volume measurements of the upper limb computed from circumferential measurements based on distance from fingertips and anatomic landmarks was compared with that of measurements obtained from volume displacement (criterion validity). The hypothesis was that volumes calculated from circumferences related to anatomic landmarks would be more accurate than volumes calculated from circumferences at fixed distances from fingertips. The reasoning was that the latter method often involves a segment across the elbow joint (which is not very conical in shape) and, because the length of the arm differs from woman to woman, fixed distances from fingertips would be in different positions relative to the anatomy in different women; this situation has implications for the validity of comparisons of grouped data. Subsidiary objectives were to assess the interrater reliability of measurements of volumes of the upper limb and to examine construct validity by comparing groups. Three groups of women were included in the study: control subjects, women with a clinical diagnosis of lymphedema in the ipsilateral arm after axillary surgery for breast cancer, and women without a clinical diagnosis of lymphedema in the ipsilateral arm after axillary surgery.

* Pero-System Messgeräte, Am Tescher Busch 9 D-42327, Wuppertal, Germany.

Method

Subjects

This study was a cross-sectional comparison of the validity and reliability of measurements of arm volumes in 66 subjects from 3 groups. There were 25 subjects in a control group, 22 subjects in the breast cancer group without a diagnosis of arm lymphedema, and 19 subjects with diagnosed arm lymphedema following breast cancer surgery. Subjects were recruited from the Breast Centre and the Lymphedema Clinic at Westmead Hospital (Sydney, New South Wales, Australia), and the control subjects were staff and volunteers at Westmead Hospital. All subjects were adult women who gave written informed consent.

Both arms of each subject were measured twice by 2 raters using 2 methods of circumferential measurement (at fixed distances from fingertips and in relation to anatomic landmarks) and water displacement. Five different raters took measurements during the study period. All raters were either occupational therapists or nurses who had previous experience with arm measurements in patients with lymphedema and who also received special training for the study, especially in measurement by water displacement, which is not routinely used at the Lymphedema Clinic.

Measurements

Arm measurements were obtained by multiple circumference measurements with a measuring tape (at various distances from fingertips and anatomic landmarks) and calculation of volumes and by water displacement. Each subject was measured once each by 2 raters. For the comparative volume estimations, the segment of the limb between the wrist and an upper boundary on the upper arm was used. This strategy was used because calculation of volumes of the hand by circumferential measurements is imprecise and a conical assumption is obviously not valid. An upper boundary was used because women are not able to submerge their entire limb for volume displacement measurements. The upper boundary selected was the level to which all women (in a pilot study) were able to submerge their arm for the water displacement measurements, and it was expressed as a relative value to take into account variations in limb lengths. The upper boundary was 65% of the distance from the elbow (olecranon) to the shoulder tip (acromion).

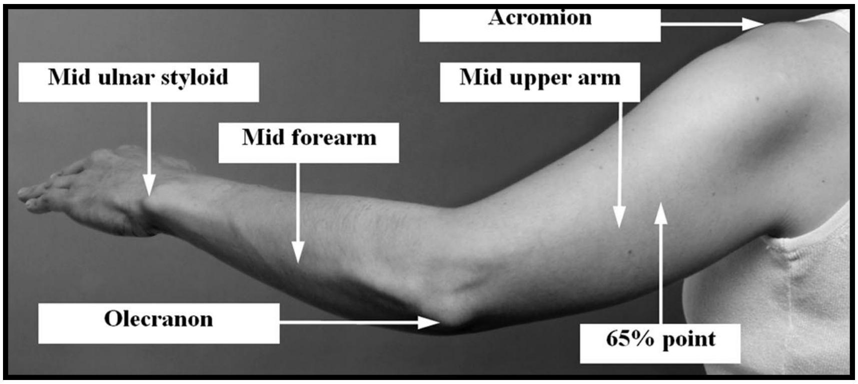

Figure.
Anatomic landmarks used in arm volume computations.

The volume for a segment between 2 adjoining measured segments was computed by assuming a truncated cone. Four truncated cones were summed to measure arm volumes based on both distance from the tip of the middle finger (wrist to 30 cm, 30–40 cm, 40–50 cm, and 50 cm to the 65% mark) and anatomic landmarks (wrist to mid forearm, mid forearm to elbow, elbow to mid upper arm, and mid upper arm to the 65% mark) (Figure). Circumferential tape measurements were taken at each of these points using a thin, flexible plastic tape.

This method is based on that described by Sander et al,[7] who found that the frustum assumption produced the smallest SEM. For circumferences computed on the basis of distances from fingertips, the first truncated cone was taken from the wrist (styloid) to 30 cm from the fingertips, and the last truncated cone was taken from the adjacent most proximate circumferential measurement, which was 50 cm because the upper boundary was less than 60 cm for most of the women and it was little more than 60 cm for the other women, to the upper boundary (described above). For circumferential measurements based on anatomic landmarks, the lower boundary was the wrist (styloid), and the upper boundary was as previously described. Total limb volume for the segment between the wrist and the upper boundary was obtained by adding the volumes of the truncated cones between these points. The volume of a truncated cone is calculated as follows[11]:

$$(1) \qquad V = h(C_1^2 + C_1C_2 + C_2^2)/12\pi$$

where V is the volume of the segment, C_1 and C_2 are the circumferences at the ends of the segment, and h is the distance between them (segment length).

For the water displacement method, volume for the hand and volume for the total limb (to the upper boundary) were recorded separately. This goal was accomplished by first measuring the water displacement for the hand (to the wrist) and then measuring the total displacement for the limb to the designated upper boundary. A special device was constructed with an overflow spout; displaced water was caught and measured in a second container. The volume of the arm segment between the wrist and the upper boundary was obtained by subtracting the hand volume recorded from the total limb volume recorded. Laterality (right, left), dominance (handedness), and whether the arm was ipsilateral (affected) or contralateral (unaffected) to the breast cancer surgery site were noted.

Data Analysis

Interrater reliability. Interrater reliability was computed by use of the intrasubject correlation, which is the correlation between measurements obtained by 2 independent raters of the same subjects across a number of different subjects. The intrasubject correlation was derived from an analysis of variance (ANOVA)[12–14] and multilevel modeling.[15] When different sets of raters evaluated each subject, the intrasubject correlation coefficient (ρ) was derived from an ANOVA model across raters, as follows:

$$(2) \qquad \rho = (MSS - MSE)/[MSS + (N-1)MSE]$$

where MSS is the mean square between subjects, MSE is the mean square error, and N is the number of raters per subject.

In multilevel modeling, the intrasubject correlation (ρ) is defined as the correlation between 2 independent ratings of the same subject.[13] The correlation between 2 such ratings is calculated as follows[15]:

$$(3) \qquad \rho = \frac{\text{between-subject variance}}{\text{total variance}} = \frac{\text{between-subject variance}}{\text{between-rater variance} + \text{between-subject variance}}$$

The reliability of ratings (r_m) or interrater reliability was computed by substituting the intrasubject correlation coefficient (ρ) and the average number of raters (N) in the Spearman-Brown formula, as follows[13,16]:

$$(4) \qquad r_m = N\rho/[1 + (N-1)\rho]$$

In the present study, both ANOVA and multilevel modeling were used to compute the reliability of circumferential measurements and volumes. The intrasubject correlations of circumferential measurements calculated from the techniques of ANOVA and multilevel modeling were identical; therefore, confidence in the results was enhanced. Circumferential measurements by 2 raters (level 1) were nested within subjects (level 2), and multilevel models, which extend linear regression to deal with variability at different levels, were used to obtain variance estimates to compute reliability values. Multilevel models were fitted for each arm with circumferential measurements as response variables to obtain estimates of between-subject and between-rater variances.[17] Intrasubject correlations were computed by substituting the estimates of between-subject and between-rater variances in equation 3. The confidence interval of the mean volume was calculated to examine differences in means. The coefficient of variation, which is defined as the standard deviation expressed as a percentage of the mean, was used to compare the consistency or variability of 2 or more series of measurements.

Criterion validity. The criterion validity of circumferential measurements was determined by comparing volumes computed from circumferential measurements with a reference standard of volume measurements obtained by water displacement. Mean volumes were compared by use of paired sample t tests, and Pearson correlation coefficients with confidence intervals were computed for volumes obtained from circumferential measurements compared with those obtained from water displacement. Volumes obtained from the 2 methods for the same arm were compared; therefore, the characteristics of the arm (eg, affected, dominant) were not important in this analysis.

Limits of agreement (LOA). By computing LOA, we could determine whether volumes calculated from the methods of circumferential measurements and water displacement were interchangeable.[4,18,19] The LOA calculations showed how closely the 2 methods agreed with each other in numeric values.[4] The LOA procedure involves computing the mean (d) and standard deviation (s) of differences in all pairs of measurements for 2 methods. If the differences are normally distributed, then 95% of differences will lie between d–2s and d+2s. This range is considered the 95% confidence interval of the LOA. The difference between the circumferential measurement and water displacement methods can be expected to vary between d–2s and d+2s.[18] The question of interchangeability should be based on a clinical criterion and the application of the measurements.

SEM and minimum detectable change (MDC). The SEM or the error associated with a single measure was computed with the following formula[20,21]:

$$SEM = SD\sqrt{1-R} \qquad (5)$$

where SD is the standard deviation of the volumes and R is the test-retest reliability coefficient. A type 2,1 ICC was used to estimate the test-retest reliability coefficient.[22,23]

The MDC was defined as follows[22]:

$$\text{MDC} = 1.64\sqrt{2}\ \text{SEM} \qquad (6)$$

where 1.64 is the 2-sided tabled z value for the 90% confidence interval.

Construct validity. The construct validity of volume measurements of the upper limb as an indicator of lymphedema was assessed by comparing volumes of affected and unaffected arms for women with breast cancer, with and without diagnosed lymphedema, and also by comparing left and right arms and dominant and nondominant arms. Volumes of affected and unaffected arms for women with breast cancer, with and without diagnosed lymphedema, were compared before and after adjustment for covariates (right arm and dominant arm) by use of multilevel modeling with arm volumes as response variables. Multilevel models were fitted with MLwiN (version 1.1).[17,†]

Results

Reliability

Circumferential measurements. The interrater reliability values based on 2 raters were .98 to .99 for circumferential tape measurements at 30 to 60 cm from the fingertips (Tab. 1). The interrater reliability values for 5 circumferential measurements taken in relation to 3 bony landmarks (wrist at styloid, elbow at olecranon, and shoulder at acromion) were .97 to .99 (Tab. 1).

Volumes. The interrater reliability values based on 2 raters for arm volumes measured by water displacement or calculated from circumferential measurements were ≥.95 for subjects with lymphedema, ≥.98 for subjects with breast cancer and without lymphedema, and ≥.94 for control subjects (for dominant and nondominant arms assessed separately) (Tab. 2). There were no differences between mean volumes, as indicated by confi-

[†] Centre for Multilevel Modeling, Institute of Education, 20 Nedford Way, London, United Kingdom WC1H 0AL.

Table 1.
Interrater Reliability of Circumferential Measurements

Position	Interrater Reliability	
	Left Arm	Right Arm
From fingertips (cm)		
60	.99	.99
50	.99	.99
40	.99	.99
30	.99	.98
Anatomic landmark[a]		
Wrist	.98	.97
Midpoint between elbow and wrist	.98	.98
Elbow	.99	.99
Midpoint between elbow and shoulder	.99	.99
Upper boundary	.99	.98

[a] Wrist at styloid process of ulna, elbow at olecranon of ulna, and shoulder at acromion of scapula. Upper boundary: 65% of the distance from the elbow (olecranon) to the shoulder tip (acromion).

dence intervals and coefficients of variation, for different measurement methods (Tab. 3).

LOA

The magnitude of the difference between volumes from water displacement and volumes from circumferential measurements was lower for the right arm. The mean difference from water displacement can be expected to vary between –260 mL and 110 mL for volumes determined from the distance from fingertips and between –217 mL and 159 mL for volumes determined from anatomic landmarks for the right arm (Tab. 3). Thus, the 95% confidence interval for the LOA suggests that the volume difference would not be acceptable clinically, because both circumferential volumes overestimate water volume by more than 110 mL. The wide LOA suggest that the 2 methods should not be used interchangeably. If a clinician is interested in the exact volume, circumferential volumes cannot be used to provide an accurate estimate.

SEM and MDC

The MDC of volumes or the error associated with a single measure for data obtained from anatomic landmarks was lower than that for data obtained from distance from fingertips (Tab. 3). The range of the SEM for the arm data was 64.5 to 81.7 mL (Tab. 3). The volumes calculated from anatomic landmarks had the smallest SEM. A difference of up to 150 mL (MDC) is considered to be measurement error.

Criterion Validity

Volumes calculated from circumferences were higher than those obtained from water displacement by up to 5% ($P<.001$). Differences from water displacement volumes were smaller for volumes computed from circumferences based on anatomic landmarks (up to 2%) ($P<.001$) than for those calculated from distance from fingertips (Tab. 3). Comparison between methods indicated that there was a high correlation (.98) between volumes obtained from water displacement and volumes calculated from circumferential measurements (Tab. 3).

Construct Validity

The volume of the affected arm for the lymphedema group was 11% higher than that of the unaffected arm ($P<.001$), but the difference was slight (0.2%) and not significant for the breast cancer group without diagnosed lymphedema (Tab. 4). The results were unchanged when adjusted for arm dominance and left or right arm by multilevel modeling. Before and after adjustment for dominant arm and right arm by multilevel modeling, the differences in volumes between the breast cancer groups with and without diagnosed lymphedema (interaction) remained statistically significant (Tab. 5). The volume of the right arm was higher (~2%) than that of the left arm, although the difference did not reach statistical significance (Tab. 3). For control subjects, the volume of the dominant arm was higher than that of the nondominant arm when determined by circumferential measurements or water displacement (~1.5%), but this difference did not reach statistical significance (data not shown).

Discussion and Conclusions

Measurement of arm volumes presents a number of difficulties. Water displacement has the highest logical validity, but it is not easy to ensure that arms are submerged to the same level. For comparability, a standard level for submersion must be relative to the length of the subject's arm rather than an absolute distance from the fingertips. When volumes determined from water displacement are compared with volumes computed from circumferential measurements, it is necessary to exclude the hand because its shape and irregularities do not correspond at all to the truncated cone used for calculating volumes from circumferences. This situation requires a 2-stage measurement (first the hand and then the hand and arm) to deduce arm volume without the hand. Despite these possible difficulties, there was high interrater reliability between different observers for water displacement determinations of arm volumes in this study.

Circumferential measurements obtained with a tape measure could be affected by the positions on the arm chosen by different observers to measure circumferences, although the observers followed the same guidelines for distance from fingertips or anatomic landmarks. Nevertheless, ICCs for interrater reliability of these circumferential measurements and volumes calculated from these measurements for the arm, excluding

Table 2.
Interrater Reliability (Intraclass Correlation Coefficients) of Arm Volumes Obtained From Different Measurement Methods

Group	N	Volume From Circumferential Measurements: Distance From Fingertips	Volume From Circumferential Measurements: Anatomic Landmarks	Volume From Water Displacement
All subjects	66			
Left arm		.98	.98	.98
Right arm		.98	.97	.97
Subjects with breast cancer				
Affected arm				
Lymphedema	19	.99	.99	.97
No lymphedema	22	.98	.99	.99
Unaffected arm				
Lymphedema	19	.98	.95	.96
No lymphedema	22	.98	.99	.98
Control subjects	25			
Dominant arm		.97	.96	.98
Nondominant arm		.98	.97	.94

the hand, were high (mostly ≥.98). The ICCs for reliability of volumes determined from water displacement also were high (mostly ≥.98). These results are similar to those of Sander et al[7] (ICC=.99) and Megens et al[4] (ICC=.99) for volumes calculated from distance from fingertips and from water displacement for a comparable group (affected arms in subjects with lymphedema).

Validity assessed by Pearson coefficients of correlation between volumes determined from water displacement and volumes calculated from circumferential measurements was very high (.98) and also similar to that found by Sander et al[7] for a comparable group (affected arms in subjects with lymphedema). Further study of criterion validity indicated that volumes computed from circumferential measurements were higher than volumes measured by water displacement, although the difference was not large (<5%). Water displacement measurements have higher validity than circumferential measurements because the latter assume that arm segments are truncated cones, but they are not. The arm in cross section is an ellipse more than a circle, and there are surface irregularities not captured by spaced measurements. Furthermore, the thickness of the tape measure, although it is tiny, makes a measurement fractionally higher than the real circumference.

The use of anatomic landmarks to determine sites for circumferential measurements produced volumes more accurate (up to 2% larger than water displacement volumes) than those obtained from measurements at distances from fingertips (up to 5% larger than water displacement volumes). Unlike the findings of our study, Sander et al[7] found volumes calculated from circumferential measurements for the arm (minus fingers) to be smaller than those determined by water displacement, as did Karges et al[24] and Pani et al.[25] However, in studies by Stranden[26] and Megens et al,[4] the volumes determined from water displacement were smaller than the volumes calculated from surface measurements for edema in the leg and upper extremity, respectively. The techniques of water displacement differed somewhat between the studies. The greater accuracy of circumferential measurements based on anatomic landmarks than on fixed distances from fingertips was hypothesized prior to data collection, because measurements related to anatomic sites take into consideration relative lengths of the arm and produce segments that correspond more to truncated cones because segments do not span the elbow. Such measurements are more comparable in different women with different arm lengths. However, obtaining measurements in relation to anatomic landmarks requires more expertise and training in observers than does using distances from fingertips because knowledge of surface anatomy is needed and therefore may be more difficult to implement in clinical practice. Nevertheless, validity, assessed by correlation coefficients, was high for both methods of circumferential measurement, compared with water displacement.

For adequate construct validity, it would be expected that the difference in arm volumes would be greater in women with clinical lymphedema following breast cancer than in those without clinical lymphedema. Women with clinical lymphedema had significantly greater differences between affected and unaffected arms than women with breast cancer but without diagnosed lymphedema. The findings of larger volumes for right arms than for left arms and dominant arms larger than nondominant arms also are consistent with expectations and confirm construct validity.

Table 3.
Arm Volumes Obtained From Circumferences and Water Displacement[a]

Measure (n=132)	Right Arm Volume: Water Displacement	Right Arm Volume: Circumferential Measurements: Distance From Fingertips	Right Arm Volume: Circumferential Measurements: Anatomic Landmarks	Left Arm Volume: Water Displacement	Left Arm Volume: Circumferential Measurements: Distance From Fingertips	Left Arm Volume: Circumferential Measurements: Anatomic Landmarks
Volume (mL)						
Mean	1,970	2,045	1,999	1,929	2,024	1,964
95% CI for mean	1,888–2,051	1,964–2,126	1,920–2,079	1,848–2,010	1,938–2,111	1,885–2,043
Range	1,135–3,664	1,174–3,396	1,158–3,622	1,027–3,625	1,138–3,765	1,129–3,496
CV (%)	24	23	23	24	25	23
MDC (SEM)	189.5 (81.7)	154.4 (66.6)	151.6 (65.4)	154.3 (66.5)	164.8 (71.0)	149.7 (64.5)
Difference from water displacement (%)		+75 (3.8)	+29 (1.5)		+95 (4.9)	+35 (1.8)
Mean LOA[b] (95% CI)		−75.4 (−260.2 to 110.2)	−29.4 (−216.8 to 158.8)		−94.8 (−309.0 to 119.0)	−34.7 (−241.0 to 171.0)
Difference bewteen circumferential methods (%)		+46 (2.3)			+60 (3.5)	
Correlation coefficient[c] (CI)		.98 (.979–.981)	.98 (.979–.981)		.98 (.979–.981)	.98 (.979–.981)

[a] Paired *t* tests for volume differences were carried out for circumferential measurements compared to water displacement (difference from water displacement) and for circumferential measurements based on distances from fingertips compared with those based on anatomic landmarks (difference between circumferential methods). All of these comparisons showed significant differences ($P<.001$). CI=confidence interval; CV=coefficient of variation, calculated as 100×(SD/Mean); MDC=minimum detectable change; SEM=standard error of measurement; LOA=limits of agreement.
[b] LOA with water displacement.
[c] Pearson coefficient of correlation with water displacement.

The volumes determined from anatomic landmarks had the lowest SEM (range=64.5–65.4 mL), followed by those determined from distance from fingertips (range=66.6–71.0 mL) and water displacement (range=66.5–81.7 mL) (Tab. 3). Sander et al[7] reported SEMs of 116 mL for volumes determined by the 9-cm frustum method and 117 mL for volumes determined by water displacement. Clinicians require that a decision can be made concerning how much change from a previous reading is clinically significant and may require a change in treatment. Minimum detectable change is the way of determining that the change is not attributable to chance variation or measurement error. The results of our study suggest that a change of less than 150 mL (MDC) should be treated as no change by clinicians[21]; a change of more than 150 mL is not likely to be attributable to chance variation or measurement error.

These results indicate high reliability of the measurements of arm circumferences and volumes and are similar to those of both Sander et al[7] and Megens et al.[4] We also noted that volumes calculated from circumferential measurements relative to anatomic landmarks were more accurate than and different from those calculated from segments defined from distances from fingertips. However, both circumferential methods produced high coefficients of correlation with water displacement. Volumes calculated from circumferential measurements based on anatomic landmarks had smaller mean LOA with water displacement and displayed smaller MDC and SEM than did those calculated from circumferential measurements based on distances from fingertips (Tab. 3). In addition, the difference in volumes between the right arm and the left arm determined by circumferences based on anatomic landmarks was closer to that determined by water displacement than to that determined by circumferences based on distances from fingertips (data not shown). Both circumferential methods systematically overestimated the volumes in relation to water displacement and should not be used interchangeably with water displacement. However, they may be used individually.

Accurate measurement of arm volumes is required in studies of women after various types of breast cancer treatment, including axillary dissection, and for assessing arm outcomes in clinical trials of different treatment procedures. These studies also permit the examination of predictors of lymphedema. Furthermore, accurate measurement of arm volumes is required for assessing methods for the treatment of lymphedema either in observational case series or in prospective clinical trials.

Valid and repeatable measures of arm volumes are required for both studies of causes and antecedents of arm lymphedema and trials of therapy for lymphedema.

Table 4.
Arm Volumes for Women With Breast Cancer[a]

Measure	Diagnosed Lymphedema (n=19) Affected Arm	Diagnosed Lymphedema (n=19) Unaffected Arm	No Diagnosed Lymphedema (n=22) Potentially Affected Arm	No Diagnosed Lymphedema (n=22) Unaffected Arm
Volume (mL)				
Mean	2,326	2,094	2,036	2,031
SEM	89	89	73	71
Range	1,434–3,370	1,346–3,622	1,158–2,998	1,129–3,004
Difference from unaffected arm[b] (%)	+232[c] (11)		+5[d] (0.2)	
Unadjusted[e] difference from unaffected arm (%)	+221.7[c] (11)		+4.9[d] (0.2)	
Adjusted[e] difference from unaffected arm (%)	+266.9[c] (13)		+43.1[d] (2)	

[a] Volumes were computed from circumferential tape measurements obtained at anatomic landmarks. SEM=standard error of measurement.
[b] The difference in volume between the affected arm and the unaffected arm was determined by use of a paired t test.
[c] Significant at P>.001.
[d] Not sigfnificant at P>.05.
[e] Volumes unadjusted or adjusted for right and dominant arms were estimated by use of the fitted multilevel equation.

Table 5.
Regression Coefficients Determined by Multilevel Modeling for Arm Volumes Obtained From Anatomic Positions of Women With Breast Cancer to Examine Construct Validity of Measurements[a]

Parameter	Unadjusted Model Estimate	Adjusted Model Estimate
Main effects		
Mean	−0.155	−0.190
Dominant arm		−0.581 (0.593)
Right arm		−0.438 (0.593)
Affected arm	0.004 (0.071)	0.132 (0.191)
Lymphedema group	0.102 (0.296)	0.018 (0.304)
Interaction effects		
Dominant and right arm		1.227 (1.132)
Dominant and affected arm		0.271 (0.398)
Dominant and lymphedema group		0.016 (0.100)
Right and affected arm		−0.500 (0.398)
Right arm and lymphedema group		0.000 (0.000)
Affected arm and lymphedema group	0.459[b] (0.100)	0.451[b] (0.098)
Random effects		
Variance between subjects	0.845[b] (0.192)	0.806[b] (0.183)
Variance between raters	0.095[b] (0.012)	0.090[b] (0.012)
Total variance	0.940	0.896

[a] The explanatory variables considered in the analyses are lymphedema group (1=lymphedema, 0=no lymphedema), dominant arm (1=dominant, 0=nondominant), handedness (1=right, 0=left), and affected arm (1=affected, 0=unaffected). Normalized arm volume was used as the response variable, and the explanatory variables were included in the models with their interactions. Values in parentheses are standard errors.
[b] P<.001.

Water displacement is too cumbersome and messy to be used in routine clinical practice. This study has shown that arm circumferential measurements in relation to anatomic landmarks are reliable and valid measurements of arm volumes, can be used with confidence, and appear to be much more efficient clinically. Future research could address the validity of data obtained with newer methods for measuring arm volumes, such as perometry.

References

1 Collard V. Understanding lymphoedema and the occupational therapist's role. *The Australian Occupational Therapy Journal.* 1990;37:109–111.

2 Kissin MW, Querci della Rovere G, Easton DF, Westbury G. Risk of lymphoedema following the treatment of breast cancer. *Br J Surg.* 1986;73:580–584.

3 Pain SJ, Purushotham AD. Lymphoedema following surgery for breast cancer. *Br J Surg.* 2000;87:1128–1141.

4 Megens AM, Harris SR, Kim-Sing C, McKenzie DC. Measurement of upper extremity volume in women after axillary dissection for breast cancer. *Arch Phys Med Rehabil.* 2001;82:1639–1644.

5 Petrek JA, Pressman PI, Smith RA. Lymphedema: current issues in research and management. *CA Cancer J Clin.* 2000;50:292–307.

6 Kaulesar Sukul DM, den Hoed PT, Johannes EJ et al. Direct and indirect methods for the quantification of leg volume: comparison between water displacement volumetry, the disk model method and the frustrum sign model method, using the correlation coefficient and the limits of agreement. *J Biomed Eng.* 1993;15:477–480.

7 Sander AP, Hajer NM, Hemenway K, Miller AC. Upper-extremity volume measurements in women with lymphedema: a comparison of measurements obtained via water displacement with geometrically determined volume. *Phys Ther.* 2002;82:1201–1212.

8 *Perometer.* Wuppertal, Germany: Pero-System Messgeräte GmbH; 1999.

9 Stanton AW, Northfield JW, Holroyd B, et al. Validation of an optoelectronic limb volumeter (Perometer). *Lymphology*. 1997;30:77–97.

10 Casley-Smith JR, Casley-Smith JR. *Modern Treatment for Lymphoedema*. Malvern, South Australia, Australia: The Lymphoedema Association of Australia. 1994:90–98.

11 Casley-Smith JR, Casley-Smith JR. Modern treatment of lymphoedema, I: complex physical therapy—the first 200 Australian limbs. *Australas J Dermatol*. 1992;33:61–68.

12 Winer BJ. *Statistical Principles in Experimental Design*. New York, NY: McGraw-Hill; 1971.

13 Jayasinghe UW, Marsh HW, Bond N. A multilevel cross-classified modelling approach to peer review of grant proposals: the effects of assessor and researcher attributes on assessor ratings. *J R Stat Soc A*. 2003;166:279–300.

14 Jayasinghe UW, Marsh HW, Bond N. Peer review in the funding of research in higher education: the Australian experience. *Education Evaluation and Policy Analysis*. 2001;23:343–364.

15 Goldstein H. *Multilevel Statistical Models*. London, United Kingdom: Edward Arnold; 1995.

16 Marsh HW, Ball S. Interjudgmental reliability of review for the *Journal of Educational Psychology*. *J Educ Psychol*. 1981;73:872–880.

17 Rabash J, Browne W, Goldstein H, et al. *A User's Guide to MLwiN*. London, United Kingdom: Centre for Multilevel Modelling, Institute of Education, University of London; 2001.

18 Portney LG, Watkins MP. *Foundations of Clinical Research Applications to Practice*. Englewood Cliffs, NJ: Prentice-Hall; 2000.

19 Bland JM, Altman DG. Statistical methods for assessing agreement between two methods of clinical measurement. *Lancet*. 1986;i:307–310.

20 Streiner DL, Norman GR. *Health Measurement Scales: a Practical Guide to Their Development and Use*. New York, NY: Oxford University Press; 1995.

21 Stratford PW, Finch E, Solomon P, et al. Using the Roland-Morris questionnaire to make decisions about individual patients. *Physiother Can*. 1996;48:107–110.

22 Binkley JM, Stratford PW, Lott SA, Riddle DL; The North American Orthopaedic Rehabilitation Research Network. The Lower Extremity Functional Scale (LEFS): scale development, measurement properties, and clinical application. *Phys Ther*. 1999;79:371–383.

23 Shrout PE, Fleiss JL. Intraclass correlations: uses in assessing rater reliability. *Psychol Bull*. 1979;86:420–428.

24 Karges JR, Mark BE, Stikeleather SJ, Worrell TW. Concurrent validity of upper-extremity volume estimates: comparison of calculated volume derived from girth measurements and water displacement volume. *Phys Ther*. 2003;83:134–145.

25 Pani SP, Vanamail P, Yuvaraj J. Limb circumference measurement for recording edema volume in patients with filarial lymphedema. *Lymphology*. 1995;28:57–63.

26 Stranden E. A comparison between surface measurements and water displacement volumetry for the quantification of leg edema. *J Oslo City Hosp*. 1981;31:153–155.