Title

Assessment of Arm Volume Using a Tape Measure Versus a 3D Optical Scanner in Survivors with Breast Cancer-Related Lymphedema



Journal

Lymphatic Research and Biology, 20(1)

ISSN

1539-6851

Authors

Mastick, Judy
Smoot, Betty J
Paul, Steven M
et al.

Publication Date

2022-02-01

DOI

10.1089/lrb.2020.0119

Peer reviewed

LYMPHATIC RESEARCH AND BIOLOGY
Volume 20, Number 1, 2022

DOI: 10.1089/lrb.2020.0119

# Assessment of Arm Volume Using a Tape Measure Versus a 3D Optical Scanner in Survivors with Breast Cancer-Related Lymphedema

Judy Mastick, RN, MN,[1] Betty J. Smoot, DPTSc, MAS,[2] Steven M. Paul, PhD,[1] Kord M. Kober, PhD,[1] Bruce A. Cooper, PhD,[1] Lori K. Madden, RN, PhD,[3] Yvette P. Conley, PhD,[4] Niharika Dixit, MD,[2] Marilyn J. Hammer, RN, PhD,[5] Mei R. Fu, RN, PhD,[6] Merisa Piper, MD,[2] Sarah P. Cate, MD,[7] John Shepherd, PhD,[8] and Christine Miaskowski, RN, PhD[1]

## Abstract

***Background:*** Lymphedema (LE) is a significant clinical problem for breast cancer survivors. While the water displacement test and circumferential assessment using a tape measure (TM) are common methods to assess differences in arm volumes, faster and more reliable methods are needed. Study purposes, in breast cancer survivors ($n = 294$), were to compare the average total arm volumes and interlimb volume ratios for women with and without a history of LE, using a TM and three-dimensional (3D), whole-body surface scanner (3D scan); compare the level of agreement between arm volumes and interlimb volume ratios obtained using the two devices; and evaluate the percent agreement between the two measures in classifying cases of LE using three accepted thresholds.
***Methods and Results:*** Measurements were done using a spring-loaded TM and Fit3D ProScanner. Paired *t*-tests and Bland–Altman analyses were used to achieve the study aims. For circumference and volume comparisons, compared with the 3D scan, values obtained using the TM were consistently smaller. In terms of level of agreement, the Bland–Altman analyses demonstrated large biases and wide limits of agreement for the calculated arm volumes and volume ratios. In terms of the classification of caseness, using the 200-mL interlimb volume difference criterion resulted in 81.6% overall agreement; using the >10% volume difference between the affected and unaffected arms resulted in 78.5% overall agreement; and using the volume ratio ≥1.04 criterion resulted in 62.5% overall agreement. For all three accepted threshold criteria, the percentage of cases was significantly different between the TM and 3D scan techniques.
***Conclusions:*** The 3D technology evaluated in this study has the potential to be used for self-initiated surveillance for LE. With improvements in landmark identification and software modifications, it is possible that accurate and reliable total arm volumes can be calculated and used for early detection.

**Keywords:** lymphedema, arm volume, 3D optical scanner, circumference measures

## Introduction

Lymphedema (LE) is a significant clinical problem for breast cancer survivors. The water displacement test (WD) and circumferential assessment using a tape measure (TM) are common methods to assess differences in arm volumes.[1] While WD is a reference standard in research, tape measurement is the standard method used in clinical practice because of its reliability, low cost, and portability.[1] WD is cumbersome, both methods are time-consuming, and neither

[1]School of Nursing and [2]School of Medicine, University of California San Francisco, San Francisco, California, USA.
[3]UC Davis Health, Sacramento, California, USA.
[4]School of Nursing, University of Pittsburgh, Pittsburgh, Pennsylvania, USA.
[5]Dana Farber Cancer Institute, Boston, Massachusetts, USA.
[6]School of Nursing, Boston College, Boston, Massachusetts, USA.
[7]ICAHN School of Medicine at Mount Sinai, New York, New York, USA.
[8]University of Hawaii Cancer Center, Manoa, Hawaii, USA.

method is error free. Faster, accurate, and more reliable methods are needed to assess LE.

Recently, camera imaging and infrared opto-electric volumetry (perometry), using three-dimensional (3D) technology, have been used to assess LE.[1] When compared with WD, limb volumes obtained using existing medical-grade, multicamera imaging systems[2,3] and infrared imaging[4] are accurate and reliable. However, significant costs limit their use in clinical practice.

Taking advantage of the low-cost, 3D imaging capability of the commercially available Kinect sensor (Microsoft, Redmond, WA), three studies evaluated differences in arm volumes obtained using a 3D scan compared with WD[5,6] or circumference measurements (CMs).[5,7] In a study of cancer survivors with LE,[5] no significant differences in arm volumes were found among the three methods. In another study of healthy participants, which used WD as the comparator,[6] Bland–Altman bias (mean) was −9.9 mL and 95% limits of agreement (LOAs) ranged from −49.6 to 29.8 mL. Test–retest reliability for both methods and standard errors of measurement were similar. In the most recent study of patients who were assessed before and following surgery,[7] while a positive linear correlation was found between the 3D scan and TM, the 3D volumes were larger. The Bland–Altman analysis demonstrated that with larger arm volumes, the level of agreement between the two methods was lower.

While early detection has decreased the prevalence of LE,[8] numerous barriers (e.g., time constraints) impede routine monitoring.[9] As a result, prospective surveillance of LE may not be included in routine care. A survivor-initiated, LE surveillance strategy that can be implemented in the community (e.g., fitness center) may improve early detection and reduce associated morbidity.

One viable approach to improve surveillance may be to use automated, 3D whole-body scanners (3D scan) found in fitness centers. However, only one study of 39 healthy adults evaluated obesity anthropometrics using a community-based 3D scan compared with dual-energy X-ray absorptiometry and TM.[10] The 3D scan's CMs of the waist and hips strongly correlated with the two-criterion methods. Given the potential of this 3D scan and paucity of research on these devices for LE, the purposes of this study (in a sample of breast cancer survivors) were to (1) compare the average total arm volumes and interlimb volume ratios for women with and without a history of LE, using a TM and 3D scan; (2) compare the level of agreement between arm volumes and interlimb volume ratios obtained using the two devices; and (3) evaluate the percent agreement between the two measures in classifying cases of LE using three accepted threshold criteria. A secondary purpose was to evaluate the agreement between two repeated measurements within the TM and 3D scan measurements in a subset of survivors.

## Materials and Methods

### *Participants*

Survivors were recruited from the general population in the San Francisco Bay Area. Participants were female, ≥18 years of age, had unilateral breast cancer, and were included ≥6 months after completion of treatment. Survivors' self-report of a past or current diagnosis of LE determined group assignment (i.e., with or without LE). Survivors were excluded if they had a bilateral mastectomy; history of primary LE; and a condition that would limit their ability to participate in testing of upper extremity strength, movement, and sensation. University of California, San Francisco (UCSF), Institutional Review Board approved the study.

For 24 hours before the study visit, survivors were instructed to stay hydrated and not participate in vigorous exercise. Two hours before the appointment, compression garments were removed. At the time of the study visit, participants provided written informed consent.

### *Circumference measurements*

As was done in our previous studies,[11,12] the survivor was positioned supine, with their straight arms abducted and palms down. A measuring board marked in 10-cm increments was placed under the arm. The center of the ulnar process was marked and aligned with the 0 mark on the measuring board. Ten-centimeter increments were marked up to the outer aspect of the arm using the ruler to confirm alignment with the measuring board. Using a spring-loaded TM, the circumference at each 10-cm segment (0, 10, 20, 30, and 40 cm) was measured on each arm.

### *3D scan measurements*

The 3D scanner system (Fit3D ProScanner, Redwood City, CA) generates automated CMs across the entire body. The device consists of a rotating platform and a 3D, optical light-coding camera mounted in a tower that is located 2 m from the center of the platform. For the test, participants wore form-fitting exercise shorts and a sports bra and longer hair was secured above the neck. While standing on the platform, the mounted handles were grasped, with arms straight, relaxed, and abducted from the body. When prompted, participants depressed buttons on the scanner handles and a 360-degree 3D image was taken in ∼40 seconds. Fifty participants were scanned twice. For the second scan, women stepped off the platform and then repositioned themselves when they stepped back on the platform.

While the head, hands, and feet are excluded, 476 measurements from the neck down to the wrists and ankles are programmed into the device. These measurements were automatically derived and stored in a proprietary database. While a routine 3D scan report would have only 11 measurements, in an effort to match the landmarks used for the TM, 10-cm segment CMs starting at the wrist were obtained. Data were stored on a secure database at Fit3D and securely transferred and stored at UCSF.

### *Data analyses*

Statistical analyses were done using SPSS, version 23.[13] Descriptive statistics were calculated for demographic and clinical characteristics. Analyses were done for the entire sample and for comparisons between women with and without a self-reported history of LE.

Each of the right and left arm measurements were recoded to indicate the affected and unaffected arms. Arm volumes for the 0 to 40 cm locations were calculated using the formula for a truncated cone.[14] Total volumes for the affected and unaffected arms as well as mean and percent differences between arms (affected–unaffected) for TM and 3D scan

measurements were compared using paired *t*-tests. Volume ratios were calculated as affected/unaffected arm and mean differences in volume ratios between the two methods were examined using paired *t*-tests.

Bland–Altman analyses were employed to determine the nature and extent of the agreement between the TM and 3D scan volumes and volume ratios (i.e., estimates of both systematic and random errors).[15] Difference scores for the volume and volume ratios between the two methods (i.e., TM vs. 3D scan) were calculated and bias was calculated as the average of the difference scores. Paired *t*-tests of the volume and volume ratios for the TM and 3D scan were done to determine bias and standard deviations (SDs) of the difference scores and to test if the average difference scores were significantly different from zero.

Scatter plots for the affected and unaffected arms with TM total volumes on the x-axis and 3D scan total volumes on the y-axis were generated with a line of equality. The process was repeated for comparison of volume ratios. Bland–Altman plots for both the affected and unaffected total volumes and volume ratios were created with the mean of both methods on the x-axis and the difference scores for the two methods on the y-axis. With the mean of both methods representing the best estimate of the true value of the measured variable, patterns in the differences of both methods can be observed.[15]

Bias represents the systematic error between the two methods. Lower and upper LOAs were calculated (i.e., lower LOA = bias $-1.96 \times$ SD; upper LOA = bias $+1.96 \times$ SD) and indicate where $\sim$95% of the differences would lie from the mean difference. To determine the degree of confidence that can be placed on the estimate of bias, a 95% confidence interval (CI) for the bias was calculated.[15]

To evaluate percent agreement for both the TM and 3D scan measurements, LE was classified as present (YES) or absent (NO) using three accepted threshold criteria: ≥200 mL difference between the affected and unaffected arms; a 10% difference in volume ([unaffected–affected] $\times$ 100); and a volume ratio (affected/unaffected) ≥1.04.[1] For each analysis, the number of cases versus noncases identified was compared between the two devices and reported as percent agreement. Kappa coefficients were calculated to evaluate agreement for categorical data and interpreted using published criteria.[16]

Data from 50 survivors who had two sequential 3D scans were analyzed to assess the LOAs between the two repeated measures using the two methods. The mean differences between the first and second TM measurements (i.e., TM1–TM2) and 3D scan measurements (i.e., SM1–SM2) were calculated and paired *t*-tests were performed. Correlation coefficients and precision values (square root of the sum of the squared T1 and T2 differences divided by the number of participants) were calculated.[17] Precision values of zero indicate perfect repeatability.

## Results

### *Demographics and clinical characteristics*

A total of 294 survivors with ($n = 109$) and without ($n = 185$) a history of LE were included in this analysis (Table 1). Average age was 61.1 (±10.1) years and BMI was 26.3 (±5.2) kg/m$^2$. Compared with women without LE, a higher proportion of women with LE had a mastectomy and axillary lymph node dissection.

### *Differences in arm volumes and ratios*

CMs obtained with the two methods are presented in Table 2. For the total sample, while the two methods were highly correlated ($r = 0.71$–$0.96$), significant differences were found between the two methods for each of the incremental measurements. For each location, compared with the 3D scan, the TM values were smaller. Except for the 10 cm measurement of the affected arm of the survivors with LE, the same relationships and differences were found for survivors with and without LE.

For total arm volumes, significant differences were found between the two methods for the total sample and for the two LE groups ($p < 0.001$; Table 3). For all of these comparisons, the TM volumes were smaller. For comparisons of the mean interlimb volume differences (affected–unaffected limb) and the interlimb volume ratios (affected/unaffected), no

Table 1. Differences in Demographic and Clinical Characteristics Between Survivors with and Without a History of Lymphedema

| | *Total sample* n = *294* | *Hx of LE* n = *109* | *No Hx of LE* n = *185* | |
|---|---|---|---|---|
| *Characteristic* | *Mean (SD)* | *Mean (SD)* | *Mean (SD)* | *Statistics* |
| Age (years) | 61.1 (10.1) | 61.2 (9.3) | 61.1 (10.6) | $t = 0.11$, $p = 0.916$ |
| Weight (kg) | 71.1 (16.0) | 72.9 (13.2) | 70.0 (17.4) | $t = 1.48$, $p = 0.141$ |
| Height (cm) | 163.7 (6.9) | 164.2 (6.1) | 163.4 (7.3) | $t = 0.86$, $p = 0.392$ |
| BMI (kg/m$^2$) | 26.3 (5.2) | 27.1 (4.9) | 25.8 (5.4) | $t = 1.96$, $p = 0.052$ |
| Months since surgery | 91.4 (79.4)<br>Median = 63.0 | 96.7 (79.6)<br>Median = 67.8 | 88.2 (79.4)<br>Median = 58.4 | $U = 0.226$ |
| | n *(%)* | n *(%)* | n *(%)* | |
| Type of surgery | | | | FE, $p < 0.001$ |
| Mastectomy | 94 (32.4) | 51 (47.2) | 43 (23.6) | |
| BCS | 196 (67.6) | 57 (52.8) | 139 (76.4) | |
| SLNB (% yes) | 211 (74.0) | 71 (67.6) | 140 (77.8) | FE, $p = 0.069$ |
| ALND (% yes) | 136 (48.1) | 75 (69.4) | 61 (34.9) | FE, $p < 0.001$ |

ALND, axillary lymph node dissection; BCS, breast-conserving surgery; BMI, body–mass index; FE, fisher's exact; Hx, history; LE, lymphedema; SD, standard deviation; SLNB, sentinel lymph node biopsy.

Table 2. Differences in 10-cm Incremental Circumference Measures and Correlations Between the Measures for the Unaffected Arm and Affected Arm Using the Tape Measure and Three-Dimensional Scan (Scan)

| | *Total sample* n = *294* | | *Hx of LE* n = *109* | | *No Hx of LE* n = *185* | |
|---|---|---|---|---|---|---|
| *Measurement location and device* | *Mean (SD)* | r | *Mean (SD)* | r | *Mean (SD)* | r |
| Unaffected arm 0 cm | | | | | | |
| TM | 15.84 (0.99)[a] | 0.71 | 15.94 (0.98)[a] | 0.63 | 15.77 (1.00)[a] | 0.76 |
| Scan | 18.68 (1.71) | | 18.93 (1.83) | | 18.54 (1.63) | |
| Affected arm 0 cm | | | | | | |
| TM | 15.97 (1.17)[a] | 0.74 | 16.33 (1.29)[a] | 0.73 | 15.76 (1.05)[a] | 0.74 |
| Scan | 18.84 (2.13) | | 19.41 (2.72) | | 18.50 (1.61) | |
| Unaffected arm 10 cm | | | | | | |
| TM | 20.85 (2.20)[a] | 0.86 | 21.09 (1.96)[a] | 0.82 | 20.70 (2.33)[a] | 0.87 |
| Scan | 21.33 (2.16) | | 21.56 (1.94) | | 21.20 (2.27) | |
| Affected arm 10 cm | | | | | | |
| TM | 21.24 (2.74)[a] | 0.87 | 22.36 (2.98) | 0.90 | 20.57 (2.36)[a] | 0.83 |
| Scan | 21.55 (2.50) | | 22.49 (2.67) | | 21.00 (2.22) | |
| Unaffected arm 20 cm | | | | | | |
| TM | 24.40 (2.15)[a] | 0.89 | 24.72 (1.91)[a] | 0.84 | 24.21 (2.26)[a] | 0.90 |
| Scan | 26.31 (2.28) | | 26.60 (2.08) | | 26.14 (2.38) | |
| Affected arm 20 cm | | | | | | |
| TM | 24.80 (2.65)[a] | 0.93 | 25.91 (2.98)[a] | 0.94 | 24.14 (2.19)[a] | 0.90 |
| Scan | 26.64 (2.80) | | 27.82 (3.05) | | 25.95 (2.40) | |
| Unaffected arm 30 cm | | | | | | |
| TM | 27.11 (3.45)[a] | 0.95 | 27.75 (3.36)[a] | 0.94 | 26.73 (3.45)[a] | 0.95 |
| Scan | 28.88 (3.80) | | 29.56 (3.65) | | 28.48 (3.84) | |
| Affected arm 30 cm | | | | | | |
| TM | 27.50 (3.91)[a] | 0.96 | 28.94 (4.26)[a] | 0.96 | 26.64 (3.41)[a] | 0.95 |
| Scan | 29.13 (4.14) | | 30.70 (4.41) | | 28.20 (3.68) | |
| Unaffected arm 40 cm | | | | | | |
| TM | 30.90 (3.99)[a] | 0.86 | 31.66 (3.83)[a] | 0.87 | 30.46 (4.03)[a] | 0.86 |
| Scan | 32.54 (4.17) | | 33.29 (4.05) | | 32.10 (4.18) | |
| Affected arm 40 cm | | | | | | |
| TM | 31.14 (4.15)[a] | 0.93 | 32.18 (4.24)[a] | 0.94 | 30.53 (3.98)[a] | 0.92 |
| Scan | 32.68 (4.45) | | 33.82 (4.48) | | 32.01 (4.31) | |

The 0 cm location is at ulnar styloid. Each additional 10 cm location is proximal to 0 cm.
[a]Significance of the paired differences in TM and scan measures—$p$-value <0.001.
TM, tape measure.

differences were found between the two methods for the total sample and for those survivors with and without a history of LE (Table 3).

### *Level of agreement*

Bland–Altman analysis results for bias, CI, and LOA are presented in Table 4. For the total sample and for the two LE groups, results were very similar for the affected and unaffected total arm volumes and volume ratios. In addition, the TM and 3D scan volumes for the affected arms of survivors with ($r = 0.953$) and without ($r = 0.934$) LE were highly correlated (both $p < 0.001$). Unaffected arm volumes from the TM and 3D scan measurements for survivors with ($r = 0.930$) and without ($r = 0.943$) a history of LE were highly correlated ($p < 0.001$). Volume ratios from the TM and 3D scan for survivors with a history of LE were moderately correlated ($r = 0.732$, $p < 0.001$). Volume ratios from the TM and 3D scan for survivors without a history of LE were not correlated ($r = 0.066$, $p < 0.371$).

Figure 1A shows the scatter plot for the total volumes of the affected arm from the TM versus 3D scan measurements, including the line of equality for the entire sample. The two methods were highly correlated ($r = 0.948$, $p < 0.001$). The Bland–Altman plot (Fig. 1B) shows the distribution of the affected arm total volume difference scores for the entire sample around the line of equality. The bias (difference) of −243.0 (±171.3) was significantly different from 0 ($p < 0.001$) and had a very wide LOA (95% LOA: −578.8 to 92.8).

Figure 1C shows the scatter plot for the total volume of the unaffected arm from the TM versus 3D scan measurements, including the line of equality for the entire sample. The two methods were highly correlated ($r = 0.939$, $p < 0.001$). The Bland–Altman plot (Fig. 1D) shows a bias (difference) of −257.4 (±162.5), was significantly different from 0 ($p < 0.001$), and had a very wide LOA (95% LOA: −575.9 to 61.1). Total volume average bias for both the affected and unaffected arms was large and negative, indicating that a 3D scan value was typically higher than the TM value, which is consistent with the findings presented in Table 2.

Table 3. Differences in Total Arm Volumes in Milliliters Between Values Obtained Using a Tape Measure Compared with a Three-Dimensional Scan for the Area Between the Wrist and 40 cm Proximal to the Wrist

| | *Entire sample* n = *294* | *Hx of LE* n = *109* | *No Hx of LE* n = *185* |
|---|---|---|---|
| *Measure and location* | *Mean (SD)* | *Mean (SD)* | *Mean (SD)* |
| Affected limb | | | |
| TM | 1958.5 (481.2)[a] | 2142.0 (541.3)[a] | 1850.5 (406.4)[a] |
| 3D scan | 2201.6 (533.8) | 2402.9 (586.0) | 2082.9 (462.7) |
| Unaffected limb | | | |
| TM | 1898.5 (403.8)[a] | 1964.0 (383.0)[a] | 1859.9 (411.7)[a] |
| 3D scan | 2155.8 (464.1) | 2227.4 (432.0) | 2113.7 (478.1) |
| Affected–unaffected | | | |
| TM | 60.1 (215.1) | 178.0 (302.6) | –9.42 (82.9) |
| TM % difference | 3.0% (10.1%) | 8.7% (1.4%) | –0.4% (4.3%) |
| 3D scan | 45.74 (311.5)[ns] | 175.6 (374.8)[ns] | –30.73 (237.1)[ns] |
| Scan % difference | 2.4% (13.6%)[ns] | 8.0% (1.6%)[ns] | –0.9% (10.4%)[ns] |
| Ratio calculated as affected/unaffected | | | |
| TM | 1.03 (0.10)[ns] | 1.09 (0.14)[ns] | 1.00 (0.04)[ns] |
| 3D scan | 1.02 (0.14) | 1.08 (0.16) | 0.99 (0.10) |

Significance of differences in TM versus 3D scan.
[a]$p < 0.001$.
3D, three-dimensional; ns, not significant.

Figure 1E shows the scatter plot for volume ratios from the TM versus 3D scan measurements, including the line of equality for the entire sample. The two methods were moderately correlated ($r = 0.597$, $p < 0.001$). The Bland–Altman plot (Fig. 1F) shows that the bias (difference) of 0.0056 (±0.1108) was not significantly different from 0 ($p = 0.387$) and had a fairly wide LOA (95% LOA: –0.2116 to 0.2228).

Although the volume ratio average bias for the entire sample was close to zero (Table 4), for any one individual, this finding was typically not true. The range of differences was wide and for approximately half the time, the TM value was higher than the 3D scan, and for the other half of the time, the 3D scan value was higher than the TM.

### *Agreement between methods in classifying cases of LE*

Classification of LE versus non-LE cases for the two devices, using the 200-mL interlimb volume difference criterion, resulted in 81.6% overall agreement, with a kappa coefficient of 0.406. Using the criterion of >10% volume difference between the affected and unaffected arms, classification of LE versus non-LE cases for the two devices resulted in 78.5% overall agreement, with a kappa coefficient of 0.359. For the volume ratio ≥1.04 criterion, classification of LE versus non-LE cases for the two devices resulted in 62.5% overall agreement, with a kappa coefficient of 0.189 (Table 5). For all three accepted threshold criteria, the

Table 4. Results of Bland–Altman Analyses Between the Tape Measure and Three-Dimensional Scan for Affected and Unaffected Total Arm Volumes and Affected/Unaffected Volume Ratios

| *Total arm volume* | *Bias Mean (SD)* | *95% CI of the bias* | *95% LOA Lower, Upper* |
|---|---|---|---|
| Analysis for total sample | | | |
| Affected arm | –243.0 (171.3)[a] | –262.7 to –223.4 | –578.8 to 92.8 |
| Unaffected arm | –257.4 (162.5) | –276.0 to –238.7 | –575.9 to 61.1 |
| Data for Hx of LE | | | |
| Affected arm | –260.9 (177.9)[a] | –294.7 to –227.1 | –609.6 to 87.8 |
| Unaffected arm | –263.4 (160.1) | –293.8 to –233.0 | –577.2 to 50.4 |
| Data for no Hx of LE | | | |
| Affected arm | –232.5 (166.9)[a] | –256.7 to –208.3 | –559.6 to 94.6 |
| Unaffected arm | –253.8 (164.2) | –227.6 to –230.0 | –575.6 to 68.0 |
| Volume ratios: affected/unaffected | | | |
| Total sample | 0.0056 (0.1108)[ns] | –0.0071 to 0.0183 | 0.2228 to –0.2116 |
| Hx of LE | 0.0072 (0.1130)[ns] | –0.0142 to 0.0287 | 0.2333 to –0.2177 |
| No Hx of LE | 0.0047 (0.1097)[ns] | –0.0113 to 0.0206 | 0.2198 to –0.2104 |

[a]$p < 0.001$.
CI, confidence interval; LOAs, limits of agreement.

percentage of YES responses, or conversely the percentage of NO responses, was significantly different between the TM and 3D scan measures, as indicated by the significant McNemar tests.

### *Level of agreement within methods*

Overall, TM measurements had better test–retest reliability results than 3D scan measurements (Table 6). The 3D scan paired difference *t*-tests resulted in significant *p*-values on the left side at 0, 10, and 20 cm. The precision values, in which lower scores represent greater precision, ranged from 0.38 cm on the right side at 20 cm to 0.91 on the left side at 0 cm. Correlation coefficients were ≥0.84. The mean differences for the two TM measurements ranged from 0.02 at multiple sites to 0.16 on the right side at 30 cm. The paired difference *t*-test resulted in significant *p*-values on the left side at 20 cm and on the right side at 30 cm. The precision values ranged from 0.19 on the right side at 0 cm to 0.46 on the left side at 30 cm. Correlation coefficients were ≥0.98.

## Discussion

Given the need for reliable self-initiated methods of surveillance for LE,[9,18] in this study, a community-based, automated, 3D whole-body scan was utilized to determine its accuracy and reliability as a potential approach to evaluate breast cancer-related LE compared with measurements done using a TM.

### *Agreement between measures*

While previous reports found that compared with WD and TM, a 3D scan overestimated total arm volume in healthy participants[2] and survivors with LE,[3,5,7] we extend these findings by performing CMs and analyses at five different locations on the affected and unaffected arms of women with and without a history of LE. Consistent with previous reports,[5,7] for all of our CM comparisons, values were always higher for the 3D scan. A possible explanation for this finding could be that the 3D scan algorithm processes data collected from the projected direct path of light to the extended arm.[19] In contrast, the TM accommodates the irregular shape of the arm[5] and adds slight compression. The position of the body in different planes (i.e., supine vs. standing) creates different measurement geometries and may contribute to differences in values between the two devices. While the TM measures a D-shaped arm that is not loaded axially by gravity, the 3D scan is axially loaded by gravity. In addition, differences in detection of anatomic landmarks may influence agreement. For example, the lowest Pearson correlation was for the 0 cm location. While a researcher can palpate for the ulnar prominence when using a TM, the 3D scan software parameters search for the wrist. The algorithm-guided scanner may not be able to reliably identify this landmark if it is obscured by swelling.[20] As noted in a previous study that identified similar discrepancies with landmark acquisition,[10] the authors suggested that software updates were needed to resolve these issues.

For both methods, the total volumes, using the equation for a truncated cone, showed significant differences and large SDs. Consistent with our CM values and previous reports,[2,3] the 3D scan volumes were larger than those calculated for the TM. However, the mean differences in volumes and volume ratios were not different between the two methods (Table 4). These findings suggest that given the challenge with landmark identification, mean volume differences or volume ratios may be more appropriate values to use for comparative purposes than the total volumes. However, our Bland–Altman analyses demonstrated large biases and wide LOAs for the calculated volumes and volume ratios. These findings suggest that the TM and 3D scan measures are not interchangeable.

### *Caseness*

Using three benchmark criteria for LE,[1] our hypothesis was that the ratio, which is not directly related to the absolute volume in ml, would show the highest percent agreement between the two devices. However, this criterion was the least favorable (62.5%). While the ≥200 mL difference between arm volumes is the highest threshold to achieve because actual arm and body size are not considered as part of the cut point, it had the highest percent agreement (81.6%). While our findings warrant replication, they suggest that comparison of ratios may not be the best criterion to use in future studies of the 3D scan.

### *Test–retest*

The wide range in 3D scan precision levels (i.e., 0.38–0.91 cm) may be related to variability in positioning between the first and second measurements and misidentification of landmarks. Consistent with a previous report,[3] while the generic-sized foot placement template and handles are fixed, the second scan may have been influenced by intraindividual variations in the survivor's repositioning of her body. As noted in another study that used an on-screen positioning template,[7] visual positioning feedback could be incorporated into commercially available devices to improve the consistency of the measurements. Equally important, software modifications are warranted to improve landmark identification.[10]

Circumferential measurements using a TM are challenging to perform. The position and tension of the TM, especially in the upper arm, are difficult to standardize. In research studies, data collection procedures need to specify the smallest increment of

**FIG. 1.** **(A)** Scatter plot of total volume of affected arm ($n = 294$) measured by TM on the x-axis versus 3D scan on the y-axis, including the line of equality. **(B)** Bland–Altman plot distribution of affected arm total volume difference scores ($n = 294$) with the mean of both TM and 3D scan methods on the x-axis and the difference scores for the two methods on the y-axis. **(C)** Scatter plot of total volume of unaffected arm ($n = 294$) measured by TM on the x-axis versus 3D scan on the y-axis, including the line of equality. **(D)** Bland–Altman plot distribution of unaffected arm total volume difference scores ($n = 294$) with the mean of both TM and 3D scan methods on the x-axis and the difference scores for the two methods on the y-axis. **(E)** Scatter plot of volume ratios (affected/unaffected) ($n = 294$) measured by TM on the x-axis versus 3D scan on the y-axis, including the line of equality. **(F)** Bland–Altman plot distribution of volume ratio (affected/unaffected) difference scores ($n = 294$) with the mean of both TM and 3D scan methods on the x-axis and the difference scores for the two methods on the y-axis. 3D, three-dimensional; TM, tape measure.

measurement and conduct inter-rater reliability testing on a periodic basis. In this study, research nurses did two measurements on both arms. If any value was substantially different from the previous value, a third measurement was done. The higher correlations and precision values for repeated measurements using the TM may be related to these rigorous study procedures.

*Limitations*

While arm dominance was used as a covariate in some studies, we did not account for this variable in our analyses. However, analyses for both methods were consistent. While we did not have direct access to the 3D scan software, the

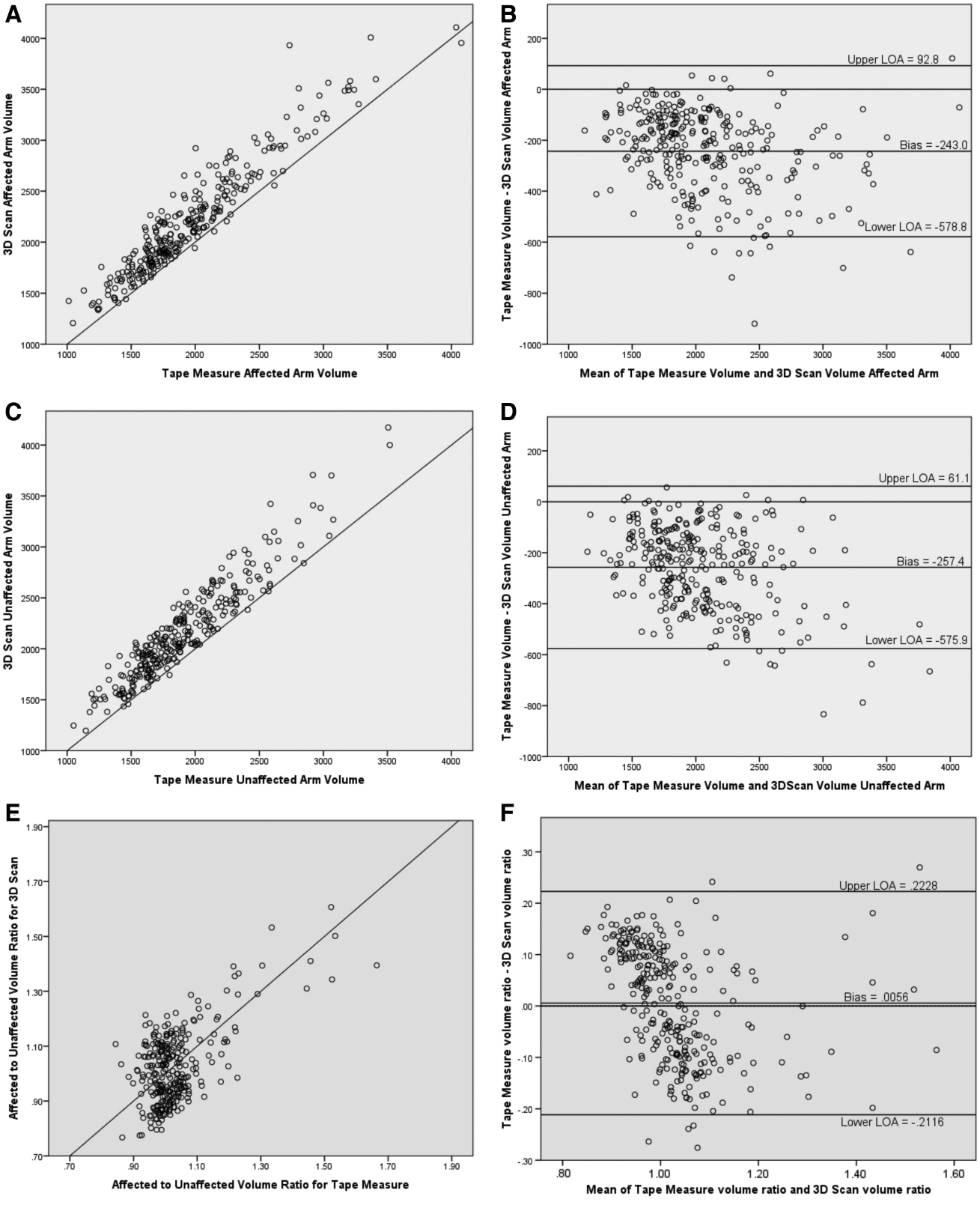

Table 5. Percent Agreement Between the Tape Measure and Three-Dimensional Scan to Determine Lymphedema Cases Using Three Clinically Meaningful Thresholds

| | *3D Scan* | | | | | |
|---|---|---|---|---|---|---|
| | *YES (*n, *%)* | *NO (*n, *%)* | *Total* | *Percent agreement* | *McNemar* p-*value* | *Kappa coefficient* |
| ≥200 mL volume threshold | | | | | | |
| TM | | | | | $p < 0.001$ | 0.406 |
| YES (%) | 27 (9.2%) | 7 (2.4%) | 34 (11.6%) | Yes 27/294 = 9.1% | | |
| NO (%) | 47 (16.0%) | 213 (72.4%) | 260 (88.4%) | No 213/294 = 72.4% | | |
| Total | 74 (25.2%) | 220 (74.8%) | 294 (100%) | Total 240/294 = 81.6% | | |
| 10% volume difference threshold | | | | | | |
| TM | | | | | $p < 0.001$ | 0.359 |
| YES (%) | 28 (9.5%) | 9 (3.1%) | 37 (12.6%) | Yes 28/294 = 9.5% | | |
| NO (%) | 54 (18.4%) | 203 (69.0%) | 257 (87.4%) | No 203/294 = 69.0% | | |
| Total | 82 (27.9%) | 212 (72.1%) | 294 (100%) | Total 231/294 = 78.5% | | |
| ≥1.04 volume ratio threshold | | | | | | |
| TM | | | | | $p < 0.001$ | 0.189 |
| YES (%) | 46 (15.6%) | 34 (11.6%) | 80 (27.2%) | Yes 46/294 = 15.6% | | |
| NO (%) | 76 (25.9%) | 138 (46.9%) | 214 (72.8%) | No 138/294 = 46.9% | | |
| Total | 122 (41.5%) | 172 (58.5%) | 294 (100%) | Total 184/294 = 62.5% | | |

Table 6. Test–Retest Reliability Results for Two Repeated Measures Using the Tape Measure and Three-Dimensional Scan ($N = 50$)

| *Arm measure location (cm)* | *Measure 1 Mean (SD)* | *Measure 2 Mean (SD)* | *Measure 1–Measure 2 Mean (SD)* | t-*Test* p-*value* | r | *Precision (cm)* |
|---|---|---|---|---|---|---|
| 3D scan measurement | | | | | | |
| Left 0 | 18.77 (1.58) | 19.03 (1.53) | –0.26 (0.88) | 0.04 | 0.84 | 0.91 |
| Left 10 | 21.55 (1.98) | 21.81 (1.87) | –0.26 (0.82) | 0.03 | 0.91 | 0.85 |
| Left 20 | 26.51 (2.47) | 26.81 (2.36) | –0.30 (0.83) | 0.01 | 0.94 | 0.88 |
| Left 30 | 29.02 (3.69) | 29.26 (3.69) | –0.24 (0.70) | 0.19 | 0.98 | 0.74 |
| Left 40 | 32.67 (3.89) | 32.87 (3.84) | –0.20 (0.88) | 0.18 | 0.97 | 0.90 |
| Right 0 | 18.28 (2.05) | 18.33 (2.36) | –0.05 (0.61) | 0.57 | 0.97 | 0.61 |
| Right 10 | 21.05 (2.08) | 21.12 (1.97) | –0.07 (0.51) | 0.33 | 0.97 | 0.51 |
| Right 20 | 26.00 (2.64) | 26.01 (2.56) | –0.06 (0.38) | 0.29 | 0.99 | 0.38 |
| Right 30 | 28.49 (3.99) | 28.51 (3.85) | –0.18 (0.49) | 0.80 | 0.99 | 0.49 |
| Right 40 | 32.34 (4.06) | 32.30 (4.04) | 0.04 (0.51) | 0.56 | 0.99 | 0.51 |
| TM measurements | | | | | | |
| Left 0 | 15.70 (0.86) | 15.67 (0.88) | 0.02 (0.19) | 0.56 | 0.98 | 0.22 |
| Left 10 | 20.92 (1.89) | 20.83 (1.89) | 0.08 (0.30) | 0.06 | 0.99 | 0.34 |
| Left 20 | 24.50 (2.02) | 24.42 (1.98) | 0.08 (0.24) | 0.03 | 0.99 | 0.38 |
| Left 30 | 27.01 (3.07) | 26.98 (3.15) | 0.02 (0.35) | 0.63 | 0.99 | 0.46 |
| Left 40 | 31.13 (3.70) | 31.04 (3.56) | 0.09 (0.41) | 0.11 | 0.99 | 0.40 |
| Right 0 | 15.90 (1.12) | 15.88 (1.05) | 0.02 (0.22) | 0.45 | 0.98 | 0.19 |
| Right 10 | 21.26 (2.41) | 21.20 (2.40) | 0.07 (0.34) | 0.18 | 0.99 | 0.31 |
| Right 20 | 24.90 (2.53) | 24.84 (2.54) | 0.06 (0.38) | 0.26 | 0.98 | 0.25 |
| Right 30 | 27.20 (3.52) | 27.05 (3.42) | 0.16 (0.44) | 0.01 | 0.99 | 0.35 |
| Right 40 | 31.23 (3.67) | 31.19 (3.76) | 0.04 (0.40) | 0.46 | 0.99 | 0.42 |

developers modified the software parameters to obtain the same circumferential measurements that were done using the TM. At the current time, these values would not be available for women who use this 3D scanner in the community. For the 3D scan, total arm volumes were calculated using the data obtained with the current algorithms and not using the formula for a truncated cone. Changes in the algorithms may improve this calculation.

## Conclusions

The 3D technology evaluated in this study has great potential for breast cancer survivors to self-initiate surveillance for LE. With improvements in landmark identification and software modifications, it is possible that accurate and reliable total arm volumes can be calculated and used for early detection. However, for the 3D scanner to be clinically useful, the software parameters need to provide more precise total arm volumes. This goal can be achieved with the addition of incremental arm measurements and more specific positioning on the scanner platform.

## Acknowledgments

Recruitment was facilitated by the Dr. Susan Love Research Foundation Army of Women® Program. C.M. is an American Cancer Society Clinical Research Professor.

## Disclaimer

The contents of this article are solely the responsibility of the authors and do not necessarily represent the official views of the NIH.

## Author Disclosure Statement

No competing financial interests exist.

## Funding Information

This study was funded by a grant from the National Cancer Institute (NCI, CA187160). This project was supported by the National Center for Advancing Translational Sciences, National Institutes of Health, through UCSF-CTSI Grant No. UL1 TR000004.

## References

1. Levenhagen K, et al. Diagnosis of upper quadrant lymphedema secondary to cancer: Clinical Practice Guideline From the Oncology Section of the American Physical Therapy Association. Phys Ther 2017; 97:729–745.
2. Erends M, et al. Validity and reliability of three-dimensional imaging for measuring the volume of the arm. Lymphat Res Biol 2014; 12:275–281.
3. Hameeteman M, et al. 3D stereophotogrammetry in upper-extremity lymphedema: An accurate diagnostic method. J Plast Reconstr Aesthet Surg 2016; 69:241–247.
4. Deltombe T, et al. Reliability and limits of agreement of circumferential, water displacement, and optoelectronic volumetry in the measurement of upper limb lymphedema. Lymphology 2007; 40:26–34.
5. Ohberg F, Zachrisson A, Holmner-Rocklov A. Three-dimensional camera system for measuring arm volume in women with lymphedema following breast cancer treatment. Lymphat Res Biol 2014; 12:267–274.
6. Buffa R, et al. A new, effective and low-cost three-dimensional approach for the estimation of upper-limb volume. Sensors (Basel) 2015; 15:12342–12357.
7. White BN, et al. An infrared 3D scanning device as a novel limb volume measurement tool in breast cancer patients. World J Surg Oncol 2020; 18:278.
8. Shah C, et al. The impact of early detection and intervention of breast cancer-related lymphedema: A systematic review. Cancer Med 2016; 5:1154–1162.
9. Rafn BS, et al. Self-measured arm circumference in women with breast cancer is reliable and valid. Phys Ther 2019; 99: 240–253.
10. Ng BK, et al. Clinical anthropometrics and body composition from 3D whole-body surface scans. Eur J Clin Nutr 2016; 70:1265–1270.
11. Smoot B, et al. Differences in limb volume trajectories after breast cancer treatment. J Cancer Surviv 2016; 10:772–782.
12. Miaskowski C, et al. Lymphatic and angiogenic candidate genes predict the development of secondary lymphedema following breast cancer surgery. PLoS One 2013; 8:e60164.
13. IBM SPSS for Windows (Version 23). Armonk, NY: SPSS, Inc.; 2015.
14. Smoot BJ, Wong JF, Dodd MJ. Comparison of diagnostic accuracy of clinical measures of breast cancer-related lymphedema: Area under the curve. Arch Phys Med Rehabil 2011; 92:603–610.
15. Szaflarski NL, Slaughter RE. Technology assessment in critical care: Understanding statistical analyses used to assess agreement between methods of clinical measurement. Am J Crit Care 1996; 5:207–216.
16. Landis JR, Koch GG. The Measurement of Observer Agreement for Categorical Data. Biometrics 1977; 33:159–174.
17. Altman D. *Practical Statistics for Medical Research*. Boca Raton: Chapman & Hall/CRC; 1999.
18. Binkley JM, et al. Assessing arm volume in people during and after treatment for breast cancer: Reliability and convergent validity of the LymphaTech System. Phys Ther 2020; 100:457–467.
19. Daanen HAM. 3D whole body scanners revisited. Displays 2013:270–2756.
20. Verhulst AC, et al. Accuracy and reproducibility of a newly developed tool for volume measurements of the arm using 3D stereophotogrammetry. J Plast Reconstr Aesthet Surg 2017; 70:1753–1759.

Address correspondence to:
*Christine Miaskowski, RN, PhD*
*School of Nursing*
*University of California*
*2 Koret Way—N631Y*
*San Francisco, CA 94143-0610*
*USA*

*E-mail:* chris.miaskowski@ucsf.edu