The Breast 28 (2016) 121–129

Contents lists available at ScienceDirect

The Breast

journal homepage: www.elsevier.com/brst

Review

The accuracy of breast volume measurement methods: A systematic review

S.B. Choppin [a, *], J.S. Wheat [a], M. Gee [b], A. Goyal [c]

[a] Centre for Sports Engineering Research, Faculty of Health and Wellbeing, Sheffield Hallam University, Sheffield S10 2NA, UK
[b] Centre for Health and Social Care Research, Faculty of Health and Wellbeing, Sheffield Hallam University, Sheffield S10 2BP, UK
[c] Department of Surgery, Royal Derby Hospital, Derby DE22 3NE, UK

ARTICLE INFO

Article history:
Received 2 March 2016
Received in revised form
20 April 2016
Accepted 21 May 2016
Available online 9 June 2016

Keywords:
Volume measurement
Systematic review
Accuracy
Mammography
MRI
3D imaging

ABSTRACT

Breast volume is a key metric in breast surgery and there are a number of different methods which measure it. However, a lack of knowledge regarding a method's accuracy and comparability has made it difficult to establish a clinical standard. We have performed a systematic review of the literature to examine the various techniques for measurement of breast volume and to assess their accuracy and usefulness in clinical practice. Each of the fifteen studies we identified had more than ten live participants and assessed volume measurement accuracy using a gold-standard based on the volume, or mass, of a mastectomy specimen. Many of the studies from this review report large (>200 ml) uncertainty in breast volume and many fail to assess measurement accuracy using appropriate statistical tools. Of the methods assessed, MRI scanning consistently demonstrated the highest accuracy with three studies reporting errors lower than 10% for small (250 ml), medium (500 ml) and large (1000 ml) breasts. However, as a high-cost, non-routine assessment other methods may be more appropriate.

Introduction

Breast volume has been identified as a key metric [1] in breast surgery [2,3]. As a clinically relevant [4] objective measure, knowledge of breast volume helps a surgeon to select protocols, choose appropriate implant sizes [5] and achieve breast symmetry [6]. It can be used to plan aesthetic [3,7–13] and breast conservation surgeries [7,9,12,14–19] and in the diagnosis of breast oedema [20,21]. The use of breast volume may lead to reductions in repeat surgeries (around 1 in 3 women in the UK are not satisfied with aesthetic outcome [22]) and better diagnosis of breast oedema.

The breast is a three-dimensional (3D) structure and difficult to assess accurately. Variations in patient pose [23], breast shape [2] and in identifying the breast boundary [24] (external and posterior chest wall) cause variability in volume measurement.

Several methods have been proposed to assess breast volume through the use of medical imaging technology [25], devices based on geometric measurement [26], water displacement techniques [27–29] and breast casts [30,31]. There is no 'accepted' technique for measurement of breast volume due to a lack of clear information regarding the accuracy and comparability of each method. This has limited the use of breast volume measurement methods in routine clinical practice. Large errors negatively impact a surgeon's ability to determine, for example, the appropriate size of breast implant or the quantity of tissue to be removed. In addition, ease of use, cost and complexity [12] cannot be dismissed.

Many advocates of particular methods of volume measurement describe them as 'accurate' without assessing or quantifying error [26,27,29,30]. Several authors have, however, made comparisons to determine accuracy. We performed a systematic review of the literature to examine the various techniques for measurement of breast volume and to assess their accuracy and usefulness in clinical practice.

Two other systematic reviews which assess breast volume measurement have been identified. Xi et al. [46] reviewed methods of breast measurement (volume, shape and surface area) with regards to cost, suitability and accuracy. However, accuracy was not dealt with in detail (focussing on reliability as the coefficient of variation) and papers were not excluded based on the quality of gold-standard. O'Connell et al. [47] focused on 3D surface imaging methods used in breast volume assessment (referred to as 3D

* Corresponding author. Tel.: +44 0114 225 5717.
E-mail address: s.choppin@shu.ac.uk (S.B. Choppin).

http://dx.doi.org/10.1016/j.breast.2016.05.010
0960-9776/

scanning in this study). While accuracy is discussed, other methods of volume measurement are not considered.

Due to the difficulties in finding a consistent method of assessing accuracy, Xi et al. proposed to assess a method's potential for accurate measurement by its definition of the external breast boundary and the internal posterior wall. According to this assessment they identified 3D scanning, MRI and CT methods as being most 'accurate'.

In this review we have focused on accuracy with regards to error and uncertainty of measurement. By identifying an established gold-standard, obtaining studies' data and performing data simulation we have been able to numerically quantify the error and uncertainty of eight different methods of breast volume measurement.

Methods

Search strategy

A search strategy combining the title/abstract words ("breast volume" OR "breast shape") in proximity to the title/abstract words (measur* OR accura* OR valid* OR estimat*) was used to locate all papers relating to breast volume measurement. The searches were run in the following databases from their inception to Dec, 2014: CINAHL Plus with Fulltext (via EBSCOHost), The Cochrane Library (Cochrane Database of Systematic Reviews, Cochrane Central Register of Controlled Clinical Trials (CENTRAL)), Database of Abstracts of Reviews of Effects (DARE), Cochrane Methodology Register (CMR), Health Technology Assessment Database (HTAD), and NHS Economic Evaluation Database (NHS EED), Embase (via NHS Evidence Search), IEEE Xplore, Medline (via EBSCOHost), Scopus (Elsevier), SPORTDiscus (via EBSCOHost), Web of Science (all Databases).

Additionally, the British Library Main Catalogue (http://explore.bl.uk/) was searched using the strategy: Main Title contains "breast volume" OR "breast shape" AND Abstract contains accura* OR measur* OR valid* OR estimat*/All materials, all dates.

We checked the reference lists from eligible studies to identify further relevant studies. Citation (forwards) searches were also carried out on included studies to identify further relevant studies.

Data collection

All citations were organised in RefWorks software and duplicates were removed. Remaining records were compiled into a Microsoft Access database within which remaining study selection was independently performed by SC and JW.

In the title/abstract screening, records were included if it was evident that breast volume was measured and a measure of accuracy/validity was taken. The remaining records proceeded to full review, where the papers were judged against the following inclusion criteria:

1. Accuracy of breast volume is assessed
2. At least ten participants included in the study
3. Human participants used (no mannequins or phantom breast objects)
4. Published, peer-reviewed study
5. A suitable 'gold-standard' was used – the volume or mass of resected breast tissue (or fluid)

All studies which met the inclusion criteria proceeded to data extraction.

Data extraction

Information was extracted independently by SC and JW using predefined fields. Specifically, the type of gold-standard, number of participants, range of data, the comparator breast volume measurement method(s), the statistical method(s) of assessing accuracy and the associated value(s).

Because the objective of the review was to establish the accuracy of volume measurement, additional raw data were sought using three methods: directly from the publication (full disclosure in a results table), directly from the authors (for publication dates within 10 years) and extraction from published figures (when possible, using image processing). Two authors published the entirety of their data [32,33] in data tables, which was used in further statistical analyses. Two authors sent us their original datasets [12,34] and in two studies [35,36] data was obtained by calculating the position of plotted data points using image processing techniques. The centroids of each data point were obtained from digital images (in image co-ordinates) and these were transformed to scaled data values by calculating the scale from the X and Y axes.

Data analysis

As an intuitive representation of a method's accuracy we used a linear regression to calculate expected error at three different breast volumes: 250, 500 and 1000 ml – representing the typical range from studies in this review. The uncertainty of each measurement at these values was not assessed – it was not given in the majority of cases.

For raw data obtained, we performed a Bland–Altman analysis [37] to calculate the limits of agreement and linear regression in order to model measurement error at different breast sizes. In cases where proportional error was apparent, the data were de-trended prior to Bland–Altman analysis, in cases where heteroscedasticity was apparent, the Bland–Altman data were processed as percentage values (and is presented as such).

A large number of studies reported correlation coefficients. To gain further insight we used Monte-Carlo simulation to estimate measurement uncertainty. We used the r value, number of samples and range of the data in the following way. Assuming error was normally distributed and homoscedastic, we created 1000 randomly generated datasets *for each study*. Each dataset had the same number of data points and nominal-range as the study it represented. In each case we adjusted the standard deviation of the error until the Pearson's r of each simulated data-set matched reported values (results are presented in Table 2). This gave an estimate of the 95% confidence intervals of each measurement.

At all stages, any disagreements were discussed and consensus reached through a discussion/investigation of the literature.

Results

See Fig. 1 for a document flow chart. The database searches yielded 701 records, and 238 unique records after removing duplicates. After title/abstract screening, 71 records proceeded to full-text screening, from which 13 records met the inclusion criteria. A further 2 records were identified from reference and citation searches, resulting in 15 studies for this review.

Table 1 summarises the results presented in the studies included in this review. The measurement method, and studies which assessed it, are identified. For each study we present the gold standard used (volume or mass), the size of the study (n) and all available information regarding accuracy. Accuracy information is split into three categories: mean error ± 1.96 standard deviations, expected errors at 3 breast sizes (250, 500 and 1000 ml) and a correlation coefficient *r*.

Table 2 presents the results from our data simulation of *r* values for measurement methods assessed in different studies. Simulated uncertainty (95% uncertainty bounds) is shown with the values used to calculate them: number of samples, range of values and Pearson's *r*.

Methods of measurement

The following measurement methods were identified from included studies: anthropometric models [12,13], immersion in water (referred to as 'Archimedes' [12]), breast casts [12,32], the Grossman–Roudner cone [12], computed tomography [38], magnetic resonance imaging [32,34,39], mammography [11,12,36,40,41] and 3D surface scanning [33,35,42–44].

Assessing error

In all but two of the reviewed studies [36,42], measuring accuracy or assessing a breast-volume measurement tool was a primary objective of the study. A variety of statistical measures were used to assess agreement between volume measurement methods, with differences between studies. Eight studies [11,36,38–41,43,44] reported Pearson's correlation *r* (Kayar et al.'s [12] *r* was calculated using intraclass correlation), six [13,32,34,35,39,42] reported coefficient of determination r^2, seven [11,35,39–42,44] reported regression equations and three studies [39,43,44] used Bland–Altman (Table 1).

Fig. 2 illustrates that the magnitude of uncertainty of a method often exceeds its mean error. The values in this figure come directly from Bland–Altman's limits of agreement or mean differences. Of the measurement methods used in this review, 3D scanning and mammography were the most favoured with five studies in both cases. It is worth noting that three 3D scanning studies assess differential, as opposed to absolute, volumes and one of the mammography studies [41] provided *r* values only, making in-depth analysis difficult.

In some cases, assessing correlation with the gold standard was the only statistical tool used [13,36,38]. In these studies, 'good' correlation (e.g. Pearson's $r > 0.9$) was used to justify the use of the measurement technique.

Two studies assessed measurement repeatability with multiple raters. Mailey et al. [33] had two raters independently assess the change in volume of the left and right breast after aesthetic breast implants reporting an intraclass correlation coefficient of 0.97 and 0.95 respectively. Losken et al. [43] used two raters to measure breast volume prior to mastectomy and reported intraclass correlation of 0.97. The measurement method in both of these studies was optical 3D scanning.

Gold standard measures

We deemed a 'gold standard' comparator to be the mass or volume of removed tissue or fluid. Of the fifteen included studies, ten used a volume-based gold standard while 5 used mass.

Studies which used a volume-based gold standard took three forms. Seven studies used the Archimedes method to measure

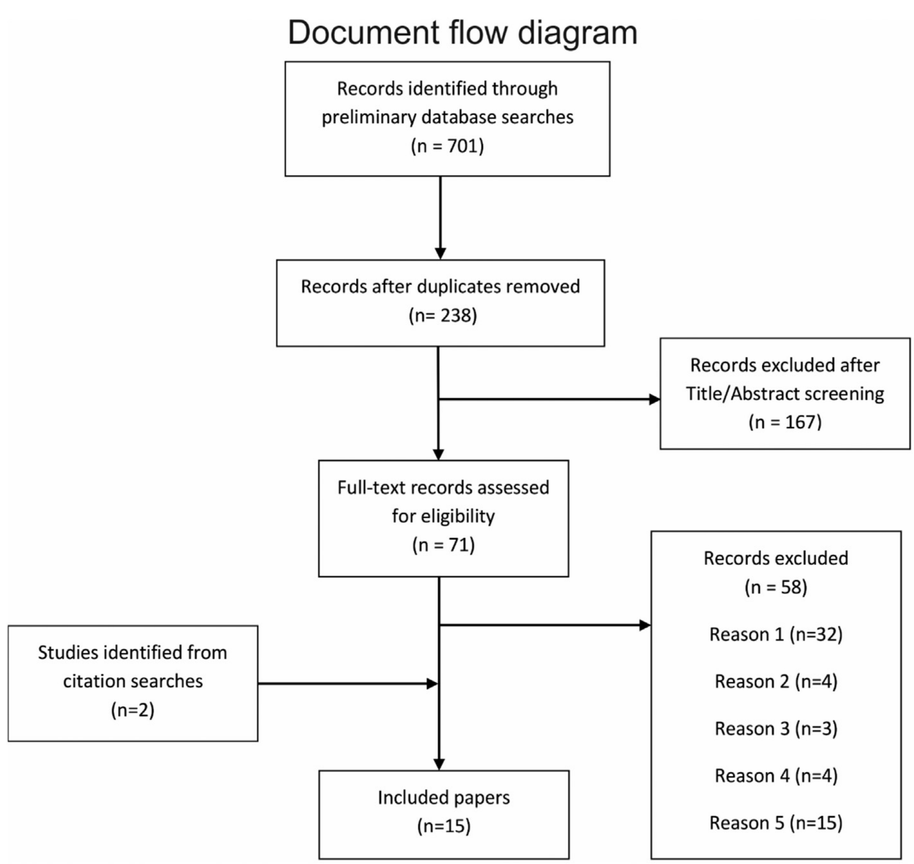

Fig. 1. From 701 identified records, 15 were eligible for full review. Ninety-four records were excluded for the following reasons. **Reason 1**: The study did not include an assessment of breast volume. **Reason 2**: The study used fewer than 10 participants. **Reason 3**: The study did not use human participants. **Reason 4**: The study was not peer-reviewed. **Reason 5**: The study did not use a sufficient gold-standard.

Table 1
A summary of the error information reported in each study. Instead of regression information the expected error at three breast sizes (250, 500, 1000 ml) is reported.

Measurement method	Study	Standard	N	Expected error (ml)				Corr/Det (r/**r²**)
				Mean error ± 1.96 SD[a]	Small (250 ml)	Medium (500 ml)	Large (1000 ml)	
Anthropometric	Kayar et al. [12]	V	30	22.0 ± 151.4	−5.6/−2.2%	−11.8/2.4%	46.7/4.7%	0.975
	Longo et al. [13]	M	108	18.4% ± –	–	–	–	**0.73**
Archimedes	Kayar et al. [12]	V	30	−39.7 ± 27.2	−12.5/−5.0%	−32.8/−6.6%	−73.6/−7.4%	0.989
Casting	Rha et al. [32]	M	20	−61.4 ± 311.0	−119.16/−47.7%	−40.6/−8.1%	116.5/11.6%	0.629
	Kayar et al. [12]	V	30	−78.8 ± 83.7	−24.93/−10.0%	−70.7/−14.1%	−162.2/−16.2%	0.946
Grossman–Roudner cone	Kayar et al. [12]	V	30	−7.1% ± 20%	−8.9/−3.6%	−47.5/−9.5%	−124.7/−12.5%	0.934
Computed tomography	Fuji et al. [38]	V	11	29.2 ± 28.1	–	–	–	0.985
Magnetic resonance imaging	Liu et al. [39]	V	46	57.8 ± 93.1	26.5/10.6%	47.8/9.6%	90.3/9.0%	0.936
			27	59.2 ± 110.7	−18.5/−7.4%	18.2/3.6%	91.7/9.2%	
	Yoo et al. [34]	M	101	−11.2 ± 141.4	−19.3/−7.7%	8.8/1.8%	65.0/6.5%	0.880
	Rha et al. [32]	M	20	33.8 ± 124.6	17.2/6.9%	32.0/6.4%	61.6/6.2%	0.945
Mammography	Katariya et al. [40]	V	30	31.5 ± *17.9*	−39.0/−15.6%	−34/−6.8%	−24/−2.4%	0.975
	Hoe et al. [36]	V	17	−173.6 ± 439.0	−146.4/−58.6%	−155.8/−31.2%	−174.6/−17.5%	0.930
Method 1	Kalbhen et al. [11]	M	36	–	21.0/8.4%	180.0/36%	498.0/49.8%	0.924
Method 2		M	36	–	53.5/21.4%	137.0/27.4%	304.0/30.4%	0.920
Method 3		M	36	–	–	–	–	0.926
Method 4		M	36	–	−59.5/−23.8%	−62.0/−12.4%	−67.0/−6.7%	0.938
Method 5		M	36	–	−51.3/−20.5%	−5.5/−1.1%	86/8.6%	0.896
Method 6		M	36	–	−39.0/−15.6%	22.0/4.4%	144.0/14.4%	0.931
EBV-E	Fung et al. [41]	M	83	–	–	–	–	0.977
EBV-C		M	83	–	–	–	–	0.952
	Kayar et al. [12]	V	30	−7.7 ± 52.7	−17.6/−7.1%	−10.8/−2.2%	2.8/0.3%	0.997
3D scanning	Daly & Hartmann [42]	Vmlk[b]	257	–	−27.5/−11.0%	−62.5/−12.5%	−132.5/−13.2%	**0.86**
	Losken et al.[c] [43]	V	19	−2.6% ± 13%	–	–	–	–
				−1.8% ± 16%				
	Yip et al. [44]	V	39	46.7 ± 353.3	63.1/25.2%	55.6/11.1%	40.6/4.1%	0.950
	Mailey et al.[d] [33]	Vimp[b]	22	−5.4% ± 27%	–	–	–	0.665
				−6.2% ± *31%*				0.646
	Lewis et al. [35]	Vliq[b]	17	38.3 ± 87.2	–	–	–	**0.219**

Gold Standards: V = volume of removed tissue, Vmlk = volume of expressed milk, Vimp = volume of inserted implant, Vinj = volume of injected liquid, M = mass of removed tissue. The bold figures correspond to r^2 values.
[a] All confidence intervals represent 1.96 standard deviations of measurement other than italicized values which represent standard error of measurement.
[b] Volume measured was differential, not absolute.
[c] Results given by two independent raters.
[d] Left and right breast respectively.

the volume of tissue excised from the breast [12,36,38–40,43,44], one study used the volume of a surgical implant [33] and two studies measured the volume of milk expressed from [42] or liquid injected into [35] the breast. It is worth noting that these studies [33,35,42] measured differential rather than absolute volume.

In studies where the mass of tissue was measured, it was converted to volume by assuming a density value(s) of the removed tissue. Of the five authors using this method, two assumed a constant density of 0.958 g/cm^3 [32,41], two didn't convert mass to volume [13,34], and one [11] used a radiologist to assess individual proportions of fat/water in the breast (based on the classification method of Stomper et al. [45]), varying the density accordingly (range 0.916 g/cm^3 to 1.000 g/cm^3).

Discussion

Methods of measurement

Table 3 gives a summary of our findings from this review which is discussed in more detail below according to the specific method of measurement.

Mammography

Volume measurement in mammography uses geometric formulae to transform measurements made from the mammogram into 3D shapes. Kalbhen et al. [11] tested 6 different models and showed that the choice of formula has a significant effect on error, which in the worst cases is nearly 50%. The earliest use of mammograms to measure breast volume was by Katariya et al. [40] who used a conical breast model. Other authors have assessed the same model and the nature of error has varied between studies. Hoe et al. [36] and Kalbhen et al. [11] (given as method 2 in Tables 1 and 2) found marked proportional error, while Katariya et al. [40] demonstrated it to be predominantly systematic. Fung et al. [41] assessed the same method but reported *r* values and an univariate regression model only. Kalbhen et al. posited that modern mammography techniques rendered this model of volume calculation redundant due to higher pressures on the breast during scanning (violating the assumption of a conical breast shape). Fung et al. [41] suggested that the base of large breasts could be missed from the view – leading to an underestimation of breast volume. A modified model of breast measurement was proposed by Kalbhen et al. based on a half-elliptical cylinder (given as method 4 in Tables 1 and 2) – also adopted by Kayar et al. [12]. Kalbhen et al.

Table 2
A summary of the size, range and 'simulated uncertainty' associated with each study (where possible) a Monte-Carlo simulation was used to approximate the 95% confidence intervals of errors representative of Pearson's *r* coefficient.

Measurement method	Study	Standard	N	Range of data (ml)	r	Simulated uncertainty (ml)
Anthropometric	Kayar et al. [12]	V	30	150–1490	0.975	±180.25
	Longo et al. [13]	M	108			
Archimedes	Kayar et al. [12]	V	30	150–1490	0.989	±118.12
Casting	Rha et al. [32]	M	20	201–910	0.629	±550.76
	Kayar et al. [12]	V	30	150–1490	0.946	±272.31
Grossman–Roudner cone	Kayar et al. [12]	V	30	150–1490	0.934	±300.43
Computed tomography	Fuji et al. [38]	V	11	180–700	0.985	±57.94
Magnetic resonance imaging	Liu et al. [39]	V	46	180–1300	0.938	±243.94
			27	225–1400	0.936	±241.31
	Yoo et al. [34]	M	101	88–743	0.880	±202.16
	Rha et al. [32]	M	20	201–910	0.945	±147.28
Mammography	Katariya et al. [40]	V	30	125–1125	0.975	±134.30
	Hoe et al. [36]	V	17	350–2120	0.930	±743.75
Method 1	Kalbhen et al. [11]	M	36	100–2300	0.924	±529.98
Method 2		M	36	100–2300	0.920	±551.41
Method 3		M	36	100–2300	0.926	±530.20
Method 4		M	36	100–2300	0.938	±474.53
Method 5		M	36	100–2300	0.896	±645.98
Method 6		M	36	100–2300	0.931	±511.28
EBV-E	Fung et al. [41]	M	83	100–1300	0.977	±150.01
EBV-C		M	83	100–1300	0.952	±222.20
	Kayar et al. [12]	V	30	150–1490	0.997	±60.70
3D scanning	Daly & Hartmann [42]	Vmlk†	257	–		
	Losken et al. [43]	V	19	200–1200	–	
	Yip et al. [44]	V	39	158–2612	0.950	±471.13
	Mailey et al. [33]	Vimp†	22	–	0.646	
					0.665	
	Lewis et al. [35]	Vliq†	17	–		

† Volume measured was differential, not absolute

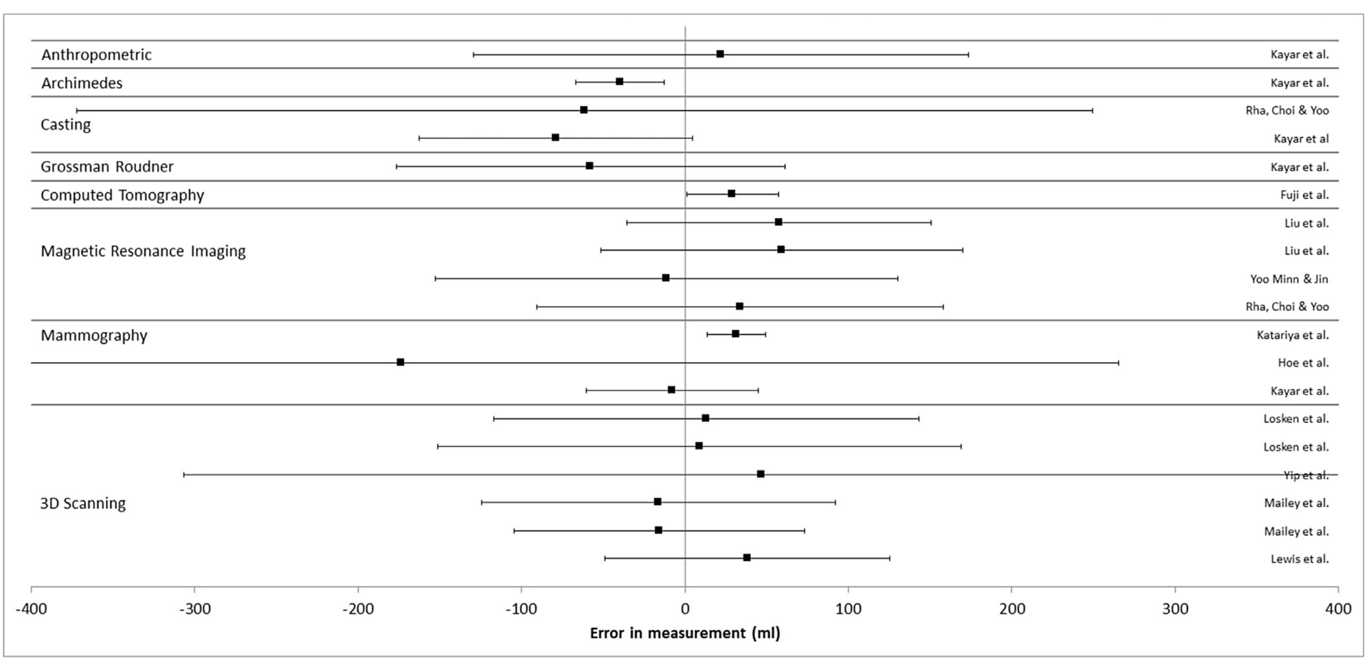

Fig. 2. A summary of the systematic error and uncertainty associated with different methods of breast volume measurement.

suggest that this model largely eliminates the proportional error reported for the earlier model but it still contains large amounts of random error according to the given r value, Table 2. Kayar et al. [12] report much lower errors and uncertainties associated with this model (errors in the region of 20 ml rather than 60 ml and uncertainties of around ±60.0 ml as opposed to ±470.0 ml). It is not clear why the reported differences in error are so large, but it is possible the error analysis of Kalbhen et al. was affected by the much larger range of tested breast sizes (100–2300 ml as opposed to 150–1490 ml).

Three-dimensional scanning

Three-dimensional scanning systems are relatively new (Daly and Hartmann's study [42] is the earliest and dates from 1995) and use a wide variety of underlying technologies (O'Connell et al.'s review gives a more detailed appraisal of this technology [47]). Daly and Hartmann [42] used Moiré topography [48] to obtain surface scans which is limited in view-range and only suitable for women with particular breast shape/size. Yip et al. [44] used a full-body Cyberware WBX (Cyberware, USA) scanning system which utilised four laser-based scanning heads with data capture taking 15 s. The lengthy data capture process could explain the high uncertainties associated with this study. The likelihood of postural sway over this time is high and the amount of sway will vary randomly between participants. Losken et al. [43], Lewis et al. [35] and Mailey et al. [33] use cameras-based scanning technologies in which the distortion of a projected texture pattern (3DMD systems, GA, USA, used by Losken and Lewis) or the differential hues from a light source (Portrait 3D, axisthree, FL, USA, used by Mailey) are used to discern surface information. These systems tend to have very short capture times (3DMD systems capture in around 50 ms) minimising the effects of postural sway. Of the three studies using camera based scanning systems, Losken et al. measure absolute breast volume while Mailey et al. and Lewis et al. report differential volumes. The absolute volume measurements exhibit lower systematic errors (below 3%) but higher uncertainties (13% and 16% for separate raters) than the studies assessing differential volume. With surface scanning systems, the difficulty in measuring absolute volume is in defining the invisible chest wall. When assessing differential volumes, the requirement to define a chest wall disappears and the uncertainty associated with this process reduces accordingly.

Computed tomography

Fuji et al. demonstrated that computed tomography (CT) scanning has low systematic and random error 29.2 ± 28.1 ml, suggesting that this method is the most accurate of the tested methods. However, as the only study to assess CT scanning it has a small sample size (11) and data range (180–700 ml) making a definitive conclusion difficult. Furthermore, it is not pragmatic because of irradiation exposure to the women.

Magnetic resonance imaging

The (absolute) systematic error of MRI volume measurement is lower than 60 ml in all studies with uncertainty lower than ±150 ml – the most accurate method in this review featured in multiple studies. An advantage of MRI (and CT) is the ability to delineate the breast region internally and externally through techniques such as 3D maximum intensity projection [39,49]. MRI avoids the X-ray dosage of CT scanning, but the process is expensive [50] and performed only for a small number of women in routine practice.

Low-cost methods

This review featured several clinical measurement techniques which are simple to implement and low-cost [12]: plaster cast measurement, Archimedes (water displacement), the Grossman–Roudner cone and anthropometric models. All of these methods were assessed in a single study [12], only plaster casting [32] and anthropometric measurement [13] were assessed by other authors. The advantages of these techniques are simplicity of the approach, low-cost and ease by which they can be incorporated into a treatment pathway. Their disadvantages are the problems which exist in defining the breast boundary. The Grossman–Roudner cone has difficulty in including the axillary tail in larger breasts [51]. Casting methods require the breast boundary to be manually delineated on the inside of the cast's surface and the breast boundary resulting from the Archimedes method is dependent on the level of submersion of the breast (creating a flat/planar extremity).

Despite potential problems, very high accuracy was reported with the Archimedes method, for a simple, low-cost method of measurement this is perhaps surprising. Bulstrode et al. [9] have also assessed the Archimedes method, reporting poor agreement with mammography. Unfortunately, volumes from mammography

Table 3
A summary of the methods assessed in this review with regards to general accuracy and practical considerations, according to the method of measurement.

Method	Considerations	Associated accuracy
3D scanning	Many possible systems exist, shorter (<1 s) capture times are best. The method used to recreate the internal chest wall can affect accuracy.	High accuracy when measuring absolute volumes but larger uncertainties. Uncertainties are lower when measuring differential volume
Anthropometric measurement	A simple method favoured by patients and surgeons.	Good accuracy for a low-cost method but high levels of uncertainty. Only one study in this review.
Archimedes, water displacement	The method inherently assumes a planar, internal chest wall.	Very mean accuracy for a low-cost method but only one study in this review. It has shown poor agreement with mammography in previous studies.
Casting	Low-cost but time-consuming. Results can vary depending on the boundary of the breast chosen by the surgeon.	Lower accuracy and large uncertainty. Has correlated poorly with other methods (MRI, 3D scanning, anthropometric) in other studies.
Computed tomography	High cost, radiation exposure, internal/external breast delineation.	Very mean accuracy and uncertainty values. Only one study in this review
Grossman–Roudner cone	Misses the axillary tail in larger breasts.	Good accuracy for a low-cost method but high uncertainty. Only one study in this review
Magnetic resonance imaging	High cost, internal/external breast delineation.	The most accurate method in this review featured in multiple studies
Mammography	A variety of geometric formulae available, the choice has a significant effect on error. A half-elliptical cylinder is recommended for modern mammography.	Mixed reports. Errors vary greatly in magnitude and nature (proportional, systematic). The large errors in some studies may be due to the larger breast sizes assessed.

also have high (>±200 ml) associated uncertainties [11,36] making a comparison between the studies difficult.

Kovacs et al. [2] compared anthropometric measurement techniques and thermoplastic casting with 3D scanning and MRI measurement techniques. Correlation between thermoplastic casting and all other methods was poor, Pearson's *r* for, MRI, 3D scanning and anthropometric measurement was 0.762, 0.727 and 0.669 respectively. While correlation is no way to quantify error, correlations between the three other methods exceeded 0.9 in all cases, suggesting that thermoplastic casting was particularly poor when accounting for differences in breast volume.

Assessing error

This review aims to assess the error associated with breast volume measurement methods – to establish how confidently measurements can be made and which methods may be preferable. Despite 87% of studies stating accuracy assessment as a main objective, many authors continue to use the correlation coefficient as the sole analytical tool, or as justification for the accuracy of their approach. The inappropriate use of correlation to validate measurement methods has been explored by numerous authors [37,52,53]. We used a data-simulation technique to estimate measurement uncertainty for a given *r* value (Table 2) which showed a general decrease in uncertainty as *r* value increases; uncertainty only consistently drops below ±200 ml for values above 0.97.

Failing to account for the nature of the error in a measurement system can lead to misrepresentation. Yip et al. [44] reported a mean error of 46.7 ml with high uncertainty of ±353.3 ml (at 95% confidence). Visual examination of the Bland–Altman suggests heteroscedasticity – the magnitude of random error increases with the size of measurement. It is reasonable to presume that with a proper treatment of the data, the errors of Yip et al. would be much more favourable. Losken et al. [43] recognise the proportional nature of their measurement errors and present their Bland–Altman error and limits of agreement as percentage values. Yoo et al. [34] present a 'difference percentage value' as a proportion of the weight of excised tissue.

Clinical efficacy and random error

When determining 'clinical efficacy', a measurement system's random error may be of more significance than its proportional and systematic components. Two studies in this review used a linear model to correct the readings of their measurement methods [34,44]. However, when random error is large, correction is ineffective – the surgeon or clinician must decide what is acceptable [54]. Losken et al. [43] found that their 3D scanning method is within 10% of their gold-standard 80% of the time, stating 10% as a limit of clinical efficacy in this case. However, falling outside of these bounds one time in every five is unlikely to be acceptable in clinical practice. Probst et al. [55] performed an adapted Delphi consensus study to define minimum standards with regards to volume accuracy and resolution. In a consensus group that included oncoplastic breast surgeons, reconstructive surgeons and lymphoedema specialists the study suggested a minimum accuracy of ±5% and a resolution of at least 25 cc. Many of the methods in this review only achieve these standards for breasts larger than 500 cc.

Authors identified many practical limitations which contribute to random error. These include the ability of a mammogram to fully capture breast tissue [41], 3D scanning systems having problems with ptotic breasts [43] and irregularities in plaster casting [32]. There are also limitations in gold-standard methods such as a failure to account for different breast types and densities [34]. Finding consistent ways of positioning the patient and identifying the breast boundary are important areas of research that will also contribute towards improving measurement accuracy [12,32–34].

Several authors don't base the efficacy of a measurement method on quantified performance. Fuji et al. [38] stated that their tested method (3D computed tomography, 3D CT) "appears to be sufficiently accurate to provide significant clinical benefit …" however, they present no information regarding the nature of their system error or how closely 3D CT can measure volume according to a gold-standard. Other authors give similar conclusions. Hoe et al. [36] state "the method used in this study was validated against a water displacement technique and shown to be accurate" based on a single *r* value (0.93) given by 17 samples.

Accuracy gold standards

While the community accepts water displacement volumes of excised breast tissue as gold-standard, the associated uncertainties should not be dismissed. The weight and volume of breast tissue removed during mastectomy depends on the surgeon's technique and interpretation of breast boundaries, and is likely to be variable. Hughes and Lau [56] used water displacement to assess the volume of accurately machined cylinders of 168 and 336 ml. They found the error of measurement to be 6.50 and 6.00 ml (3.9 and 1.8%) respectively. Any analysis of agreement should ideally reflect this uncertainty by allowing for error in the standard and comparator (using a Bland–Altman analyses [37] or Ordinary least products [57] regression for example).

When using breast mass, the value was converted to volume using a single, or variable density value. The legitimacy of these approaches was not explored in detail although other authors have explored the relationship between breast weight and volume. Yip et al. [44] correlated the weight and volume of mastectomy specimens (in a range 158–2612 ml) finding extremely high correlation ($r = 1.00$, $p < 0.001$). This suggests that a single density value for multiple participants can be used effectively. Parmar et al. [58] compared the weight and volume of 69 breast specimens finding an average density of 1.07 g/cm^3. This anomalously high value was not directly discussed although it was stated that previous literature states the value should lie between 0.92 and 1.00 g/cm^3. Given that breast tissue is predominantly water and fat, a value over 1.00 g/cm^3 is unlikely to be correct. The ease and accuracy to which mass can be measured (i.e. on an electronic scale) may compensate for the imprecision of assuming constant density values – making it comparable to direct volume measurement as a gold standard.

Acceptability

Bulstrode et al. [9] compared five different volume measurement techniques but was excluded from this review due to a lack of a sufficient gold-standard (reason 5, Fig. 1). However, their assessment of the acceptability of different volume measurement techniques is particularly relevant to this review. They found that mammography was particularly disagreeable to patients while MRI and the Archimedes method were not favoured by surgeons. Of the tested methods, anatomical measurements were favoured by both groups (a simple to execute and non-intrusive procedure). Other authors have been eager to point out advantages of other techniques, with Kovacs et al. pointing out quick, non-contact 3D scanning is particularly agreeable to the patient. The feasibility of a measurement technique comes not only from its accuracy but whether it is accepted by those expected to use it.

There remains no consensus on the appropriate measurement method to assess breast volume. Previous studies have stated mammography as a comparator as it *was* the only method which had been compared to mastectomy specimens [9] –prior to 2005 only mammographic techniques had been assessed against mastectomy specimens. However, this review shows that while mammography is often part of a patient's treatment pathway, it has high associated uncertainty, is often not suitable for assessing larger breasts and has low patient acceptability. Given the recent interest in breast volume measurement (10 of the 16 studies in this review have been published since 2010) there is a need to ensure appropriate statistical techniques are used to assess accuracy and to establish acceptable levels of error.

Conclusions

The available methods to measure breast volume are associated with large (>±200 ml) uncertainty in breast volume, much of this may come from the variability in data acquisition rather than the method itself. Consistent patient positioning and pose and reliable methods of segmenting breast volume from the chest is essential in order to lower measurement uncertainty.

The error in measurement of breast volume is complex. It will often contain proportional error and may be heteroscedastic. Appropriate data processing should be used to ensure the accuracy of a method is best represented. Performing log-transforms and presenting values as percentages can account for types of error which change with the size of measurement.

Clinicians must agree on a useful limit of measurement error – inaccuracies can lead to poor decision making and practice. Losken et al. [43] suggest around 10% of breast volume as a useful limit. Probst et al. [55] found consensus accuracy values of ±5% and resolution of measurement of 25 cc. Some studies [12,32,34,39,43] demonstrate accuracy to within 10% of gold-standard for small, medium and large breast sizes. A 5% to 10% measurement error appears to be acceptable for decision making in practice. Of the methods assessed, MRI scanning consistently demonstrated the highest accuracy with three studies reporting errors lower than 10% for small (250 ml), medium (500 ml) and large (1000 ml) breasts. However, as a high-cost, non-routine assessment other methods may be more appropriate in clinical practice.

Conflict of Interest Statement

None declared.

Acknowledgements

The authors of this paper would like to acknowledge the authors of “Five methods of breast volume measurement: a comparative study of measurements of specimen volume in 30 mastectomy cases” and “Magnetic resonance imaging-based volumetric analysis and its relationship to actual breast weight” for providing their original data and enabling further data analysis.

References

[1] Tepper OM, Unger JG, Small KH, Feldman D, Kumar N, Choi M, et al. Mammometrics: the standardization of aesthetic and reconstructive breast surgery. Plast Reconstr Surg 2010;125:393–400. http://dx.doi.org/10.1097/PRS.0b013e3181c4966e.
[2] Kovacs L, Eder M, Hollweck R, Zimmermann A, Settles M, Schneider A, et al. Comparison between breast volume measurement using 3D surface imaging and classical techniques. Breast 2007;16:137–45. http://dx.doi.org/10.1016/j.breast.2006.08.001.
[3] Hansson E, Manjer J, Ringberg A. Reliability of plastic cups to measure breast volume. J Plast Surg Hand Surg 2014;48:254–8. http://dx.doi.org/10.3109/2000656X.2013.870908.
[4] Beattie S, Siebert J, Stevenson J, Yap L. Assessment tools for breast surgery based on 3D surface anatomy imaging. Proc 6th Int Work Digit Mammogr.. 2003. p. 530–2.
[5] Brody GS. Breast implant size selection and patient satisfaction. Plast Reconstr Surg 1981:68.
[6] Eder M, Waldenfels FV, Swobodnik A, Klöppel M, Pape A-K, Schuster T, et al. Objective breast symmetry evaluation using 3-D surface imaging. Breast 2011;21:152–8. http://dx.doi.org/10.1016/j.breast.2011.07.016.
[7] Palin WEJ, Fraunhofer JA, Smith DJJ. Measurement of breast volume: comparison of techniques. Plast Reconstr Surg 1986;77:253–4.
[8] Chou S, Chua C, Ng W, Fang W, Saw CA, Seah CS. Non-contact mammary volume measurement. Automedica 1998;16:75–81.
[9] Bulstrode N, Bellamy E, Shrotria S. Breast volume assessment: comparing five different techniques. Breast 2001;10:117–23. http://dx.doi.org/10.1054/brst.2000.0196.
[10] Bulstrode NW, Shrotria S. Prediction of cosmetic outcome following conservative breast surgery using breast volume measurements. Breast 2001;10:124–6. http://dx.doi.org/10.1054/brst.2000.0197.
[11] Kalbhen CL, McGill JJ, Fendley PM, Corrigan KW, Angelats J. Mammographic determination of breast volume: comparing different methods. Am J Roentgenol 1999;173:1643–9.
[12] Kayar R, Civelek S, Cobanoglu M, Gungor O, Catal H, Emiroglu M. Five methods of breast volume measurement: a comparative study of measurements of specimen volume in 30 mastectomy cases. Breast Cancer (Auckl) 2011;5:43–52. http://dx.doi.org/10.4137/BCBCR.S6128.
[13] Longo B, Farcomeni A, Ferri G, Campanale A, Sorotos M, Santanelli F. The BREAST-V: a unifying predictive formula for volume assessment in small, medium, and large breasts. Plast Reconstr Surg 2013;132:1–7.
[14] Clarke D, Martinez A, Cox RS. Analysis of cosmetic results and complications in patients with stage i and ii breast cancer treated by biopsy and irradiation. Int J Radiat Oncol 1983;9:1807–13. http://dx.doi.org/10.1016/0360-3016(83)90348-6.
[15] Van Limbergen E, van der Schueren E, Van Tongelen K. Cosmetic evaluation of breast conserving treatment for mammary cancer. 1. Proposal of a quantitative scoring system. Radiother Oncol 1989;16:159–67. http://dx.doi.org/10.1016/0167-8140(89)90016-9.
[16] Vrieling C, Collette L, Fourquet A, Hoogenraad WJ, Horiot J-C, Jager JJ, et al. The influence of patient, tumor and treatment factors on the cosmetic results after breast-conserving therapy in the EORTC “boost vs. no boost” trial. Radiother Oncol 2000;55:219–32. http://dx.doi.org/10.1016/S0167-8140(00)00210-3.
[17] Olivotto IA, Rose MA, Osteen RT, Love S, Cady B, Silver B, et al. Late cosmetic outcome after conservative surgery and radiotherapy: analysis of causes of cosmetic failure. Int J Radiat Oncol 1989;17:747–53. http://dx.doi.org/10.1016/0360-3016(89)90061-8.
[18] Pezner RD, Patterson MP, Hill LR, Vora N, Desai KR, Archambeau JO, et al. Breast retraction assessment: an objective evaluation of cosmetic results of patients treated conservatively for breast cancer. Int J Radiat Oncol 1985;11:575–8. http://dx.doi.org/10.1016/0360-3016(85)90190-7.
[19] Taylor ME, Perez CA, Halverson KJ, Kuske RR, Philpott GW, Garcia DM, et al. Factors influencing cosmetic results after conservation therapy for breast cancer. Int J Radiat Oncol 1995;31:753–64. http://dx.doi.org/10.1016/0360-3016(94)00480-3.
[20] Gerber LH. A review of measures of lymphedema. Cancer 1998;83:2803–4.
[21] Gjorup C, Zerahn B, Hendel HW. Assessment of volume measurement of breast cancer-related lymphedema by three methods: circumference measurement, water displacement, and dual energy X-ray absorptiometry. Lymphat Res Biol 2010;8:111–9. http://dx.doi.org/10.1089/lrb.2009.0016.
[22] Jeevan R, Cromwell D, Browne J, van der Meulen J, Pereira J, Caddy C, et al. National mastectomy and breast reconstruction audit. NHS Inf Cent 2011.
[23] Henseler H, Ju X, Ayoub A, Ray AK. The importance of the pose in three-dimensional imaging of the ptotic breast. J Plast Reconstr Aesthet Surg 2013;66:1551–6. http://dx.doi.org/10.1016/j.bjps.2013.06.040.
[24] Kovacs L, Eder M, Hollweck R, Zimmermann A, Settles M, Schneider A, et al. New aspects of breast volume measurement using 3-dimensional surface imaging. Ann Plast Surg 2006;57:602–10. http://dx.doi.org/10.1097/01.sap.0000235455.21775.6a.
[25] Herold C, Reichelt A, Stieglitz LH, Dettmer S, Knobloch K, Lotz J, et al. MRI-based breast volumetry – evaluation of three different software solutions. J Digit Imaging 2010;23:603–10. http://dx.doi.org/10.1007/s10278-009-9264-y.
[26] Grossman AJ, Roudner LA. A simple means for accurate breast volume determination. Plast Reconstr Surg 1980;66:851–2.
[27] Tezel E, Numanoglu A. Practical do-it-yourself device for accurate volume measurement of breast. Plast Reconstr Surg 2000:1019–23.
[28] Wilkie T, Ship AG. Volumetric breast measurement during surgery. Aesthetic Plast Surg 1977;1:301–5. http://dx.doi.org/10.1007/BF01570265.
[29] Tegtmeier R. A quick, accurate mammometer. Ann Plast Surg 1978;1(6):625–6.
[30] Campaigne BNN, Katch VLL, Freedson P, Sady S, Katch FI. Measurement of breast volume in females: description of a reliable method. Ann Hum Biol 1979;6:363–7. http://dx.doi.org/10.1080/03014467900003741.
[31] Edsander-Nord A, Wickman M, Jurell G. Measurement of breast volume with thermoplastic casts. Scand J Plast Reconstr Surg Hand Surg 1996;30:129–32.
[32] Rha EY, Choi IK, Yoo G. Accuracy of the method for estimating breast volume on three-dimensional simulated magnetic resonance imaging scans in breast reconstruction. Plast Reconstr Surg 2014;133:14–20. http://dx.doi.org/10.1097/01.prs.0000436813.03838.09.

[33] Mailey B, Freel A, Wong R, Pointer DT, Khoobehi K. Clinical accuracy and reproducibility of portrait 3D surgical simulation platform in breast augmentation. Aesthet Surg J 2013;33:84–92. http://dx.doi.org/10.1177/1090820X12469807.
[34] Yoo A, Minn KW, Jin US. Magnetic resonance imaging-based volumetric analysis and its relationship to actual breast weight. Arch Plast Surg 2013;40:203–8. http://dx.doi.org/10.5999/aps.2013.40.3.203.
[35] Lewis PG, Mattison GL, Kim HY, Gupta SC. Evaluation of 3D photographic imaging as a method to measure differential volumes in reconstructed breast tissue. Annu Meet Plast Surg Res 2014.
[36] Hoe AL, Guyer PB, Taylor I, Mullee MA, Royle GT. Breast size and prognosis in early breast cancer. Ann R Coll Surg Engl 1993;75:18–22.
[37] Bland JM, Altman DG. Statistical method for assessing agreement between two methods of clinical measurement. Lancet 1986;i:307–10.
[38] Fujii T, Yamaguchi S, Yajima R, Tsutsumi S, Asao T, Kuwano H. Accurate assessment of breast volume by computed tomography using three-dimensional imaging device. Am Surg 2012;78:933–5.
[39] Liu Q, Ye JM, Xu L, Duan XN, Zhao JX, Liu YH. Correlation between dynamic contrast-enhanced MRI and histopathology in the measurement of tumor and breast volume and their ratio in breast cancer: a prospective study. Chin Med J Engl 2012;125:3856–60. http://dx.doi.org/10.3760/cma.j.issn.0366-6999.2012.21.021.
[40] Katariya R, Forrest AP, Gravelle I. Breast volumes in cancer of the breast. Br J Cancer 1974;29:270–3.
[41] Fung JTK, Chan SWW, Chiu ANK, Cheung PSY, Lam SH. Mammographic determination of breast volume by elliptical cone estimation. World J Surg 2010;34:1442–5. http://dx.doi.org/10.1007/s00268-009-0283-0.
[42] Daly SEJ, Hartmann PE. Infant demand and milk supply. Part 2: The short-term control of milk synthesis in lactating women. J Hum Lact 1995;11:27–37. http://dx.doi.org/10.1177/089033449501100120.
[43] Losken A, Seify H, Denson DD, Paredes AA, Carlson GW. Validating three-dimensional imaging of the breast. Ann Plast Surg 2005;54:471–6. http://dx.doi.org/10.1097/01.sap.0000155278.87790.a1.
[44] Yip JM, Mouratova N, Jeffery RM, Veitch DE, Woodman RJ, Dean NR. Accurate assessment of breast volume: a study comparing the volumetric gold standard (direct water displacement measurement of mastectomy specimen) with a 3D laser scanning technique. Ann Plast Surg 2012;68:135–41.
[45] Stomper PC, Souza DJD, Dinitto PA, Arredondo MA. Analysis of parenchymal density on mammograms in 1353 women 25–79 years old. Am J Roentgenol 1996;167:1261–5.
[46] Xi W, Trisliana A, Yeesiang P. Objective breast volume, shape and surface area assessment: a systematic review of breast measurement methods. Aesthetic Plast Surg 2014:1116–30. http://dx.doi.org/10.1007/s00266-014-0412-5.
[47] O'Connell RL, Stevens RJG, Harris PA, Rusby JE. Review of three-dimensional (3D) surface imaging for oncoplastic, reconstructive and aesthetic breast surgery. Breast 2015;24:331–42. http://dx.doi.org/10.1016/j.breast.2015.03.011.
[48] Porto F, Gurgel JL, Russomano T, Farinatti PDT. Moiré topography: characteristics and clinical application. Gait Posture 2010;32:422–4. http://dx.doi.org/10.1016/j.gaitpost.2010.06.017.
[49] Wallis JW, Miller TR. Volume rendering in three-dimensional display of SPECT images. J Nucl Med 1990;31:1421–8.
[50] Caruso MK, Guillot TS, Nguyen T, Greenway FL. The cost effectiveness of three different measures of breast volume. Aesthetic Plast Surg 2006;30:16–20. http://dx.doi.org/10.1007/s00266-004-0105-6.
[51] Ward C, Harrison B. The search for volumetric symmetry in reconstruction of the breast after mastectomy. Br J Plast Surg 1986;39:379–85.
[52] McLaughlin P. Testing agreement between a new method and the gold standard-how do we test? J Biomech 2013;46:2757–60. http://dx.doi.org/10.1016/j.jbiomech.2013.08.015.
[53] Altman DG, Bland JM. Measurement in medicine: the analysis of method comparison studies. Stat 1983;32:307. http://dx.doi.org/10.2307/2987937.
[54] Atkinson G, Nevill AM. Measurement error (reliability) in variables relevant to sports medicine. Sports Med 1998;26:217–38.
[55] Probst H, Choppin SB, Wheat JS, Harrison M, Goyal A. The development of a low cost 3D surface imaging system to measure breast volume: defining minimum standards using an adapted Delphi consensus study. J Plast Reconstr Aesthetic Surg 2015;68(12):1770–2.
[56] Hughes S, Lau J. A technique for fast and accurate measurement of hand volumes using Archimedes' principle. Australas Phys Eng Sci Med 2008;31:56–9. http://dx.doi.org/10.1007/BF03178454.
[57] Ludbrook J. Statistical techniques for comparing measurers and methods of measurement: a critical review. Clin Exp Pharmacol Physiol 2002;29:527–36. http://dx.doi.org/10.1046/j.1440-1681.2002.03686.x.
[58] Parmar C, West M, Pathak S, Nelson J, Martin L. Weight versus volume in breast surgery: an observational study. JRSM Short Rep 2011;2:87. http://dx.doi.org/10.1258/shorts.2011.011070.