Multimodal assessment of mammogram, ultrasound and clinical palpation in relation to pathological size of breast carcinoma

Fadzli F 1 *, Rahmat K 1 , Mohd-Shah MN 1 , Muridan R 1 , Nawawi O 1 , Rozalli FI 1 , Looi LM 2 and Yip CH 3

1 Departments of Biomedical Imaging, University of Malaya Medical Centre, Malaysia
2 Departments of Biomedical Pathology, University of Malaya Medical Centre, Malaysia
3 Departments of Biomedical Surgery, University of Malaya Medical Centre, Malaysia

Abstract

Objectives: Ultrasonography, mammography and clinical examination (conventional triple assessment modalities) are compared to the gold standard of pathological size to assess their accuracy in measuring size of palpable breast cancer lesions. This knowledge has an important role in the patient’s further management: for staging, choice of surgical technique and prognostication.

Patients and methods: 71 tumours were studied in 70 patients who were prospectively collected over a 2 year period. Pearson's correlation test and Bland-Altman's plot were used to analyse tumour size by clinical palpation, ultrasound and mammogram.

Results: Histopathological examination revealed 60 invasive carcinomas (58 ductal type, 2 lobular type), 4 ductal carcinoma in situ, 2 mixed and 5 special types. The pathological size varied from 0.5cm to 9.5cm. Ultrasound had the best correlation with pathological size, with r= 0.845 (p< 0.001), but with a tendency to underestimate. Size correlation for mammogram and clinical palpation were similar and statistically significant. However, the standard deviation of mammographic size was more compared to size on clinical examination (1.3 (mammogram) vs. 2.2 (clinical)). Clinical palpation was inclined to overestimate, whereas mammogram neither over- nor under-estimated size. Measurements using ultrasonography produced the lowest standard deviation, thus a lower variability from the mean.

Conclusion: This study demonstrated that ultrasound is the more accurate modality compared to mammogram and clinical palpation size for the measurement of palpable breast cancer lesions. However, ultrasound may downstage tumours due to its tendency to underestimate size.

Introduction

Breast cancer is globally the most common cancer in women [1], and poses a significant health problem in Malaysia, being the most common malignancy in the female population. It constitutes 29.9% of new cancers in women with an age standardised incidence rate of 39.3 per 100,000 women [2].

Patients may present for screening purposes or with breast symptoms. In the presence of a symptom such as a lump, triple assessment is undertaken, which includes clinical evaluation, mammography, ultrasound and pathological examination (cytology with or without histology) [3]. This study aims to evaluate the accuracy of these 3 modalities (clinical palpation, ultrasound and mammogram) in measuring the size of palpable breast cancer lesions compared to pathological measurement, which is the accepted gold standard. Size plays an influential role in breast cancer assessment and management in 3 main aspects, clinical staging, choice of surgical technique, feasibility and planning of non-surgical technique [4]. Clinical staging aided by imaging drives the decision for mastectomy or breast conserving surgery, both of which are significant life-changing procedures for the patient. Since the 1990s research has been addressing the issue of which imaging modality correlates best with pathological size (Table 1) [4-17]. These studies variably assessed size via clinical, ultrasound, mammogram and pathological examination. The literature predominantly demonstrated superior accuracy of ultrasound compared to other modalities [6,10,11,14,16]. However, ultrasound tended to underestimate [5,6,15,16], whereas clinical palpation tended to overestimate size [5,6,16] and mammography was equivocal [5].

Materials and methods

A prospective study to compare ultrasound, mammogram and clinical breast tumour size measurements with pathological measurement was carried out at University Malaya Medical Centre (UMMC). The study subjects were selected from the operation list of the Breast Surgery Division. Patients with palpable tumours and had mammogram performed in the Biomedical Imaging Department were chosen.

Mammography

Mammograms were performed with Mammmomat Novation digital mammography machine (Siemens, Erlangen, Germany). Typically,
30kVp was used with tungsten/rhodium targets and 360mAs. Two views were obtained, craniocaudal and mediolateral oblique, and occasionally magnification or compression views were added. Both views were studied by a radiologist specialized in breast imaging, and measurements were taken on the view in which the lesion appeared larger (Figure 1). The largest measurement was estimated excluding any spiculations extending from the mass. Measurements were performed on IMPAX picture archiving and communication system (PACS) network (Agfa, Ridgefield Park, NJ) with Barco monitor (Brussels, Belgium). Tabar Patterns and BIRADS density on mammogram were also assessed.

**Ultrasonography**

Ultrasound images were performed with Phillips iU22 unit (Phillips Medical Systems, Bothell, WA) utilizing a Philips L12-5 and L17-5 high-frequency linear ultrasound probe. The frequencies of the probes varied from 5 to 12 MHz with 50mm aperture and 5 to 17 MHz with 38mm aperture. Most images were taken with a cine view, to enable subsequent review to accurately measure the largest diameter of the lesion (Figure 2). Other images were conventional 2 plane views of the lesion. Areas of posterior shadowing and partial voluming were carefully excluded from the measurement. Lesions of larger size compared to the footprint of the probe required use of widescan and

### Table 1. Summary of previous literature comparing imaging modality/ clinical palpation pathological size measurement

<table>
<thead>
<tr>
<th>Year</th>
<th>Author</th>
<th>No of cases</th>
<th>Clinical</th>
<th>USS</th>
<th>Mammo</th>
<th>Conclusion of study</th>
</tr>
</thead>
<tbody>
<tr>
<td>1993</td>
<td>Madjar [14]</td>
<td>100</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>USS most accurate</td>
</tr>
</tbody>
</table>

(Key: Mammo: Mammogram, USS: Ultrasound, under: underestimate, Over: overestimate, neither: neither overestimate or underestimate, -: Over or underestimation not specified)

![Figure 1 (A,B,C): These mammogram images (A, C (CC view); B (MLO view)] were taken from a 64 year old patient with invasive ductal carcinoma (IDC) grade 2. The view which demonstrates the lesion at its maximal size was chosen, CC view (A) in this case. Two measurements perpendicular to each other are obtained using callipers provided by IMPAX(C), with the measurement representing the longest diameter being chosen for analysis, 2.1cm.

![Figure 2 (A,B): Ultrasound images from the same patient as in Figure 1, selected from cine views of the lesion. The imaging plane in which the lesion appears largest is chosen, and two measurements are performed perpendicular to each other (B). The largest diameter obtained on ultrasound was 2.1cm.](image-url)
panoramic settings. The depth of the tumour from the skin was also measured. Images were reviewed on OsiriX software (Version 3.9.1; 64 bit; OsiriX Foundation, Geneva) on a Mac (Apple Incorporated, USA) workstation and Barco monitors.

Clinical measurement

Clinical examination was done with the patient supine on the operating table, and measurement was obtained by placing a ruler on the patient’s breast and measuring the longest diameter of the palpated tumour.

Histopathological measurement

All patients underwent either a lumpectomy or mastectomy. The specimen was sliced at 5mm intervals and the tumour measured, in cm, in 3 planes. The specimens were fixed in buffered formalin and subsequent samples of the tumour, up to 1.5 cm width and about 3-4 mm thickness, were processed into paraffin blocks. Microtomed 4-micron thick sections from the blocks were stained using haematoxylin and eosin for histological examination. For tumours more than 1cm in largest dimension by gross examination, the size was determined by naked eye measurement using a ruler. For tumours less than 1 cm in largest dimension on gross examination, the whole tumour was sampled for histology. The size of the tumour was then determined through microscopic examination using the microscope stage scale [18]. For the purpose of this study, the largest dimension of the tumour, whether determined by naked eye measurement or microscope measurement, as the case may be, was taken as its pathological size and served as the gold standard.

Statistical analysis

Measurement and demographic data was inserted into Microsoft Excel (Redmond, WA) spreadsheets and analysed using two computerized statistical software, SPSS version 17 (Illinois, USA) and MedCalc (Mariakerke, Belgium). Bland-Altman analysis was obtained from MedCalc software. The rest of the analysis: descriptive statistics, Pearson’s correlation, Spearman’s correlation, one-way ANOVA and Wilcoxon Signed Ranks test were performed using SPSS. These analyses were also repeated after division of data according to pathological T-staging.

Results

Seventy patients with 71 malignant breast tumours were included in the study from December 2009 until December 2011. Several patients were excluded from the initially recruited list: 8 patients due to histopathologically confirmed benign diagnosis, 1 patient had surgery postponed indefinitely, 1 had a lesion size >10cm, 1 due to immeasurable size on histopathology and 1 due to single palpable lesion clinically but 2 nodules were found on imaging and pathology.

The study population encompassed the 3 main ethnic groups in Malaysia - Malays, Chinese and Indians. All were women, aged between 35 to 84 years, with a mean of 57.2 years. All clinical measurements were performed within 24 hours of surgery. The mean number of days between imaging and surgery for ultrasound and mammogram were 12 (range:1- 62) and 27 (range:1-64) respectively.

All patients underwent either lumpectomy or mastectomy. There were 17 lumpectomies and 54 mastectomies. The histopathological types of tumours were: 58 invasive ductal carcinoma (IDC), 4 ductal carcinoma in-situ (DCIS), 2 invasive lobular carcinoma (ILC), 2 mixed IDC/Mucinous and 5 special types. Of the special type, 2 were mucinous carcinomas, 1 neuroendocrine carcinoma, 1 invasive cribriform carcinoma and 1 borderline phyllodes tumour.

Size comparisons

All study patients had clinical, ultrasound and mammographic measurements. Collectively, for these study modalities, the largest size measured was 20.0cm by clinical examination, and the smallest size was 0.9cm via ultrasound. In contrast, the maximum pathological size was 9.5cm and the minimum was 0.5cm (Table 2). Ultrasound measurements were closest to the mean of the pathological size.

The differences between the study modalities and pathological size were plotted against pathological size using Bland-Altman analysis (Figure 3). Ultrasound had the lowest standard deviation (0.95), followed by mammogram (1.30) and clinical (2.06). This analysis also demonstrated that ultrasound tended to underestimate and clinical palpation tended to overestimate size. When the Wilcoxon Signed Ranks test was applied, these differences were statistically significant (p<0.01 for ultrasound and p<0.001 for clinical). Mammogram neither over- nor under-estimated pathological size, and this was not statistically significant using the Wilcoxon test.

The scatter plots for correlation data are presented in Figure 4. Ultrasound size and pathological size showed the best correlation (r=0.845). This would correspond with very good correlation [19]. The correlation coefficients for clinical measurement and mammogram were 0.717 and 0.713. These correlations can be categorized as moderate to good [19]. From the lines representing 95% confidence intervals of the means, the size values of all modalities deviated from the line of true fit as lesions increase in size. All the coefficients were statistically significant (p< 0.001 for ultrasound, mammogram and clinical). The equation for the regression lines of each modality are presented with the scatter plots. When the 3 modalities: ultrasound, mammogram and clinical examination are analysed using stepwise linear regression, the overall formula obtained was Pathology= 1.048*US+0.28.

Utilizing the mean of the difference between a measurement and the pathological size showed that on average, ultrasound and mammogram varied from pathological size by less than 1cm. Ultrasound was the closest at 0.72cm. Clinical palpation deviated more from pathological size, with a mean of 1.75cm. When these data was applied as a percentage, ultrasound diverged the least by 23.5%. Clinical palpation had a percentage variation of 68.7%.

Other parameters

No significant correlation was found between age of patient, depth of lesion, Tabar Pattern or BIRADS density when compared with differential measurements by ultrasound, mammogram and clinical palpation when Spearman’s correlation or one-way ANOVA tests were used (where appropriate). However, statistically significant difference was found between mammogram and ultrasound measurements and pathology type, particularly in the ILC histological types. In terms of mammography measurement, the difference was significantly (p<0.001) less accurate when ILC was compared to all other histological types, whereas for ultrasound measurement, the significance (p<0.01)

Table 2. Means, minimum and maximum of modality and pathology measurements

<table>
<thead>
<tr>
<th>Modality</th>
<th>Mean (cm)</th>
<th>Minimum (cm)</th>
<th>Maximum (cm)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ultrasound</td>
<td>2.8</td>
<td>0.9</td>
<td>9.3</td>
</tr>
<tr>
<td>Mammogram</td>
<td>3.3</td>
<td>1.0</td>
<td>10.0</td>
</tr>
<tr>
<td>Clinical Examination</td>
<td>4.6</td>
<td>1.0</td>
<td>20.0</td>
</tr>
<tr>
<td>Pathology</td>
<td>3.2</td>
<td>0.5</td>
<td>9.5</td>
</tr>
</tbody>
</table>
Fadzli F (2018) Multimodal assessment of mammogram, ultrasound and clinical palpation in relation to pathological size of breast carcinoma

Integr Cancer Sci Therap, 2018      doi: 10.15761/ICST.1000267

Figure 3. Bland-Altman plot showing the limit of agreement between differential modality and pathology measurements with pathology size (in cm)

Figure 4. Scatter plots showing modality measurements plotted against pathology measurements (in cm). Linear regression equations for each modality are also displayed.
was only seen when compared to IDC and the special types group. No significant difference was seen with clinical measurements. There was also statistically significant difference between differential measurement of clinical palpation, and this was noted selectively in the retroareolar group with a larger difference in measurement in contrast to other groups (p<0.001).

Pathological stage

The study patients were divided into 2 groups: pathology T1 (≤2cm) size and T2 (>2cm to 5 cm) or larger pathology size. Pearson’s correlation was applied to both groups. In both groups, ultrasound demonstrated the highest correlation with pathology size, with r = 0.623 (p=0.001) in the T1 group and r = 0.786 (p<0.001) in the T2 group. Better correlation was seen with larger tumours. The correlation coefficients of the other modalities in the larger tumour groups were all statistically significant; 0.692 for clinical measurement (p<0.001) and 0.588 for mammogram (p<0.001). However, the correlations were much poorer for clinical examination and mammogram in the T1 group, with r values of 0.068 (p=0.753) and 0.401 (p=0.052) respectively.

Analysis of tumours measuring more than 5cm (T3) by pathological size showed higher accuracy of ultrasound and clinical examination in relation to mammography, with statistically significant values. The correlation coefficients were 0.805 for ultrasound (p=0.003), 0.630 for clinical examination (p=0.038) and 0.451 for mammogram (p=0.164).

Ultrasound (67.6%/48 patients) had the most number of cases correctly staged. Clinical examination overstaged 25 cases (35%) whereas ultrasound understaged the most number of cases (21%/15 patients). Of 25 cases overstaged by clinical palpation, 1 was overstaged from T1 to T3.

Discussion

Breast cancer in the Malaysian population tends to be diagnosed at later stages, usually with a palpable tumour mass. Hisham et al reported that between 1998 to 2001, 50 to 60% of new cases presented at Stage 3 or 4, and only 5% were impalpable tumours. Also, the mean tumour size was 5.4cm with a range of 1 to 20cm [20]. Our study showed a similar range of clinical size and a mean of 4.6cm. A wide range of lesional sizes were encountered, from 0.5cm to 9.5cm on histopathological measurement, and up to 20.0cm on clinical measurement, reflecting that variation in clinical T staging of these lesions is very much dependent on modality. The time period between surgery and imaging of the lesions were also wide especially for mammography. The doubling time for breast cancer tumours in patients aged 50-70 years was estimated to be 157 days [21]. In view of the increased time intervals between mammography and surgery, one would expect significant overestimation of pathology size. However, Wilcoxon Signed ranks test did not show significant difference between pathology size and mammogram size, which essentially means that there is insufficient evidence to say that the values differ. The p value in this test for mammography versus Pathology size was also close to significance level (0.059), and coupled with the larger standard deviation values, made further analysis necessary.

Bland-Altman analysis was additionally applied to the measurements of imaging modalities against histopathology measurements as were also performed by Allen et al and Dummin et al. [5,8]. Bland and Altman argued that correlation coefficients were often incorrectly applied to studies when comparing methods of measuring the same variable. The correlation coefficient (r) is a measure of the strength of a relationship between two measurements, and not the agreement between the two. This analysis is a graphical representation of data agreement, which may help to corroborate the correlation results by excluding erroneously good correlation. This can occur when measurements are consistently higher or lower than the compared measurement by an approximate constant, effecting an apparent ‘good’ correlation, despite the true deviation of value by that constant. Correlation is also influenced by the range of the quantity in a sample, in which a wider range would give a better correlation [22]. Ultrasound measurements had the lowest mean, standard deviation and best correlation, which translates to better accuracy. This result is comparable to studies by Forrange and Pierie et al in which ultrasound also showed the best correlation with similar r values [10,16]. Mammogram measurement revealed a larger standard deviation compared to ultrasound, but the largest standard deviation was noted with clinical measurements, thus the least accurate. Using Bland-Altman and Wilcoxon Signed Rank test, it was clear that ultrasound and clinical measurements tended to under and overestimate the lesions respectively which was statistically significant. Other studies have also shown the proneness of ultrasound to underestimate [4-6,8,15-17]. In all the modalities, it can be deduced that as the size of the lesion increases, error in size measurement magnifies. The analysis of clinical T staging versus pathological T staging showed that most modalities predicted the pathological staging correctly. Ultrasound had the highest percentage of correctly predicted staging, whereas clinical examination had the lowest percentage. However, ultrasound had the highest number understaged and clinical examination had the highest number overstaged. These conclusions echo the earlier results of Wilcoxon Signed Rank test.

The clinical significance of correlating size to the pathological gold standard size lies in 3 main aspects, staging of the patient, surgical technique and in planning non-invasive surgical therapy [4]. It is also important prior to initiating neoadjuvant chemotherapy as an adjunct to surgery [12]. Clinical staging of the patient is usually performed with the guidance of examination and imaging, and allows prognostication prior to surgery and pathology measurement. This initial staging also guides the surgeon to the decision of mastectomy or breast conserving therapy. In our centre, guidelines dictate that in patients with multifocal or multicentric disease, large ratio of tumour to breast causing poor cosmesis if lumpectomy is performed and patients who cannot undergo radiotherapy for any reason are contraindicated for breast conserving surgery [23]. It has been demonstrated that in patients with small (<4cm) tumours in which total excision with disease free margins is achieved, there is no significant difference in survival [24,25]. However, the choice of surgery has important consequences for the patient, as it influences psychological, social and marital-sexual wellbeing [26]. With regards to tumour ablation, this is an upcoming technique for the treatment of breast tumours likely of early stage, in which imaging will be the sole guide to planning and execution of treatment [4,12].

There are multiple reported reasons for the varying difference between imaging and pathological size. Size of the tumour is one factor, with generally lower correlation manifesting at the extremes of tumours size (smaller then 2cm or larger than 5cm). This was particularly seen with clinical and mammography measurements, whereas ultrasonography maintained a reasonably high correlation regardless of size.

In order to reduce inter-operator variability and standardize measurement, the ultrasound scans were repeated by the same operator mostly within 24 hours of surgery. 11 out of the 71 cases
were larger than 5 cm, which is a recognised problem with ultrasound due to the tumour being larger than the field of view or footprint of the probe [7,27]. This issue was addressed in this study by using wide scan and panoramic views, and also by excluding very large tumours measuring more than 10 cm. Larger tumours may also be more difficult to penetrate for posterior margin visualisation and if coupled with posterior shadowing may cause further complexity [7,27].

Clinical palpation was performed by different surgeons in this study, contributing to error. Clinical palpation is known to overestimate measurement of tumours and out of the 9 studies in literature that included clinical palpation as a comparison, only 1 study found it to be most accurate [13]. One of the reasons for overestimation is the skin folds and subcutaneous tissue between the examining fingers and both sides of the lesion, particularly in smaller lesions. Reactive changes of the subcutaneous tissue secondary to the tumour may also be misconstrued as tumour by the palpating hand [7]. Other factors that affect clinical palpation include obesity [9].

Based on the results in this study, mammograms were more accurate compared to clinical palpation in measuring tumours but less accurate compared to ultrasound. There is an inherent slight magnification present in this modality which will cause apparent enlargement of the tumour. Underestimation of the tumour size can occur if the lesion was imaged obliquely to the x-ray plane. Dense breast parenchymal tissue will also blur the margins of the tumour [7].

In all modalities, size of infiltrative tumours will be difficult to estimate due to the poorly defined margins. Despite pathological size being considered a gold standard, it has quandaries of its own. The size of tumours with irregular tumour margins are difficult to measure, more so if the tumour is surrounded by fibrous tissue, spicules or glandular tissue [7]. The radiologist also has to slice the tumour according to the longest dimension based on palpation of the tumour. Thus, the longest dimension may not always be included in a slice and achieving it may be even more difficult when the margins are ill-defined [5,9]. Other possible causes of error are shrinkage of the tumour during formalin fixation and paraffin embedding [5].

Another possible source of error would be the inherent nature of the tumour to grow, although the reported doubling times for breast cancer is 157 days for patients 50-70 years of age and 80 days for those younger than 50 years [21]. Clinical examination were all performed within 24 hours prior to surgery, however, times between mammography and surgery were longer than the other modalities, with a mean of 27 days (range between 1 to 64 days). However, another study also had comparable imaging and surgical time ranges [5].

The overall stepwise regression equation: Pathology = 1.048*US + 0.28 could be used in the clinical setting in our centre to estimate pathological size. It does not include mammogram or clinical measurements due to the high correlation between ultrasound, mammogram and clinical measurements. Since ultrasound had the best correlation, it was included in the equation.

The patient’s age, Tabar pattern, breast density, pathology type, ultrasound depth and lesion position in the breast were several factors compared with the differential measurements of the imaging modalities. The measurements of ILC lesions were found to be significantly less accurate when measured using ultrasound and mammogram measurements. This finding is consistent with conclusions from other studies. Skaane et al suggested that assessment of ILC is more difficult compared to other types of tumour, particularly with mammography.

This may be due to similar appearance of the tumour to adjacent breast tissue and lack of features such as calcification to assist in the differentiation [17,28,29]. Thus, there have been more cases of false negative mammograms with ILC compared to other pathology types [29]. In this study, all the modalities underestimated the size of the tumour. However, it is also notable that ILC tumours tend to be of larger size compared to non ILC tumours [30]. A study of 95 patients with ILC demonstrated ultrasound and mammography measurements markedly underestimate the size of tumours, as seen in this study [17]. However, the small number of patients (n=2) in this group is also an important factor.

Significant difference was also seen between lesions in the retroareolar region when compared with differential measurement of clinical palpation. This may be due to the location posterior to the nipple which causes increased likelihood of including it in the measurement. The number of patients in this group was also small, which may further accentuate inaccuracies.

In conclusion, this study identified ultrasound as the most accurate modality for the assessment of breast cancer size, followed by mammogram and clinical palpation. An overall equation to estimate pathological size from ultrasound size was derived from the data which may help in future clinical staging in our centre. However, these results do not spell the end of mammography or clinical examination, but further establishes the combined role of all three modalities in assessing these tumours.

Acknowledgement

This study was approved by the Medical Ethics Committee of UMMC. The authors gratefully acknowledge support from University of Malaya Research Grants (RG178/09HTM and RG390/11HTM).

Financial disclosure

All the authors do not have any financial interests related to the material in this manuscript to disclose.

References


Integr Cancer Sci Therap, 2018      doi: 10.15761/ICST.1000267

Copyright: ©2018 Fadzli F. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.

Copyright: ©2018 Fadzli F. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original author and source are credited.