

Evaluation of conventional training in Clinical Breast Examination (CBE)

Veitch, Daisy; Goossens, Richard; Owen, Harry ; Veitch, James ; Molenbroek, Johan; Bochner, Melissa

DOI

[10.3233/WOR-192899](https://doi.org/10.3233/WOR-192899)

Publication date

2019

Document Version

Final published version

Published in

Work

Citation (APA)

Veitch, D., Goossens, R., Owen, H., Veitch, J., Molenbroek, J., & Bochner, M. (2019). Evaluation of conventional training in Clinical Breast Examination (CBE). *Work*, 62(4), 647-656.
<https://doi.org/10.3233/WOR-192899>

Important note

To cite this publication, please use the final published version (if applicable).
Please check the document version above.

Copyright

Other than for strictly personal use, it is not permitted to download, forward or distribute the text or part of it, without the consent of the author(s) and/or copyright holder(s), unless the work is under an open content license such as Creative Commons.

Takedown policy

Please contact us and provide details if you believe this document breaches copyrights.
We will remove access to the work immediately and investigate your claim.

Green Open Access added to TU Delft Institutional Repository

'You share, we take care!' – Taverne project

<https://www.openaccess.nl/en/you-share-we-take-care>

Evaluation of conventional training in Clinical Breast Examination (CBE)

Daisy Veitch^{a,*}, Richard Goossens^a, Harry Owen^b, James Veitch^c, Johan Molenbroek^a
and Melissa Bochner^d

^a*Medisign, Delft University of Technology (TU Delft), Delft, The Netherlands*

^b*Clinical Skills and Simulation Unit, School of Medicine, Flinders University, Adelaide, SA, Australia*

^c*Thermofisher Scientific Inc, CA, USA*

^d*Breast Endocrine and Surgical Oncology Unit, Royal Adelaide Hospital, Port Road, Adelaide, SA, Australia*

Received 3 November 2018

Accepted 28 January 2019

Abstract.

BACKGROUND: Clinical Breast Examination (CBE) is the examination of a women's breasts by a healthcare professional, such as a breast surgeon, family physician or breast-care nurse who is trained to recognise many different types of abnormalities and warning signs in the breast [1]. CBE is particularly important in rural areas and developing countries who have limited access to technology such as mammography. CBE needs to be taught to health professionals like any other clinical skill used by medical professionals in the workplace. CBE in part involves palpation of the breast, that is, determining by touch which breast lumps are normal and which are suspicious in feeling. The gold standard for assessing tactile skills in CBE is seeing whether students can accurately identify and discriminate between different breast lumps also known as masses (IDBM) on actual patients in a clinical setting. However, this is not practical in a medical education setting. Usually the testing methods 'go through the motions' of feeling the breast as part of CBE. So the students' technique is examined either using unrealistic simulation models or using an intimate examination associate (IEA), an actor/volunteer who permits students to examine their intimate body parts such as breast or genitals for teaching purposes. These volunteers do not have any abnormalities so this teaching does not include the actual detection of suspicious lumps. We undertook a study of clinical skill with 10 medical students to examine different methods of assessing novice student clinical skills after a brief training in CBE.

OBJECTIVES: This study aims to evaluate the effectiveness of current training and assessment of novice students in CBE and their capacity to identify and discriminate breast masses (IDBM) on actual patients.

METHODS: We assessed each student's IDBM ability in an actual clinical situation, a breast clinic with a mixture of eight IEAs and one real patient with a large, easily palpable, putative breast cancer. We recruited 10 clinically inexperienced medical students, who were trained for 30 minutes by two breast surgeons using an IEA. Students were tested in a simulated clinical setting, a breast clinic where each examined 4 IEAs and one patient. The students were blind to who was the real patient and who was an IEA. Patients were examined by a breast surgeon in private prior to the commencement in the study. The breast surgeon recorded any clinical finding on the patients during the initial examination. The surgeon coached each patient on how to mark the students and showed the patient their results so the patients had a benchmark. After each examination was finished the students had four different assessments: 1) patients marked each student, 2) students were independently proctored – that is, marked by an expert, 3) students recorded their clinical findings and 4) students recorded how confident they were that they had the correct findings. Results from different kinds of student assessments were compared.

*Address for correspondence: Daisy Veitch, Delft University of Technology (TU Delft), Landbergstraat 15, 2628 CE Delft,

The Netherlands. Tel.: +61 8 83700202; E-mail: daisy.veitch@gmail.com.

RESULTS: A chi-square test for independence between true positive or negative masses versus student-assessed positive or negative masses was not significant at $\alpha = .05$. This means that there was no statistical association in the indication of positive or negative presence of masses versus whether such masses were actually present or absent. By comparison, experts (breast surgeons) were able to detect normal and abnormal breast masses by palpation alone 100% of the time and rate their confidence level as 'certain'. Unlike the experts, student self-reported confidence was unrelated to their competence score (CS). Proctoring was inversely related to the students' CS.

CONCLUSIONS: The main conclusion is that novice students do not seem to be able to accurately detect breast masses in a clinical setting even after training. On the basis of these results, we believe that a comprehension component in the current CBE testing is needed in addition to the current methods of testing.

Keywords: Medical training, medical education, simulation training, student testing, palpation, Medisign

List of abbreviations

| | |
|--------|--|
| BCS/MD | Bachelor of Clinical Sciences/Doctor of Medicine |
| CBE | Clinical Breast Examination |
| CS | Competence Score |
| IDBM | Identification and discrimination of breast masses |
| IEA | Intimate Examination Associate |

1. Introduction

Breast cancer is still the most prevalent cancer in women, and the second most common cancer diagnosed worldwide [2].

The gold standard for breast cancer detection is the triple test with more than 99% sensitivity [3, 4]. As suggested by the name, the triple test has three different parts: Clinical Breast Examination (CBE), imaging and biopsy: each of which can detect breast cancer in different ways. Each part of the triple test uses different equipment and needs a different skill set to administer. Clinical Breast Examination (CBE) is defined as the examination of a woman's breasts by a healthcare professional, such as a breast surgeon, family physician or breast care nurse who is trained to recognise many different types of abnormalities and warning signs in the breast [1]. Out of these three tests, CBE is the first line, and requires clinician training but not specialist equipment such as mammography. Thus CBE is important as a screening and diagnostic tool for rural and remote communities with limited access to expensive technological resources. Even in first-world countries with all available imaging techniques, CBE remains important for detecting interval cancers (those that become apparent between image-based appointments) and those

that that are mammographically occult (negative on a mammogram but palpable) [4, 5]. In addition, lost CBE skills among health professionals at work are an issue in 50% of physician-delayed diagnosis (where the clinician falsely reassured the patient that the lump being palpated was benign), while delayed diagnosis after a mass has been palpated is the leading cause of malpractice litigation related to breast cancer [6, 7].

CBE requires taking a patient history and completing a physical examination. The physical examination requires breast palpation. Breast palpation skills are taught and assessed in medical school. The gold standard for assessing tactile skills and CBE is determining whether a student can accurately identify and discriminate between different breast lumps on actual patients in a clinical setting. However, this is not practical in a medical education setting. Student in medical school assessment conditions are usually assessed on IEA actors who do not have putative cancer thus IDBM findings are always normal. Alternatively, students can be assessed on simulation models where there are always positive IDBM findings. Thus students in absence of real patients, some of whom have cancer and some not, can not be assessed on IDBM in a realistic way. Instead students are usually tested on their technique alone. One recent analysis of testing techniques compare 4 different tests: 1) technical skills using checklists and global rating scales, 2) assessment by a trained actor, sometimes called a teaching associate or intimate examination associate (IEA), whose breast is normal, 3) student questionnaire and self-assessment, and 4) questionnaire on the frequency of performing intimate examinations skills during internship [8]. Sometimes student testing involves proctoring where a breast nurse observes the student and assesses them for physical manoeuvres, asking 'did the student examine all areas of the breast?', 'did they

use appropriate pressure?', 'did the student introduce themselves?' etc.

These skills are important and testing for them is necessary, however, we are not aware of any routine tests in medical school that involve detecting and discriminating breast lumps in a simulated clinical environment with real patients. In addition, we are not aware of any objective, standardised physical testing for students in identifying and discriminating between different breast masses (IDBM) as part of medical student training. Given the importance of CBE in the diagnosis of breast cancer, this seems to be a significant lack in medical training.

Chalabian and Dunnington (1998) argue that "to certify competence in clinical breast evaluation or any physical examination skills, we must remind ourselves that demonstration of a physical examination manoeuvre does not equate with ability to detect physical examination findings. To ignore lump detection as part of our construct of breast examination is as remiss as leaving out reading comprehension from a test of reading ability" [9].

Up to 85% of medical students report that they feel they need further training in detecting breast masses [10, 11]. This indicates that while CBE is a core part of the curriculum, what is taught is not sufficient to produce competency in a large percentage of students. By competency we mean the accurate identification and discrimination of breast masses (IDBM), which is not only the essential goal of CBE, but also of the triple test (CBE, imaging, and biopsy) and all combinations of the various technologies and series of manoeuvres developed to detect breast cancer.

2. Methods

This paper explores test results of 10 novice students (year 2 of 6 in a medical degree) after a 30-minute training in CBE. The training was conducted by two breast surgeons using IEAs or teaching associates (real women as training subjects) and students were given an opportunity to practise.

All students were given a code according to the order they enrolled in the study. None of the students had prior experience in breast examination although one had received informal training from a parent. There were five males and five females.

This paper will compare existing conventional methods of assessing student's skill in CBE with results of testing students in a clinical setting with a

mixture of patients and actors. The tests were administered after a 30-minute training session in clinical breast examination (CBE). The methods are 1) student self-confidence, 2) proctoring, and 3) patient feedback, compared to the gold standard of examining a breast patient in a clinical setting.

The trial was conducted in the Women's Health Clinic which is attached to the Royal Adelaide Hospital for four hours on a single afternoon, during normal breast clinic times. The trial was supervised by the regular clinic staff composing of four breast surgeons and two breast nurses. These staff acted as the proctors and marked the students. Students believed that all the patients were real patients and were not told that all but one were actors. The real patient was recruited in a GP clinic the day before the trial and was on her way to have an MRI breast scan to investigate a large suspicious putative cancer that was easily palpable. The patient also had bilateral breast implants and inframammary scars that were potential clinical findings. The actors were examined by the breast specialists and found to be normal prior to the commencement of the trial. Medical students ($N_S = 10$) who were novice (pre-clinical) were recruited from Flinders Medical Centre. Students (S) were tested in a clinical setting with a blinded mixture of patients (P) ($N_P = 1$) and actors ($N_A = 8$) making the examinations - (E). Each student conducted five examinations in a random order. There was a total of 50 examinations ($N_E = 50$).

Student performances in CBE were evaluated in four ways. Clinical findings recorded were recorded by the student after each physical examination on a questionnaire where they also rated how confident they were that they had the clinically correct finding. While the student was filling in their questionnaire the supervising breast specialist recorded a proctored score for each student, rating them for coverage, pressure and appropriate hand technique. The patient also ranked each student's performance.

Student performance was compared in the clinical setting with real patients so absolute truth, as determined privately by a breast surgeon, that is, palpable putative breast cancer 'present' or 'not present' was known for each patient. The four assessment methods were students; 1) conducting a CBE and recording clinical findings of putative cancer or not, 2) proctored by breast surgeons, 3) as self-assessed by themselves (confidence), and 4) as assessed by the patient. We wanted to assess the relationship between confidence and competency and devised a Competence Score as described next:

Table 1
Matrix to describe possible competency scores

| | Diagnosis no lump (0) | Diagnosis with lump (1) |
|-----------------------------|-----------------------|-------------------------|
| Actor (0) Truth = absent | Truth (0,0) | False (0,1) |
| Patient (1) Truth = present | False (1,0) | Truth (1,1) |

2.1. Student conducting a CBE and recording clinical findings of putative cancer or not

The students had a possibility of four combinations of answers in the clinical setting, see Table 1.

We calculated a difference score (DS) for each student ($n = s_1, s_2, s_3, \dots, s_{10}$):

$$DS(n) = \left(\frac{n_{0,0}}{4} - 1\right)^2 + \left(\frac{n_{0,1}}{4}\right)^2 + \left(\frac{n_{1,0}}{1}\right)^2 + \left(\frac{n_{1,1}}{1} - 1\right)^2$$

We divided by 4 so that we could equally weight the penalty for being wrong with a real lump versus no real lumps. This was arbitrary and in real life we might want to weight it differently according to a cost-benefit analysis but we went for simplicity here.

To make a Competence Score (CS) we looked at deviation from the truth where the truth was defined as the diagnosis made when the breast surgeon examined the patient. The worst possible DS score was 4. So the CS was calculated by subtracting the student’s DS from 4.

$$CS_n = 4 - DS_n \tag{1}$$

Table 2 shows the range of scores. Lines 3,4 and 5 show the scores from 3 possible scenarios, 2 of directed guesses and the last line shows the scores if there was a random guess, which is theoretical and is of course impossible to divide the one patient into.5 as results are binary.

Table 2
Range of scores from best to the worst possible score

| *Actual number of patients | Truth (0,0) no lump | False (0,1) positive | False (1,0) negative | Truth (1,1) lump | DS | CS |
|------------------------------------|---------------------|----------------------|----------------------|------------------|----|----|
| Best possible score | 4* | 0 | 0 | 1* | 0 | 4 |
| Worst possible score | 0 | 4 | 1 | 0 | 4 | 0 |
| Directed guess everyone has a lump | 0 | 4 | 0 | 1 | 2 | 2 |
| Directed guess no one has a lump | 4 | 0 | 1 | 0 | 2 | 2 |
| Random guess | 2 | 2 | 0.5 | 0.5 | 1 | 3 |

2.2. Student assessment by an expert – ‘Proctoring’

Proctoring for this trial was one usually used in medical schools [8]. Students were each supervised by an expert (1 of 5 breast surgeons) who assessed them during each clinical examination. They were assessed for appropriate hand positioning, whether they examined all areas of the breast and used appropriate pressure. They were assessed on a 6-point scale where 0 was ‘not at all’ and 5 was ‘always’. The proctors also scored them on a pass/fail scale as part of their coursework. All of these students passed.

2.3. Student self-assessment — Confidence

Students completed a questionnaire (Q1) prior to the start of the study, after training, before the clinical trial (Q2), and also an exit questionnaire (Q8). Students were asked the following questions three times and asked to rate themselves on a six point scale from 0 = False to 5 = true:

- A “Knowledge perspective (I feel I have a good grounding in theory of CBE)”
- B “Practical perspective (I feel comfortable conducting a physical examination on a patient)”
- C “Practical perspective (I feel I can competently apply the theory of CBE in practice)”

The results presented here averages the self-reported confidence scores of each student (Q8) recorded after each of the five patients were examined.

2.4. Student assessment by each patient

Students were assessed by each patient they examined at the completion of each examination. Patients had previously been examined by a breast surgeon and had been coached in technique and how to rate or assess the student. Patients were asked, “Do you think they (the student) carried out the physical examina-

Table 3
Descriptive statistics for the student's scores

| | N | Minimum | Maximum | Mean | Std. Deviation |
|--------------------------------------|----|---------|---------|-------|----------------|
| Competence Score (CS) | 10 | 0.0 | 3.5 | 1.800 | 1.0750 |
| Confidence (self-reported) | 10 | 1.4 | 4.0 | 2.960 | 0.7412 |
| Proctor (educator's view) | 10 | 2.0 | 4.1 | 2.990 | 0.5953 |
| Patient's view of student competence | 10 | 1.8 | 4.4 | 2.880 | 0.8702 |
| Valid N (listwise) | 10 | | | | |

Table 4
Student scores

| Breast lump | True positive | True negative | Total |
|--------------------|---------------|---------------|-------|
| Indicated Positive | 4 | 27 | 31 |
| Indicated Negative | 6 | 13 | 19 |
| Total | 10 | 40 | 50 |

tion competently?'. Patients recorded the student's mark on a 6-point scale where 0 was 'not at all' and 5 was 'always'.

We used a Pearson's chi-square statistic to test the relationship between the different types of student test.

3. Results

3.1. Clinical assessment CBE - CS versus student assessment by an expert – 'Proctoring'

3.1.1. The specialist results

The eight actors and one real patient were each examined by a breast surgeon prior to the start of the study. The breast surgeon filled out the student answer sheet (Q3) and this became the marking sheet. As expected none of the eight actors had any suspicious breast lumps. The patient however had bilateral breast implants, inframammary scars, nodularity periareolar and a large suspicious lump on top of one of her implants that was clearly palpable and diagnosed the day before. She participated in our trial and then went immediately to her specialist appointment as it was a putative cancer. Two of the actors had inverted nipples but no other signs. All of the specialists marked their confidence as 5 (I'm absolutely certain), i.e., that they gave the right recommendations.

3.1.2. The student results

Of the 50 student examinations, 33 were incorrect (66%) and 17 were correct (34%). The students found a breast mass that they thought needed to

be referred in 31 (62%). However, there were only four student referrals of the true positives (8%). The actual incidence of suspicious breast masses in this study was exactly 20%. They found no suspicious breast masses in 19 (38%). However, the actual incidence of negative findings in the study was 80%.

3.1.3. Real patient – false negatives

There were 10 examinations of the real patient and 4 students correctly identified the putative cancer, although only one of these students noted the scar and none noted the breast implant. So there were 6 false negatives, which in clinical practice would directly relate to 'physician delayed diagnosis' [6, 7].

3.1.4. Actors – false positives

There were 40 examinations in total of the actors. Of these 13 (32.5%) were correctly identified as negative. However, 27 (67.5%) examinations were false positives also known as 'ghosts'. The students largely referred the false positive patients for specialist follow up, so in practice this would result in specialist resources being occupied with normal patients.

Table 4 shows the results. The answers are correct in the upper left quadrant where there were 4 true positives and lower right quadrant where there were 13 true negatives.

The Chi-square test of relationship between truth and indicated anomalies (breast lesions) was done. Chi-square = 2.5679 with a p value of 0.109052. This is not significant at alpha = .05. There does not appear to be a relationship between the true state and the student assessed state. The student assessment appears to be arbitrary. If we look at the marginals, student

Table 5

Student competency scores (CS) are arranged from worst to best out of a possible maximum of 4. Student self-reported confidence, the proctor's score (educator's view) and patient's assessment are in the remaining columns

| Number of patients shaded grey | Truth (0,0) no lump | False (0,1) positive | False (1,0) negative | Truth (1,1) lump | CS | Confidence (self-reported) | Proctor (educator's view) | Patients' assessment of student competence |
|--------------------------------|------------------------|-------------------------|-------------------------|---------------------|-----|----------------------------|---------------------------|--|
| Maximum score | | | | | 4 | 5 | 5 | 5 |
| Student 5 | 0 | 4 | 1 | 0 | 0.0 | 4 | 2 | 1.8 |
| Student 8 | 1 | 3 | 1 | 0 | 0.9 | 2.8 | 4.1 | 3.4 |
| Student 1 | 1 | 3 | 1 | 0 | 0.9 | 3.8 | 3.3 | 3.4 |
| Student 2 | 2 | 2 | 1 | 0 | 1.5 | 1.4 | 2.9 | 1.8 |
| Student 3 | 2 | 2 | 1 | 0 | 1.5 | 2.2 | 3.7 | 4.4 |
| Student 4 | 3 | 1 | 1 | 0 | 1.9 | 3.2 | 2.6 | 3 |
| Student 9 | 0 | 4 | 0 | 1 | 2.0 | 3 | 2.7 | 2.6 |
| Student 7 | 1 | 3 | 0 | 1 | 2.9 | 3 | 3.1 | 3.8 |
| Student 10 | 1 | 3 | 0 | 1 | 2.9 | 3.2 | 2.8 | 2.4 |
| Student 6 | 2 | 2 | 0 | 1 | 3.5 | 3 | 2.7 | 2.2 |

Table 6
Pearson Correlations

| | | Competence Score (CS) | Confidence (self-reported) | Proctor (educator's view) | Patient's view of student competence |
|--------------------------------------|---------------------|-----------------------|----------------------------|---------------------------|--------------------------------------|
| Competence Score (CS) | Pearson Correlation | 1 | -0.192 | -0.063 | 0.019 |
| | Sig. (2-tailed) | | 0.594 | 0.864 | 0.958 |
| | N | 10 | 10 | 10 | 10 |
| Confidence (self-reported) | Pearson Correlation | -0.192 | 1 | -0.379 | -0.057 |
| | Sig. (2-tailed) | 0.594 | | 0.280 | 0.877 |
| | N | 10 | 10 | 10 | 10 |
| Proctor (educator's view) | Pearson Correlation | -0.063 | -0.379 | 1 | 0.731* |
| | Sig. (2-tailed) | 0.864 | 0.280 | | 0.016 |
| | N | 10 | 10 | 10 | 10 |
| Patient's view of student competence | Pearson Correlation | 0.019 | -0.057 | 0.731* | 1 |
| | Sig. (2-tailed) | 0.958 | 0.877 | 0.016 | |
| | N | 10 | 10 | 10 | 10 |

*Correlation is significant at the 0.05 level (2-tailed).

indication of positive when true positive is 40%, and student indication of positive when true negative is 67.5%. Any potential relationship between student indication and truth would be an inverse relationship. This is in contrast to the breast surgeons who achieved 100% accuracy.

Student scores are shown in descending order by CS score in Table 5. The proctors' view of their ability is also shown, along with the patients' assessment of student competence.

Student 5 scored zero, i.e., of the five 'patients' examined, all four actors were referred for unnecessary follow-up and the actual patient was considered normal. However, Student 5 was confident enough to self-report a score of 4 out of 5 that the right recommendations had been made. Interestingly the proctor score for this student was the lowest although this was

still a pass (all the students passed when assessed for CBE physical manoeuvres). The five patients rated this student at an average of 1.8 out of a possible 5.

The lowest Proctor Score was 2 and highest was 4.1. This is not a very broad range out of 5, possibly due to central tendency bias.

The proctor's competency rating was compared to the CS, see Table 3. The correlation ($r = -0.063$) was not significant indicating that the score on the competency exam is not indicative of actual competency. However, examination of the two variables in a plot revealed a possible outlier. See Fig. 1 below.

Student proctored marks ranged from 2 to just above 4, however Subject 5 seems to be an outlier. The correlation was recalculated without Subject 5 ($R = -0.619$) and was significant at $\alpha = .10$. This indicates a potential inverse relationship between

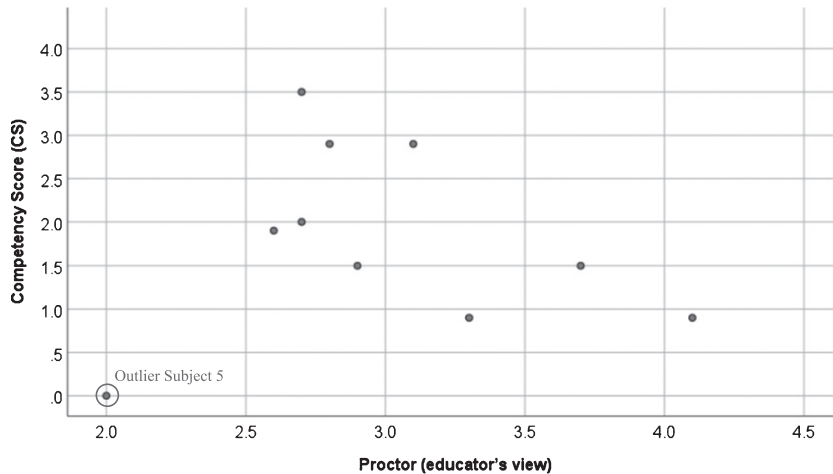


Fig. 1. Proctored marks given by breast surgeons on the y-axis and Competence scores (CS) on the x-axis.

competency evaluation and actual competency. In other words, a good score on the competency exam (physical maneuverers) would indicate a poor score in actual competency.

3.2. Clinical assessment CBE- CS versus student self-assessment — Confidence

Scores from the final self-reported confidence questionnaire were compared to competence scores (CS). The Pearson Correlation between student self-reported confidence in their competence and their actual competence ($r = -0.192$) was not significant. See Table 6.

3.3. Clinical assessment CBE - CS versus student assessment by the patient

A comparison was done of the patients' view of student competence and their CS. The correlation was not significant indicating the patient assessment of competence was not related to actual competence.

3.4. Patient assessment versus proctor assessment

The strongest correlation was between patient assessment of student competency and proctor assessment ($r = 0.731$) which was significant at $\alpha = 0.05$. This indicates that the competency assessment examination may be testing student interpersonal skills plus CBE physical manoeuvres rather than actual competency in IDBM.

4. Discussion

The students did no better than chance in the clinical setting for IDBM with a Chi-square score of 2.5679 and p value of 0.109052. This was in contrast to the experts who achieved 100% accuracy. The students were correct in their clinical findings 34% of the time and incorrect in 66% of their clinical examinations in the four possible combinations of answers: true-positive, false-positive, true-negative and false-negative. The actual incidence of true-positive in the study was 20%. Although the students reported positive findings 62% of the time, the actual incidence of the students reporting a true-positive finding was only 8%. The data suggest that the students were prone to over-diagnose, thus there was a high incidence of false-positives. However, even with this bias it was not enough to prevent the all-important false-negatives that, if they occurred in practice would result in physician-delayed diagnosis. The incidence of false negatives in this study was very large at 12%.

There was no correlation or relationship between proctor ($= -0.063$), student ($= -0.192$) and patient ($= 0.019$) assessment scores and CS. However, there was a strong correlation between proctor and patient assessment scores ($= 0.731$). This seems to indicate two things: 1) current education is not demonstrating competency at actual IDBM and 2) the existing test is only evaluating how comfortable the students are during CBE. The proctor marked the student on a 6-point scale for appropriate positioning, examining all areas of the breast including the axilla, using appropriate hand positioning and pressure with '0'

being not at all and '5' always. That is, the proctor was marking the student on physical examination manoeuvres. The clinical findings are the same as the CBE physical examination findings which equates to CS. This paper supports results of previous studies that demonstration of physical examination manoeuvres do not equate with the ability to detect physical examination findings for CBE [9]. The proctor and the patient were both assessing the student on these maneuvers hence the strong correlation.

The borderline negative correlation (when subject 5 is omitted) between proctors' mark and actual competence might potentially have been caused by student inexperience combined with an effort to look professional. A novice student may have 1) taken longer, 2) expressed uncertainty and 3) conducted parts of the examination repeatedly, all of which might have appeared to look unprofessional. Thus, students who showed uncertainty (because they were uncertain) were potentially being penalised by the proctor marking which encourages the student to ignore the clinical findings in favour of 'looking professional'. So current assessment for CBE novice students appears to encourage them to go through the motions of CBE at the expense of the clinical findings. The proctor marking in this research passed all the students probably due to the educator giving the student the benefit of the doubt. Potentially it is difficult to fail a student if the fail criteria is missing a putative cancer when there are no abnormal patients thus no putative cancers to identify in the test. This type of test may be beneficial for becoming comfortable examining patients but not demonstrating their ability in IDBM.

5. Conclusion

Current training for CBE based in medical schools appears to be inadequate and actual training of IDBM is needed. Despite the best efforts of the breast surgeons teaching the students, the tests do not indicate competency in actual anomaly (breast lesion) detection. One of the reasons may be a lack of real cancer patients for demonstration and teaching purposes. The IEAs used do not actually have any suspicious breast lumps, which limits the students' exposure to abnormalities such as those experienced by experts in clinical situations. Also, with a single teaching associate or IEA the student can't experience the range of normal for softness or lumpiness. Current simulation

models do not show these problems in the complex and nuanced way required for translation of the skill into actual clinical practice.

Current CBE marking schemes favour students who appear confident and professional by going through the motions of the physical examination. This means students are potentially more focussed on inter-personal skills than clinical findings. While looking professional results in a pass mark for the student, it is not sufficient to demonstrate competence in their capacity to IDBM. Looking professional is not sufficient for patients, for whom clinical competence is the primary requirement. If the core skills of IDBM are not being tested, then there can be no assurance of competence in CBE of medical students and graduates.

Teaching and marking that only involves IEA (teaching associates) where pathology is absent or simulated models that do not have the complex multi-layered realism needed by beginner students are not sufficient to teach and test IDBM. The IEA teaching that our novice students received in this trial was not sufficient for them to pass a clinical test. However using the current assessment methods all of the students would have passed, despite their lack of competence in IDBM. Existing student assessment is relevant and important and needs to remain; yet this trial suggests there needs to be an additional test for IDBM comprehension. This implies the teaching needs to be supplemented with an additional component that adds the physical skill of palpation and interpretation of what is felt as a component of the overall teaching.

The results indicate there is no comprehension component in the current CBE testing whilst this is clearly needed. If there was effective testing for IDBM the quality of training could be tested and immediately be improved, especially in relation to the poorest performing students.

Creation of a teaching package with simulation models and a standardized testing tool calibrated to actually test for the students' ability to identify and discriminate between breast masses (IDBM) both normal and pathological would greatly facilitate the teaching of CBE, making it more cost-effective and efficient, thereby saving lives. It is cost-effective because under the current system the overworked hospital doctors and staff have to catch each inadequately trained student at intern stage and individually remediate them. Efficient because if the hospital staff were freed up from teaching the

basics to those students who missed them the first time around, they could concentrate on refining and improving the skills of all students, making the identification of breast masses more accurate. This would then save many unnecessary false-negative referrals and free up much needed resources for the true-positives, thereby improving overall patient care.

This study only included 10 students and further testing is needed to verify these results and also to derive a satisfactory training process or methodology. We were extremely lucky to have recruited a real patient at all considering the timeline between patient diagnosis with a breast lump and further testing can be as short as a day. Our patient only had 24 hours between diagnosis and her hospital follow-up. She stopped to participate in the study on her way to hospital. In addition, this can be a very emotional time for a patient with suspected (putative) breast cancer. Nonetheless, we would recommend extending any future study and recruiting more actual patients for future work.

Acknowledgments

We wish to acknowledge the input from the many people who helped to make this study a success – in particular the staff and doctors at the Breast Clinic, Womens Health Centre, Royal Adelaide Hospital who generously gave their time and allowed us to run the study during clinic hours. Special thanks to Dr Janne Bingham for teaching the students. Thank you to the many volunteers who participated as ‘patients’ and IEAs, to Anne Veitch, Sharlene Lynch and Verna Blewett who came along and helped organise on the day, Dr Henry Fellner for recruiting the actual patient, thanks to the patient who so generously stopped by between medical appointments (you know who you are!), Lilian Fellner, who coordinated and recruited the medical students from Flinders University and lastly thank you to the Flinders University BCS/MD(2020) students for participating in the study.

Authors’ contributions

Daisy Veitch conceived and designed the analysis, organised the logistics, collected the data and wrote the paper. Melissa Bochner co-conceived the study, provided medical expertise, recruited the medical

staff, provided the facilities, conducted the teaching, trained the patients and proctored the students as well as editing the paper. Harry Owen contributed the original idea and gave valuable suggestions to the design of the study. James Veitch conceived the competency score statistic and gave valuable guidance in addition to editing the paper. Richard Goossens and Johan Molenbroek supported the project and edited the final paper.

Conflict of interest

None to report.

Ethics approval

This research has ethics approval SA HREC 34.13.

Funding

The research has been self-funded by the authors.

References

- [1] National Breast Cancer Foundation Inc. What’s The Difference Between A Breast Self-Exam And A Clinical Breast Exam? National Breast Cancer Foundation, Inc.; 2018 [updated 2018. Available from: <https://www.nationalbreastcancer.org/clinical-breast-exam>.] accessed 3 November 2018.
- [2] Worldwide data: World Cancer Research Fund; [Available from: <http://www.wcrf.org/int/cancer-facts-figures/world-wide-data>.] accessed 30 January 2016.
- [3] Ahmed I, Nazir R, Chaudhary MY, Kundi S. Triple assessment of breast lump. *Journal of the College of Physicians and Surgeons—Pakistan: JCPSP*. 2007;17(9):535-8.
- [4] Irwig L, Macaskill P, Houssami N. Evidence relevant to the investigation of breast symptoms: The triple test. *The Breast*. 2002;11:215-20.
- [5] Haakinson DJ, Stucky CC, Dueck AC, Gray RJ, Wasif N, Apsey HA, et al. A significant number of women present with palpable breast cancer even with a normal mammogram within 1 year. *The American Journal of Surgery*. 2010;200(6):712-8.
- [6] Goodson WH. Clinical Breast Examination and Breast Self-Examination. In: Sauter ER, Daly MB, editors. *Breast Cancer Risk Reduction and Early Detection*, 2010.
- [7] Goodson WH, Moore DH. Causes of physician delay in the diagnosis of breast cancer. *Archives of Internal Medicine*. 2002;162(12):1343-8.
- [8] Hendrickx K, Winter BD, Tjalma W, Avonts D, Peer-aer G, Wyndaele J-J. Learning intimate examinations with

- simulated patients: The evaluation of medical students' performance. *Medical Teacher*. 2009;31:e139-e47.
- [9] Chalabian J, Dunnington G. Do Our Current Assessments Assure Competency in Clinical Breast Evaluation Skills? *The American Journal of Surgery*. 1998;175(6):497-502.
- [10] Chalabian J, Formenti S, Russell C, Pearce J, Dunnington G. Comprehensive needs assessment of clinical breast evaluation skills of primary care residents. *Annals of Surgical Oncology*. 1998;5(2):166-72.
- [11] Saslow D, Hannan J, Osuch J, Alciati M, Baines C, Barton M, et al. Clinical Breast Examination: Practical Recommendations for Optimizing Performance and Reporting. *CA Cancer Journal for Clinicians*. 2004(54):327-44.