Detection and Description of Small Breast Masses by Residents Trained Using a Standardized Clinical Breast Exam Curriculum

Elizabeth Steiner, MD, Donald F. Austin, MD, MPH, and Nancy C. Prouser, MS

OHSU Cancer Institute Breast Health Education Program, Oregon Health & Science University, Portland, OR, USA.

OBJECTIVES: We evaluated the effect of standardized clinical breast examination (CBE) training on residents' ability to detect a 3-mm breast mass in a silicone breast model.

METHODS: In this nonrandomized controlled trial, 75 first year residents (R1s) at 8 family medicine, internal medicine, and obstetrics and gynecology training programs received the intervention and second year residents (R2s) did not. Trained residency faculty taught R1s vertical strip, three-pressure method (VS3PM) CBE using a standardized curriculum, including a 1- to 2-hour online self-study with video and 2.5-hour practicum using silicone models and a trained patient surrogate.

RESULTS: Solitary mass detection: 84% by R1s, 46% by R2s (RR=1.82, 95%CI=1.36, 2.43, $P<0.0001$). Of those finding a mass, 62% of R1s and 10% of R2s used at least 5 of 8 standardized descriptors (RR=6.19, 95%CI=2.06, 18.59, $P=0.001$). R1s false positive findings were not statistically different from R2s ($P=0.54$). Both the use of VS3PM and total time spent on CBE were independently highly predictive of finding the mass in either group.

CONCLUSIONS: Most untrained primary care residents are not proficient in CBE. Standardized VS3PM CBE training improves the ability to detect and describe a small mass in a silicone breast model. Better CBE training for residents may improve the early detection of breast cancer.

KEY WORDS: breast cancer; medical education; clinical breast examination; early detection.

J Gen Intern Med 23(2):129-34

DOI: 10.1007/s11606-007-0444-5

© Society of General Internal Medicine 2007

BACKGROUND

Breast cancer is the second leading cause of cancer death in women and affects 1 in 8 women in her lifetime.^{1,2} Long-term survival from breast cancer is improved when breast cancer is detected early.³ Evidence in the literature and our experience in

training over 700 primary care providers in high-quality CBE suggests that primary care physicians may underestimate the importance of CBE and overestimate the efficacy of mammography in the early detection of breast cancer.⁴ The best available evidence on the efficacy of these 2 tests demonstrates that in routine clinical practice CBE sensitivity is 26–35%^{5,6} and mammography sensitivity ranges from 90% in women over 70 years of age down to 60–70% in women under 45 years of age.^{5,7} Although there is a paucity of evidence regarding the effectiveness of screening CBE, the Canadian National Breast Screening Studies 1 and 2 (CNBSS) demonstrated that in women 40–49 and 50–59 years old, CBE alone in a 10- to 15-minute exam resulted in equal 7-year mortality rates to CBE plus mammography.^{8–11} Whereas these breast cancer studies were not structured to assess the effectiveness of screening CBE, their results do suggest that high-quality CBE may play an important role in the early detection of breast cancer.

Most physicians do not screen appropriately for breast cancer.^{12–15} Recently, the Centers for Disease Control and Prevention (CDC) and the American Cancer Society (ACS) published consensus recommendations on the best method of performing CBE, along with a standard lexicon to use in documenting CBE findings.^{16,17} The consensus conclusion is that the vertical strip, three-pressure method (VS3PM) of CBE is the most sensitive technique and most likely to detect breast cancer at its earliest stages. This method contrasts with the circle or wedge, single-pressure methods commonly used in clinical practice. The lexicon includes terms for 8 descriptors: location in 3 dimensions (clock face, distance from nipple, and depth), size, shape, mobility, texture, and margins. The publications recommend standardizing training for CBE and documentation. We developed such a training program in 1999 and have used it to retrain practicing clinicians in Oregon since then (Fig. 1). The training is designed to improve performance in 'the 4 Ds' of breast cancer early diagnosis: detection, description, discrimination, and decision-making. Pre/post training skills assessment on over 700 practicing clinicians has shown an increase in the ability to find masses in a silicone model from 10% to 60% of masses. This is consistent with Fletcher et al., who found that untrained practicing physicians detected only 14% of 3-mm masses in silicone models.¹⁸

We expected it would be more effective and efficient to include standardized CBE training in graduate medical education rather than waiting to attempt retraining when in practice. We wanted to evaluate the effectiveness of training primary care residents with a standardized VS3PM CBE curriculum. To accomplish this, we used a 'train the trainer' model to implement our standardized CBE training in 8 family medicine, internal medicine, and obstetrics and gynecology

The work described in this project was originally presented at the Annual Meeting of the North American Primary Care Research Group, Quebec City, Quebec, Canada, in October 2005.

Received November 22, 2006

Revised May 14, 2007

Accepted September 7, 2007

Published online December 4, 2007

- Self-Study: 1-2 hours. *Full content available from the corresponding author*
 - Breast Anatomy: normal and abnormal
 - Breast Pathology: all types of cancer, benign lesions, skin changes, etc.
 - Breast Imaging: screening vs. diagnostic mammography (BIRAD classification system), ultrasound, magnetic resonance imaging
 - Training Video: full review of entire exam, demonstrated on surrogate patients
 - Management of Abnormal Findings: palpable masses, non-palpable masses on imaging, nipple discharge (includes algorithms)
 - Screening Average Risk Women: taking a complete history, benefits & disadvantages of teaching breast self-exam, counseling women about imaging & risk reduction
 - Supplemental Materials: algorithms for screening and diagnosis, relevant literature, Web resources
 - Knowledge post-test
- Practicum: 2-2.5 hours (One CBE faculty, One CBE surrogate patient, One or Two learners)
 - Baseline skill assessment with two breast models (10-15 minutes)
 - Review of Knowledge post-test (5-10 minutes)
 - Practice core skills with breast models: Perimeter, Pattern, Palpation, Pressure, Description of masses (30 minutes)
 - Practice skills with surrogate patient: communicating about exam with patient, thorough lymph node exam, proper positioning of patient (Cahan position) (45 minutes)
 - Evaluation of skills with surrogate patient: learners perform full exam of one breast without ongoing feedback. Faculty records time taken for exam and thoroughness of coverage. Patient instructor evaluates for performance of palpation at superficial, medium, and deep pressures. (20 minutes)
 - Post-training skill assessment with two breast models (15-20 minutes)
 - Feedback to learners about progress in skills during training (10 minutes)

Figure 1. Training program developed in 1999 used to retrain practicing clinicians in Oregon.

residency programs in Oregon. We hypothesized that (1) CBE-trained residents would perform better on at least 3 of the 4 Ds (detection, description, and decision-making) than residents receiving usual training and (2) a 'train the trainer' model would integrate this educational intervention effectively into multiple types of residency programs. Using a nonrandomized controlled trial study design, we trained some faculty and residents in each program and objectively evaluated the skills of trained versus untrained residents.

METHODS

This project was reviewed by the OHSU Institutional Review Board (IRB) and determined to be IRB exempt as an evaluation of an educational intervention. All the residency program directors supported the project.

Study Design

We conducted a nonrandomized controlled trial in which the study group received standardized CBE training in the first year of a primary care residency, and the control group (second year residents) did not receive training. We elected not to randomize the training among both first (R1) and second (R2) year residents because we believed it would increase the risk of contaminating the control group and because a core goal for this project was the introduction of an educational enhancement into residency training. We enrolled all internal medicine, family medicine, and obstetrics and gynecology residency programs in Oregon (8 programs total with 99 first year residents annually). We developed and provided training for 1 to 4 faculty members in each program in 3 phases.

In phase 1, interested faculty members participated in the existing CBE training program. They then used VS3PM CBE in their clinical practices for 2 to 4 months. Phase 2 required the faculty to participate in group sessions during which we reviewed the training session in detail and provided the information necessary to teach the standardized CBE training sessions consistently. In phase 3, an experienced CBE program faculty mentored all newly trained residency faculty members as they trained residents in their own programs.

Trained residency faculty taught CBE to 76% of their first year residents from September 2003–January 2004 (Table 1). Nonbiased clinical duty assignment resulted in 24% of residents not attending the required training. Experienced CBE faculty mentored residency faculty at each session; they collected data about the new faculty performance using an instrument developed to train faculty for our existing program.¹⁹ Mentors assisted with the teaching when faculty did not address key points or learners had questions the faculty could not answer. The mentors gave the faculty constructive feedback immediately after each session.

From February to May 2004, we administered a single-station Objective Structured Clinical Exam (OSCE) to 70% of trained R1s and 70% of R2s (Table 1). All testing occurred at residents' primary clinical sites 3 to 6 months after R1 training. The OSCE required residents to (1) read a subjective presentation of a well woman in her 40s for routine preventive care; (2) examine a silicone breast model (MammaCare® California C, a model not used during the training) in a standard orientation; and (3) write the objective, assessment, and plan portions of the note. Residents were told the model might contain zero, one, or several discrete masses. The test model contained one 3-mm 40 durometer¹ fixed nodule and an area of approximately 25% simulating normal nodularity.

We scored residents' notes according to whether the solitary nodule was detected and, if so, was described thoroughly and considered concerning requiring further evaluation. Study staff observed all the OSCEs and recorded data about the use of individual exam components and time taken performing the CBE. To prevent observer bias, attending physicians from the residency programs did not observe the OSCEs. All residents completed the CBE in less than 10 minutes.

Evaluation Design

We compared OSCE performance between the R1s and R2s. Our primary outcome measures were the ability to (1) detect the lump in the breast model; (2) describe the findings using the recommended lexicon; (3) discriminate between concerning and nonconcerning masses; and (4) decide on an appropriate follow-up plan for management of abnormal findings. We judged performance on outcomes 2, 3, and 4 against CDC guidelines for documentation and management of abnormal breast findings.^{16,17,20} We scored the key components individually in accordance with the CBE training module and the CDC guidelines.

To ensure inter-observer consistency, we pilot-tested the OSCE on 7 R3s sequentially with all observers scoring simultaneously. In an iterative process, the observers compared scoring and discussed discrepancies, achieving approximately 95% inter-observer consistency. One investigator (E.S.) scored all the written notes.

Statistical Analysis

We used descriptive statistics to characterize participants. We analyzed whether training status predicted performance on all 4 outcome measures: (1) ability to detect 3-mm mass; (2) usage of all 8 descriptors; (3) discrimination between normal nodu-

¹The durometer is a measure of firmness; a durometer of 40 is hard.

Table 1. Residents Trained and Tested, by Specialty and Year of Training

	Trained/eligible (%)		Tested/trained (%)	
Program type	R1		R1	R2
Internal medicine	50/68 (74)		32/50 (64)	48/68 (71)
Family medicine*	20/25 (80)		14/20 (82)	11/19 (58)
Obstetrics and gynecology	5/6 (83)		4/5 (80)	5/6 (83)
Total	75/99 (76)		50/75 (70)	65/93 (70)

Some residents not trained or tested because of conflict with rotation schedules. Nonbiased assignment. *Three trained R1s and 6 R2s not tested because of program's distance from main study site.

larity and a solitary mass; and (4) documentation of assessment and education about risk factors, plans for imaging, biopsy, and follow-up. We compared the proportions of all residents who found the mass within a given exam time. We used life-table survival analysis to plot the time spent on the exam by residents in each of the 2 groups by mass found/not found. We performed univariate logistic regression to identify which exam components were predictors of the primary outcome measure (lump detection). To test for a difference in proportions between the trained and untrained residents with regard to number of descriptors, we used the chi-square test. We computed relative risks and 95% confidence intervals. All statistical analyses were run on SAS version 9.0 (SAS Corporation, Cary, NC, USA).

RESULTS

On the OSCE testing, 84% of R1s and 46% of R2s found the 3-mm mass. This proportional difference held across family medicine and internal medicine; all of the 4 obstetrics and gynecology R1s tested found the mass. R1s were significantly more likely to find the solitary mass than R2s (RR=1.82, 95% CI=1.36–2.43, $P<0.0001$). Regardless of the training status, use of deep pressure during CBE also correlated strongly with finding the mass (RR=1.91, 95%CI=1.44–2.54, $P<0.0001$), as did using the vertical strip pattern (RR=1.89, 95%CI=1.34–2.47, $P=0.0001$), circling downward (RR=2.23, 95%CI=1.56–3.17, $P<0.001$), and having adequate overlap of coverage during the exam (RR=1.67, 95%CI=1.33–2.08, $P<0.0001$) (Table 2). Gender was not associated with mass detection.

Table 2. Univariate Predictors of Finding the 3 mm Mass

Variable	Relative risk	95% Confidence interval		P value*
Training (trained R1s versus R2s)	1.82	1.36	2.43	<.001
Vertical strip pattern	1.89	1.34	2.47	<.001
Adequate overlap	1.67	1.34	2.08	<.001
Slides fingers	1.48	1.16	1.90	.002
Circles downward	2.23	1.56	3.17	<.001
Does deep palpation	1.91	1.44	2.54	<.001
Exam time (per 15 s added)	1.31	1.14	1.47	<.001

*Statistical significance via chi-square test.

Figure 2. Percent of subjects detections mass by time spent in performing exam.

The time spent performing the exam affected mass detection significantly. Regardless of training level, residents who took <90 seconds to perform the exam (39 of 115) found the mass 36% of the time. For those who took 90–149 seconds (33 of 115), mass finding increased to 58%. The largest increase in proportion of residents finding the mass occurred among those who took >150 seconds (43 of 115) in which 90% of residents found the mass (Fig. 2). Figure 3 plots the exam times for each group according to whether they found the mass or not. Although the time spent doing the exam was the most significant variable, using the vertical strip pattern was independently associated with mass finding and whereas the two factors were somewhat confounded, they were identifiably independent, as demonstrated in Table 3.

False positive findings were not statistically different between the 2 groups. Among R2s, 13% found false positive masses compared to 21% of R1s (RR=1.61, $P=0.54$).

Figure 3. Plot of the exam times for each group according to whether they found the mass or not.

Table 3. Vertical Strip Versus Time Taken (<120 vs 120+ s) Effect on Mass Finding

		Vertical strip		Total finding mass
		Yes	No	
Time	Long (120+ s)	39/44 (89%)	8/10 (80%)	87%
	Short (<120 s)	9/14 (64%)	19/50 (38%)	44%
Total finding mass (%)		83%	45%	64%

Pearson correlation coefficient (vertical strip versus time)=0.51.

Analysis of the ability to document findings consistent with the recommended lexicon showed that among those who found the mass R1s were significantly more likely than R2s to include documentation of the majority of the 8 descriptors (Table 4). Documentation of assessment and plan was essentially equivalent between R1s and R2s who found the mass with both groups equally likely to include or omit each of the 5 required assessment and plan items.

DISCUSSION

Regardless of the training status, residents using components of VS3PM were significantly more likely to find a 3-mm mass in a silicone model. Residents formally trained in this method performed better than more experienced but untrained residents. In both groups, those finding the mass spent more time on the exam than those who did not. These results are consistent with previous research showing increased CBE sensitivity with additional exam time,^{18,21} but also suggest that formal training in VS3PM and use of this specific method of CBE enhances the sensitivity of CBE. The value of taking this additional time presents a challenge to busy clinicians who must address multiple needs in the course of routine health screening exams, but the dramatic improvement in sensitivity suggests this time would be well spent.

We hypothesized that our standardized VS3PM CBE training method would enable residents to improve performance of the 4 Ds (detection, description, discrimination, and decision-making). We found improvements with regard to detection and description, and equal capacity with regard to discrimination and decision-making. Particularly significant results include the association of core VS3PM components (consistent search pattern, deep palpation, circling downward, and adequate

overlap of coverage) with detecting the mass; and the markedly improved documentation of findings by R1s.

Unlike the work of Campbell et al.,²² we found an insignificant difference in false positives between R1s and R2s. Although this study was underpowered to detect a moderate increase in false positives, the pre/post training skill assessment revealed an absolute 2% decrease in false positives among R1s. The R1 false positive rate on the OSCE was less than half that of the post training assessment, suggesting that clinical experience in the interval between training and testing may decrease false positives. Some residents interpreted areas representing normal nodularity as abnormal masses.

The highly standardized approach to training both faculty and residents makes this program very transportable to other residency programs whose trainees will use CBE in clinical practice, similar to the transportability of trainings such as the Advanced Cardiac Life Support. The interobserver consistency in scoring the OSCE supports the validity of the results. Finally, the OSCE-tested skills are relevant to actual clinical practice, including careful documentation of findings, assessment, and plan; and appropriate clinical decision-making based on history and physical exam findings.

These results must be evaluated in the light of several possible limitations. Although this study was not randomized, important strengths do exist in nonrandomized controlled studies.²³ If any systematic difference in skills existed between R1s and R2s before training, the advantage would likely fall to the more experienced group, counter to our hypothesis. None of the residency programs changed their resident selection criteria between the R2 and R1 classes; it is therefore likely that the two classes were similar in overall skill and knowledge of CBE at entrance to residency. R1s and R2s performed equally on the assessment and plan section of the OSCE. It could be that our training did not enhance this ability or that the advanced experience of R2s gave them equivalent skill. As the R1s had exposure during training to the type of silicone models used for the OSCE, whereas the R2s did not, the R2s may have been limited in their ability to detect the mass. This is less concerning as R2s who did use components of VS3PM performed better than those who did not.

In addition, a significant percentage of eligible residents did not receive training and of those trained, were not tested. Discussions with the faculty at each program showed that this "drop-out" was nonbiased and based solely on clinical rotation at the times of training or testing. However, it does point out one of the challenges of implementing such a curriculum in residency programs.

Those scoring performance of the actual exam were not blinded to training status (R1 or R2) and, therefore, could have been biased in interpreting and recording exam technique. This could not have affected the difference between groups in detecting the mass, in the time taken in the exam, or the written findings by the residents. In addition, the data on exam technique are very consistent with the detection results, providing internal validity.

Another potential limitation is that the study depended on performance on an OSCE using a silicone breast model and a written history rather than evaluation of a live patient. Review of the literature shows that OSCE performance generally does translate to performance in real clinical settings,²⁴⁻²⁷ although there is no direct evidence that performance of CBE on a silicone model represents performance on a patient.

Table 4. Number of Descriptors Used in Documenting Mass and the Relative Risk for Using Numbers of Descriptors, by Residency Group

Number of descriptors*	R1, N=42 (%)	R2, N=30 (%)	Relative risk	95 % Confidence limits for relative risk	P value
At least 3	39 (93)	21 (70)	1.33	1.03 1.69	.026
At least 4	33 (79)	10 (33)	2.36	1.39 4.01	.002
At least 5†	26 (62)	3 (10)	6.19	2.06 18.59	.001
At least 6†	18 (43)	1 (3)	12.86	1.81 91.13	.011

*Any of the eight: clock face, depth, distance from nipple, size, consistency, mobility, margins, shape.

†Because of small cell size (<5), exact RR confidence limits were calculated.

Delay of diagnosis of breast cancer is among the leading costs of malpractice claims in the United States.²⁸ A standardized clinical breast examination curriculum for primary care residents and retraining of practicing physicians could improve the early detection of breast cancer and more appropriate management of abnormal findings, thereby decreasing this type of malpractice claim.

Whereas this study demonstrated the effectiveness of this training, it did not explicitly demonstrate the benefit of training in residency as opposed to retraining in clinical practice. These results should influence practice only if CBE plays a role in early diagnosis of breast cancer. No trials to date have explicitly compared early diagnosis of breast cancer in patients screened using CBE compared to no CBE. All trials to date have included CBE in both the study and control groups. Determining the true role of CBE in the early detection of breast cancer will require further studies comparing the early detection of breast cancer among women who receive high-quality CBE compared to standard CBE compared to no CBE (with or without mammography in each case to allow determination of the independent effect of high-quality CBE).

Future Studies

There is now a need for studies to follow clinicians trained in residency compared to those retrained later to determine actual CBE skills in practice. Future studies to better assess the role of high-quality CBE in the early detection of breast cancer will better inform physicians on this issue.

CONCLUSION

CBE is an integral part of the early detection of breast cancer. Currently, most residents do not enter graduate medical training with adequate CBE skills and should be trained in residency. This study demonstrates the feasibility and advantage of implementing a standardized curriculum for teaching CBE skills during postgraduate medical education. This standardized approach to teaching CBE shows clear benefit in significantly enhancing CBE sensitivity and appropriate documentation of findings without increasing false positive findings when used with residents. We expect that these improved skills will enhance the early detection of breast cancer, an important diagnosis in primary care practice. Primary care clinicians who perform CBE should be encouraged to use the vertical strip, three-pressure method and to spend at least 2 minutes per breast performing the examination to optimize the early detection of breast cancer.

Acknowledgements: This work was funded in part by a grant from Northwest Health Foundation and supported by the OHSU Department of Family Medicine; Cancer Institute; and Department of Public Health & Preventive Medicine. Northwest Health Foundation had no role in any aspect of the study or the preparation of the manuscript. The Department of Family Medicine, Department of Public Health & Preventive Medicine, and the Cancer Institute did not require that they approve the manuscript.

We are grateful for the assistance of Deanne Neth BA—OSCE observing, data entry, manuscript preparation; Scott Fields MD, Rebecca Rdesinski MSW, and the Working On Research Collabora-

tively group in the Department of Family Medicine—editing; Sara Rogers BS—database design; and the program coordinators in the residency programs. Biostatistical support was provided by Solange Mongoue-Tchokote MS of the Biostatistics Shared Resource of the OHSU Cancer Institute (P30 CA69533).

Conflict of Interest: None disclosed.

Corresponding Author: Elizabeth Steiner, MD; OHSU Cancer Institute Breast Health Education Program, Oregon Health & Science University, 3181 SW Sam Jackson Park Road, Mail Code FM, Portland, OR 97239, USA (e-mail: steinere@ohsu.edu).

REFERENCES

1. Society AC. Estimated new cancer cases and death by sex for all sites, US, 2004.
2. Society AC. Probability of developing invasive cancers over selected age intervals, by sex, US, 1998–2000.
3. Surveillance E, and End Results (SEER) Program. SEER*Stat Database: Incidence—SEER 17 Regs Public—Use, Nov 2005 Sub (1973–2003 varying). National Cancer Institute, DCCPS, Surveillance Research Program, Cancer Statistics Branch. <http://www.seer.cancer.gov>. Accessed May 3, 2007, 2007.
4. Randal J. High expectations for mammography at heart of many breast cancer malpractice cases. *J Natl Cancer Inst*. 2004;96(6):429–30.
5. Bancej C, Decker K, Chiarelli A, Harrison M, Turner D, Brisson J. Contribution of clinical breast examination to mammography screening in the early detection of breast cancer. *J Med Screen*. 2003;10(1):16–21.
6. Park B, Suent-li K, Mi Hye K, Eun-Kyung K, Se Ho P, Kyong Sik L. Clinical breast examination for screening of asymptomatic women: the importance of clinical breast examination for breast cancer detection. *Yonsei Med J*. 2000;41:312–8.
7. Carney P, Miglioretti D, Yankaskas B, et al. Individual and combined effects of age, breast density, and hormone replacement therapy use on the accuracy of screening mammography. *Ann Intern Med*. 2003;138(3):168–75.
8. Baines CJ, Bassett AA. Physical examination. Its role as a single screening modality in the Canadian National Breast Screening Study. *Cancer*. 1989;63(9):1816–22.
9. Miller A, Baines C, To T, Wall C. Canadian National Breast Screening Study: 1. Breast cancer detection and death rates among women aged 40 to 49 years. *Can Med Assoc J*. 1992;147(10):1459–76.
10. Miller A, Baines C, To T, Wall C. Canadian National Breast Screening Study: 2. Breast cancer detection and death rates among women aged 50 to 59 years. *Can Med Assoc J*. 1992;147(10):1477–88.
11. Baines CJ, Miller AB. Mammography versus clinical examination of the breasts. *J Natl Cancer Inst Monogr*. 1997;22:125–9.
12. Lane D, Burg M. Promoting physician preventative practices: needs assessment for CME in breast cancer detection. *J Contin Educ Health Prof*. 1989;9:245–56.
13. Warner S, Sorden J, Solomon L. Physician interest in breast cancer screening education. A survey of Vermont family physicians. *J Fam Pract*. 1989;29:281–5.
14. Costanza M, Luckmann R, Quirk M, Clemow L, White M, Stoddard A. The effectiveness of using standardized patients to improve community physician skills in mammography counseling and clinical breast exam. *Prev Med*. 1999;29:241–8.
15. Fletcher S. Following up abnormal breast cancer screening results: lessons for primary care clinicians. *J Am Board Fam Pract*. 2000;13(2):152–4.
16. Saslow D, Hannan J, Osuch J, et al. Clinical breast examination: practical recommendations for optimizing performance and reporting. *CA Cancer J Clin*. 2004;54(6):327–44.
17. McDonald S, Saslow D, Alciati M. Performance and reporting of clinical breast examination: a review of the literature. *CA Cancer J Clin*. 2004;54(6):345–61.
18. Fletcher SW, O'Malley MS, Bunce LA. Physicians' abilities to detect lumps in silicone breast models. *JAMA*. 1985;253(15):2224–8.
19. Vetto JT, Petty JK, Dunn N, Prouser N, Austin D. Structured clinical breast examination (CBE) training results in objective improvement in CBE skills. *J Cancer Educ*. 2002;17(3):124–7.
20. Oregon Breast and Cervical Cancer Early Detection Program. Algorithms for evaluation of abnormal findings. <http://egov.oregon.gov/DHS/ph/bcc/diag/index.shtml>. Accessed March 31, 2005.

21. **Barton M, Harris R, Fletcher S.** Does this patient have breast cancer? The screening clinical breast examination: should it be done? How? *JAMA*. 1999;282:1270–80.
22. **Campbell HS, Fletcher SW, Pilgrim CA, Morgan TM, Lin S.** Improving physicians' and nurses' clinical breast examination: a randomized controlled trial. *Am J Prev Med*. 1991;7(1):1–8.
23. **Concato J, Shah N, Horwitz RI.** Randomized, controlled trials, observational studies, and the hierarchy of research designs. *N Engl J Med*. 2000;342(25):1887–92.
24. **Durning S, Cation L, Markert R, Pangara L.** Assessing the reliability and validity of the mini-clinical evaluation exercise for internal medicine residency training. *Acad Med*. 2002;77(9):900–4.
25. **Dupras D, Li J.** Use of an objective structured clinical examination to determine clinical competence. *Acad Med*. 1995;70:1029–34.
26. **Cuddy M, Dillon G, Clauser B, et al.** Assessing the validity of the USMLE step 2 clinical knowledge examination through an evaluation of its clinical relevance. *Acad Med*. 2004;79(10 Suppl):S43–5.
27. **Carney PA, Dietrich AJ, Freeman DH Jr, Mott LA.** A standardized-patient assessment of a continuing medical education program to improve physicians' cancer-control clinical skills. *Acad Med*. 1995;70(1):52–8.
28. Data Sharing Project P. PIAA Breast Cancer Claims Study: Physicians' Insurance Association of America; 2002. PIAA online store. <http://www.shop.piaa.us>.