PREVENTIVE MEDICINE 24, 255–258 (1995)

Assessing Breast Self-Examination

POLLY A. NEWCOMB, PH.D.,*,[2] SHARON J. OLSEN, M.S.R.N.,*,[3] FELICIA D. ROBERTS, M.S.,*
BARRY E. STORER, PH.D.,* AND RICHARD R. LOVE, M.D.*

University of Wisconsin Comprehensive Cancer Center, 1300 University Avenue, Madison Wisconsin 53706

Background. **Women practice breast self-examination (BSE) according to their own schedule and skill. This variation in how BSE is performed has complicated the interpretation of studies of BSE efficacy and utilization.**

Methods. **We compared two methods commonly used to assess BSE competency, self-report of practice and ability to detect lumps in a model, among 81 women participating in a controlled toxicity trial of tamoxifen. Subjects were postmenopausal, were under 65 years of age, and had a history of breast cancer within the past 10 years but were currently free of disease. Women were asked to describe their usual BSE practice and were assigned a score (0–10) based on the number of recommended techniques and positions mentioned. Subjects were then instructed to examine a silicone breast model embedded with lumps (HealthEdCo) and to report any abnormalities.**

Results. **Overall proficiency was low by both measures. No significant correlation ($r = 0.16$, $P = 0.15$) was found between the two measures of proficiency. Higher correlations were observed among older women and among those practicing BSE monthly.**

Conclusions. **This study suggests that the two techniques are not comparably evaluating BSE proficiency. Further, both of these methods are likely to be poor measures of true BSE practice.**

INTRODUCTION

Unlike most early detection techniques, the use of breast self-examination (BSE) cannot be standardized in practice. Since the skill of individual practitioners will vary, any evaluations of BSE efficacy and utilization must consider differences in how BSE is performed. Due to the nature of how and when BSE is practiced, a truly valid assessment of BSE proficiency may be impossible to obtain. Therefore, all evaluation studies will necessarily depend upon an indirect appraisal of actual BSE performance. In descriptive studies of BSE, proficiency has been evaluated by observation as a subject demonstrated her personal technique (1–3), through subjects' examination of silicone models (4, 5), or by verbal or written reports of practice (6). Temporal considerations, cost, and convenience argue for the use of self-reported practice, yet these approaches have limitations.

Only a few studies have attempted to evaluate the validity or reliability of reported BSE performance. In these studies, investigators assessed concordance between self-report and observed performance (7) and self-report and model demonstration (8, 9). Two of these studies are limited, however, because the study groups were not typical of the population at risk for breast cancer (7, 8). The common use of a reliable and valid tool to evaluate BSE is important for future studies of efficacy and utilization. In this report, we describe a study designed to assess BSE proficiency by comparing two common methods of BSE performance evaluation: verbal description and competency in detecting lumps in a breast model.

METHODS

Women participating in a randomized controlled toxicity trial of the antiestrogen tamoxifen were invited to participate in this evaluation of BSE (10). These postmenopausal women all had a previous diagnosis of node-negative breast cancer within the past 10 years and were less than 65 years of age. Women with bilateral mastectomies were excluded from the study. Participants returned to the study clinic at regular intervals for examination and educational sessions. A BSE educational session was scheduled for the second study visit at 3 months. Prior to the BSE instruction session, each woman's practice of BSE was evaluated.

Women were first asked if they ever practiced BSE. For practitioners, two assessments were made. A ver-

[1] Supported in part by grants from the American Cancer Society (IN 352829 and PDT 302A and B) and Public Health Service Grant CA 14520 from the National Cancer Institute, National Institutes of Health.

[2] To whom reprint requests should be addressed at Department of Human Oncology, UW Comprehensive Cancer Center, 1300 University Ave., Ste. 4760 Madison, WI 53706-1532.

[3] Current address: Johns Hopkins School of Nursing, 1830 E Monument, Baltimore, MD 21025.

0091-7435/95 $6.00

bal description of a woman's current practice of BSE was elicited by an open-ended question. Responses to this question were scored according to normative criteria recommended by national cancer organizations, minimizing the reporting bias resulting from a woman's recognition of specific performance components. This approach and method of assessing the adequacy of examination were modeled after the National Cancer Institute's (NCI) National Survey on Breast Cancer (11) and have been used with some modifications in previous investigations of BSE (6, 7, 9, 12). In the present study, the question was posed similarly to the one in the NCI survey, but did not include the leading probes of the NCI instruments and allowed the assignment of a summary score totaling 10 points based on the explicit mention of specific examination steps. In the absence of evidence regarding the relative importance of each of these steps, all criteria were equally weighted in determining a final score.

A breast model (prepared by HealthEdCo) was presented to each subject with instructions to examine the model and report any abnormalities detected. This model, prepared specifically for this study, contained six fixed and movable lumps of varying sizes and depths in a silicone matrix. The same model was used by all participants. A woman's score was the number of lumps correctly identified. In addition to these evaluations, subjects were asked to respond to questions regarding specific techniques and frequency of BSE practice and limited demographic characteristics.

Competency scores according to subject characteristics were evaluated using standard tests for comparisons of means. The relationship between the two measures of competency was evaluated with Pearson's correlation coefficient. In turn, the square root of this quantity was used to estimate the validity coefficient for these measures (13), i.e., the correlation between the measure and "true" BSE proficiency (defined as the probability of detecting a clinically significant lesion with BSE as usually practiced).

RESULTS

Of the 97 trial participants completing the BSE education session, 16 women were excluded because they did not practice BSE or were not evaluated using the standard study breast model. Of the 81 subjects assessed, the mean age was 58 years and one-third had some college education (Table 1). On average, subjects were diagnosed with breast cancer 3½ years before this study. Ninety percent of subjects practiced BSE at least once per month; nearly half reported performing BSE more often, usually daily. Overall proficiency scores were low: more than half of the verbal scores were below 5 (on a scale of 1–10) and two-thirds identified less than two lumps (out of a total of six).

The mean competency score, based on the verbal description of technique, was 3.8 (SD 2.0) (Table 2). Increased competency, as measured by this self-report, was significantly associated with higher level of education ($P = 0.02$) and longer duration of time (in minutes) reported doing BSE ($P < 0.001$). The mean number of lumps detected in the six-lump model was 1.3 (SD 1.5). The ability to correctly detect lumps in the model was statistically significantly influenced by higher educational level ($P = 0.003$). Increasing age also significantly increased a woman's ability to detect lumps in a model.

Scores from the two measures of competency were not statistically significantly correlated ($r = 0.16$, $P = 0.15$) in the total study group (Table 3). However, among subgroups of women with higher skill, such as women who reported practicing BSE for a longer duration of time, the two measures were significantly correlated. High BSE frequency, often considered a measure of overall skill, was not associated with a higher correlation between the measures.

TABLE 1
Characteristics of Study Participants

	N	%
Age		
45–54 years	18	22.2
55–59 years	33	40.7
60–65 years	30	37.0
Education		
<12 years	5	6.1
12 years	34	41.9
13 or more years	42	51.8
Time since diagnosis		
<1 year	10	12.3
1–2 years	25	30.8
3–5 years	32	39.5
>5 years	14	17.2
Frequency of BSE		
More than monthly	39	48.1
Monthly	34	41.9
Less than monthly	8	9.8
Time spent doing BSE		
≤2 min	28	34.5
2.5–5 min	36	44.4
>5 min	17	20.9
BSE score		
0	4	4.9
1	6	7.4
2–3	27	33.3
4–5	28	34.5
6+	16	19.7
No. of lumps detected		
0	33	40.7
1	20	24.6
2–3	18	22.2
4+	10	12.3

DISCUSSION

Among women in this study, BSE proficiency was low, whether measured by self-report or lump detec-

TABLE 2

BSE Proficiency According to Selected Subject Characteristics

	Competency score			Lumps detected		
	Mean	(SD)	*P* Value	Mean	(SD)	*P* Value
All subjects	3.8	(2.0)		1.3	(1.5)	
Age						
45–54 years	3.1	(1.8)	0.25	1.8	(1.4)	0.03
55–59	4.0	(2.1)		0.8	(1.2)	
60–65 years	4.0	(2.0)		1.6	(1.7)	
Education						
≤12 years	3.3	(1.7)	0.02	0.8	(1.1)	0.003
13+ years	4.3	(2.1)		1.8	(1.7)	
Frequency of BSE						
More than monthly	3.4	(2.0)	0.09	1.1	(1.4)	0.16
Monthly or less	4.2	(2.0)		1.5	(1.6)	
Time spent doing BSE						
0–2 min	2.7	(1.7)	<0.001	1.5	(1.4)	0.13
2.5–5 min	4.1	(2.0)		1.0	(1.2)	
>5 min	4.9	(1.7)		1.8	(2.1)	
Years since diagnosis						
0–1.5	4.0	(1.9)	0.11	1.3	(1.4)	0.16
1.6–4.9	3.4	(1.8)		1.1	(1.3)	
5+	4.6	(2.3)		1.9	(1.9)	

tion in a breast model. No significant correlations were observed between these two common methods of assessment which we considered *a priori* to be of similar accuracy, as well as adequate surrogates for measuring actual BSE skill.

Three previous reports have specifically evaluated the concordance between various BSE assessment modalities. In a study similar to ours, Stefanek and Wilcox compared procedural assessment (an expanded verbal evaluation) and lump detection in a silicone model among a group of women at high risk for breast cancer (due to a family history of breast cancer) (9). Howe (8) also compared two assessment modalities: verbal description of BSE technique and observed demonstration on silicone models. Among this group of 160 healthy women, no correlation was observed between the two measures. In contrast, Mamon and Zapka (7) observed a significant correlation, 0.62 ($P < 0.01$), between self-report and observed performance among university students participating in a BSE demonstration project. In that study, however, proficiency as measured by written self-report actually underestimated a woman's demonstrated BSE practice.

TABLE 3

Correlation between Verbal Report and Lump Detection

	Pearson's correlation coefficient	*P* value[a]	Validity coefficient[b]
All subjects	0.16	0.15	0.40
Age			
44–54 years	−0.01	0.97	0.10
55–59 years	0.11	0.55	0.33
60–65 years	0.48	0.01	0.70
Education			
≤12 years	0.10	0.55	0.32
13+ years	0.14	0.38	0.37
Frequency of BSE			
More than monthly	−0.09	0.59	0.30
Monthly or less	0.39	0.01	0.62
Time spent doing BSE			
0–2 min	−0.19	0.34	0.44
2.5–5 min	0.40	0.01	0.63
>5 min	0.45	0.07	0.67
Time since diagnosis			
0–1.5 years	0.23	0.32	0.48
1.6–4.9 years	0.09	0.56	0.30
5 or more years	0.21	0.44	0.46

[a] From Fisher's z test.
[b] $\sqrt{r}$ estimates the validity coefficient of both measures if they are assumed to be equally valid.

In our study, the absence of association between the two methods likely reflects a number of factors, including the nature of BSE, the skills of the study population, and the two assessment methods chosen for study. In effect, BSE is a form of physical examination usually performed without the requisite training and monitoring provided to health professionals. Thus, the quality of the self-examination likely varies within one subject's performance of the technique as well as across all subjects' performance. Compounding the difficulties presented by variability in technique, the actual performance of BSE is not easily observed and cannot be directly measured. Certain dynamic components of BSE (e.g., time spent, pressure applied) cannot be reliably measured and therefore provide weak objective criteria against which to check performance. Further, there is no evidence that the more objective measures of BSE performance (e.g., positions, frequency) correlate with true practice. Finally, since complete and accurate BSE ostensibly includes a complex number of positions, techniques, and pressure levels, the simple assessments used in this study may not capture the entirety of the self-examination. Also, the ability to detect lumps during a single evaluation may not accurately reflect a woman's usual proficiency during her own self-examination. The adequacy and validity of using simulated breast models as a measure of personal BSE competency require further investigation. Only one study has demonstrated that improved detection of lumps in silicone models results in improved detection of lumps in the human breast (15).

All subjects in this study had a history of cancer, a group we assumed to have a high level of motivation and BSE proficiency as a result of each woman's post-diagnosis education, which included information about BSE. In fact, women in this study had proficiency as low as or lower than that of other cancer patient series (12, 14) or healthy women (3, 9, 8, 15). Certain subgroups of women had lower competence than others. For example, the women claiming very frequent prac-

tice in this study (usually every day) had lower proficiency. A similar observation was made in another report, where women who claimed to practice the most frequently had the lowest scores on a verbal assessment, likely indicating some form of reporting bias (12). Because competence overall was so low, this study was limited in its ability to detect significant correlations between measures. In the study by Mamon and Zapka (7) subjects had higher scores on the verbal measure than did subjects either in our study or in Stefanek and Wilcox (9), and their results indicate a positive correlation between self-report and demonstration of BSE technique.

While data demonstrating the intermethod reliability of these measures would have been reassuring, the extent to which each technique is valid is the principal concern. The validity of measures of BSE proficiency cannot be directly estimated since there is no practical "true" measure of proficiency with which to correlate these indirect assessments. However, under certain conditions the intermethod reliability coefficient will estimate the product of the validity coefficients of the two methods (13), providing practical information for the appropriate use of these assessments. To interpret these coefficients, however, the errors of the measurements must be (a) uncorrelated with each other and (b) uncorrelated with the true value of BSE proficiency. While these conditions might plausibly hold for the two methods evaluated, it is not possible to characterize further the validity of the measures without additional assumptions about the measures' relative bias (likely a function of true proficiency and error variances). Thus, while it is possible to estimate that the more accurate of the two methods has an overall validity coefficient of at least 0.40, it is not possible to determine which method this is.

The absence of reliability and (presumably) validity observed in this study warrants reconsideration of data regarding both BSE efficacy and utilization. By far the most common evaluation of BSE practice is simply frequency, despite the lack of evidence on optimal interval (16). Additionally, evaluating competency appears to improve the assessment of BSE both intuitively and empirically (1, 12). However, based upon the simple methods evaluated here, it is clear that both approaches are inexact means of measuring ability to identify a breast tumor if one had been present.

Received March 21, 1994
Revision requested June 15, 1994
Accepted October 17, 1994

ACKNOWLEDGMENTS

The authors gratefully acknowledge the staff of the Cancer Prevention Clinic for study support, Tanya Surawicz for statistical assistance, and Mary Pankratz for technical support.

REFERENCES

1. Huguley C, Brown R. The value of breast self-examination. *Cancer* 1981; **47**(5):989–995.
2. Laughter DC, Kean TJ, Drean KD, et al. The breast self-examination practices of high risk women: Implications for patient education. *Patient Couns Health Educ, 3rd Q* 1981; 103–107.
3. Chie WC, Cheng KW, Fu CH, Yen LL. A study of women's practice of breast self-examination in Taiwan. *Prev Med* 1993; **22**:316–324.
4. Alagna SW, Morokoff PJ, Bevett JM, Reddy DM. Performance of breast self-examination by women at high risk for breast cancer. *Women Health* 1987; **22**(2):29–46.
5. Assaf A, Cummings M, Graham S, *et al.* Comparison of three methods of teaching women how to perform breast self-examination. *Health Educ Q* 1985; **12**(3):259–272.
6. Celentano DD, Holtzman D. Breast self-examination competency: An analysis of self-reported practice and associated characteristics. *Am J Public Health* 1983; **73**(11):1321–1323.
7. Mamon J, Zapka J. Determining validity of measuring the quality of breast self-examination: Comparison of self-report with observed demonstration. *Eval Health Prof* 1985; **8**:55–68.
8. Howe HL. Proficiency in performing breast self-examination. *Patient Couns Health Educ, 4th Q* 1980; **2**:151.
9. Stefanek ME, Wilcox P. Breast self-examination among women at increased risk: Assessment of proficiency. *Cancer* 1990; **1**:591–596.
10. Love RR, Newcomb PA, Wiebe DA, Surawicz TS, Jordon VC, Carbone PP, DeMets DL. Lipid and lipoprotein effects of tamoxifen therapy in postmenopausal patients with node-negative breast cancer. *J Natl Cancer Inst* 1990; **82**:1327–1332.
11. National Cancer Institute: National Survey on Breast Cancer. A measure of progress in public understanding. Washington, DC: Govt. Printing Office, 1980; DHHS publication no. (NIH) 81-2306.
12. Newcomb PA, Weiss NS, Storer BE, Scholes D, Young B, Voight LF. Breast self-examination in relation to the occurrence of advanced breast cancer. *J Natl Cancer Inst* 1991; **83**:260–265.
13. Armstrong BK, White E, Saracci R. Principles of Exposure Measurement in Epidemiology. Oxford: Oxford Univ. Press, 1992: 51–83.
14. Haughey BP, Marshall JR, Nemoto T, Kroldart K, Mettlin C, Swanson M. Breast self-examination: Reported practices, proficiency, and stage of disease at diagnosis. *Oncol Nurs Forum* 1988; **15**(3):315–319.
15. Hall DC, Adams CK, Stein GH, Stephenson HS, Goldstein MK. Pennypacker HS. Improved detection of human breast lesions following experimental training. *Cancer* 1980; **56**:508–514.
16. Morrison AS. Screening in Chronic Diseases. Oxford: Oxford Univ. Press, 1992.