Overall Clinical Breast Examination as a Factor in Delayed Diagnosis of Breast Cancer

William H. Goodson III, MD; Dan H. Moore II, PhD

Hypothesis: A standardized description of clinical breast examination (CBE) can predict the risk of delayed diagnosis of breast cancer.

Design: Survey of case series.

Setting: Breast surgery referral practice.

Patients: Consecutive sample of 371 women with 386 breast cancers of any stage for whom overall characteristics of CBE were recorded at the initial consultation.

Intervention: None.

Main Outcome Measures: Overall breast “durity” (from Latin duritia, meaning “hardness”) was recorded as the inverse of whether rib edges could be felt through breast tissue in the most “dur” (firm or hard) part of the breast, usually the upper outer quadrant adjacent to the areola. “Nodularity” was recorded in this same area by means of an ordinal scale ranging from “surface is smooth” to “coarse nodularity.” Delayed diagnosis was tabulated if the patient was told that cancer was not present when there was a sign of cancer on CBE, mammogram, and/or pathology slides. Relative risk of delayed diagnosis was determined within categories of nodularity and durity and within nodularity and durity categories combined.

Results: Diagnosis was delayed for 35 (9.1%) of cancers. Delay was least common (2 [2.2%] of 92) for less dur and less nodular breasts (relative risk, 1.0), most common (18 [13.5%] of 133) for less dur and more nodular breasts (relative risk, 6.23; 95% confidence interval, 3.58-10.22), and intermediate for other descriptions (χ² = 9.08; P = .03). Neither nodularity alone nor durity alone correlated with delay.

Conclusions: A standardized system to describe CBE will alert physicians to an increased risk of delayed diagnosis of breast cancer (especially for women with less dur and more nodular breasts), help improve interpretation of CBE, and reduce delayed diagnosis of breast cancer.

Arch Surg. 2002;137:1152-1156

MISINTERPRETATION of clinical breast examination (CBE) is the leading cause of delayed diagnosis of breast cancer. The leading cause of delay is more commonly stated to be inappropriate physician reassurance that a mass is benign without a biopsy.1-3 However, presumably the delay is unintentional and physicians believe that their advice is correct. This implies that primary findings on examination have been interpreted and, in that process, misinterpreted by physicians.

We believe that misinterpretation is influenced by the variations of the characteristics of overall CBE of individual women. To test this hypothesis, we sought a relationship between delayed diagnosis and overall results of CBE. Testing for this relationship requires (1) criteria for tabulating delay and (2) a system to describe the overall CBE.

We recently described clinical factors related to delay in which we defined delay as a physician action that completed an episode of care without diagnosing cancer of which there was a sign.4 Such a sign might be on CBE, a mammogram, or a pathology slide.

See Invited Critique at end of article

We describe CBE by means of a system that considers nodularity and “durity” (from Latin duritia, meaning hardness; see the “Comment” section) of the breast as independent characteristics.5 This system is different from most other systems that have focused only on presence of degrees of nodularity. We developed this system out of earlier experience through
which we realized that there are differences between the breasts of individual women that cannot be conceptualized as a simple ordinal progression with more or less of a single characteristic. Our work follows that of Rasmussen and Tobiassen, who used independent scales to describe “palpable structure” and nodularity, and the work of Swann et al, who observed significant variations in the resistance of breasts to compression. We emphasize that “durability” in our system refers to resistance to compression during palpation; thus, durability is different from and bears little relationship to radiodensity of tissue as visualized with a mammogram.

Using our CBE system, we asked whether delayed diagnosis was more common for patients with any specific combination of breast nodularity and durability. Finding a relationship would support our hypothesis that the overall characteristics of the breast as palpated during CBE are a factor in misinterpretation of CBE and/or delayed diagnosis.

### PATIENTS AND METHODS

From January 1, 1992, through December 31, 1999, we prepared short clinical abstracts of 433 consecutive patients with breast cancer of any stage referred from multiple sources to one surgeon (W.H.G.). Data included how the patient became aware of a breast abnormality and steps leading to diagnosis. We prepared abstracts in a database and analyzed redacted records with standard statistical programs. We have previously reported the relationship of clinical factors to delayed diagnosis in this set of patients.

Descriptors of the patient’s overall findings on CBE were recorded at the initial physical examination for 371 women with 386 cancers by means of a previously described, 4-point, ordinal system with independent scales to describe durability and nodularity separately. Breast durability was scored as the inverse of the ability to feel ribs through breast tissue in the most “dur” part of the breast—usually the upper outer quadrant. Scores were based on descriptions ranging from “rib edges easily felt” without interference by interposed breast tissue to “rib edges cannot be felt” because of breast tissue and tissue cannot be deformed (Table 1). Breast nodularity was scored as the more or less nodular character of breast tissue in the same, most-dur area of the breast. Despite a justifiable aversion to food names as descriptors, nodularity scores can probably best be understood as ranging from “no nodules at all” through sequential similarities to rice, peas, and beans (Table 1). The description of the contralateral breast was used if the index breast was distorted because of cancer or previous biopsy.

#### Table 1. Ordinal Scales Used to Describe Durability and Nodularity

<table>
<thead>
<tr>
<th>Category</th>
<th>Descriptors</th>
<th>Combined Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>Durability</td>
<td></td>
<td></td>
</tr>
<tr>
<td>a</td>
<td>Ribs easily felt through breast tissue</td>
<td>Less dur</td>
</tr>
<tr>
<td>b</td>
<td>Ribs felt through breast tissue with difficulty</td>
<td>More dur</td>
</tr>
<tr>
<td>c</td>
<td>Ribs cannot be felt through breast tissue</td>
<td>More nodular</td>
</tr>
<tr>
<td>d</td>
<td>Ribs cannot be felt and tissue cannot be deformed</td>
<td></td>
</tr>
<tr>
<td>Nodularity</td>
<td></td>
<td></td>
</tr>
<tr>
<td>A</td>
<td>No nodules at all</td>
<td>Less nodular</td>
</tr>
<tr>
<td>B</td>
<td>Fine nodularity—“rice”*</td>
<td></td>
</tr>
<tr>
<td>C</td>
<td>Prominent nodularity, but not coarse—“peas”*</td>
<td></td>
</tr>
<tr>
<td>D</td>
<td>Course, prominent nodularity—“beans”*</td>
<td></td>
</tr>
</tbody>
</table>

*Foodstuff descriptors are included for simplicity; see “Patients and Methods” section. Also see “Comment” section for description of “durability.”

### RESULTS

Four hundred thirty-four women and 1 man (without delay) had surgery for 434 breast cancers during the study interval. Description of the overall CBE was recorded for 371 women with 386 breast cancers, which are the subject of this article (15 women had synchronous, bilateral cancer). No woman had bilateral delay. For analysis, we consider each breast with cancer as one case.

Durability and nodularity results are given in Table 2. Few breasts had extremes of either durability or nodularity. For this analysis, durability scores were dichotomized into...
less or more dur according to whether rib edges could be felt or could not be felt, respectively. Similarly, nodularity was dichotomized according to whether nodularity was prominent. Of 386 breasts, 225 (58.3%) were less dur and 161 were more dur. One hundred eighty-five breasts (47.9%) were less nodular and 201 were more nodular. There was an inverse relationship between durability and nodularity (χ² = 10.71; P = .001; Table 2).

Delay occurred for 35 (9.1%) of these 386 cancers (Table 3). (There were 42 delays [9.3%] in the entire group of 454 cancers reported previously.) Sixteen women (4.1% of cancers) were inappropriately reassured that a lump was benign. Thirteen (3.4%) had mammography reports, but review of the original slides found incompletely excised ductal carcinoma in situ. One of these women died of invasive cancer identified by a second biopsy of residual calcifications in the same cluster. One had a diagnosis of lobular carcinoma in situ, but review of original slides found incompletely excised ductal carcinoma in situ. Five women had poorly performed FNA. Of these, 4 had no malignant cells obtained by FNA of a palpable mass that was subsequently diagnosed by FNA performed by a physician with specific training in FNA.† The fifth patient had a suspicious mammogram and then had a benign palpation-guided FNA of an area her physician assumed was the same as the area of suspicion on the mammogram; however, subsequent image-guided biopsy found cancer, ie, her physician had performed an FNA in the wrong area.

Analysis of relative risk (Table 3) shows that the fewest delays occurred when patients’ breasts were less dur and less nodular (used as the reference category, with relative risk of 1.0). The most delays occurred when breasts were less dur and more nodular (relative risk, 6.23; P = .03; Table 3). The subset of palpation-related (ie, CBE-related) delays was also related to the same categories of CBE as was the set of all delays (Table 3).

Two hundred thirty-nine women were aged 50 years or older; 47% of 228 postmenopausal women were using hormone replacement therapy at the time of diagnosis. Neither age (P = .20) nor current use of hormone replacement therapy (P = .12) related to the interaction of nodularity and durity (data not shown). Neither nodularity alone (P = .12) nor durity alone (P = .57) was associated with delay (data can be extracted from Table 3).

The risk of delayed diagnosis of breast cancer correlates with the overall characteristics of a woman’s CBE. By inference, therefore, delay relates to interpretations placed on what is felt and how these interpretations are understood in relation to the overall characteristics of a patient’s CBE.

The nodularity and durity scales in this study evolved from our work and the work of others. In textbooks, for example, Haagensen5,12,13 suggested recording nodularity with a sketch, and recently, Gadd and Souba10 noted that nodularity is “generally not abnormal.” However, we have found no textbooks that suggest distinguishing various degrees of nodularity to differentiate between patients.

Ordinal nodularity scales have been used to index response to drugs used to treat benign breast conditions, with most subjects selected on the basis of persistent breast pain.5,11,14 For example, one scale records nodularity as “none, single, few, [or] numerous.” However, a single scale for nodularity is insufficient to describe all breasts.

We had also initially focused on nodularity15,16; however, after we used “hard, confluent, or irregular tissue” in the same ordinal scale as nodularity,17 we realized that breasts have at least 2 characteristics that vary independently.† This independent variation of different charac-

Table 3. Summary of Delays*

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>No delay, No.</td>
<td>92</td>
<td>68</td>
<td>93</td>
<td>133</td>
<td>386</td>
</tr>
<tr>
<td>Delay, No. (% of total)</td>
<td>90 (2.2)</td>
<td>63 (7.4)</td>
<td>83 (10.8)</td>
<td>115 (13.5)</td>
<td>351</td>
</tr>
<tr>
<td>Inappropriate reassurance†</td>
<td>2 (2.2)</td>
<td>5 (7.4)</td>
<td>10 (10.8)</td>
<td>18 (13.5)</td>
<td>35</td>
</tr>
<tr>
<td>Misread mammogram</td>
<td>0 (0)</td>
<td>3 (5)</td>
<td>3 (3)</td>
<td>2 (2)</td>
<td>11</td>
</tr>
<tr>
<td>Inappropriate reassurance and</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>2 (2)</td>
<td>2</td>
</tr>
<tr>
<td>misread mammogram†</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>1 (2)</td>
<td>2 (3)</td>
<td>3</td>
</tr>
<tr>
<td>Poorly performed FNA†</td>
<td>1 (0.3)</td>
<td>0 (0)</td>
<td>0 (0)</td>
<td>4 (5)</td>
<td>5</td>
</tr>
<tr>
<td>RR of delay (95% CI)</td>
<td>1.0 (1.0)</td>
<td>3.38 (1.06-8.29)</td>
<td>4.95 (2.30-9.51)</td>
<td>6.23 (3.58-10.22)</td>
<td>21</td>
</tr>
</tbody>
</table>

*FNA indicates fine-needle aspiration; RR, relative risk; and CI, confidence interval. †Inappropriate reassurance is reassuring the patient that a mass is benign on the basis of palpation alone without biopsy (see “Patients and Methods” section).

**FNA indicates fine-needle aspiration; RR, relative risk; and CI, confidence interval. †Inappropriate reassurance is reassuring the patient that a mass is benign on the basis of palpation alone without biopsy (see “Patients and Methods” section).
Characteristics had previously been observed by Tobiassen et al., who found that, after treatment with danazol, “palpable structure” (see third paragraph following) decreased and “palpation at depth became possible,” but nodularity often did not change.

Single nodularity scales tend to be interpreted to mean that change toward the low end of the scale correlates with a benefit, although this may not be true. For example, tamoxifen citrate—which is effective for breast pain—may cure the pain but leave nodularity unchanged. Since tamoxifen reduces the risk of breast cancer, this is a benefit without decreased nodularity.

Most single nodularity scales index the number of nodules. However, in our observations, the principal difference between patients is not in the number of nodules present, but rather in the overall prominence of nodules. Therefore, a nodularity score is a problem of subjective scaling. We have demonstrated that our scale for the prominence of nodularity is reproducible when the same patients are examined after an interval of 2 to 24 months.

Rasmussen and Tobiassen and Tobiassen et al. first described breasts with independent scales for “palpable structure” and nodularity. They defined palpable structure as the “filling” of breasts by ducts, glands, etc. We agree with their overall assumption of what constitutes palpable structure. However, we believe that using the term palpable structure fosters an inaccurate concept that one can identify specific tissue components—and possibly diagnose—by palpation alone.

We use the word durity, from the Latin duritia, meaning hardness, instead of palpable structure. We chose an uncommon word to emphasize that our defined durity scale is not the same as a statement that a breast is firm or hard, because durity elicits few preconceived ideas, and because durity conveys a specific physical characteristic without interpretation as to presumed cause. Because the breast is palpated against the chest wall during CBE, we index durity according to how clearly rib edges—consistent structures of the underlying chest wall—can be felt through the breast.

Differences in compressibility of the breast, the approximate inverse of durity, were also observed by Swann et al., who asked technicians to estimate the compressibility of breasts during mammography. They found, as we did, that resistance to compression did not correlate with radiodensity visualized on a mammogram.

Since we found neither a relationship of delay to durity (which has been postulated), nor a relationship of delay to nodularity, we did not expect delay to relate to the interaction of nodularity and durity. However, because we have previously shown that nodularity and durity have an inverse relationship, we tested how categories based on interaction of durity and nodularity—without an assumed ordinal relationship of the categories—would relate to delay.

When we reason from observation to explanation, the relationship of delay to the characteristics of overall CBE is plausible. Delays were most common when the breast was more nodular and less dury. In this situation, the breast consists of many nodules grouped together in the breast with the underlying rib edges readily palpated at the same time. The plethora of nodules and irregularities makes it difficult to discern a discrete mass. To complicate matters, this is a common breast pattern, and an examiner can establish a habit of deciding that specific nodules feel benign and usually “get away with it” because the prior probability of cancer in any specific nodule, in any specific patient, is very low.

Fewest delays occurred when the breast was less nodular and less dury. In this setting, there is little nodularity to confuse the examiner and minimal durity to obscure any distortion of tissue by a cancer; it is easy to define the entire shape of ribs so that they are not misinterpreted as nodules. Frequency of delay was intermediate when the breast was both more nodular and more dury, possibly because the examination was obviously difficult and the examiner had a higher index of suspicion.

Technologies such as mammograms, ultrasound, and magnetic resonance imaging all have errors. Consequently, CBE will remain an integral part of breast cancer detection for the foreseeable future. We believe that CBE should be taught in the context of durity and nodularity as independent characteristics found in all breasts. At present, CBE is taught as a sequence of hand and body motions with little consideration of the overall CBE characteristics of an individual patient. Lack of a system forces the clinician to conceptualize each breast finding without reference to the range of variation among breasts.

We have demonstrated the potential of systematic description of overall CBE to alert physicians to increased risk of misinterpretation of CBE, especially for women with less dury and more nodular breasts. If overall characteristics of a woman’s CBE are not recorded systematically, useful information is being discarded. Delayed diagnosis of breast cancer will be reduced when CBEs are based on systematic evaluation of breast characteristics and clinical findings are interpreted in the context of these characteristics.

This study was presented in part as a poster at the San Antonio Breast Cancer Conference, San Antonio, Tex, December 11, 2001.

We thank Brian Mayall, MD, for invaluable criticism during the preparation of the manuscript.

Corresponding author and reprints: William H. Goodson III, MD, 2100 Webster, Suite 401, San Francisco, CA 94115 (e-mail: whg3@cooper.cmcm.org).

REFERENCES

The standardization of the description of the breast physical examination has received little attention to date, particularly in the surgical literature. This study represents a welcome first step in this direction. The authors have coined the term durnity for what appears to be the firmness of the breast to examination and have applied a relatively easily understood threshold for this. The categorical definitions for nodularity, on the other hand, are much more subjective. In an earlier study the authors reported concordance data on the assessment of breast nodularity, but 38.6% of women who were classified as having a nodularity grade of A or B at first examination in that study were thought to have a nodularity grade of C or D at a blinded second examination. It is not clear whether the same or a different observer performed the second examination, or what the average duration between examinations was.

The idea that the CBE characteristics of the breast can be separated into 2 discrete dimensions, durnity and nodularity, which have separate impacts on the likelihood of diagnostic delay, is intuitively reasonable. However, of the 18 instances of delayed diagnosis in the most risky category of more nodular and less durn, 3 were related to misread mammogram, 2 to misread pathology report, and 4 to poorly performed FNA. Thus, only half of the delays in diagnosis are reasonably attributable to the CBE characteristics of the breast, and if the odds ratios were recalculated on this basis, the findings would certainly be much weaker. In addition, it is not possible to conclude whether the effect of durnity and nodularity is a statistically independent effect, since only univariate results are presented.

Finally, for a system such as this to be useful in identifying women at risk for delay in breast cancer diagnosis on the basis of their CBE findings, interobserver and intraobserver variability data are essential. Although it is commendable that data on a reasonably large number of women were collected prospectively, the next step in assessing the applicability of a system like this must involve the demonstration that different observers can reproducibly categorize the palpable texture of the breast.

Seema A. Khan, MD
Chicago, Ill


Correction

Error in Table. In the Original Article by Goodson and Moore titled “Overall Clinical Breast Examination as a Factor in Delayed Diagnosis of Breast Cancer,” published in the October issue of the ARCHIVES (2002;137:1152-1156), an error occurred in Table 2 on page 1153. In that table, the parenthetical dichotomized data should have appeared in a third, unheaded column in each quadrant. The corrected table is reprinted herein. The journal regrets the error.

Table 2. Distribution of Nodularity and Durity Scores*

<table>
<thead>
<tr>
<th>Increasing Nodularity</th>
<th>a</th>
<th>b</th>
<th>c</th>
<th>d</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>A</td>
<td>13</td>
<td>18</td>
<td></td>
<td></td>
<td>40</td>
</tr>
<tr>
<td>B</td>
<td>26</td>
<td>35</td>
<td>(92)</td>
<td></td>
<td>125</td>
</tr>
<tr>
<td>C</td>
<td>63</td>
<td>66</td>
<td>(133)</td>
<td></td>
<td>196</td>
</tr>
<tr>
<td>D</td>
<td>3</td>
<td>1</td>
<td></td>
<td></td>
<td>5</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Increasing Durity, No.</th>
</tr>
</thead>
<tbody>
<tr>
<td>Total</td>
</tr>
<tr>
<td>105</td>
</tr>
<tr>
<td>120</td>
</tr>
<tr>
<td>(138)</td>
</tr>
<tr>
<td>386</td>
</tr>
</tbody>
</table>

*Obtained with 4-point scales and as dichotomized (in parentheses) for the analysis in this report. χ² for 16 categories = 24.13, P = .004; χ² for 4 categories = 10.71, P = .001.