Beyond the Mammography Quality Standards Act: Measuring the Quality of Breast Cancer Screening Programs

Garth H. Rauscher[1]
Anne Marie Murphy[2]
Jennifer M. Orsi[2]
Danielle M. Dupuy[2]
Paula M. Grabler[3]
Christine B. Weldon[4]

OBJECTIVE. A high-quality screening mammography program should find breast cancer when it exists and when the lesion is small and ensure that suspicious findings receive prompt follow-up. The Mammography Quality Standards Act (MQSA) guidelines related to tracking outcomes are insufficient for assessing quality of care. We used data from a quality improvement project to determine whether screening mammography facilities could show that they met certain quality benchmarks beyond those required by MQSA.

MATERIALS AND METHODS. Participating facilities provided aggregate data on screening mammography examinations performed in calendar year 2009 and corresponding diagnostic follow-up, including patients lost to follow-up, timing of diagnostic imaging and biopsy, cancer detection rates, and the proportion of cases of cancer detected as minimal and early-stage tumors.

RESULTS. Among the 52 participating institutions, the percentage of institutions meeting each benchmark varied from 27% to 83%. Facilities with American College of Surgeons or National Consortium of Breast Centers designation were more likely to meet benchmarks pertaining to cancer detection and early detection, and disproportionate share facilities were less likely to meet benchmarks pertaining to timeliness of care.

CONCLUSION. The results suggest a combination of quality of care issues and incomplete tracking of patients. To accurately measure the quality of the breast cancer screening process, it is critical that there be complete tracking of patients with abnormal screening mammography findings so that results can be interpreted solely in terms of quality of care. The MQSA guidelines for tracking outcomes and measuring quality indicators should be strengthened for better assessment of quality of care.

Keywords: breast cancer, disparities, mammography, quality improvement, screening

DOI:10.2214/AJR.13.10806

Received February 20, 2013; accepted after revision May 21, 2013.

Supported by a grant (1 R01 HS018366-01A1) to the University of Illinois at Chicago from the Agency for Health Research and Quality and by grants to the Metropolitan Chicago Breast Cancer Task Force from Susan G. Komen for the Cure and the Avon Foundation for Women.

[1]Division of Epidemiology and Biostatistics, School of Public Health, University of Illinois at Chicago, Chicago, IL.

[2]Metropolitan Chicago Breast Cancer Task Force, 1645 W Jackson Blvd, Ste 450, Chicago, IL 60612. Address correspondence to A. M. Murphy (annemarie_murphy@rush.edu).

[3]Robert H. Lurie Comprehensive Cancer Center, Northwestern University Feinberg School of Medicine, Chicago, IL.

[4]Executive Frameworks and the Center for Business Models in Healthcare, Chicago, IL.

AJR 2014; 202:145–151

0361–803X/14/2021–145

In the United States, non-Hispanic black women are more likely than non-Hispanic white women to die of breast cancer despite being less likely to have the disease diagnosed. In Chicago, this disparity is especially large: Breast cancer mortality in Chicago is 61% higher for African-American women, representing one of the highest documented disparities in the United States [1]. There are many potential contributors to the disparity, including established differences in tumor aggressiveness [2–4], access to and use of mammography [5], and timeliness and quality of treatment [6]. Central to our study was whether variation in the quality of mammography and its effectiveness could also contribute to this disparity [7]. A task force was established in Chicago in 2007 to explore this possibility [8] in addition to examining the aforementioned factors.

Indications of problems with the quality of mammography were first seen in the mid 1980s. A study known as the Nationwide Evaluation of X-Ray Trends (NEXT) [9] conducted by state radiation control agencies in cooperation with the U.S. Food and Drug Administration (FDA) showed that image quality in perhaps as many as one third of the facilities was less than desirable. The Mammography Quality Standards Act (MQSA) was instituted in 1992 in response and as an attempt to improve the quality of breast cancer screening with mammography nationwide. It set out basic standards that a facility needed to meet to be certified under MQSA. These included standards related to mammography machine calibration, maintenance and quality control, and qualifications of staff. MQSA was reauthorized in 1999 to update experience and continuing education requirements for medical physicists and radiologic technologists and to

clarify equipment standards. With that update, each facility was additionally required to have a system in place to ensure that mammography results were communicated to patients in a timely manner and in terms that a nonprofessional would understand [10]. Federal regulations derived under the authority of MQSA effective in 2002 spelled out further quality assurance measures. Included was a requirement for a medical outcomes audit to follow up on the disposition of all abnormal mammography findings and correlation of pathologic results with the interpreting physician's findings [11]. All interpreting physicians at a facility are required to perform these outcomes analyses individually and collectively.

Although the regulations state that the provisions are designed to ensure the reliability, clarity, and accuracy of the interpretation of mammograms, in truth, the regulations do not require rigorous patient tracking for several reasons. First, a facility is required only to obtain pathologic and surgical reports and to review screening and diagnostic mammograms in cases that subsequently become known to the facility. There is no requirement for due diligence in actively determining whether a patient with abnormal mammography findings subsequently has breast cancer diagnosed. Second, the regulations require analysis only of mammograms interpreted as suspicious or highly suggestive of malignancy (BI-RADS categories 4 and 5) rather than all abnormal results, including those designated incomplete (BI-RADS 0). Thus FDA guidance documents acknowledge that a screening program that never classifies a lesion as BI-RADS 4 or 5 (e.g., facilities that only conduct screening and those that do not use BI-RADS categories 4 and 5 to interpret their screening mammograms) will have no required patients to track and thus no patients to include in the required audit. Third, MQSA law and regulations do not require facilities to separate screening from diagnostic mammography results, without which measures of screening quality become meaningless [11]. Fourth, MQSA law and regulations do not specify which quality metrics must be included in medical audits of facilities or individual radiologists, leaving this to the discretion of each facility and radiology practice [11].

Some organizations have recommended strategies in addition to MQSA guidelines for improving the quality of mammography. The National Cancer Policy Board (NCPB) was asked by the U.S. Congress to review the adequacy of MQSA, which was due for reauthorization in 2007. NCPB published a report through the Institute of Medicine in 2005 recommending that the required medical audit component of MQSA be standardized and that institutions voluntarily participate in a more advanced medical audit [12]. NCPB reiterated the importance of separating screening mammography data from diagnostic mammography data. The American College of Radiology also recommends that facilities meet certain additional quality benchmarks above and beyond the MQSA guidelines pertaining to proportion of abnormal screening mammograms (recall rate), timeliness of follow-up, extent of screen-detection (i.e., cancer detection rate for screening mammograms), and ability to detect small and early-stage tumors [12]. None of these recommendations have been incorporated into MQSA guidelines.

It is unclear the extent to which institutions are tracking or attempting to track additional data beyond those necessary to meet minimum MQSA requirements and whether the data are of sufficient quality to make statements about screening mammography quality. Little research has been done to examine how institutions go about tracking patients whose mammography results are abnormal, especially institutions performing screening mammography but not generally performing biopsies, and the level of success in obtaining follow-up information from other institutions. We used data collected from screening mammography facilities as part of a quality improvement project and determined whether facilities were able to collect data to calculate certain quality metrics and to show that they could meet certain benchmarks pertaining to quality of the mammography process.

Materials and Methods

Data for analysis were collected by the Chicago Breast Cancer Quality Consortium, which is a project of the Metropolitan Chicago Breast Cancer Task Force [8]. The consortium was created in 2008 in an effort to address breast cancer mortality disparity in Chicago through quality improvement. The aim of the consortium was to recruit institutions that screen for, diagnose, and treat breast cancer and to measure and improve the quality of breast care provided. Participating facilities included lower-resourced facilities and those serving predominantly minority or underserved patients. Expert advisory boards were established for both mammography screening quality and breast cancer treatment quality to identify which measures were both high priority and could be feasibly estimated through collection of aggregate data from facilities. All participating institutions obtained a signed data-sharing agreement and institutional review board approval for the study. Institutions that lacked a review board used the Rush University institutional review board. Electronic data collection forms were designed for collecting data on the screening mammography process pertaining to screening mammograms obtained during calendar year 2009. A series of webinars were conducted to familiarize staff at each institution with the data collection form and the submission process and to emphasize specific points pertaining to quality data. Emphasis was placed on submission of patient-level counts as opposed to procedure-level counts and on explicitly accounting for missing data. These issues had arisen during pilot data collection on screening mammogram data for calendar year 2006.

The screening mammography process was defined as the entire process from the initial screening mammographic examination through diagnostic follow-up imaging, biopsy, and breast cancer diagnosis. Figure 1 shows a listing of the requested aggregate counts pertaining to patients who underwent screening at each institution during calendar year 2009. The data collection instrument was created in a spreadsheet with each section of Figure 1 in bold type in its own table within the spreadsheet. Autocalculated cells and data validation checks were built into the instrument to help guide the data collection and entry process. For example, in the purely hypothetical data in Figure 1, among 1000 screened patients, 135 abnormal screens resulted in 15 biopsies and five diagnoses of breast cancer, of which four were known to be early-stage and minimal cancer (Fig. 1).

Screening Measures

From the data we estimated the following 11 measures of the screening process. Benchmarks for these measures were established by consulting American College of Radiology benchmarks and through consultation with clinical experts in these fields who participate on our mammography quality advisory board. The benchmarks also take into account population-based estimates and ranges for these measures from the Breast Cancer Surveillance Consortium [13].

The following six measures of mammogram interpretation and diagnostic follow-up were calculated for all participating facilities.

Recall rate was the proportion of screening mammograms interpreted as abnormal (BI-RADS 0, 4, or 5). The benchmark for recall rate was met if no less than 5% and no greater than 14% of screening mammograms were interpreted as abnormal. Not lost at imaging was the proportion of abnormal screening mammograms followed up with diagnostic imaging within 12 months of the screening mammogram (benchmark of 90% and above). Timely

follow-up imaging was diagnostic imaging within 30 days of an abnormal screen among those undergoing diagnostic imaging within 12 months of the screen (benchmark of 90% and above). Biopsy recommendation rate was the proportion of abnormal screening mammograms resulting in a recommendation for biopsy (benchmark of 8–20%). Not lost at biopsy was the proportion of women with a biopsy recommendation who underwent biopsy within 12 months of the abnormal screen (benchmark of 70% and above). Timely biopsy was biopsy within 60 days of the abnormal screen among those undergoing a biopsy within 12 months of the screen (benchmark of 90% and above).

The following three measures of cancer detection were calculated for facilities that reported at least 1000 screening mammograms during calendar year 2009. Cancer if abnormal screen was the proportion of patients with abnormal screens who received a breast cancer diagnosis within 12 months of the screen, also known as positive predictive value 1 (benchmark of 3–8%). Cancer if biopsied was the proportion of patients undergoing biopsy after an abnormal screen who received a breast cancer diagnosis within 12 months of the screen, also known as positive predictive value 3 (benchmark of 15–40%). Cancer detection rate was the number of cases of breast cancer detected after an abnormal screen result for every 1000 screening mammograms obtained (benchmark of 3–10 per 1000).

The following two measures of early cancer detection were calculated for facilities that reported at least 10 screen-detected cases of breast cancer during calendar year 2009. Proportion minimal was the proportion of screen-detected cases of breast cancer that were either in situ or no larger than 1 cm in largest diameter (benchmark of > 30%). Malignant breast tumors with unknown minimal status were excluded from both numerator and denominator of this measure. Although we attempted to collect information on lymph node status for minimal tumors, many institutions were unable to provide these data reliably, so we did not include lymph node status in our definition of minimal cancer.

Proportion early stage was the proportion of screen-detected cases of breast cancer that were either in situ or stage I (benchmark of > 50%). Cases of breast cancer of unknown stage were excluded from both numerator and denominator of this measure.

We calculated each of the 11 measures separately for each institution when both numerator and denominator data were available. Facility estimates for each measure were plotted with corresponding 95% CIs for visual depiction of the range encountered and the stability of those estimates. The plot for each estimate was overlaid with a shaded area to help identify values that fell within the range of acceptable values for the benchmark (Fig. 2). For each measure, we calculated the percentage of facilities that were able to show they met each benchmark. Facilities with incomplete or missing data on a given measure were defined as not able to show they met that benchmark. These results are presented in Table 1.

ALL Screening Mammograms January 1, 2009–December 31, 2009

TABLE 1. Screening Mammograms in Calendar Year 2009

Screening BI-RADS Category	N
BI-RADS 0	100
BI-RADS 1	710
BI-RADS 2	150
BI-RADS 3	5
BI-RADS 4	20
BI-RADS 5	15
Other BI-RADS	0
Total Number of Screening Mammograms	**1000**

TABLE 2. Follow-Up Imaging

	N
There are 100 BI-RADS 0 screeners	**100**
How many of 100 BI-RADS 0 screeners received follow-up imaging within 12 months of the screen?	80
How many of 80 BI-RADS 0 screeners received their follow-up imaging within 60 days of the screen?	60
How many of 60 BI-RADS 0 screeners received their follow-up imaging within 30 days of the screen?	40
There are 35 BI-RADS 4, 5 screeners	**35**
How many of the 35 BI-RADS 4 and 5 screeners received follow-up imaging within 12 months of the screen?	30
How many of these 30 BI-RADS 4 and 5 screeners received follow-up imaging within 60 days of the screen?	25
How many of these 25 BI-RADS 4 and 5 screeners received follow-up imaging within 30 days of the screen?	20

TABLE 3. Follow-Up Biopsy Among BI-RADS 0, 4, 5 Screeners

	N
There are 135 BI-RADS 0, 4, 5 screeners	**135**
How many of these 135 received a biopsy recommendation within 12 months of the screen?	20
How many of these 20 actually received a biopsy within 12 months of the screen?	15
How many of the 15 patients received their biopsy within 60 days of the screen?	10

TABLE 4. Cancer Among BI-RADS 0, 4, 5 Screeners

	N
15 patients received a biopsy within 12 months of the screen	**15**
How many of these 15 patients who received biopsy had a cancer diagnosis within 12 months of the screen?	5
5 patients had a diagnosis of breast cancer	**5**
How many of those 5 are UNKNOWN stage?	0
How many of those 5 are STAGE 0, 1?	4
How many of those 5 are STAGE 2, 3, 4?	1
5 patients had a diagnosis of breast cancer	**5**
How many of these 5 diagnoses are UNKNOWN tumor size?	0
How many of these 5 diagnoses ARE NOT MINIMAL cancer?	1
How many of these 5 diagnoses ARE MINIMAL cancer?	4

Fig. 1—Example of data collection instrument used for counts pertaining to mammography screening processes at individual institutions. Data shown are exemplar and do not come from actual institution.

Facility Designation and Benchmarks Met

We tabulated the percentage of facilities that met each benchmark by facility designation status as assigned by the American College of Surgeons Commission on Cancer (ACS CoC) and the National Consortium of Breast Centers (NCBC). We compared 24 facilities with ACS CoC designation with 28 facilities without such designation on each benchmark. We also compared 10 facilities with NCBC designation with the 42 facilities without such designation. Facilities were also categorized by whether they were hospitals that met the criteria for being disproportionate-share hospitals ($n = 12$) or were public facilities that served predominantly uninsured patients ($n = 6$). These 18 facilities (collectively referred to as disproportionate-share facilities) were compared with the other 34 facilities on benchmarks met.

Results

Screening Results

Fifty-two mammography facilities contributed data on a total of 330,806 screening mammographic examinations (mean, 6362; range, 136–23,898). These facilities represented 27% of facilities in the six-county area providing mammography in 2009. In the city of Chicago, one half (25/49) of mammography facilities participated. There was wide variation in measures calculated across facilities (Table 1). The percentage of institutions meeting each

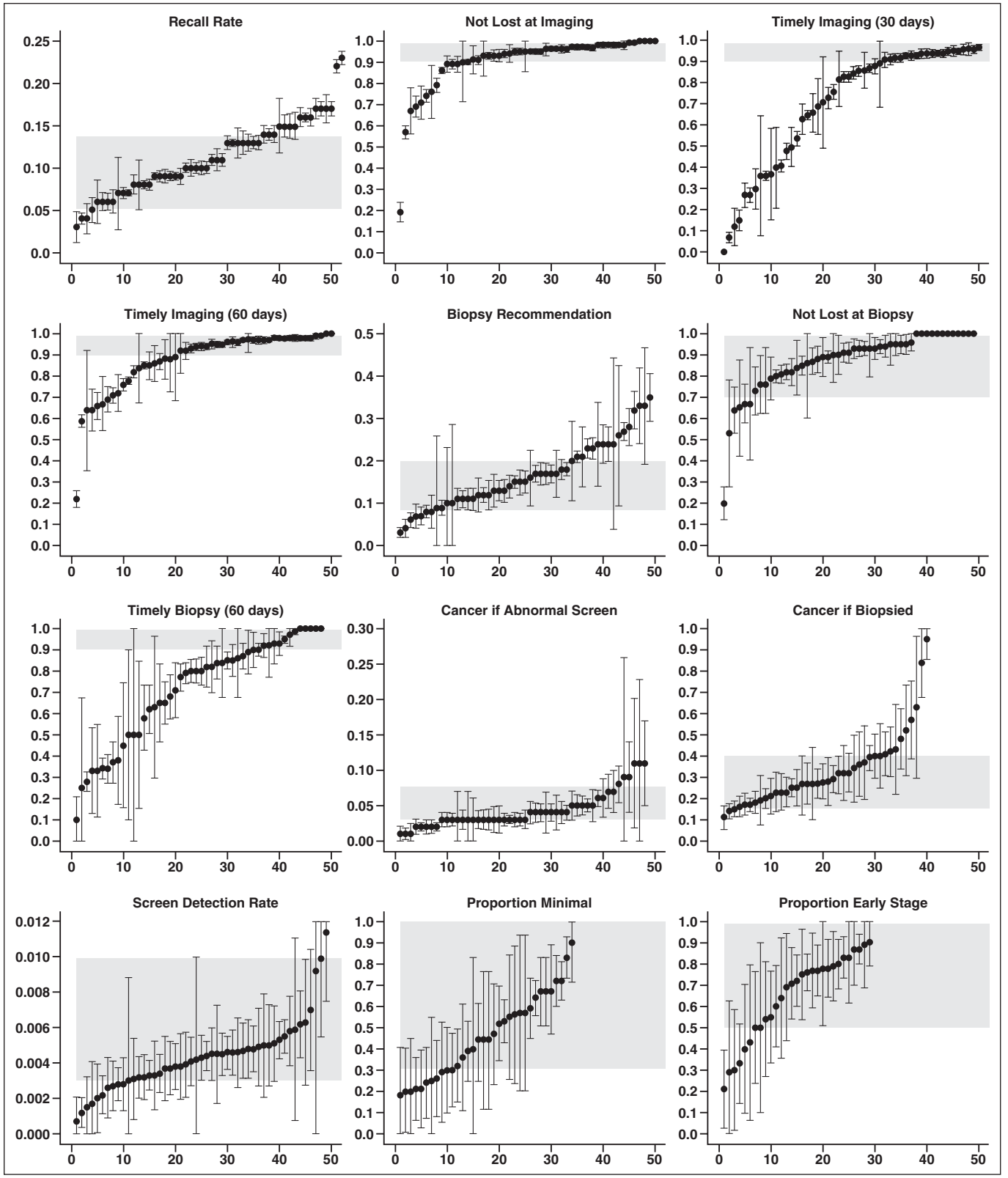

Fig. 2—Graphs show screening process measures and 95% CIs (*y*-axis) for each institution (*x*-axis) and benchmark ranges (*shaded area*). For each graph, facilities are ordered from smallest to largest value of corresponding measure; therefore, ordering of specific facilities differs from graph to graph.

benchmark varied from a low of 27% (receiving a timely biopsy, defined as within 60 days of the screen) to a high of 83% (receipt of a biopsy if recommended).

Approximately 70% of facilities met the benchmark for recall rate (Table 1). A disproportionate number of facilities had values above versus below the benchmark range (Fig. 2). The benchmarks for not lost to follow-up at diagnostic imaging and biopsy were met by 73% and 83% of facilities (Table 1).

The benchmark for proportion with timely diagnostic imaging among patients not lost to follow-up was met by 62% of facilities. The other facilities had a wide range of suboptimal values, some fairly close to meeting the benchmark but many others far below the benchmark of greater than 90% within 30 days of screening. A similar pattern was observed for timeliness of biopsy. Estimates were stable enough to suggest real deficiencies in timeliness of care for both measures (Fig. 2).

Between one half and three fourths of facilities could show that they met benchmarks for screen-detected cancers among abnormal screening results, among patients who underwent biopsy, and for all screens combined. There was a tendency for the cancer rate among abnormal screens to be below or near the low end of the benchmark for many institutions. Many institutions also were below or at the low end of the benchmark for screening cancer detection rate. Whereas approximately one half of institutions could show that they met benchmarks for early detection of breast cancer (proportion of minimal and early stage), the stability and precision of these estimates were limited by the small number of cases of cancer detected at many institutions (Fig. 2).

Facility Designation and Benchmarks Met

Twenty-four of the 52 participating facilities were accredited through the ACS CoC, and 10 of these 24 were accredited with the NCBC. Facilities accredited through either body were more likely than nonaccredited facilities to meet benchmarks related to biopsy and cancer detection (Table 2). Most notably, 79% of ACS CoC centers met the benchmark for early stage detection compared with only 14% of facilities not accredited through ACS CoC (80% vs 36% for NCBC accreditation). Disproportionate-share facilities were less likely than other facilities to meet specific benchmarks, including those related to follow-up and timeliness of imaging and timeliness of biopsy (Table 3).

Discussion

A high-quality screening mammography program should find breast cancer when it exists, find it early and when it is small so that treatments can be more effective, and ensure that when a mammogram shows something suspicious that a woman receives follow-up quickly. The goal of this study was to examine the extent to which institutions are able to collect the data required to measure mammography quality and whether institutions could show that they were performing high-quality screening mammography according to established benchmarks.

Whether a facility meets a particular benchmark can be a function of two distinct processes, namely quality of care and quality of data submitted to the consortium. Certain benchmarks such as proportion of minimal and early-stage cancers and cancer detection rate are also likely to be sensitive to the patient mix with respect to how regularly or irregularly screened the patient population is and the age distribution. However, when one looks at these measures together, it is possible to get a reasonable indication of whether quality issues are present. For instance, at a facility such as a disproportionate-share facility where the patient population is likely to be less regularly screened, one may expect the cancer detection rate to be higher. However, with such a population mix, one also expects the proportion of minimal and early-stage cancers detected to be lower. However, a low cancer detection rate and small proportion of minimal and early stage cancers may imply that some cases of cancer are being missed. Facilities do not readily have data available on how well screened their patient population is, and this is a limitation of our analysis.

The ability to accurately measure the quality of breast-related health care provided to patients depends crucially on the extent to which institutions are able to collect accurate data required to measure mammography quality. For facilities that fulfill only the minimum require-

TABLE 1: Percentage of Facilities Able to Show They Met Specific Benchmarks for Mammography Screening During Calendar Year 2009

Measure	Benchmark (%)	Met Benchmark	
		No.	%
All facilities (n = 52)			
Recall rate	5–14	36	69
Not lost at imaging	≤ 10	38	73
Timely diagnostic imaging	≥ 90	32	62
Biopsy recommendation rate	8–20	29	56
Not lost at biopsy	≥ 70	43	83
Timely diagnostic biopsy	≥ 90	14	27
Cancer detection (n = 45 facilities with at least 1000 screens)			
Cancer if abnormal screen (positive predictive value 1)	3–8	34	76
Cancer if biopsied (positive predictive value 3)	15–40	29	64
Cancer detection rate	3–10 per 1000	34	76
Early detection (n = 41 facilities with at least 10 detected cases of cancer)			
Proportion minimal	> 50	22	54
Proportion early stage	> 30	22	54

TABLE 2: Percentage of Facilities Meeting Each Benchmark by Facility Designation

Measure	American College of Surgeons Commission on Cancer Designation		National Consortium of Breast Centers Designation	
	No (n = 28)	Yes (n = 24)	No (n = 42)	Yes[a] (n = 10)
Recall rate	68	71	67	80
Not lost at imaging	71	75	71	80
Timely imaging	64	58	60	70
Biopsy recommendation	57	54	57	50
Not lost at biopsy	75	92[b]	81	90
Timely biopsy	18	38[b]	19	60[d]
Cancer among biopsied	39	75[c]	52	70
Cancer among abnormal screen results	57	79[c]	64	80
Cancer detection rate	68	79	71	80
Early-stage cancers	14	79[e]	36	80[d]
Minimal cancers	32	67[d]	40	80[c]

[a]Of the 24 American College of Surgeons Commission on Cancer accredited centers, 10 were also accredited through the National Consortium of Breast Centers.
[b]$p \leq 0.20$.
[c]$p \leq 0.10$.
[d]$p < 0.01$.
[e]$p < 0.001$.

ments of MQSA, accuracy of data collection is potentially threatened by lack of sufficient follow-up of abnormal screening results and possibly by incomplete or nonexistent differentiation between screening and diagnostic mammograms. Facilities contributing data for these analyses used a wide range of systems for tracking screening mammogram results and diagnostic follow-up, from rudimentary ad hoc methods to state-of-the-art electronic commercial mammography databases. In addition, although many institutions lacked staff time allocated to tracking abnormal mammograms, others allocated a full-time staff member for this purpose. This lack of standardization in terms of data collection systems and staffing resources would be expected to produce variation in the quality of data submitted for this or any similar quality improvement effort. Different measures have different levels of difficulty in terms of data collection and tracking, and what follows is a description of our results interpreted in the context of possible data collection inaccuracies.

Recall rate was defined as the proportion of screening mammograms interpreted as BI-RADS category 0, 4, or 5, which by definition require diagnostic follow-up imaging or biopsy. According to the American College of Radiology the percentage of patients recalled after screening mammography should be 10% or less [14]. The federal government is considering instituting a reimbursement policy in favor of recall rates not exceeding 14%. Results of some research studies suggest that recall rates of approximately 5% achieve the best tradeoff of sensitivity and positive predictive value [15, 16]. For these analyses, the benchmark for recall rate was met if no less than 5% and no more than 14% of screening mammograms were interpreted as abnormal. We observed institutions at both ends of the spectrum in terms of recall rates that were too low or too high, and results could be used to determine whether there is a need to improve the quality of mammography interpretation. For example, a low recall rate could be a function of a highly screened population but also suggests insensitive interpretation of screening mammograms. Too high a recall rate could be a function of an infrequently screened population or a patient population with less access to previous images but also can suggest that too many patients are undergoing workups, resulting in excessive morbidity and financial costs.

Lost to follow-up at either follow-up imaging or biopsy could simply reflect that a patient leaves the facility to undergo follow-up care elsewhere. However, for a facility to measure and improve the effectiveness of its screening program, obtaining this follow-up information from the other institution or health care provider is critical.

The benefits of routine screening could be diminished by long delays in receiving follow-up care. We observed institutions with apparent deficits in terms of timeliness of care. A qualitative analysis of a subset of sites revealed that all sites reported following MQSA requirements to send screening results to patients and follow-up letters to those with abnormal results, but less than one half attempted to contact by telephone all screening patients with abnormal results (Weldon CB, et al., unpublished results). This general observation could help to explain why timeliness of follow-up appeared problematic for many sites [17].

To assess the effectiveness of detection of early-stage and minimal-size malignant tumors, we asked for the proportion of cases of screen-detected breast cancer that were minimal and the proportion that were early stage, excluding from our calculations facilities detecting fewer than 10 cases of cancer. Only approximately one half of the remaining facilities were found to meet each of these two benchmarks. Insufficient rates of early detection could be a function of an infrequently screened population but also suggests that for many women, early-stage breast cancer is missed on a previous mammogram only to be detected on a subsequent mammogram at a later stage. This finding is in line with other research findings made by our group, in which previous images of women with breast cancer diagnoses were analyzed for potentially missed breast cancer. This research showed higher rates of poorer-quality imaging and potential missed detection among publicly insured women, poor women, and women with less education, indicating that these groups were accessing lower-quality mammography [18, 19].

TABLE 3: Percentage of Facilities Meeting Each Benchmark by Facility Designation as Disproportionate Share

Measure	Disproportionate Share (%)		*p*
	No (*n* = 34)	Yes (*n* = 18)	
Recall rate	76	56	0.12
Not lost at imaging	82	56	0.04
Timely imaging	76	33	0.002
Biopsy recommendation	59	50	
Not lost at biopsy	85	78	
Timely biopsy	35	11	0.06
Cancer among biopsies	62	44	
Cancer among abnormal screen results	85	33	0.001
Cancer detection rate	76	67	
Early-stage cancer	50	33	
Minimal cancer	53	39	

Note—$p > 0.20$ suppressed. Six public facilities were included in the definition of disproportionate share.

There were other limitations to our study. Only 27% of facilities in the six-county area and 50% of mammography facilities in the city of Chicago participated in this voluntary effort. Nonetheless, they represented a range of public, private, and academic facilities not typically seen in mammography quality efforts of this type, which tend to be heavily weighted toward academic and higher-resource facilities. It is possible that institutions that are more secure in the quality of their data would be more likely than others to participate, in which case these results might be overly optimistic regarding the ability of institutions more generally to meet quality benchmarks.

Conclusion

We found that most institutions are not tracking additional data beyond those necessary to meet minimum MQSA requirements. The minimum MQSA requirements in and of themselves are not useful for understanding an institution's mammography screening process from a radiologist's quality of reading perspective and from the perspective of timeliness of follow-up and linkage to biopsy services when necessary. As a result, although data were generally consistent with many quality deficits, these same data were frequently insufficient for making definitive statements about mammography quality. Measures that were likely least affected by insufficient tracking were those pertaining to recall for abnormal screening results and timeliness of diagnostic imaging and biopsy, all of which at least in theory were based on well-defined denominators. These measures strongly suggest that the recall rate is sometimes too low or too high and that timeliness is a real problem across many facilities. To differentiate issues of tracking from issues of quality of care, it is crucial that complete tracking be performed so that results can be interpreted solely in terms of quality of care. This will only happen in a general sense if MQSA is reauthorized and the guidelines for tracking outcomes and measuring quality indicators are strengthened to better reflect actual quality of care as suggested in 2005 by the National Cancer Program Board.

Acknowledgment

We thank the institutions and their staff who collected data for this effort.

References

1. Whitman S, Orsi J, Hurlbert M. The racial disparity in breast cancer mortality in the 25 largest cities in the United States. *Cancer Epidemiol* 2012; 36:e147–e151
2. Carey LA, Perou CM, Livasy CA, et al. Race, breast cancer subtypes, and survival in the Carolina Breast Cancer Study. *JAMA* 2006; 295:2492–2502
3. Cunningham JE, Montero AJ, Garrett-Mayer E, Berkel HJ, Ely B. Racial differences in the incidence of breast cancer subtypes defined by combined histologic grade and hormone receptor status. *Cancer Causes Control* 2010; 21:399–409
4. Ooi SL, Martinez ME, Li CI. Disparities in breast cancer characteristics and outcomes by race/ethnicity. *Breast Cancer Res Treat* 2011; 127:729–738
5. Sabatino SA, Coates RJ, Uhler RJ, Breen N, Tangka F, Shaw KM. Disparities in mammography use among US women aged 40-64 years, by race, ethnicity, income, and health insurance status, 1993 and 2005. *Med Care* 2008; 46:692–700
6. Elmore JG, Nakano CY, Linden HM, Reisch LM, Ayanian JZ, Larson EB. Racial inequities in the timing of breast cancer detection, diagnosis, and initiation of treatment. *Med Care* 2005; 43:141–148
7. Rauscher GH, Allgood KL, Whitman S, Conant E. Disparities in screening mammography services by race/ethnicity and health insurance. *J Womens Health (Larchmt)* 2012; 21:154–160
8. Ansell D, Grabler P, Whitman S, et al. A community effort to reduce the black/white breast cancer mortality disparity in Chicago. *Cancer Causes Control* 2009; 20:1681–1688
9. Gray JE. Mammography (and radiology?) is still plagued with poor quality in photographic processing and darkroom fog. *Radiology* 1994; 191: 318–319
10. Mammography Quality Standards Reauthorization Act of 1998. Public Law 105-248. 42 USC §263b. http://www.gpo.gov/fdsys/pkg/PLAW-105publ248/html/PLAW-105publ248.htm. Accessed August 28, 2013
11. Monsees BS. The Mammography Quality Standards Act: an overview of the regulations and guidance. *Radiol Clin North Am* 2000; 38:759–772
12. Nass S, Ball J, eds.; Committee on Improving Mammography Quality Standards, National Cancer Policy Board. *Improving breast imaging quality standards*. Washington, DC: National Academies Press, 2005
13. Breast Cancer Surveillance Consortium website. Breastscreening.cancer.gov. Accessed February 4, 2013
14. Rosenberg RD, Yankaskas BC, Abraham LA, et al. Performance benchmarks for screening mammography. *Radiology* 2006; 241:55–66
15. Schell MJ, Yankaskas BC, Ballard-Barbash R, et al. Evidence-based target recall rates for screening mammography. *Radiology* 2007; 243:681–689
16. Yankaskas BC, Cleveland RJ, Schell MJ, Kozar R. Association of recall rates with sensitivity and positive predictive values of screening mammography. *AJR* 2001; 177:543–549
17. Goel A, George J, Burack RC. Telephone reminders increase re-screening in a county breast screening program. *J Health Care Poor Underserved* 2008; 19:512–521
18. Rauscher GH, Conant EF, Khan J, Berbaum ML. Mammogram image quality as a potential contributor to disparities in breast cancer stage at diagnosis: an observational study. *BMC Cancer* 2013; 13:208
19. Rauscher GH, Khan JA, Berbaum ML, Conant EF. Potentially missed detection with screening mammography: does the quality of radiologist's interpretation vary by patient socioeconomic advantage/disadvantage? *Ann Epidemiol* 2013; 23: 210–214