


Discordance between original and central laboratories in ER and HER2 results in a diverse, population-based sample

Jennifer J. Griggs¹  · Ann S. Hamilton² · Kendra L. Schwartz³ · Weiqiang Zhao⁴ · Paul H. Abrahamse¹ · Dafydd G. Thomas¹ · Julie M. Jorns¹ · Rachel Jewell⁴ · Maria E. Sibug Saber² · Reina Haque⁵ · Steven J. Katz¹

Received: 13 October 2016 / Accepted: 19 November 2016 / Published online: 29 November 2016
© Springer Science+Business Media New York 2016

Abstract

Purpose To investigate the discordance between original and central laboratories in estrogen receptor (ER) status, in tumors originally deemed to be ER-negative, and in HER2 status in a diverse population-based sample.

Methods In a follow-up study of 1785 women with Stage I–III breast cancer diagnosed between 2005 and 2007 in the Detroit and Los Angeles County SEER registry catchment areas, participants were asked to consent to reassessment of ER (in tumors originally deemed to be ER-negative) and HER2 status on archival tumor samples approximately four years after diagnosis. Blocks were centrally prepared and analyzed for ER and HER2 using standardized methods and the guidelines of the American Society of Clinical Oncology and the College of American Pathologists. Analyses determined the discordance between original and central laboratories.

Results 132 (31%) of those eligible for ER reassessment and 367 (21%) eligible for HER2 reassessment had archival blocks reassessed centrally. ER discordance was only 6%. HER2 discordance by immunohistochemistry (IHC) was 26%, but final HER2 results—employing FISH in tumors that were IHC 2+ at the central laboratory—were discordant in only 6%. Half of the original laboratories did not perform their own assays.

Conclusions Discordance between original and central laboratories in two large metropolitan areas was low in this population-based sample compared to previously reported patient samples. Centralization of testing for key pathology variables appears to be occurring in many hospitals. In addition, quality improvement efforts may have preceded the publication and dissemination of specialty society guidelines.

Keywords Breast cancer · Estrogen receptor · HER2 · SEER registry · Pathology

Electronic supplementary material The online version of this article (doi:10.1007/s10549-016-4061-z) contains supplementary material, which is available to authorized users.

✉ Jennifer J. Griggs
jengrigg@umich.edu

- ¹ University of Michigan, 2800 Plymouth Rd., Building 16, 116W, Ann Arbor, MI 48109, USA
- ² Keck School of Medicine, University of Southern California, 2001 N. Soto St 318E, Los Angeles, CA 90089, USA
- ³ Wayne State University School of Medicine, 320 E. Canfield, Detroit, MI 48201, USA
- ⁴ The Ohio State University, 2001 Polaris Parkway, Columbus, OH 43240, USA
- ⁵ Kaiser Permanente Southern California, Research & Evaluation, 100 S Los Robles, Pasadena, CA 91101, USA

Introduction

Decisions regarding systemic treatment of invasive breast cancer in the adjuvant setting incorporate stage and tumor characteristics, particularly estrogen receptor (ER) and progesterone receptor (PR) status, and the presence or absence of human epidermal growth factor 2 (HER2) [1]. Accurate assessment of these key biologic characteristics is thus critically important in the delivery of optimal care.

With respect to hormone receptor status, false-negative results are particularly detrimental as they lead to omission of beneficial endocrine therapy [2–5] and may prompt a recommendation for adjuvant chemotherapy in a patient in whom the risk–benefit ratio is unfavorable. ER status is currently measured using immunohistochemistry (IHC) [6].

Concerns have been raised about variation in hormone receptor testing both in the United States and internationally [7, 8]. Studies of patients enrolled in clinical trials indicate that misclassification of ER status may be a widespread problem, with the percentage of false-negative tumors being as high as 77% in one study [3]. In addition, misclassification of tumors has been identified in 20–63% of laboratories [9]. In contrast, false-positive ER results are much less common (<3%) [3, 9]. The prevalence of false-negative ER results in population-based samples across diverse laboratories and practice settings has been understudied.

Accurate assessment of HER2 status is similarly essential in treatment decision-making for patients with breast cancer. False-negative HER2 status may lead to omission of anti-HER2-directed therapy, and, conversely, false-positive HER2 results may lead to needless administration of costly and prolonged treatment with no benefit. Methods for assessing HER2 status include IHC for quantification of protein expression and fluorescent in situ hybridization (FISH) for measurement of gene amplification. Current joint guidelines of the American Society of Clinical Oncology (ASCO) and the College of American Pathologists (CAP) specify that tumors with equivocal IHC (2+) results should be tested for HER2 gene amplification with FISH [10–12].

Compelling studies have indicated that HER2 testing has been poorly standardized. Variations in HER2 results may occur at multiple steps in the assessment process [11–13]. Discordance between local and central laboratory testing in patients in clinical trials has been reported to be as high as 20% for IHC and 12% for FISH [11, 12, 14–17]. As with assessment of ER, there is limited information about the quality of HER2 assessment in patients who are not enrolled in clinical trials.

In this context, we chose to investigate the discordance in ER and HER2 results between the originating laboratories and central laboratories in a diverse population-based sample of women who were not participating in clinical trials. Because the rate of false-positive ER results is extremely low [3, 9], we reassessed the ER status of only those tumors originally identified as ER-negative; all tumors were reassessed for HER2 status.

Methods

Subject identification and recruitment for assessment of ER status and HER2 status

We conducted a population-based study of women between the ages of 21 and 79 who were diagnosed with a first breast cancer (Stages 0–III) between June 2005 and February 2007 and who were reported to the registries of

the Detroit and Los Angeles County Surveillance, Epidemiology, and End Results (SEER) program. Details of the sampling and survey methods have been previously described [18, 19]. Only women with Stages I–III are included in the analytic sample for this study.

A follow-up study was conducted with these participants an average of four years after diagnosis, at which time study participants were asked to provide HIPAA authorization and informed consent for abstraction of their breast cancer medical records and provision of a tumor sample for reassessment ER and HER2 status.

The institutional review boards at the University of Michigan, the University of Southern California, Wayne State University, and participating hospitals approved the study. Laboratory and treating hospitals and providers were de-identified by the SEER Registries according to requirements of the SEER program.

Data collection

ER status and HER2 data were obtained from the SEER registries. Medical record information collected by trained abstractors was used if HER2 data were missing from the SEER registry. The method(s) for HER2 assessment (IHC, FISH, and both) and results were abstracted. The dataset was merged with SEER and survey data from the time of the original (baseline) survey, including tumor characteristics and educational attainment, marital status, employment status, insurance status and type, and income.

Archival tumor specimen retrieval

Formalin-fixed paraffin-embedded (FFPE) tissue samples were obtained from consenting patients for whom tissue was available. In Los Angeles, a SEER Registry pathologist selected a representative tissue block for reassessment; whenever possible, this was the same tumor block that was used for the original diagnosis. In Detroit, pathologists at each of the laboratories selected representative whole tissue sections that contained cancer representing areas that resembled the original diagnostic tumor sections. If applicable, blocks were prepared at the Pathology Core Facility within the Department of Pathology at the Ohio State University Wexner Medical Center. They were thereafter returned to the SEER sites and then to the laboratories by SEER personnel.

Detailed descriptions of the sample processing for IHC (for ER and HER2 status) and for FISH (for HER2) are provided in the Appendix in Supplementary material.

Central reassessment of ER status

ER expression was recorded as the percentage of staining cells and was classified as a dichotomous variable (present

or absent) using two threshold values: (1) $\geq 1\%$ staining cells or (2) $\geq 10\%$ staining cells. Tumor cells were scored for the nuclear expression of ER using the Allred scoring schema, where score intensity (0–3) and the proportion of immunoreactive cells (0, none; 1, $<1\%$; 2, 1–10%; 3, 10–33%; 4, 33–66% and 5, $>66\%$) were summed. Tumors were considered ER-positive if the Allred score was 2 or greater.

Central reassessment of HER2 status by IHC

Two pathologists at the Ohio State University Wexner Medical Center performed the HER2 assessments and assigned an IHC score of 0, 1+, 2+, or 3+ in concordance with the ASCO-CAP Guidelines of 2007 [11].

Central reassessment of HER2 status by FISH

Using image analysis as described in the Appendix in Supplementary material, results were defined into three groups: “amplified” (positive by FISH), “not amplified,” (negative by FISH) or “equivocal” for each specimen.

Measures

The primary outcome measure was discordance between ER status and HER2 status of the primary tumor at the original laboratory and the central laboratories (University of Michigan for ER and The Ohio State University for HER2). As described above, only tumors originally reported to the SEER registries as being ER-negative were reassessed for ER status. Two pathologists at each central site reviewed the slides, and the pathologists at the two sites were available to resolve any discrepancies between pathologists.

Statistical analyses

The demographic and disease characteristics of the patients in whom we successfully obtained tissue for repeat ER and HER2 assessment were compared with those in whom we did not. The percentage of tumors retrieved according to laboratory among consenting patients were also calculated. The dependent variable was a binary indicator of test discordance between the originating laboratory and the central laboratory. ER status and HER2 status were considered discordant if the results of the test determined by the originating laboratory differed from the result of the testing at the central laboratory. We calculated both discordance for HER2 by IHC and overall HER2 status using the combination of IHC and FISH. The percentage of discordant results with standard errors (SE) or 95% confidence

intervals (CI), as appropriate, were calculated for ER and HER2 (IHC, FISH, and overall HER2 results).

Results

Sample characteristics

Figure 1 shows the original sample eligibility and the number of patients who provided consent and for whom we obtained tissue for repeat ER and HER2 testing. The original sample selected from the SEER registries included 3252 patients. After exclusions and non-respondents, the analytic sample numbered 1785 patients.

Patient sample for central ER assessment

Of the 1785 patients, 428 had tumors that were ER-negative and comprised the analytic sample for ER reassessment. Of these 428, we obtained consent and tumor specimens for 132 patients (31% of the analytic sample). Consent and retrieval were significantly higher (40%) among non-Hispanic whites compared with Hispanic (26%) and black (34%) women (p value = 0.038) (data not shown). There were no differences in consent/tissue retrieval according to any of the other demographic or clinical variables. Characteristics of the patient sample are shown in Table 1. The 132 patients were treated in 48 hospitals; the tumor samples originated from 20 laboratories, and retrieval varied significantly according to laboratory ($p < 0.001$).

Patient sample for central HER2 assessment

Of the 1785 patients who considered as the analytic sample, 964 (54.0%) provided consent for medical record review and tumor reassessment of HER2 status. HER2 status was available for 761 (78.9%) of these 964 patients (42% of the entire eligible sample of patients). Tumor specimens were available for central review from 367 (48.2%) (20.5% of the entire eligible sample). Consent, medical record, and tumor retrieval were significantly higher in women with higher levels of educational attainment ($p = 0.006$), in non-Hispanic white women ($p < 0.01$), and in women with higher levels of income ($p < 0.001$) (data not shown). There were no differences in consent, medical record, or tumor retrieval according to other demographic or clinical factors. The characteristics of the final patient sample are shown in Table 2. The 367 patients received care in 83 hospitals; their tumor samples originated from 44 laboratories. Retrieval varied significantly according to laboratory ($p < 0.001$).

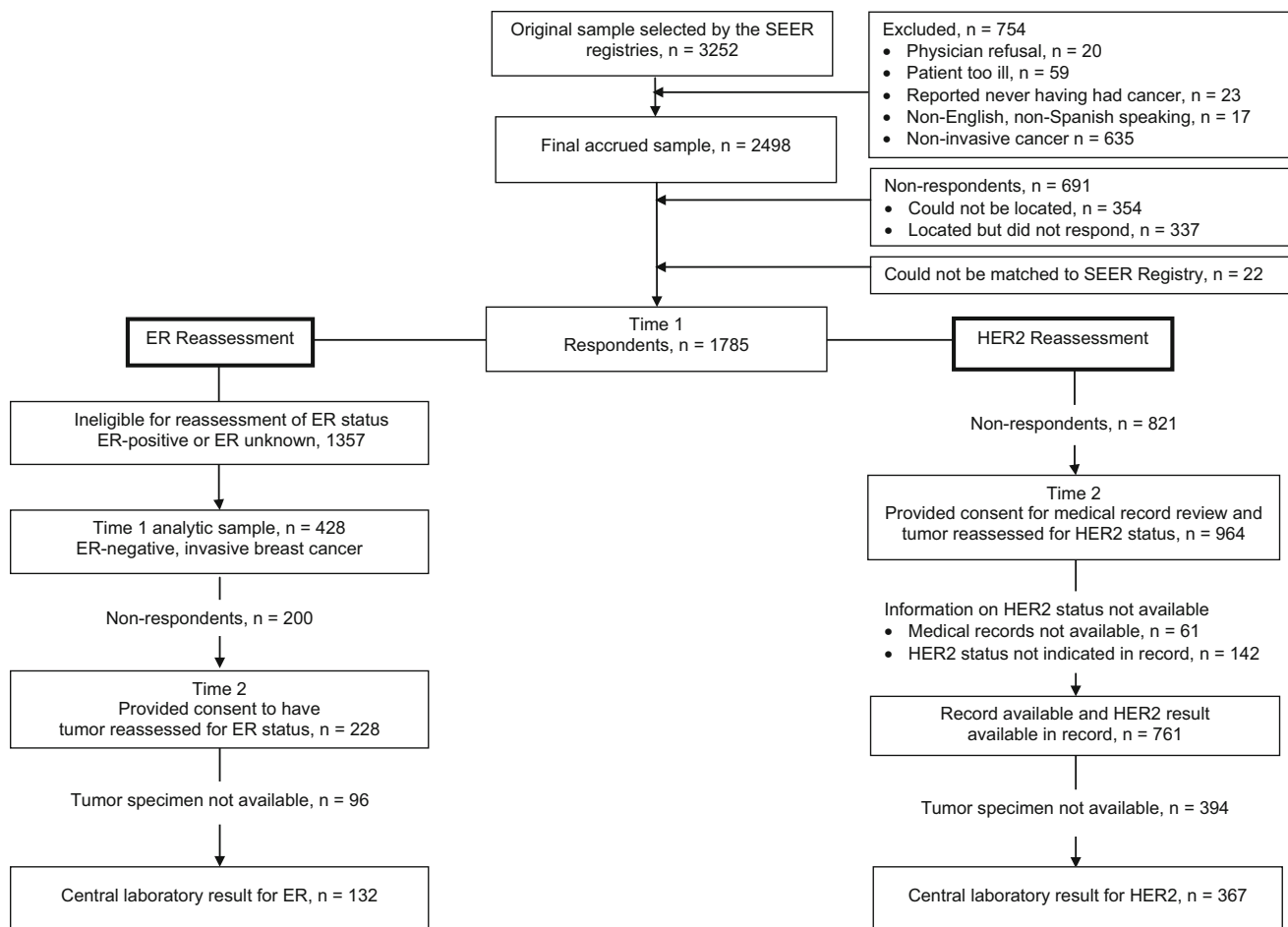


Figure. Derivation of analytic samples with reasons for exclusion.

Fig. 1 Derivation of analytic samples with reasons for exclusion

Discordance between Original and Central Laboratories

ER results

Of the 132 samples deemed to be ER-negative at the original laboratories that were adequate for repeat IHC assessment, 8 (6.0%, SE \pm 2.1%) were ER-positive by the central laboratory. Of these, one had an Allred score of 3, three an Allred score of 4, two an Allred score of 5, one an Allred score of 7, and one an Allred score of 8.

HER2 results

Of the 367 tumors obtained for repeat HER2 assessment, immunohistochemistry was performed on all but 48 at the original laboratories. In the remaining 48, only FISH was performed. IHC was performed using automated

microscopy and image analysis with the Automated Cellular Imaging System (ACIS, DAKO) [20, 21] in 32 cases. Of the 24 samples in which the IHC results were 2+, FISH was performed according to guidelines in all but three (12.5%) of samples at the original laboratories.

Discordance between original and central laboratory HER2 results

When the central laboratory results of IHC and FISH were combined using the recommended HER2 assessment algorithm across the entire sample of 367, only 22 of the (6.0%, SE \pm 2.4%) of the tumors had discordant results between the original and the central laboratories as shown in Table 3. Of these, 19 (86.4%, SE \pm 14.3) were determined to be HER2-positive (originally reported as HER2-negative), and three of the 22 (13.6%, SE \pm 14.3) were determined to be HER2-negative (originally reported as

Table 1 Sample characteristics for ER reassessment at central laboratory, *N* = 132

	<i>N</i>	%
Age group (years) missing, <i>n</i> = 1		
<40	14	11
40–49	28	21
50–59	43	33
60–69	32	24
70+	14	11
Comorbid conditions		
None	61	46
One	39	30
Two or more	32	24
Stage		
Stage I	53	40
Stage II	60	45
Stage III	19	14
Histologic grade, missing, <i>n</i> = 4		
Grade 1	1	1
Grade 2	26	20
Grade 3	101	79
Income, missing, <i>n</i> = 17		
Less than \$20,000	20	15
\$20,000 to \$69,000	63	48
\$70,000 or more	32	24
Insurance status		
None	8	6
Other	83	63
Medicaid	19	14
Medicare	22	17
Education, missing, <i>n</i> = 1		
Not HS graduate	21	16
HS graduate	24	18
Some college	54	41
College graduate	32	24
Marital status		
Not married/partnered	53	40
Married/partnered	79	60
Employment status, missing, <i>n</i> = 2		
Employed	87	67
Not employed	43	33
Race/ethnicity		
Black	44	34
Hispanic	34	26
Non-Hispanic White	52	40
Other	2	2

HER2-positive). Seven cases with positive FISH results at the original laboratories were negative (0 or 1+) by IHC at the central laboratory and thus not evaluated for gene amplification by FISH at the central laboratory.

Table 2 Sample characteristics for HER reassessment at central laboratory, *N* = 367

	<i>N</i>	%
Age group (years) missing, <i>n</i> = 1		
<40	39	11
40–49	126	21
50–59	180	33
60–69	155	24
70+	107	11
Comorbid conditions		
None	250	46
One	180	30
Two or more	178	24
Stage		
Stage I	292	40
Stage II	235	45
Stage III	76	14
Hormone receptor status, missing, <i>n</i> = 28		
ER and/or PR+	452	78
ER–/PR–	128	22
Histologic grade, missing, <i>n</i> = 12		
Grade 1	119	1
Grade 2	229	20
Grade 3	218	79
Income		
Less than \$20,000	100	15
\$20,000 to \$69,000	237	48
\$70,000 or more	187	24
Missing	84	14
Insurance status, missing = 3		
None	39	6
Other	357	63
Medicaid	58	14
Medicare	151	17
Education, missing, <i>n</i> = 8		
Not HS graduate	88	16
HS graduate	113	18
Some college	219	41
College graduate	180	24
Marital status		
Not married/partnered	258	40
Married/partnered	350	60
Employment status, missing, <i>n</i> = 15		
Employed	351	67
Not employed	242	33
Race/ethnicity		
Black	119	34
Hispanic	133	26
Non-Hispanic White	343	40
Other	13	2

Table 3 Summary of discordant HER2 results at original and central laboratories

Original		Central		Interpretation	
IHC	FISH	IHC	FISH	Original	Central
0		2+	2.37	–	+
0–1+		2+	2.01	–	Equivocal
2.5 ^a		0		+	–
2+	2.9	1+		+	–
2+	2.93	0		+	–
2–3+	3	0		+	–
3+		1+		+	–
3+		0		+	–
3+		2+	1.54	+	–
3+		0		+	–
3+	11.42	0		+	–
3+		0		+	–
3+		1+		+	–
3+		1+		+	–
3+		2+	1.8	+	–
3+	9.08	2+	0.9	+	–
4.0 ^a		0		+	–
NA	2.59	2+	0.83	+	–
NA	6.4	0		+	–
NA	1.2	2+	2.53	–	+
NA	6.89	1+		+	–
NA	2.6	1+		+	–

IHC immunohistochemistry, FISH fluorescent in situ hybridization

^a Performed using Automated Cellular Imaging System, ACIS (image analysis of IHC)

When only IHC test results were considered, discordance was much higher (Table 4). Among the 319 patients in whom IHC was performed, the overall discordance was 26.0% (95% CI 21.2–41.8%). In tumors that were negative (0 or 1+) at the original laboratory, the discordance was 16%. In the 38 tumors that were 2+, most (71.1%) were negative (0 or 1+) at the central laboratory, leading to 29.9% discordance. (As described above, according to the ASCO-CAP guidelines, FISH was not performed in these tumors at the central laboratory.) The discordance among tumors that were 3+ at the original laboratory was 41% (six were 2+, and 12 were 0 or 1+).

For 237 cases in which IHC was 0 or 1+ at the original laboratory, 37 (15.6%) were 2+ by IHC at the central laboratory. These were evaluated using FISH according to guidelines; one of these cases was positive for overamplification (FISH ratio 2.37), and one was equivocal (FISH ratio 2.01). All of the other 2+ cases were FISH-negative. Altogether, reassessment by the central laboratory applying the ASCO-CAP guidelines identified one HER2-positive

Table 4 Agreement between original laboratory and central laboratory IHC results in the 319 tumors with IHC at the original laboratories

Original laboratory	Central laboratory			
	0–1+	2+	3+	
0–1+	199	37	1	237
2+	23	11	4	38
3+	12	6	26	44
	234	54	31	Total

Bolded text indicates concordance; 26% (4.8) of results were discordant

case and one HER2-equivocal case out of 237 that were negative at the original laboratory.

Discussion

In summary, in women with invasive breast cancer drawn from population-based samples in two SEER catchment areas between 2005 and 2007, the proportion of ER-negative tumors with discordant results at the central laboratory was only 6%. For HER2, in contrast, reassessment at a central laboratory identified IHC discordance in 26% of samples. When IHC followed by FISH testing was performed using ASCO-CAP guidelines, however, HER2 discordance dropped to only 6%.

ER discordance

In those tumors that were deemed to be ER-positive at the central laboratory, none of the tumors had a score of ≤ 3 using the Allred scoring system. This suggests that the discordance between the central and original laboratories was not due merely to differences in the cut-off used to classify tumors as ER-positive. In addition, we cannot attribute discordance between the original and the central laboratory results to pre-analytic factors, such as sample ischemic time, because all of our samples were archival specimens. The literature does suggest, however, that most errors in ER assessment occur in the analytic phase [9, 22, 23]. One explanation for the difference in ER status between the original and the central laboratories is heterogeneity in ER expression within tumors [24].

The discordance between the original and central laboratories in our study was substantially lower than in a previously reported study of patients enrolled on clinical trials [3] but is similar to that in a cohort study conducted by Ma and colleagues of women diagnosed between 1994 and 1998 enrolled in the multicenter Women's

Contraceptive Reproductive Experiences (CARE) Study [25]. In the CARE study, a convenience sample of patients reported to the Los Angeles and Detroit Surveillance, Epidemiology, and End Results registries between 1994 and 1998 had their tumors reassessed for ER and PR at a central laboratory. Among 316 tumors reported to the SEER registries as being ER-negative, 28 (8%) were deemed to be ER-positive upon reassessment at a single laboratory [25].

We identified that some degree of centralization of ER assessment has taken place given that only 20 laboratories performed ER assays in patients receiving care at 48 hospitals; this may explain the low discordance in our sample.

Another explanation for the low discordance in our study compared with many of the previous reports may be due to improvements in the quality of laboratory procedures during the timeframe in which our patient sample, and that of Ma and colleagues [25], was diagnosed and treated, improvements that preceded published recommendations. The high percentage of discordance in patients enrolled in a large cooperative group trial cited above [3], which motivated our population-based study, enrolled patients between 1998 and 2003, while the quality improvement program demonstrating interlaboratory variation in ER assessment began in 1994 with published results in 2000 [9]. Systematic problems in ER assessment were well publicized before the guidelines were published [26].

HER2 discordance

For HER2 reassessment, the discordance between the original laboratories and the central laboratory was 26% for patients whose tumors were tested with IHC and was highest among tumors that were 2+ (71%) and 3+ (41%). Further, when the HER2 guidelines were applied, which recommend FISH testing for gene amplification in tumors that are 2+ by IHC, the overall HER2 discordance decreased to only 6%. The majority of cases (all but 2 of the 22) that were discordant were positive at the original laboratory but deemed to be negative by the central laboratory. Most studies investigating HER2 discordance have done so in patients participating in clinical trials of trastuzumab [14–16, 27]. Such trials require that tumor expression of HER2 by IHC be 3+. Our findings demonstrated higher discordance among patients whose tumors were originally deemed to be 3+ by IHC (41%) than in these study participants, where the discordance between original and central laboratories for IHC 3+ is 20%–25% [14–16, 27]. However, it is possible that the higher false-positive rate observed in our study could be confounded by pre-analytic factors, such as tissue handling, intratumoral heterogeneity, or protein degradation that

could have occurred given the time that elapsed between diagnosis and specimen retrieval in this study.

There is very little published work on discordance in tumors that are originally deemed to be HER2-negative [27–29]. In the trial by Reddy, discordance was 66% for tumors that were HER2 IHC 0 at the original laboratory and 46% for tumors that were IHC 1+ [27]. These proportions differ from our findings for the IHC-negative tumors, where the proportion was only 16% for 0 to 1+ tumors. Because the investigators did not present their findings using the ASCO-CAP algorithm for final HER2 results, it is not possible to compare an analogous discordance proportion with our own results. Additional information on tumors originally deemed to be HER2-negative comes from an observational clinical study designed to investigate outcomes in patients with HER2-negative metastatic disease [28]. Of the 552 samples retrieved and centrally reassessed for HER2 status from the 1267 study participants (44% retrieval rate), 22 (4%) were found to be positive for HER2 using a combination of IHC and FISH. Discordance was seen with both IHC and FISH.

Finally, published information on discordance between original and central laboratories is available from a community-based sample of patients with metastatic breast cancer treated in clinical practice in two provinces in Canada [29]. Participants diagnosed between 1999 and 2002 with metastatic HER2-positive breast cancer had tumors reassessed at one central laboratory using both IHC and FISH. Among tumors deemed positive by IHC at the central laboratory, concordance ranged between 79 and 90% (depending on the HER2 IHC method used at the central laboratory). Among tumors deemed negative by IHC at the central laboratory, concordance was high, ranging from 95 to 100%. For FISH, the concordance was 98.5%. The discordance in this study was thus lower than in ours.

We applied the ASCO-CAP guidelines in the design of this study and thus do not have FISH results in the tumor samples that were IHC 0 or 1+ or tested only with FISH at the original laboratories. As described in the Results, seven cases with positive FISH results at the original laboratories were negative (0 or 1+) by IHC at the central laboratory and were thus not evaluated for gene amplification by FISH at the central laboratory. The application of the ASCO-CAP guidelines would have led to the omission of trastuzumab in these patients if the central laboratory results were viewed as the “gold standard” and if FISH had not been performed.

As with ER testing, another explanation for our finding of low discordance is that efforts to improve the quality of IHC in the assessment of HER2 status appear to be changing practice [30], and it is possible that efforts to standardize HER2 testing were already underway at the time our patient cohort was diagnosed in 2005–2007.

Although the population impact of wholesale adoption of both IHC and FISH in all samples is likely to be low if our discordance is a true rate in population-based samples, arguments have been made that both IHC and FISH or that primary FISH should be done on all tumors in an effort to avoid misclassification of tumors and omission of trastuzumab in patients who may benefit from this drug [31, 32]. In addition, the provocative finding that some patients with low HER2 expression do indeed benefit from trastuzumab [33] and the results of NSABP B-47 could alter the interpretation of our findings if there is indeed shown to be benefit of trastuzumab in patients with low levels of HER2.

The primary limitation of this study is the lower-than-expected patient participation/tumor retrieval and the small sample size. Participants in this study may have had their tumors assessed in laboratories that differ systematically from the other laboratories, and this represents an important consideration when interpreting the findings of our study. Laboratories were de-identified so that we were unable to examine further on the characteristics of the laboratory associated with tissue retrieval rate or assess the potential bias of differential retrieval rates by laboratory participation. Our findings may not be generalizable to areas other than Los Angeles and Detroit, two large metropolitan areas.

These limitations notwithstanding, the findings of this study may indicate that progress is being made in the quality of tumor biology evaluative testing for women with breast cancer [34]. Such progress may be the result of standardization in HER2 testing methods, assay validation, and interpretation of IHC results through the use of, for example, automated image analysis [35]. Improving the precision of evaluative testing for cancer has the high potential for improving the quality of care through more accurate promulgation of treatment guidelines to individual patients. Ongoing efforts to increase participation in laboratory certification programs should be supported to continue to improve assessment of key pathology variables that drive treatment decision-making in breast cancer [12, 30].

Acknowledgements This work was funded by Grant R01 CA139014 from the National Cancer Institute to the University of Michigan (Principal Investigators Dr. Griggs and Dr. Katz). Dr. Katz was supported by an Established Investigator Award in Cancer Prevention, Control, Behavioral, and Populations Sciences Research from the National Cancer Institute (K05 CA111340). Dr. Thomas was supported by The Fashion Footwear Charitable Foundation of New York/QVC Presents Shoes on Sale. The collection of Los Angeles County cancer incidence data was supported by the California Department of Public Health as part of the statewide cancer reporting program mandated by California Health and Safety Code §103885. The National Cancer Institute's Surveillance, Epidemiology, and End Results Program under contract N01-PC-35139 was awarded to the University of Southern California Contract N01-PC-54404, and agreement 1U58DP0087-01 was awarded to the Public Health

Institute. The collection of metropolitan Detroit cancer incidence data was supported by the National Cancer Institute's Surveillance, Epidemiology, and End Results Program under contract N01-PC-35145. The authors wish to acknowledge the contributions of Polaris Core Facility technical staff at the Ohio State University.

Compliance with Ethical Standards

Conflict of interest The authors declare that they have no conflict of interest.

References

1. National Comprehensive Cancer Network (2016) <http://www.nccn.org/clinical.asp>, Accessed 13 Oct 2016
2. Barnes DM, Harris WH, Smith P, Millis RR, Rubens RD (1996) Immunohistochemical determination of oestrogen receptor: comparison of different methods of assessment of staining and correlation with clinical outcome of breast cancer patients. *Br J Cancer* 74:1445–1451
3. Viale G, Regan MM, Maiorano E, Mastropasqua MG, Dell'Orto P, Rasmussen BB, Raffoul J, Neven P, Orosz Z, Braye S, Ohlschlegel C, Thurlimann B, Gelber RD, Castiglione-Gertsch M, Price KN, Goldhirsch A, Gusterson BA, Coates AS (2007) Prognostic and predictive value of centrally reviewed expression of estrogen and progesterone receptors in a randomized trial comparing letrozole and tamoxifen adjuvant therapy for postmenopausal early breast cancer: BIG 1-98. *J Clin Oncol* 25(25):3846–3852. doi:10.1200/JCO.2007.11.9453
4. Jalava P, Kuopio T, Huovinen R, Laine J, Collan Y (2005) Immunohistochemical staining of estrogen and progesterone receptors: aspects for evaluating positivity and defining the cut-points. *Anticancer Res* 25:2535–2542
5. Rasmussen BB, Regan MM, Lykkesfeldt AE, Dell'Orto P, Del Curto B, Henriksen KL, Mastropasqua MG, Price KN, Mery E, Lacroix-Triki M, Braye S, Altermatt HJ, Gelber RD, Castiglione-Gertsch M, Goldhirsch A, Gusterson BA, Thurlimann B, Coates AS, Viale G (2008) Adjuvant letrozole versus tamoxifen according to centrally-assessed ERBB2 status for postmenopausal women with endocrine-responsive early breast cancer: supplementary results from the BIG 1-98 randomised trial. *Lancet Oncol* 9:23–28. doi:10.1016/s1470-2045(07)70386-8
6. Harvey JM, Clark GM, Osborne CK, Allred DC (1999) Estrogen receptor status by immunohistochemistry is superior to the ligand-binding assay for predicting response to adjuvant endocrine therapy in breast cancer. *J Clin Oncol* 17:1474–1481
7. Allred DC (2008) Commentary: hormone receptor testing in breast cancer: A distress signal from Canada. *Oncologist* 13:1134–1136. doi:10.1634/theoncologist.2008-0184
8. Allred DC (2008) Problems and solutions in the evaluation of hormone receptors in breast cancer. *J Clin Oncol* 26:2433–2435. doi:10.1200/JCO.2007.15.7800
9. Rhodes A, Jasani B, Barnes DM, Bobrow LG, Miller KD (2000) Reliability of immunohistochemical demonstration of oestrogen receptors in routine practice: interlaboratory variance in the sensitivity of detection and evaluation of scoring systems. *J Clin Pathol* 53:125–130
10. Carlson RW, Moench SJ, Hammond ME, Perez EA, Burstein HJ, Allred DC, Vogel CL, Goldstein LJ, Somlo G, Gradishar WJ, Hudis CA, Jahanzeb M, Stark A, Wolff AC, Press MF, Winer EP, Paik S, Ljung BM (2006) HER2 testing in breast cancer: NCCN Task Force report and recommendations. *JNCCN Suppl* 3:S1–22

11. Wolff AC, Hammond ME, Schwartz JN, Hagerty KL, Allred DC, Cote RJ, Dowsett M, Fitzgibbons PL, Hanna WM, Langer A, McShane LM, Paik S, Pegram MD, Perez EA, Press MF, Rhodes A, Sturgeon C, Taube SE, Tubbs R, Vance GH, van de Vijver M, Wheeler TM, Hayes DF (2007) American Society of Clinical Oncology/College of American Pathologists guideline recommendations for human epidermal growth factor receptor 2 testing in breast cancer. *J Clin Oncol* 25:118–145. doi:[10.1200/JCO.2006.09.2775](https://doi.org/10.1200/JCO.2006.09.2775)
12. Wolff AC, Hammond ME, Hicks DG, Dowsett M, McShane LM, Allison KH, Allred DC, Bartlett JM, Bilous M, Fitzgibbons P, Hanna W, Jenkins RB, Mangu PB, Paik S, Perez EA, Press MF, Spears PA, Vance GH, Viale G, Hayes DF (2013) Recommendations for human epidermal growth factor receptor 2 testing in breast cancer: American Society of Clinical Oncology/College of American Pathologists clinical practice guideline update. *J Clin Oncol* 31(31):3997–4013. doi:[10.1200/jco.2013.50.9984](https://doi.org/10.1200/jco.2013.50.9984)
13. Thomson TA, Hayes MM, Spinelli JJ, Hilland E, Sawrenko C, Phillips D, Dupuis B, Parker RL (2001) HER-2/neu in breast cancer: interobserver variability and performance of immunohistochemistry with 4 antibodies compared with fluorescent in situ hybridization. *Mod Pathol* 14:1079–1086. doi:[10.1038/modpathol.3880440](https://doi.org/10.1038/modpathol.3880440)
14. Paik S, Bryant J, Tan-Chiu E, Romond E, Hiller W, Park K, Brown A, Yothers G, Anderson S, Smith R, Wickerham DL, Wolmark N (2002) Real-world performance of HER2 testing—National Surgical Adjuvant Breast and Bowel Project experience. *J Natl Cancer Inst* 94:852–854
15. Roche PC, Suman VJ, Jenkins RB, Davidson NE, Martino S, Kaufman PA, Addo FK, Murphy B, Ingle JN, Perez EA (2002) Concordance between local and central laboratory HER2 testing in the Breast Intergroup Trial N9831. *J Natl Cancer Inst* 94:855–857
16. Perez EA, Suman VJ, Davidson NE, Martino S, Kaufman PA, Lingle WL, Flynn PJ, Ingle JN, Visscher D, Jenkins RB (2006) HER2 testing by local, central, and reference laboratories in specimens from the North Central Cancer Treatment Group N9831 intergroup adjuvant trial. *J Clin Oncol* 24:3032–3038. doi:[10.1200/JCO.2005.03.4744](https://doi.org/10.1200/JCO.2005.03.4744)
17. Bianchi S, Caini S, Paglierani M, Saieva C, Vezzosi V, Baroni G, Simoni A, Palli D (2015) Accuracy and reproducibility of HER2 status in breast cancer using immunohistochemistry: a quality control study in Tuscany evaluating the impact of updated 2013 ASCO/CAP recommendations. *Pathol Oncol Res* 21:477–485. doi:[10.1007/s12253-014-9852-0](https://doi.org/10.1007/s12253-014-9852-0)
18. Griggs JJ, Hawley ST, Graff JJ, Hamilton AS, Jagsi R, Janz NK, Mujahid MS, Friese CR, Salem B, Abrahamse PH, Katz SJ (2012) Factors associated with receipt of breast cancer adjuvant chemotherapy in a diverse population-based sample. *J Clin Oncol* 30:3058–3064. doi:[10.1200/JCO.2012.41.9564](https://doi.org/10.1200/JCO.2012.41.9564)
19. Friese CR, Pini TM, Li Y, Abrahamse PH, Graff JJ, Hamilton AS, Jagsi R, Janz NK, Hawley ST, Katz SJ, Griggs JJ (2013) Adjuvant endocrine therapy initiation and persistence in a diverse sample of patients with breast cancer. *Breast Cancer Res Treat* 138:931–939. doi:[10.1007/s10549-013-2499-9](https://doi.org/10.1007/s10549-013-2499-9)
20. Minot DM, Kipp BR, Root RM, Meyer RG, Reynolds CA, Nassar A, Henry MR, Clayton AC (2009) Automated Cellular Imaging System III for assessing HER2 status in breast cancer specimens: development of a standardized scoring method that correlates with FISH. *Am J Clin Pathol* 132:133–138. doi:[10.1309/ajcpjv0skaf2pcmy](https://doi.org/10.1309/ajcpjv0skaf2pcmy)
21. Hanley KZ, Siddiqui MT, Lawson D, Cohen C, Nassar A (2009) Evaluation of new monoclonal antibodies in detection of estrogen receptor, progesterone receptor, and HER2 protein expression in breast carcinoma cell block sections using conventional microscopy and quantitative image analysis. *Diagn Cytopathol* 37:251–257. doi:[10.1002/dc.20989](https://doi.org/10.1002/dc.20989)
22. Diaz LK, Sneige N (2005) Estrogen receptor analysis for breast cancer: current issues and keys to increasing testing accuracy. *Adv Anat Pathol* 12:10–19
23. Apple S, Pucci R, Lowe AC, Shintaku I, Shapourifar-Tehrani S, Moatamed N (2011) The effect of delay in fixation, different fixatives, and duration of fixation in estrogen and progesterone receptor results in breast carcinoma. *Am J Clin Pathol* 135:592–598. doi:[10.1309/ajcpb1rit5yxmrjs](https://doi.org/10.1309/ajcpb1rit5yxmrjs)
24. Chung GG, Zerkowski MP, Ghosh S, Camp RL, Rimm DL (2007) Quantitative analysis of estrogen receptor heterogeneity in breast cancer. *Lab Invest* 87:662–669. doi:[10.1038/labinvest.3700543](https://doi.org/10.1038/labinvest.3700543)
25. Ma H, Wang Y, Sullivan-Halley J, Weiss L, Burkman RT, Simon MS, Malone KE, Strom BL, Ursin G, Marchbanks PA, McDonald JA, Spirtas R, Press MF, Bernstein L (2009) Breast cancer receptor status: do results from a centralized pathology laboratory agree with SEER registry reports? *Cancer Epidemiol Biomarkers Prev* 18:2214–2220. doi:[10.1158/1055-9965.epi-09-0301](https://doi.org/10.1158/1055-9965.epi-09-0301)
26. Hammond ME, Hayes DF, Dowsett M, Allred DC, Hagerty KL, Badve S, Fitzgibbons PL, Francis G, Goldstein NS, Hayes M, Hicks DG, Lester S, Love R, Mangu PB, McShane L, Miller K, Osborne CK, Paik S, Perlmutter J, Rhodes A, Sasano H, Schwartz JN, Sweep FC, Taube S, Torlakovic EE, Valenstein P, Viale G, Visscher D, Wheeler T, Williams RB, Wittliff JL, Wolff AC (2010) American Society of Clinical Oncology/College of American Pathologists guideline recommendations for immunohistochemical testing of estrogen and progesterone receptors in breast cancer. *J Clin Oncol* 28:2784–2795. doi:[10.1200/jco.2009.25.6529](https://doi.org/10.1200/jco.2009.25.6529)
27. Reddy JC, Reimann JD, Anderson SM, Klein PM (2006) Concordance between central and local laboratory HER2 testing from a community-based clinical study. *Clin Breast Cancer* 7:153–157. doi:[10.3816/CBC.2006.n.025](https://doi.org/10.3816/CBC.2006.n.025)
28. Kaufman PA, Bloom KJ, Burris H, Gralow JR, Mayer M, Pegram M, Rugo HS, Swain SM, Yardley DA, Chau M, Lalla D, Yoo B, Brammer MG, Vogel CL (2014) Assessing the discordance rate between local and central HER2 testing in women with locally determined HER2-negative breast cancer. *Cancer* 120:2657–2664. doi:[10.1002/cncr.28710](https://doi.org/10.1002/cncr.28710)
29. O'Malley FP, Thomson T, Julian J, Have C, Crosby R, Gelmon K, Andrulis I, Whelan T (2008) HER2 testing in a population-based study of patients with metastatic breast cancer treated with trastuzumab. *Arch Pathol Lab Med* 132:61–65. doi:[10.1043/1543-2165\(2008\)](https://doi.org/10.1043/1543-2165(2008)132:61-65)
30. Middleton LP, Price KM, Puig P, Heydon LJ, Tarco E, Sneige N, Barr K, Deavers MT (2009) Implementation of American Society of Clinical Oncology/College of American Pathologists HER2 Guideline Recommendations in a tertiary care facility increases HER2 immunohistochemistry and fluorescence in situ hybridization concordance and decreases the number of inconclusive cases. *Arch Pathol Lab Med* 133:775–780. doi:[10.1043/1543-2165-133.5.775](https://doi.org/10.1043/1543-2165-133.5.775)
31. Tubbs RR, Hicks DG, Cook J, Downs-Kelly E, Pettay J, Hartke MB, Hood L, Neelon R, Myles J, Budd GT, Moore HC, Andresen S, Crowe JP (2007) Fluorescence in situ hybridization (FISH) as primary methodology for the assessment of HER2 status in adenocarcinoma of the breast: a single institution experience. *Diagn Mol Pathol* 16:207–210. doi:[10.1097/PDM.0b013e318064c72a](https://doi.org/10.1097/PDM.0b013e318064c72a)
32. Garrison LP Jr, Lalla D, Brammer M, Babigumira JB, Wang B, Perez EA (2013) Assessing the potential cost-effectiveness of retesting IHC0, IHC1+, or FISH-negative early stage breast cancer patients for HER2 status. *Cancer* 119(17):3113–3122. doi:[10.1002/cncr.28196](https://doi.org/10.1002/cncr.28196)

33. Pogue-Geile KL, Kim C, Jeong JH, Tanaka N, Bandos H, Gavin PG, Fumagalli D, Goldstein LC, Sneige N, Burandt E, Taniyama Y, Bohn OL, Lee A, Kim SI, Reilly ML, Remillard MY, Blackmon NL, Kim SR, Horne ZD, Rastogi P, Fehrenbacher L, Romond EH, Swain SM, Mamounas EP, Wickerham DL, Geyer CE Jr, Costantino JP, Wolmark N, Paik S (2013) Predicting degree of benefit from adjuvant trastuzumab in NSABP trial B-31. *J Natl Cancer Inst* 105:1782–1788. doi:[10.1093/jnci/djt321](https://doi.org/10.1093/jnci/djt321)
34. Fitzgibbons PL, Hicks DG (2014) Progress in implementing HER2 testing guidelines. *Arch Pathol Lab Med* 138:863–864. doi:[10.5858/arpa.2014-0064-ED](https://doi.org/10.5858/arpa.2014-0064-ED)
35. Holten-Rossing H, Moller Talman ML, Kristensson M, Vainer B (2015) Optimizing HER2 assessment in breast cancer: application of automated image analysis. *Breast Cancer Res Treat* 152:367–375. doi:[10.1007/s10549-015-3475-3](https://doi.org/10.1007/s10549-015-3475-3)