Supplementary Online Content

Elmore JG, Longton G, Carney PA, Geller B, Onega T, Tosteson ANA, Nelson HD, Pepe MS, Allison KH, Schnitt SJ, O'Malley FP, Weaver DL. Diagnostic concordance among pathologists interpreting breast biopsy specimens. *JAMA*. doi:10.1001/jama.2015.1405.

eFigure 1. BPATH-Dx Histology Form[1] for Data Collection on Each Case Used by Participants

Primary Key: _____
Pathologist Name: _____

Clinical History
Patient Age _____
Specimen Type: ☐ Core needle biopsy ☐ Excisional biopsy

I. Histologic Assessment: Diagnoses – Check all that apply. Choose the best fit among the options.

Non-Proliferative changes
- ☐ Non-proliferative changes only

Proliferative lesion without atypia:
- ☐ Fibroadenoma
- ☐ Intraductal papilloma without atypia
- ☐ Usual ductal hyperplasia
- ☐ Columnar cell hyperplasia /Columnar cell change
- ☐ Sclerosing adenosis
- ☐ Radial scar/complex sclerosing lesion

Atypical lesion:
- ☐ Flat epithelial atypia
- ☐ Atypical ductal hyperplasia
- ☐ Intraductal papilloma with atypia
- ☐ Atypical lobular hyperplasia

Carcinoma in situ:
- ☐ Ductal carcinoma in situ:

  Nuclear grade:
  a. ☐ Low
  b. ☐ Intermediate
  c. ☐ High

  Necrosis**:**
  a. ☐ Absent
  b. ☐ Present, focal (small foci/single cell necrosis)
  c. ☐ Present, central (expansive "comedo" necrosis)
- ☐ Lobular carcinoma in situ

(For mixed ductal & lobular features, check both DCIS & LCIS boxes and nuclear grade + necrosis)

Invasive carcinoma :
- ☐ Invasive carcinoma (ductal, lobular or other special type):
  - a. Tubule formation score: 1☐ 2☐ 3☐
  - b. Nuclear grade score 1☐ 2☐ 3☐
  - c. Mitotic activity score: 1☐ 2☐ 3☐
  - Overall Nottingham grade: ☐ Low(total score 3-5) ☐ Intermediate(6, 7) ☐ High(8, 9)

Additional comments: _____
_____
_____

II. If you considered this case borderline between two diagnoses, which diagnoses were you considering? Please check only two options: (Otherwise skip to Section III.)

Non-Proliferative changes
- ☐ Non-proliferative changes only

Proliferative lesion without atypia:
- ☐ Fibroadenoma
- ☐ Intraductal papilloma without atypia
- ☐ Usual ductal hyperplasia
- ☐ Columnar cell hyperplasia /Columnar cell change
- ☐ Sclerosing adenosis
- ☐ Radial scar/complex sclerosing lesion

Atypical lesion:
- ☐ Flat epithelial atypia
- ☐ Atypical ductal hyperplasia
- ☐ Intraductal papilloma with atypia

☐ Atypical lobular hyperplasia

Carcinoma in situ:

☐ Ductal carcinoma in situ:

☐ Lobular carcinoma in situ

Invasive carcinoma:

☐ Invasive carcinoma

What particular features made you favor the final diagnostic category you chose for the lesion?

_____

III. Additional questions regarding this case:

Please rate on the following scale your opinion of the **level of diagnostic difficulty** of this case:

1☐ 2☐ 3☐ 4☐ 5☐ 6☐

Very easy Very challenging

Please rate on the following scale your **confidence in your assessment**:

1☐ 2☐ 3☐ 4☐ 5☐ 6☐

Very confident Not at all confident

Would you ask for a **second pathologist's opinion** of this case before finalizing the report? (Assume a pathologist is available)

1. ☐ No
2. ☐ Yes, because it is our policy to get a second opinion in cases with this diagnosis.
3. ☐ Yes, because I would want a second pathologist's opinion for diagnostic reasons (e.g. challenging/borderline/uncertain).

1. Adapted from Allison KH, Reisch LM, Carney PA, et al. Understanding diagnostic variability in breast pathology: lessons learned from an expert consensus review panel. *Histopathology*. Aug 2014;65(2):240-251.

eFigure 2. Baseline B-Path Study Survey of Participants' Demographic and Clinical Practice Characteristics, and Attitudes about Breast Pathology Interpretation

SURVEY OF PATHOLOGISTS

Instructions: **This survey takes < 10 minutes to complete. It asks about your background and what we think are extremely important general questions related to research and clinical care in breast pathology.**

GENERAL PROFESSIONAL INFORMATION

1. **What is your year of birth? ___ ___ ___ ___ Year**

2. **What is your gender?**
- ☐ Male
- ☐ Female

3. **Are you affiliated with an academic medical center**
- ☐ Yes, adjunct/affiliated clinical faculty
- ☐ Yes, primary appointment
- ☐ No

4. **Have you received fellowship training in surgical or breast pathology?** (check all that apply)
- ☐ Yes, surgical
- ☐ Yes, breast pathology
- ☐ No

5. **The following questions are about your experience interpreting breast pathology cases.**

a. How many years have you been interpreting breast pathology cases (not including residency/fellowship training)?
- ☐ < 1 year
- ☐ 1-2 years
- ☐ 3-4 years
- ☐ 5-9 years
- ☐ 10-19 years
- ☐ ≥20 years

b. What percentage of your caseload includes interpreting breast specimens?
- ☐ <10%
- ☐ 10-24%
- ☐ 25-49%
- ☐ 50-74%
- ☐ ≥75%

c. Estimate the number of breast cases you interpret during an average week.
- ☐ <5 breast cases per week
- ☐ 5-9 breast cases per week
- ☐ 10-19 breast cases per week
- ☐ 20-29 breast cases per week
- ☐ 30-39 breast cases per week
- ☐ 40-49 breast cases per week
- ☐ ≥50 breast cases per week

d. Do your colleagues <u>consider you an expert</u> in breast pathology?

☐ Yes
☐ No

6. In general, **<u>how challenging</u>** do you find breast cases to interpret?

1☐ 2☐ 3☐ 4☐ 5☐ 6☐
Very easy Very challenging

7. What are your thoughts on interpreting breast pathology?

	Strongly disagree 1	**Disagree 2**	**Slightly disagree 3**	**Slightly agree 4**	**Agree 5**	**Strongly agree 6**
A. Interpreting breast pathology is enjoyable	□	□	□	□	□	□
B. Interpreting breast pathology makes me more nervous than other types of pathology.	□	□	□	□	□	□

8. In general, **<u>how confident</u>** are you in your assessments of breast cases?

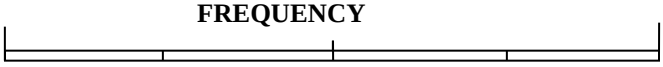
1☐ 2☐ 3☐ 4☐ 5☐ 6☐
Very confident Not at all confident

SECOND OPINION BY ANOTHER PATHOLOGST ON BREAST SPECIMENS (e.g. consultation, second read, second review)

9. Please consider the following hypothetical scenario…. You are reviewing a breast needle core biopsy from a 45 year old woman with no history of breast disease. There is an intra-ductal process that you consider to be borderline between atypical ductal hyperplasia (ADH) and ductal carcinoma in situ (DCIS), but you favor classifying as ADH.

a. In situations like this, what percentage of cases would you get a second opinion?

FREQUENCY

0%  100% ____%

(hover mouse cursor over bar to see percentage, or type a number in the box)

b. If you were to obtain a second opinion, would your second reviewer usually be blinded to your opinion on the case?

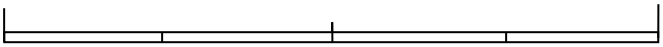
☐ Yes, they would be blinded
☐ No

c. If you obtain a second opinion and they favor DCIS, how often would you use the following methods to resolve the disagreement?

i. <u>Try to come to consensus</u> by discussing the case with the second reviewer

0% 100% ____%

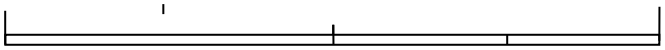
☐ Not used



ii. Diagnose according to the <u>most experienced pathologist's opinion</u>

0% 100% ____%

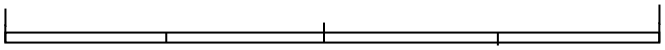
☐ Not used



iii. <u>Get a third "tie-breaker"</u> opinion or present at a consensus conference

0% 100% ____%

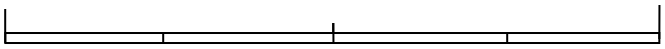
☐ Not used



iv. Diagnose as <u>borderline</u> or <u>suspicious</u> (i.e. "ADH bordering on DCIS" or "ADH suspicious for DCIS")

0% 100% ____%

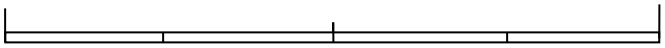
☐ Not used



v. Diagnose as DCIS to go with the <u>more severe</u> diagnosis

0% 100% ____%

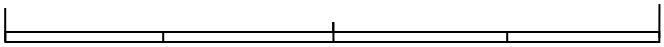
☐ Not used



vi. Diagnose as ADH to go with the <u>less severe</u> diagnosis

0% 100% ____%

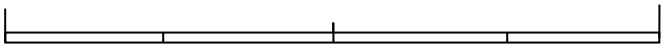
☐ Not used



vii. Other: _____

0% 100% ____%

☐ Not used



10. Some facilities have <u>policies requiring</u> a second opinion which may differ from our <u>actual practices</u> or what we think is <u>ideal for patient care</u>. Please describe your experience and thoughts on second opinions:

INITIAL DIAGNOSIS	**POLICY REQUIRED** (% of cases for which <u>my practice requires me</u> to obtain a second opinion)	**ACTUAL PRACTICE** (% of cases for which I <u>usually obtain</u> a second opinion)	**IDEAL PRACTICE FOR PATIENT CARE** (% of cases which I think <u>should ideally receive</u> a second opinion)
Invasive	**0% ---------- 100%**	**0% ---------- 100%**	**0% ---------- 100%**

DCIS	**0% ---------- 100%**	**0% ---------- 100%**	**0% ---------- 100%**
ADH	**0% ---------- 100%**	**0% ---------- 100%**	**0% ---------- 100%**
Negative (non-atypical)	**0% ---------- 100%**	**0% ---------- 100%**	**0% ---------- 100%**

11. What are your thoughts on asking another pathologist for a second opinion on cases?

	DISAGREE			**AGREE**		
	Strongly disagree	**Disagree**	**Slightly disagree**	**Slightly agree**	**Agree**	**Strongly agree**
	1	**2**	**3**	**4**	**5**	**6**
A. Improves my diagnostic accuracy	□	□	□	□	□	□
B. Takes too much time	□	□	□	□	□	□
C. Protects me from malpractice suits	□	□	□	□	□	□
D. I wish it was more available	□	□	□	□	□	□
E. I'm often hesitant to request as it may make me look less adequate as a diagnostician	□	□	□	□	□	□

DIGITIZED WHOLE SLIDE IMAGING

(Virtual microscopy is a digital process by which an electronic scanner converts histological slides into high-resolution digitized pictures known as digitized whole slide images. The term "digitized whole slides" does not refer to jpeg-style images or PowerPoint images.)

12. In what ways do you use digitized whole slides in your professional work? (check all that apply)

- ☐ Primary pathology diagnosis
- ☐ Tumor board/clinical conference
- ☐ Consultative Diagnosis
- ☐ CME/Board exams/ Teaching in general
- ☐ Archival purposes
- ☐ Research
- ☐ Other: _____
- ☐ Not at all (skip to Question 14)

13.

13a. Do you interpret digitized whole H & E slide images of breast tissue for rendering a primary diagnosis?

- ☐ No
- ☐ Yes

POP UP if YES:
I. I render a primary diagnosis in _____% of my H & E breast cases using digital whole slide imaging.
II. I render a second opinion on _____% of my second review/consultation cases using digital whole slide imaging.

III. How long have you been using digital whole slide imaging for H & E interpretation of breast cases?
☐ *≤ 6 MONTHS* ☐ *> 6 MONTHS*

13b. Do you interpret digitized whole slide images on IHC stained breast tissue slides for rendering a primary diagnosis?
☐ No
☐ Yes

POP UP if YES: I. I interpret digitized whole IHC slides in breast cases for the following (check all that apply)
☐ *Prognostic/predictive breast cancer markers (e.g., ER, HER2, other)*
☐ *Diagnostic questions (e.g., Invasive cancer vs. DCIS, E-cadherin, other)*

14. What are your thoughts on H & E digitized whole slide imaging being used for primary diagnostic purposes? (We refer to digital whole slide images as digital slides)

	Strongly disagree 1	**Disagree 2**	**Slightly disagree 3**	**Slightly agree 4**	**Agree 5**	**Strongly agree 6**
A. Accurate diagnoses can be rendered using digital slides	□	□	□	□	□	□
B. Digital slides are useful for obtaining a second opinion	□	□	□	□	□	□
C. Digital slides increase pathologist exposure to medical malpractice suits	□	□	□	□	□	□
D. It is too difficult to learn how to use digital slides	□	□	□	□	□	□
E. Overall I think the benefits of digital whole slide imaging outweigh the concerns	□	□	□	□	□	□
F. Digital slides are too slow for routine use when interpreting a case	□	□	□	□	□	□
G. I would like to adopt digital whole slide imaging or increase use of it in my personal practice	□	□	□	□	□	□

MEDICAL MALPRACTICE

15. Have you ever been named in a medical malpractice suit (including any suit filed *and* either dropped, settled out of court or gone to trial)? (*check all that apply)*
☐ No, never been sued
☐ Yes, suit(s) related to breast pathology cases
☐ Yes, suit(s) related to other pathology or other medical cases

16. Have medical malpractice concerns affected **your peer's** practice with breast cases in the following ways?

	Strongly disagree	**Disagree**	**Slightly disagree**	**Slightly agree**	**Agree**	**Strongly agree**
A. My peers order additional immunohistochemistry tests	☐	☐	☐	☐	☐	☐
B. My peers recommend						

additional surgical sampling	☐	☐	☐	☐	☐	☐
C. My peers request additional reviews (second opinion)	☐	☐	☐	☐	☐	☐
D. When a case is borderline between DCIS and ADH, my peers generally choose the more severe diagnosis of DCIS	☐	☐	☐	☐	☐	☐

17. Have medical malpractice concerns affected **<u>your own practice</u>** with breast cases in the following ways?

	Strongly disagree	**Disagree**	**Slightly disagree**	**Slightly agree**	**Agree**	**Strongly agree**
A. I order additional IHC tests	☐	☐	☐	☐	☐	☐
B. I recommend additional surgical sampling	☐	☐	☐	☐	☐	☐
C. I request additional reviews (second opinion)	☐	☐	☐	☐	☐	☐
D. When a case is borderline between DCIS and ADH, I generally choose the more severe diagnosis of DCIS	☐	☐	☐	☐	☐	☐

CONTACT INFORMATION

We will contact you in the next few weeks to schedule a convenient time for your review of the sample breast cases. Below is the contact information we have for you. Please edit and/or complete this information as needed.

Daytime phone	(Auto populate)
Email	(Auto populate)
Evening phone	(Leave blank)
Cell phone	(Leave blank)
address	(Auto populate)
City	(Auto populate)
Zip code	(Auto populate)

☐ Click this box to confirm the above information is correct

The best way to reach me is (check all that apply)

☐ Email

☐ Daytime phone

☐ Evening phone

☐ Cell phone

Thank you for participating in this exciting study. Feel free to share any additional comments:

_____

Click here to submit your survey.

eTable 1. BPATH-Dx Hierarchical Description Showing the Mapping Used to Categorize Individual Interpretations Into One of the Five Major Categories.

Diagnostic Interpretation	Primary Mapping Analysis Main BPATH-Dx Category	Alternative Mapping Analysis Main BPATH-Dx Category
Invasive (ductal or lobular or other special type)	Invasive	Invasive
Ductal carcinoma in situ (DCIS)	DCIS	DCIS
Atypical ductal hyperplasia (ADH)	Atypia	Atypia
Intraductal Papilloma with Atypia (IPA)	Atypia	Atypia
Usual Ductal Hyperplasia (UDH)	Benign without Atypia (Proliferative)	Benign without Atypia (Proliferative)
Columnar Cell Hyperplasia/ Columnar Call Change (CCH/CCC)	Benign without Atypia (Proliferative)	Benign without Atypia (Proliferative)
Sclerosing Adenosis	Benign without Atypia (Proliferative)	Benign without Atypia (Proliferative)
Radial Scar/Complex Sclerosing lesion	Benign without Atypia (Proliferative)	Benign without Atypia (Proliferative)
Flat Epithelial Atypia (FEA)	Benign without Atypia (Proliferative)	Atypia
Intraductal Papilloma w/o Atypia (IP)	Benign without Atypia (Proliferative)	Benign without Atypia (Proliferative)
Non-Proliferative only	Benign without Atypia (Non-Proliferative)	Benign without Atypia (Non-Proliferative)
Fibroadenoma (FA)	Benign without Atypia (Non-Proliferative)	Benign without Atypia (Proliferative)
LCIS*	Benign without Atypia (Non-Proliferative-Please see footnotes)[1]	DCIS
ALH*	Benign without Atypia (Non-Proliferative-Please see footnotes)[1]	Atypia

Footnote 1. The primary and alternative categorical mapping strategies differ in how four diagnostic assessments, LCIS, ALH, FEA, and FA, are assigned to BPATH-Dx categories in the analysis of reference and participant diagnoses. These lesions are not a focus of the B-Path study but were present by random chance on some slides. For primary mapping, if ALH or LCIS is present, the case maps to the other diagnoses also on the slide using the hierarchy, or is grouped with non-proliferative if no other diagnoses are noted. This allowed the analysis to focus on ADH and DCIS. For alternative mapping, ALH is grouped with ADH in the atypia category and LCIS is grouped with DCIS following traditional cancer progression schemes. For primary mapping, FA is grouped with non-proliferative if no other diagnosis is noted and is grouped with proliferative in the alternative mapping; FA is technically a proliferative lesion but has little associated risk. FEA is a lower risk lesion biologically, may be a precursor to ADH, and for primary mapping it was grouped lower than ADH in the proliferative category in the primary mapping. In the alternative mapping, FEA was grouped with ADH because FEA may lead to excision in some institutions. Analyses were performed for the primary and alternative mapping schemes.

eTable 2. Measures of Overinterpretation, Underinterpretation and Concordance when Comparing Pathologists' Interpretation to the Reference Diagnosis. Three Alternative Methods are Employed: I. Using the Alternative Mapping Scheme Described in eTable 1; II. Using the Participants' Community Standard Diagnosis for 17 Cases instead of the Expert Consensus Reference Diagnosis;[1] and III. Deleting the 17 Cases.[1]

Reference Diagnosis	Overinterpretation Rate % (95% CI)	Underinterpretation Rate % (95% CI)	Overall Concordance Rate % (95% CI)
I. Results Following Alternative Mapping Scheme Described in eTable 1			
Benign without Atypia	18% (16%, 21%)	---	82% (79%, 84%)
Atypia	19% (16%, 22%)	27% (24%, 30%)	54% (51%, 57%)
Ductal Carcinoma in situ (DCIS)	2% (2%, 4%)	10% (9%, 12%)	87% (85%, 89%)
Invasive Breast Cancer	---	4% (3%, 6%)	96% (94%, 97%)
II. Results Using Participant Majority Diagnosis as the Reference Diagnosis for 17 cases[1]			
Benign without Atypia	16% (14%, 18%)	---	84% (82%, 86%)
Atypia	18% (15%, 21%)	31% (27%, 35%)	51% (48%, 55%)
Ductal Carcinoma in situ (DCIS)	3% (2%, 4%)	12% (11%, 14%)	85% (82%, 87%)
Invasive Breast Cancer	---	1% (0%, 3%)	99% (97%, 100%)
III. Results Without the 17 Cases [1]			
Benign without Atypia	12% (10%, 14%)	---	88% (86%, 90%)
Atypia	18% (15%, 22%)	30% (27%, 34%)	52% (48%, 55%)
Ductal Carcinoma in situ (DCIS)	3% (2%, 4%)	11% (10%, 13%)	86% (84%, 88%)
Invasive Breast Cancer	---	1% (0%, 3%)	99% (97%, 100%)

1. For 223/240 (93%) cases, we considered the reference adequate as the three consensus panel members' independent interpretations agreed and/or their reference consensus diagnosis corresponded to the most frequent interpretation by the participating pathologists. For the remaining 17/240 (7%), we reanalyzed the data by substituting the most frequent participant interpretation as the reference diagnosis, or excluding the 17 cases.

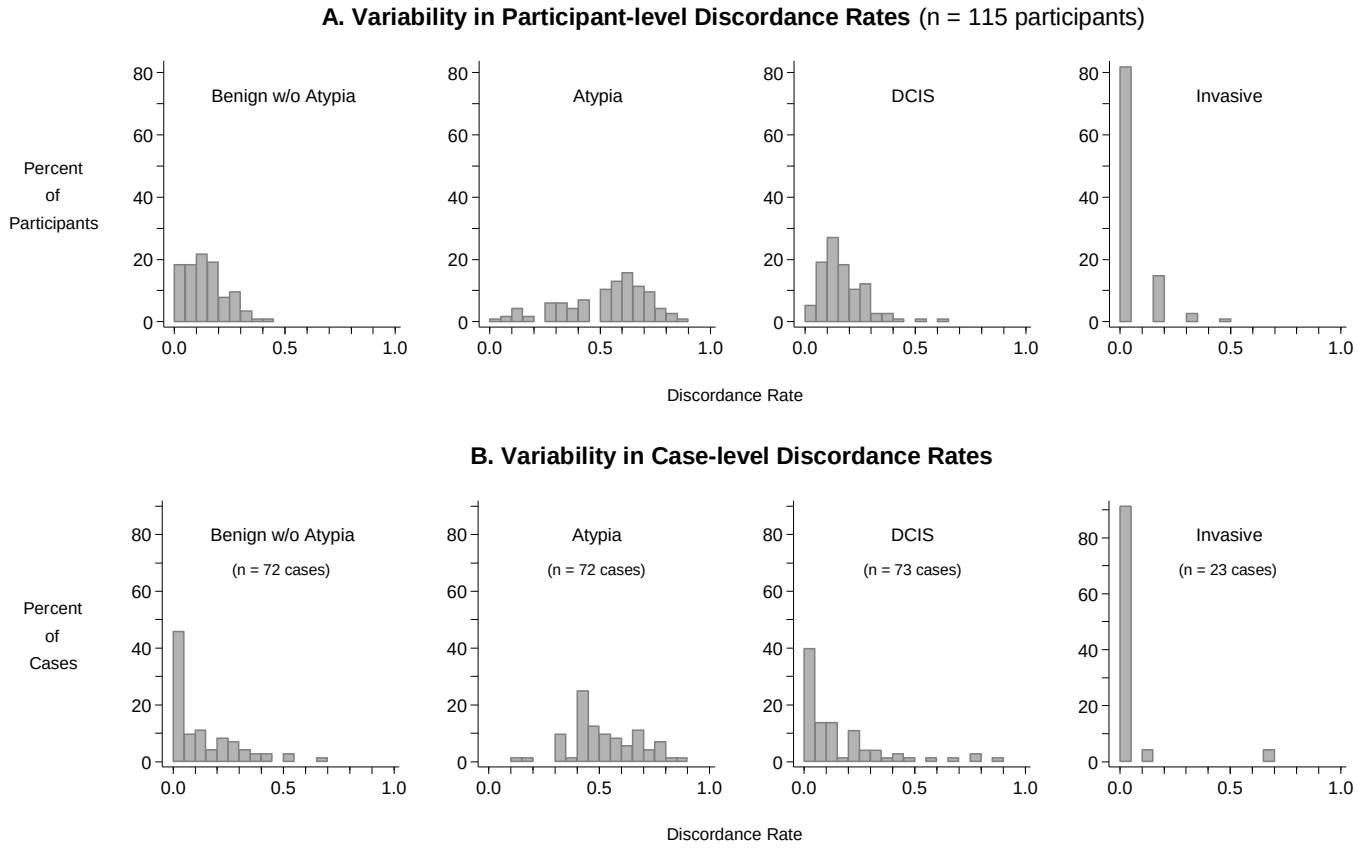
eTable 3. Multivariable Logistic Regression Model of Participant Misclassification with Respect to the Four Category Consensus Reference Diagnosis.[a]

Participant Characteristics[b]	Odds Ratio	95% CI	Z	P-value	No. misclassified cases / total no. (%) [reference category]
Breast specific case load (≥10 cases/week)	0.799	0.68, 0.94	-2.75	.006	517/2400 (21.5%) [1189/4500 (26.4%)]
Academic affiliation	0.768	0.65, 0.90	-3.24	.0012	346/1680 (20.6%) [1360/5220 (26.1%)]
Practice size (≥10 pathologists)	0.849	0.72, 1.00	-2.05	.0399	631/2820 (22.4%) [1075/4080 (26.3%)]
constant	0.502	0.40, 0.62	-	-	

[a.] Wald test statistics are based on bootstrap standard errors from 3000 bootstrap samples. Confidence intervals for the odd ratios are based on percentiles of the bootstrap sample coefficient estimates. Sampling was clustered on participating pathologist: A sample consisted of 115 participants drawn randomly with replacement from the original sample, along with all case observations for each sampled participant.

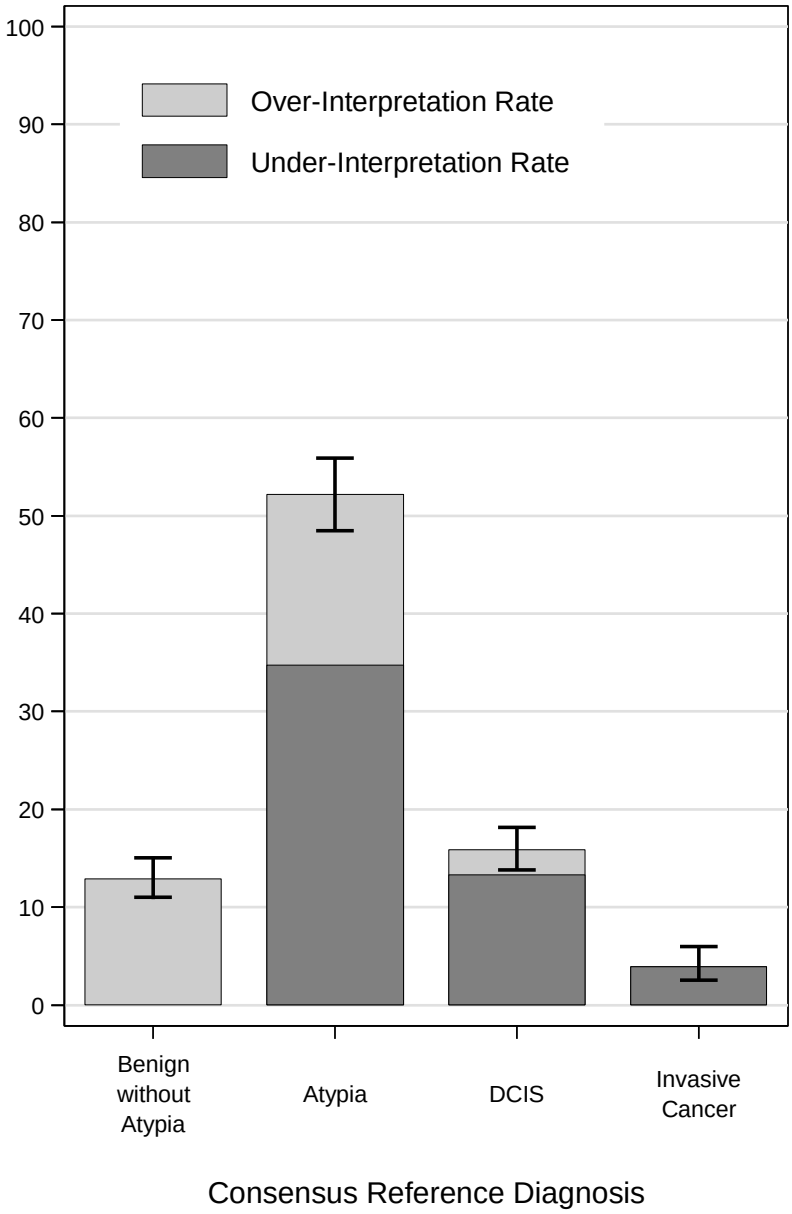
[b]Breast-specific case load is 10+ breast cases/week vs < 10; academic affiliation is any (primary or adjunct) vs none; practice size is 10+ pathologists who interpret breast cases in the same lab vs < 10.

eFigure 3. Variability in Discordance Rate at the Participant-level (N=115 pathologists, Figure A) and Case-level (N=240 cases, Figure B) by Reference Diagnostic Category.[1]



1. Substantial variability was noted in the discordance rates for individual pathologists and for individual cases. For example, although the average discordance rate of pathologists' interpretations was 0.52 for atypia cases, 9% of pathologists had discordance rates of <.20 while another 17% had discordance rates of >.70 for the atypia cases (Figure A). From the perspective of the cases deemed atypia by the reference standard (Figure B), the discordance rate was > 50% for 46% of atypia cases. For cases deemed DCIS by the reference standard, discordance rates were more than .20 for 32% of cases but complete agreement was noted with the reference standard for 15% of cases. Although there was one case of invasive cancer under-interpreted by over 60% of pathologists (a case of micro-invasion), 78% of invasive cases were correctly classified by all pathologists that read them.

eFigure 4. Over- and Underinterpretation Rates by Consensus Reference Diagnosis.



eFigure 5. Over- and Underinterpretation Rates by Participants' Rating of Specific Attributes of the Case: a) Diagnostic Difficulty of the Case; b) Their Level of Confidence in Their Assessment of the Case; c) Whether They Would Obtain a Second Opinion on the Case in Their Own Practice (Either a Required Second Opinion Due to An Existing Policy Or Because They Would Want a Second Opinion); and d) Their Assessment of Whether the Case is "Borderline" Between Two Assessments.