AUDIT
(Not the IRS!)

2013 Assessment Category Standardized for All Modalities

Assessment	Management	Likelihood of Cancer
Category 0: Incomplete—Need Additional Imaging Evaluation and/or Prior Mammograms for Comparison	Recall for additional imaging and/or comparison with prior examination(s)	N/A
Category 1: Negative	Routine screening	Essentially 0% likelihood of malignancy
Category 2: Benign	Routine screening	Essentially 0% likelihood of malignancy
Category 3: Probably Benign	Short-interval (6-month) follow-up or continued surveillance imaging	> 0% but ≤ 2% likelihood of malignancy
Category 4: Suspicious (also 4A/B/C for Mammo and US)	Tissue diagnosis	> 2% but < 95% likelihood of malignancy
Category 5: Highly Suggestive of Malignancy	Tissue diagnosis	≥ 95% likelihood of malignancy
Category 6: Known Biopsy-Proven Malignancy	Surgical excision when clinically appropriate	N/A

WHAT IS A POSITIVE SCREEN EXAM?

A positive mammogram is a 4 view (or 2 view) study that is given an assessment of BI-RADS 0,3,4 or 5 (Negative if BI-RADS 1,2)

A positive breast ultrasound exam is a 5 view study with the recording of one image per quadrant and an additional retro areolar image (ACRIN 6666 criteria) that is given an assessment of BI-RADS 0,3,4 or 5(Negative if BI-RADS 1,2)

Since the number of images and parameters for either a screening or diagnostic MRI are the same the definitions of a positive screen and diagnostic exam are the same.

WHAT IS A POSITIVE SCREEN EXAM?

NB: If the screening exam includes additional images other than those done for breast coverage issues, this is also a positive exam .For example your technologist adds a 90 degree lateral on a screening mammogram or orthogonal images of a cyst on a screening US. These exams are actually a combined screen and diagnostic even if clinically read as one exam and assigned a 1/2. For *auditing purposes* these should be separated.

TRUTH

THE DETECTION OF MALIGNANCY WITHIN A YEAR OF THE EXAM

BASIC AUDIT DEFINITIONS

Sensitivity - The percent of cancer detected from all cancers

Specificity -The percent of negative cases identified when no cancer is present

Recall Rate - The percent of screens given 0 (additional imaging evaluation)

Abnormal Interpretation rate – The percentage of all positive exams /all exams

Accuracy-The percent of cancer and negative cases identified from all cases

Positive Predictive Value$_1$ (PPV)- The percent of screening exams with a positive interpretation and cancer within a year

Positive Predictive Value $_2$ -The percentage of all positive exams with a biopsy recommended*(4/5)* and cancer within a year

Positive Predictive Value $_3$ - The percent of biopsies done with a positive interpretation *(4/5)* and a known biopsy of cancer in a year

Cancer detection rate (per thousand)- Number of cancers detected per thousand women

Percentage of minimal cancer -The amount of cancer that is </= 1cm or DCIS from all cancers detected

BASIC AUDIT CALCULATIONS

- **Sensitivity** = TP/TP+FN
- **Specificity** = TN/TN+FP
- **Accuracy** = TP +TN/TP+TN+FP+FN
- **PPV_1** = TP/TP + FP (Percentage of positive screens with a cancer diagnosis within a year)
- **PPV_2** = TP/TP + FP(Percentage of exams with a *recommendation* for biopsy and cancer within a year)
- **PPV_3** = TP/TP + FP (Percentage of known biopsies *done* for 4/5s with a diagnosis of cancer within the year)

"FLAVORS" OF PPV_2

- Classic definition is a recommendation for biopsy(4/5) given after *diagnostic* imaging evaluation and cancer within a year. Meant to evaluate diagnostic exams. (PPV_2 dx)

- *Alternate definition is ALL recommendations for biopsy both from screens and from "0"s resulting in 4/5.*Meant to indicate downstream outcomes of tissue diagnoses resulting from 4/5 on screening exams and 4/5 from 0s after a dx workup.(PPV_2 sc)

DATA: 1000 SCREENS,4 READ AS 4/5,100 READ AS CAT 0 , WORKUP OF 0s YIELDED 16 4/5s, 14 BIOPSIES DONE AND 6 CANCERS DISCOVERED

PPV_1

- 4 read as "4/5" from screen are +
- 100 read as "0" are +
- 6 cancers in the year from exam
- Thus PPV_1 = TP/TP + FP
  - PPV_1 = 6/6+98 = **5.7%**

$PPV_{2\ sc}$

- 4 recommended for biopsy from the screen are +
- 16 of 100 "0"s end up as 4/5 are +
- 6 cancers in the year
- Thus PPV_2 = TP/TP+FP
  - PPV_2 = 6/6 + 14 = **30%**

DATA: 1000 SCREENS 4 READ AS 4/5,100 READ AS 0,WORKUP OF 0s YIELDED 16 4/5s, 14 BIOPSIES DONE AND 6 CANCERS DISCOVERED

$PPV_{2\ dx}$

- 16 of 100 "0"s read as 4/5 are +
- 6 cancers discovered

- Thus PPV_2 = TP/TP + FP
  - PPV_2 = 6/6+ 10 = **37.5%**

$PPV_{3\ dx}$

- Of those recommended for BX 14 are actually done

- Thus PPV_3 = TP/TP+FP
  - PPV_3 = 6/6+8 = **42.8%**

Academic Bar Bet: I can read 3000 mammograms in 5 seconds with a 99.9% accuracy !!!!!!!!!!!!!!!!

Any takers?

I read all the mammograms as category 1 /2 , thus all negative reads

Data: 7 cancers in a year, no positive reads, all negative reads

TN=2993 TP=0, FP=0 FN=7

ACC = TP+TN/TP+FP+TN+FN

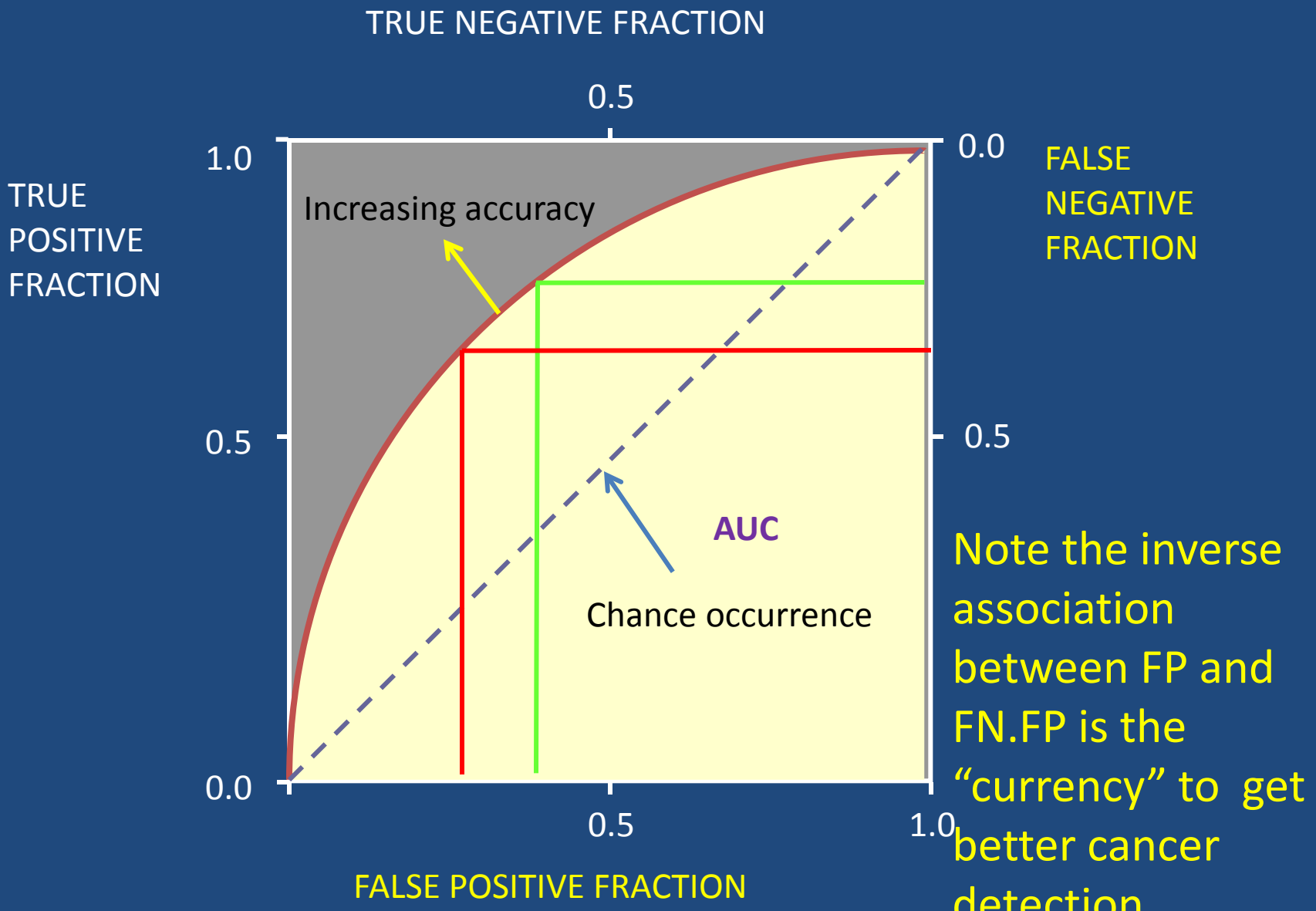
ACC= 0 + 2993/0+0 + 2993 +7

ACC = 99.9%

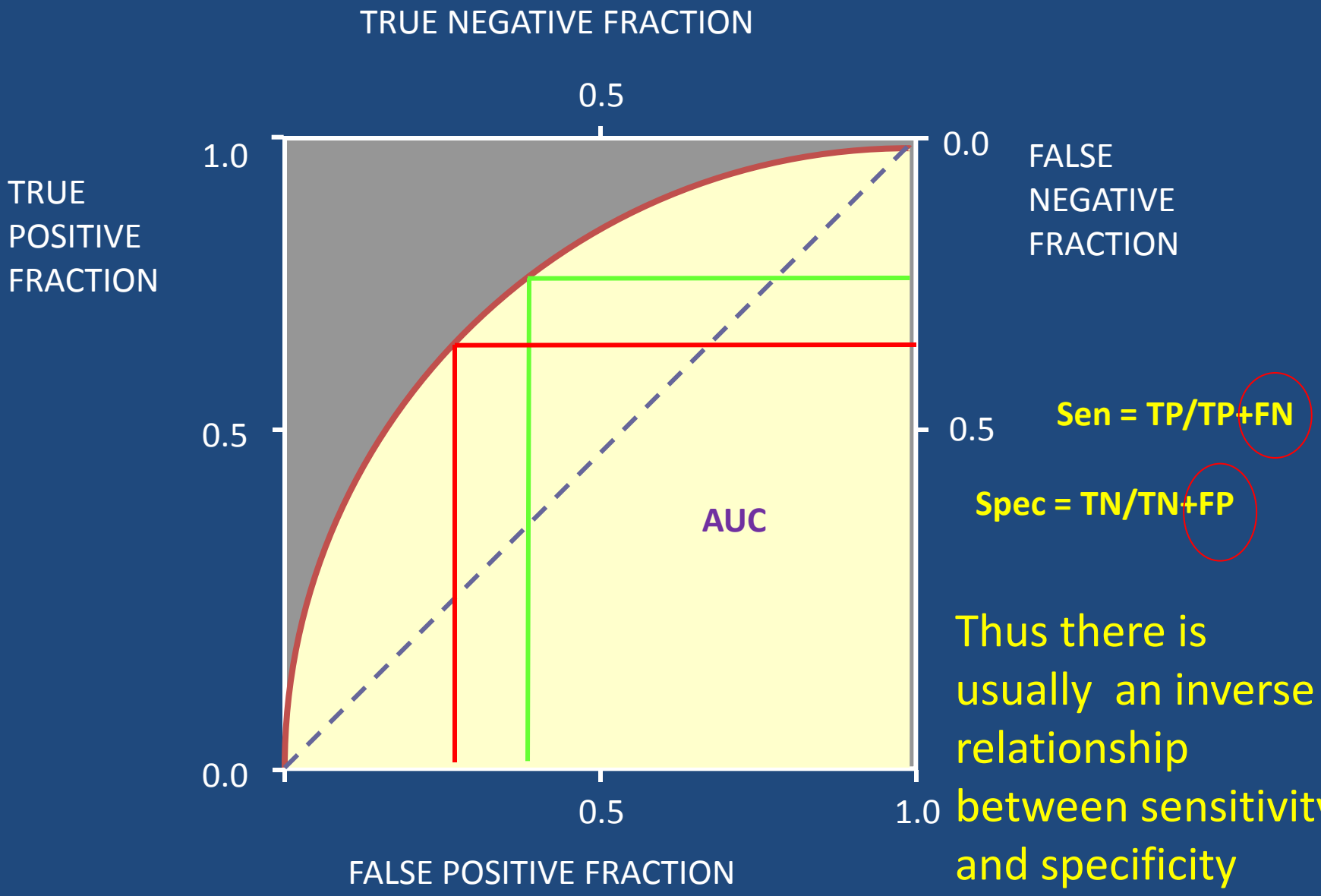
One metric in isolation can produce unwarranted assumptions

ROC

THE “NO FREE LUNCH” CURVE



Note the inverse association between FP and FN.FP is the "currency" to get better cancer detection



Thus there is usually an inverse relationship between sensitivity and specificity

What are reasonable metrics to develop a performance profile for individuals and facilities

- Since sensitivity includes the FNs and specificity includes TNs ,which require a tumor registry, these may be difficult for many facilities to calculate.
- PPV combines what you *"pay"* in FPs in order to detect a breast cancer (PPV=TP/TP+FP)
- Cancer Detection Rate (CDR) is *"how much"* breast cancer is detected
- Minimal cancer is the *"type"* of cancer detected
  (</= 1cm or DCIS)

Identifying Minimally Acceptable Interpretive Performance Criteria for Screening mammography *Carney,Sickles Monsees et al Radiology 2010;255:354-361*

MEASURE	LOW PERFORMANCE RANGE	PERCENTAGE OF BCSC RADIOLOGISTS IN LOW PERFORMANCE RANGE
SENSITIVITY		
SPECIFICITY	<88 or >*95%*	47.7
RECALL RATE		
PPV_1		
$PPV_{2,3}$		
CANCER DETECTION RATE		

Why isn't a high specificity rate not wonderful! Remember that Sp = TN/TN + FP. From ROC we know that FPs and FNs and thus Se and Sp are inversely related so this low FP may result in lower Se and higher FNs

MEASURE	LOW PERFORMANCE RANGE	PERCENTAGE OF BCSC RADIOLOGISTS IN LOW PERFORMANCE RANGE
SENSITIVITY		
SPECIFICITY		
RECALL RATE		
PPV_1	<3 or > 8%	38.4
$PPV_{2,3}$		
CANCER DETECTION RATE		

Doesn't a higher PPV mean I'm brilliant and can pick out cancers from my mammo screens? Remember PPV_1 = TP/TP +FP .Once again the lower FPs, for a screen, indicate that recalls may be too low and one is only detecting the low hanging fruit (more obvious advanced stage cancers)

MEASURE	LOW PERFORMANCE RANGE	PERCENTAGE OF BCSC RADIOLOGISTS IN LOW PERFORMANCE RANGE
SENSITIVITY		
SPECIFICITY		
RECALL RATE		
PPV_1		
$PPV_{2,3}$	< 20 0r > *40%*	34.0
CANCER DETECTION RATE		

I get a 40% or greater yield of malignancy and I'm low performing ??? This metric must be closely associated with tumor type found. If you are only sampling features that have a high possibility of malignancy again you will miss less obvious tumors that may be at an earlier stage.

What is the National Mammography Database (NMD)?

A registry for breast imaging that

allows facilities and physicians to ***monitor and improve quality*** using standardized data elements and measures consistent with BI-RADS®

National Mammography Database

- Went live in 2009
- Currently 275 registered sites; 162 contributing data
- Over 9 million exams
- Good representation across the country and across practice types and locations

DATA ELEMENTS

Patient background: includes patient age, race, ethnicity, height, weight, personal and family history of breast cancer

Exam information: includes date of exam, identifying code of interpreting radiologist, indication, breast density, assessment category, management recommendation.

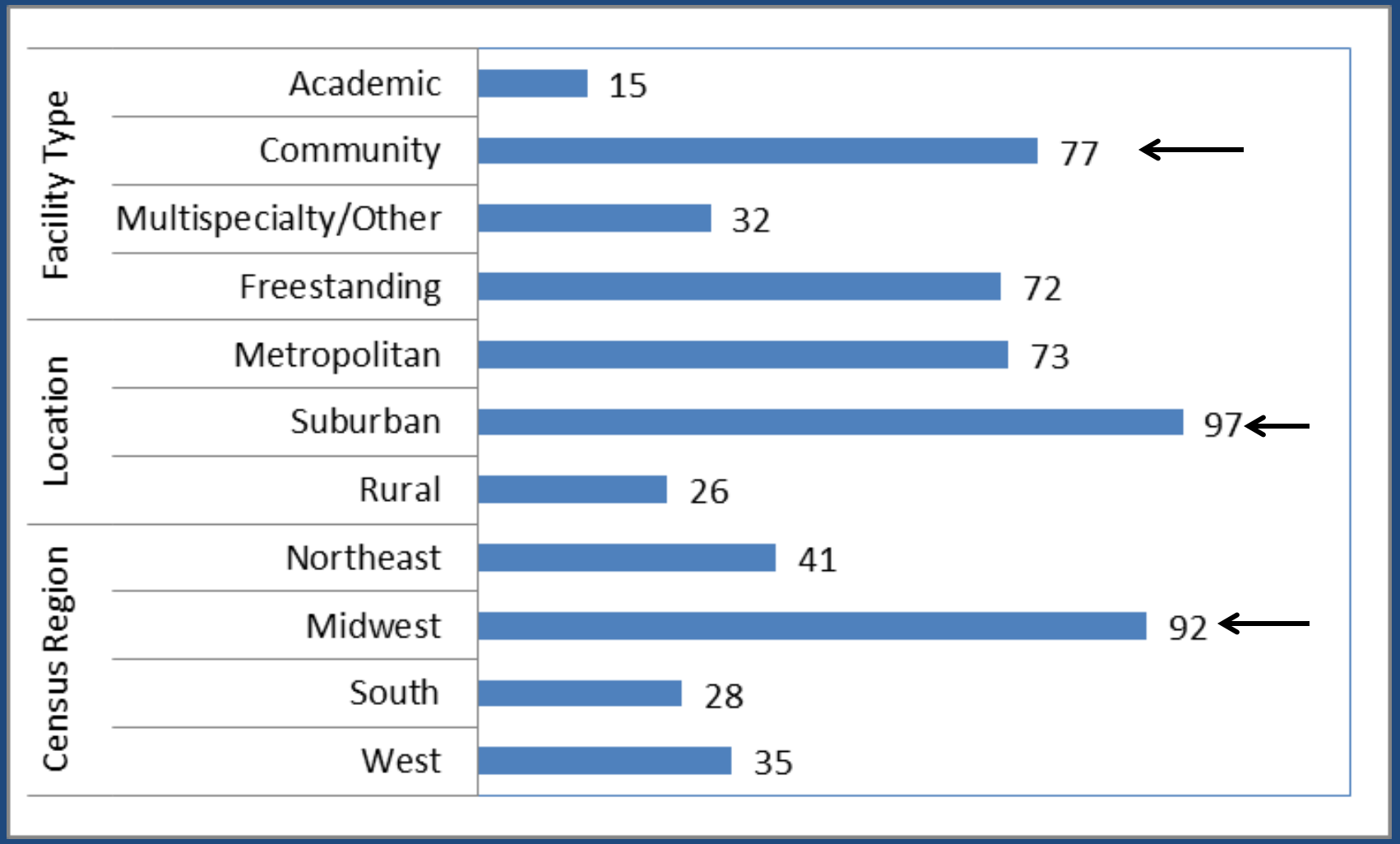
Outcomes: includes biopsy procedure and date, biopsy result, tumor size, nodal status, tumor stage.

NMD Data Submission

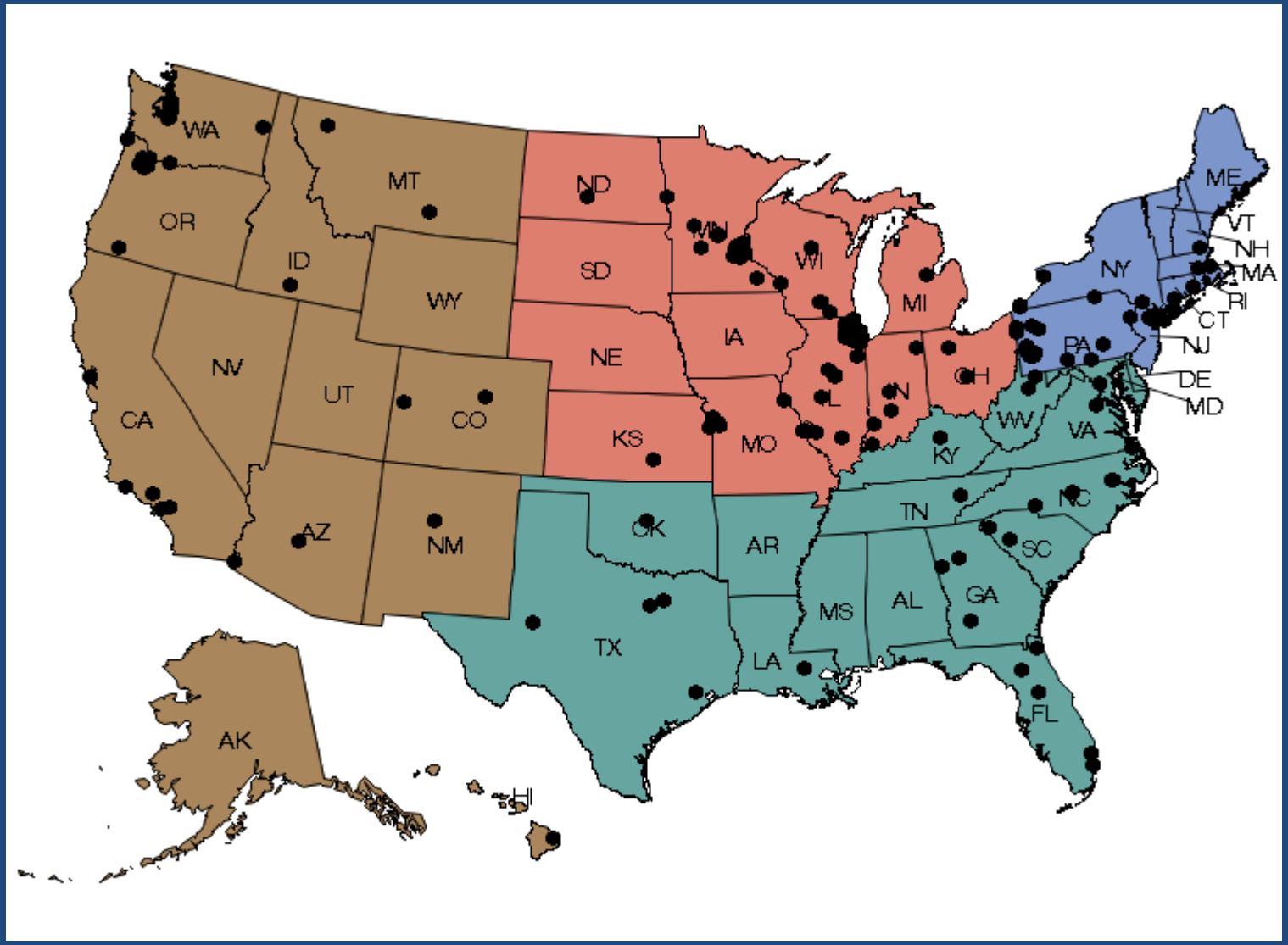
- Automated
- Data sent directly from certified vendor or through certification of home-grown software
- Currently includes mammography only; will expand to include MRI and Ultrasound in late 2015

NMD-certified software vendors

Participating facilities – 2008-2015



Participating facilities – March 2015



Source: NMD, Jan-Dec 2013

Measure	All NMD facilities	Community	BCSC benchmark
All exams	1,642,471	641,317	2,410,932
Recall rate	→ 9.9%	10.9%	11.0%
Biopsy rec	1.5%	1.6%	1.4%
Cancers and PPV2	→ 23.7%	22.9%	22.9%
Biopsy negative	75%	76%	72%
Biopsy positive	25%	24%	28%
CDR	3.6%	3.8%	4.2%
Invasive	73%	74%	76%
DCIS	27%	26%	24%
Minimal cancer	→ 66%	64%	53%
Nodal status neg	80%	83%	76%

Performance Evaluation

Metrics	Benchmarks
• Recall rate	• 5-12 %
• PPV2 sc (TP/TP+FP)	• 20 - 40%
• PPV2 dx (TP/TP+FP)	• 20 – 45%
• CDR	• 2.5/1000 or greater
• % minimal cancer (</=1cm and DCIS)	• 50%