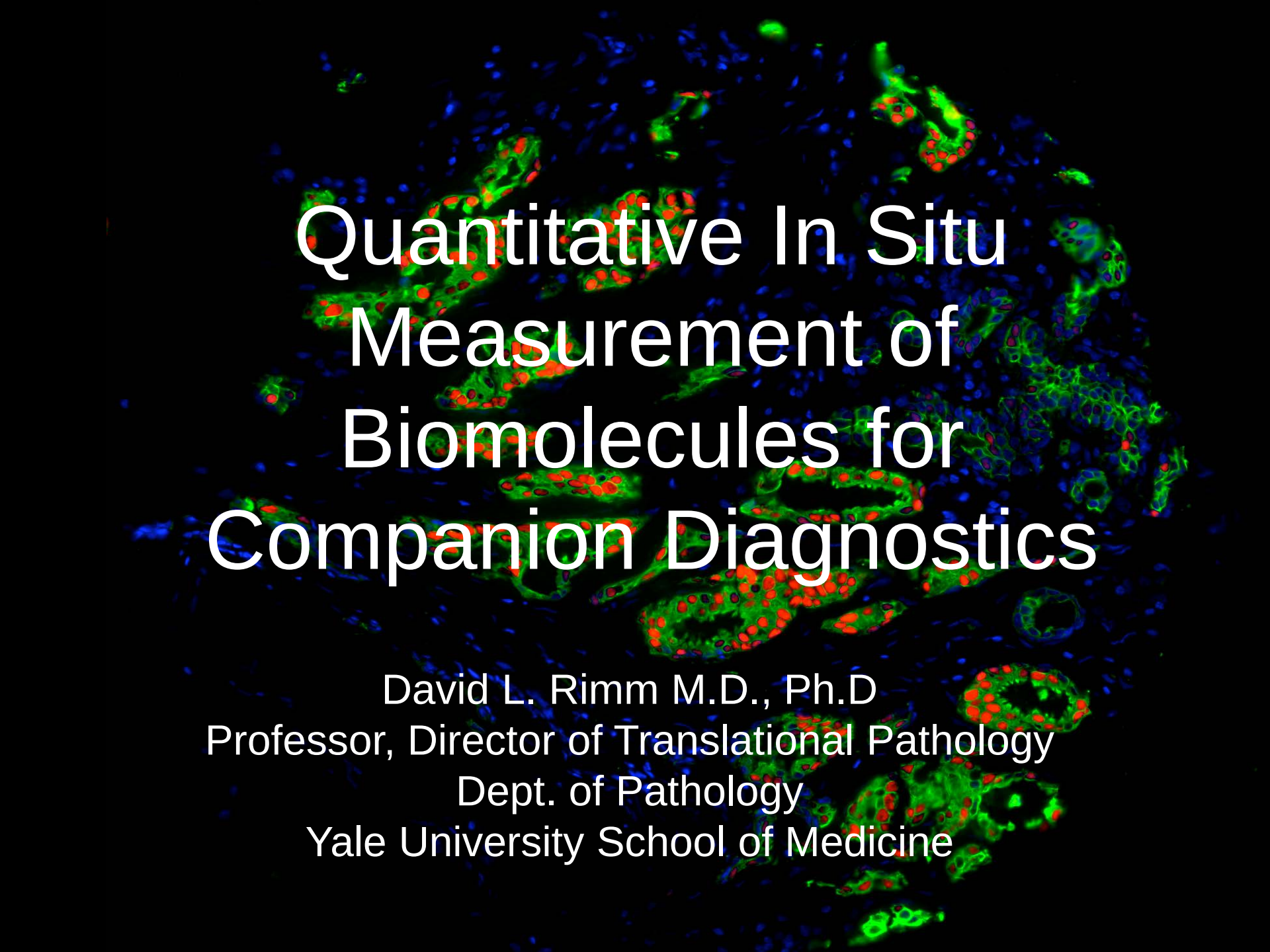


Quantitative In Situ Measurement of Biomolecules for Companion Diagnostics

David L. Rimm M.D., Ph.D
Professor, Director of Translational Pathology
Dept. of Pathology
Yale University School of Medicine

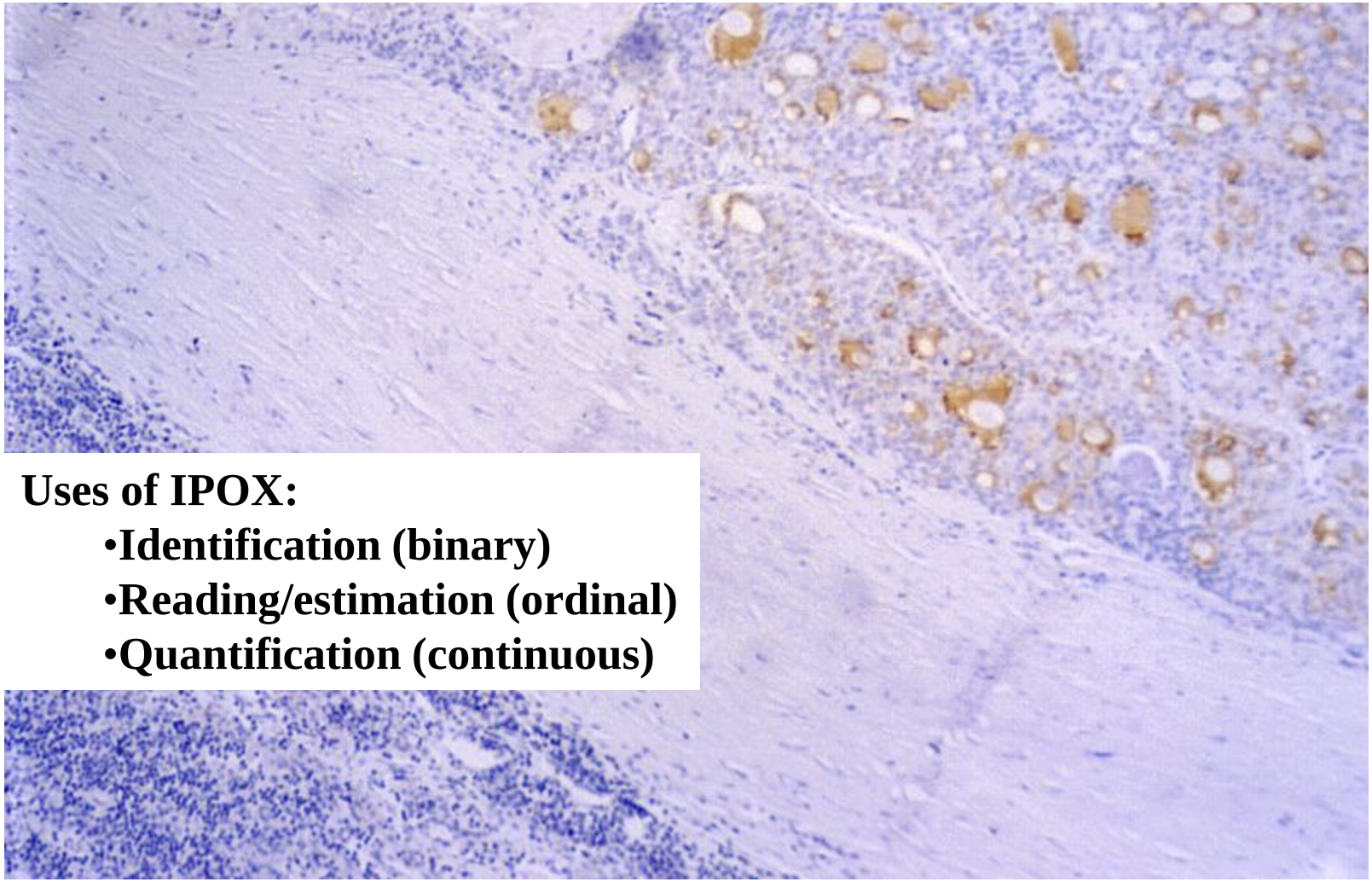
Disclosures

- I am a consultant to Amgen, Applied Cellular Diagnostics, Avida Labs, Biocept, BMS, Cernostics, Genoptix/Novartis, Metamark Genetics, MDAgree, OptraScan, and Perkin Elmer
- Cepheid, Genoptix, Gilead Sciences, Kolltan and OncoplexDx fund research in my lab.

Outline for Presentation

- Using immunohistochemistry (IHC) or quantitative immunofluorescence (QIF) to measure protein on slides
- The challenge of a continuous biomarker – the HER2 example
- The challenge of defining the threshold of detection – The ER example

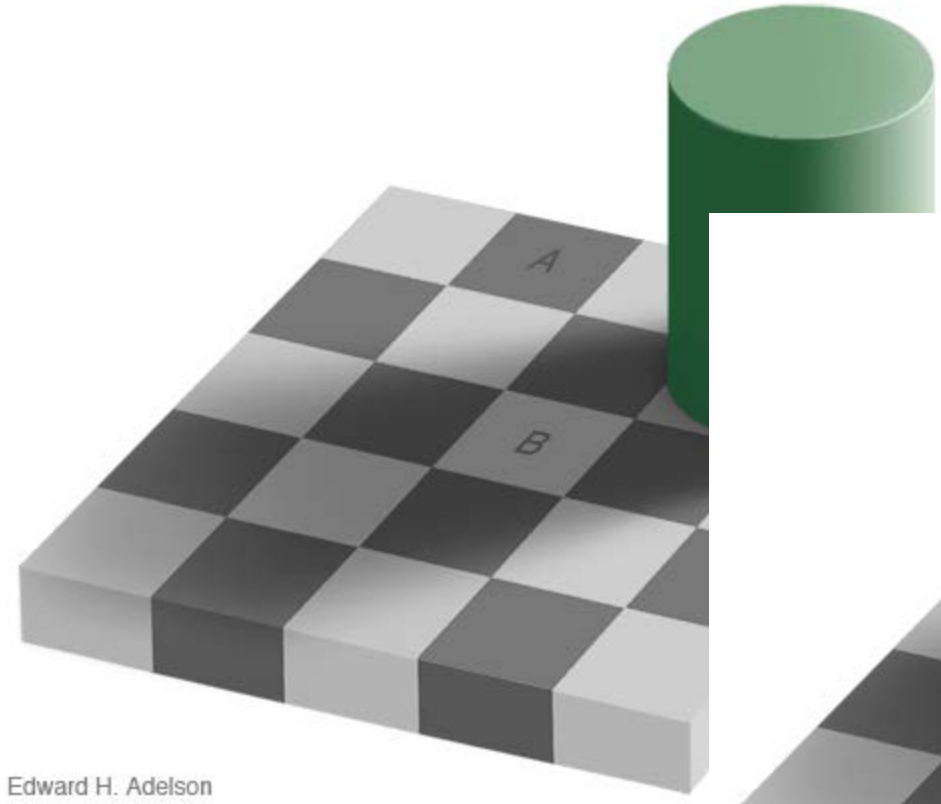
Immunoperoxidase stain



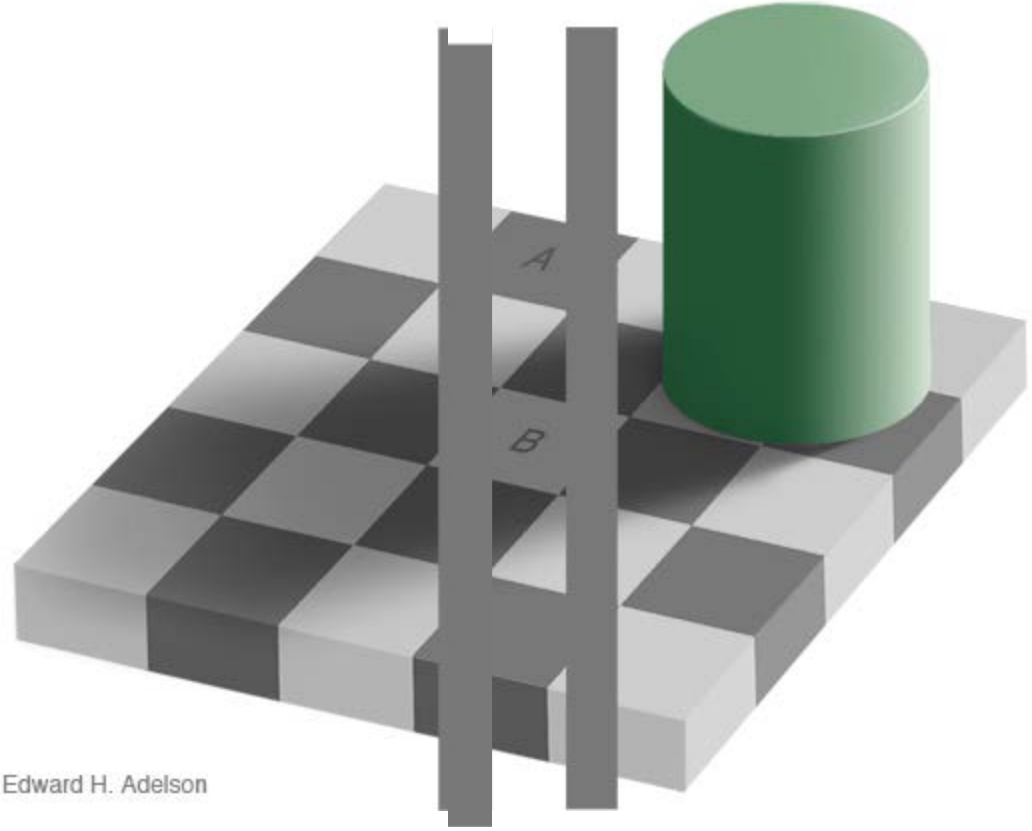
Uses of IPOX:

- **Identification (binary)**
- **Reading/estimation (ordinal)**
- **Quantification (continuous)**

The human eye is not a great tool for assessment of intensity



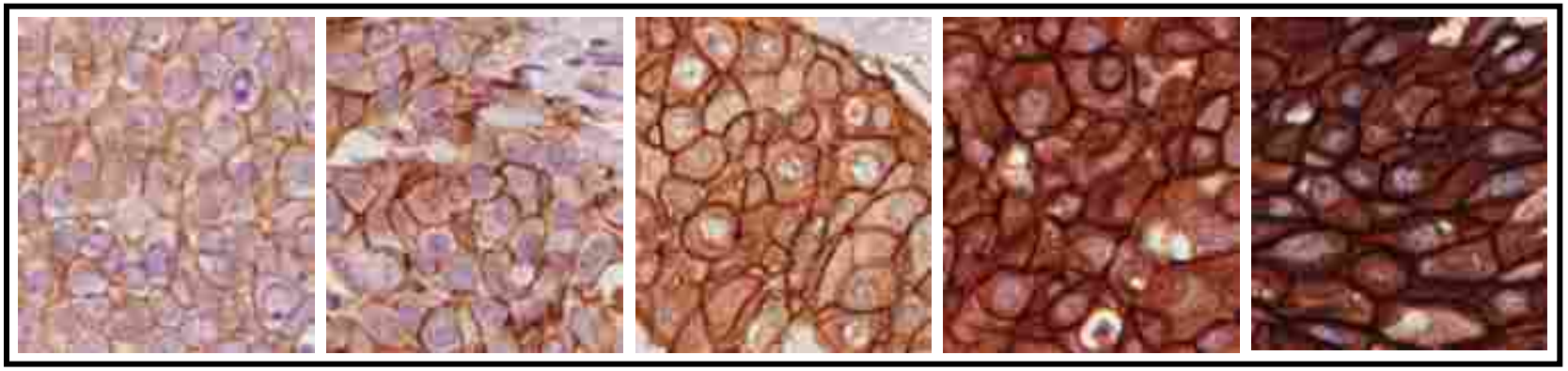
Edward H. Adelson



Edward H. Adelson

Different Intensities of HER2 IHC Staining Observed Within HER2(+) Patient Population

Intensities of HER2 IHC Staining Observed in HER2^{+} Patients From the HERA Trial



2+ **3+** **3+** **3+** **3+**

Although there was a high proportion of positive staining tumor cells in all of the 3+ samples the range of staining intensity varied.

Slide provided by Mitch Dowsett

Zabaglo L, et al. Presented at 33rd SABCS; December 8-12, 2010; San Antonio, TX. Abstract PD10-01.

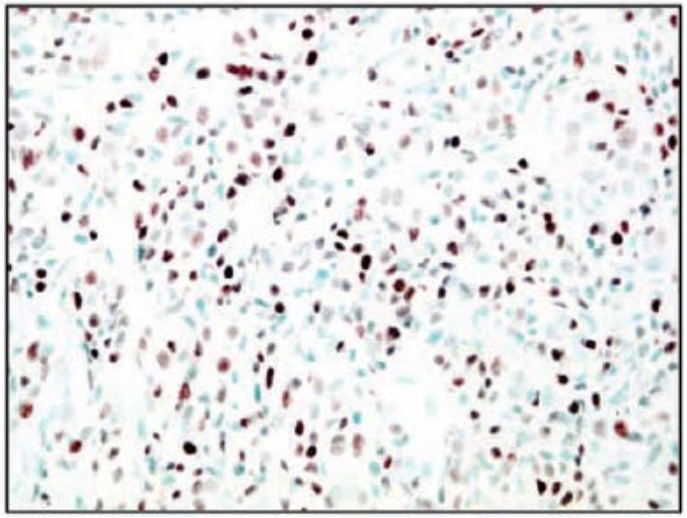
From an FDA submission for an HER2 antibody submission

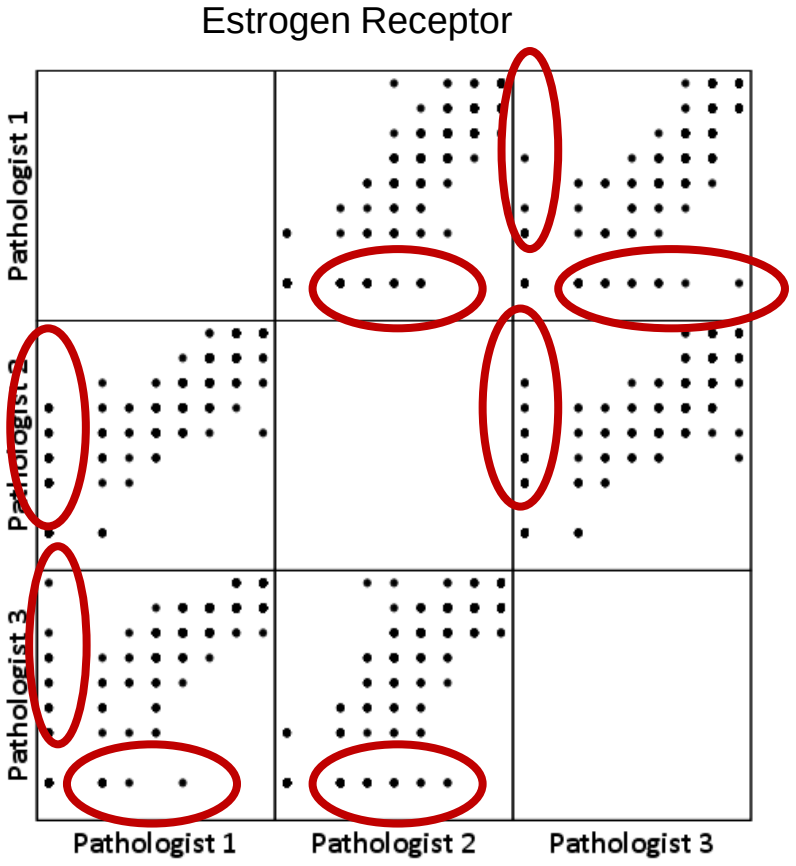
Inter-Observer Reproducibility

The inter-observer reproducibility was assessed using 40 invasive breast cancer cases (resection specimens) that were sectioned and provided to the 3 sites for staining and interpretation. The sections were blinded and randomized at each site prior to scoring. Interobserver agreement between the 2 independent study sites, was 87.5% (95% CI = 73.3% to 95.8%). The agreement between the independent study sites and LBN was 92.5% (95% CI = 79.6 to 98.4%) and 85% (95% CI = 70.1% to 94.3%), at site 1 and site 2, respectively.

Conclusion: We are willing to accept an error rate of up to 15%!?

“Real” Pathologist Reproducibility: 3 different pathologists read Allred scores on 100 cases of breast cancer





Path 1 v. Path 2: Kappa = 0.482 (p<0.001
Path 1 v. Path 3: Kappa = 0.444 (p<0.001
Path 2 v. Path 3: Kappa = 0.400 (p<0.001

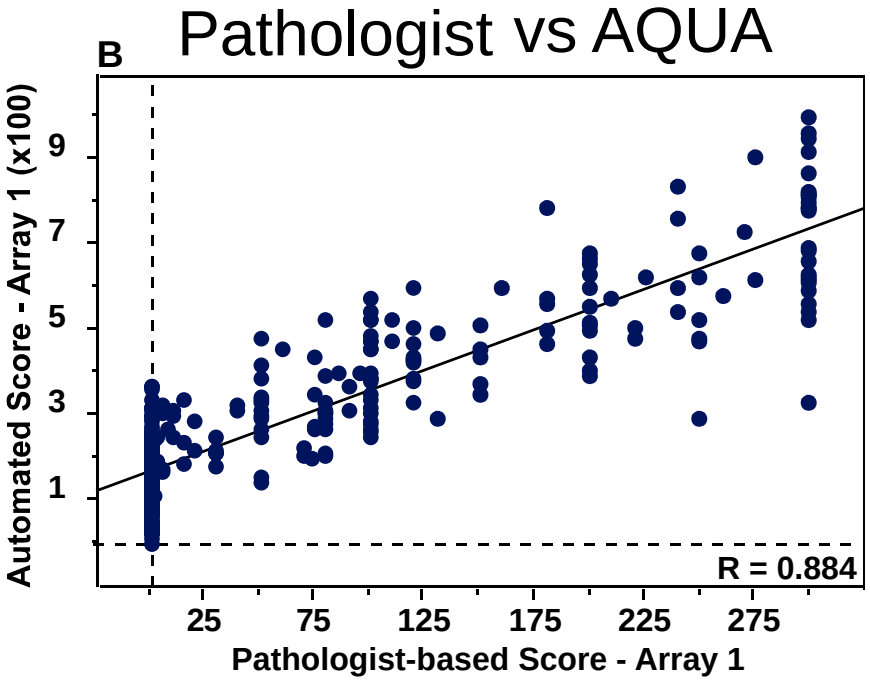
***Positive/Negative concordance: 92-95%**

In RED, 9% misclassification rate

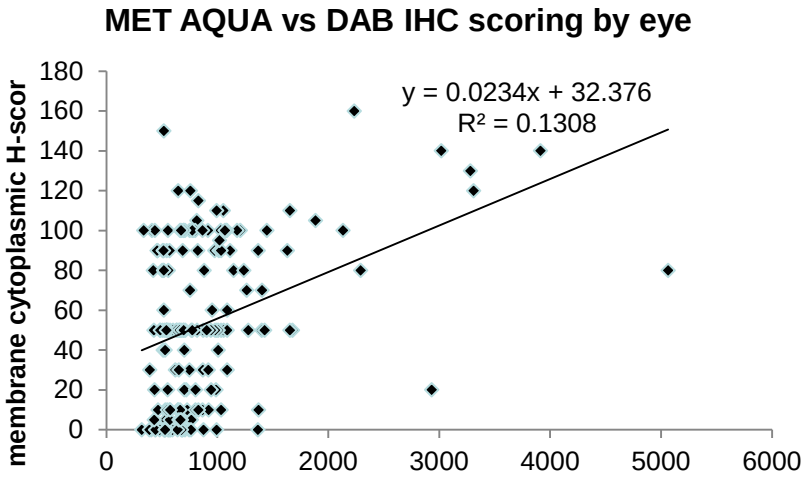
Mark Gustavson

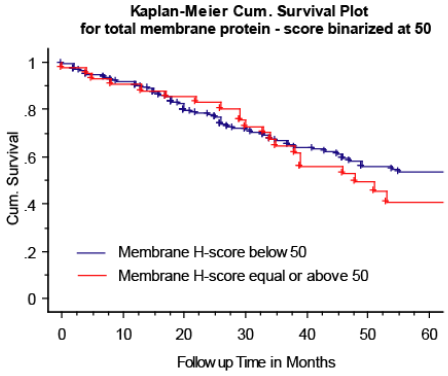
Artifactual Grouping by Pathologists

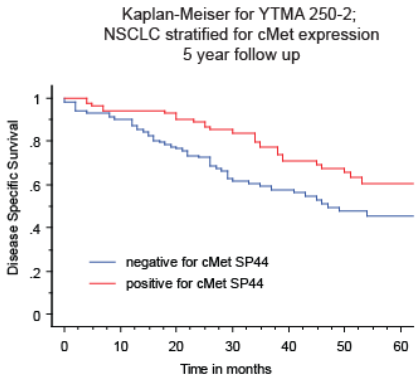
Estrogen Receptor



MET (HGF Receptor)







Commercially Available Pathology-Focused Image Analysis Platforms

- Definiens
- Visiopharm
- Perkin Elmer Vectra/INform
- Hamamatsu Nanozoomer
- Leica/Aperio
- Tissuegnostics
- Ventana (formerly Bioimagene)
- Optra SCAN
- Genoptix (formerly HistoRx) AQUA
- Many others sold as research platforms

AQUA®: objective analyte measurement on a tissue slide based on co-localization

Step 1: **Mask (define region of interest, exclude stroma, blank space, etc) = colocalization with Cytokeratin for carcinoma**

Step 2: **Define the numerator (target) and denominator (compartment)**

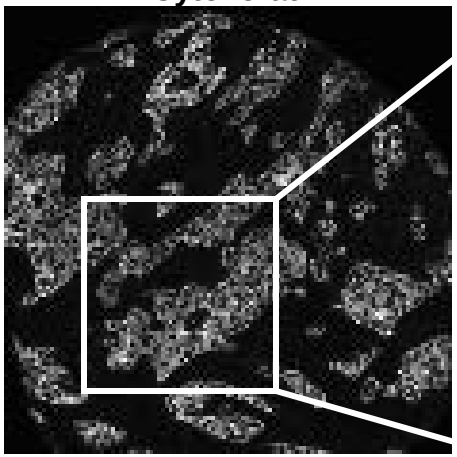
$$\text{Concentration} = \frac{\text{Numerator}}{\text{Denominator}} \Rightarrow \frac{\Sigma \text{ target intensity in compartment pixels}}{\Sigma \text{ compartment pixel area}} = \text{AQUA score}$$

Step 3: **Calculate the AQUA score**

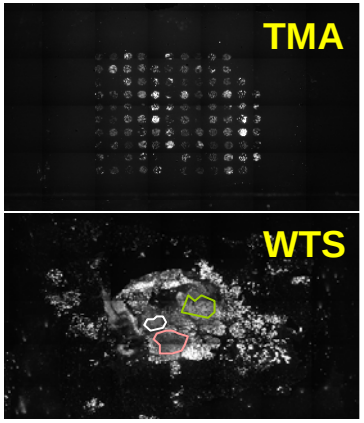
Step 4: **Convert to absolute concentration or normalize to set of uniform standards**

Generating the AQUA® score

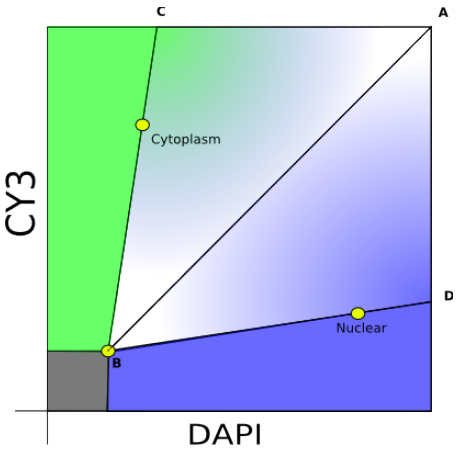
Cytokeratin



Tumor Mask

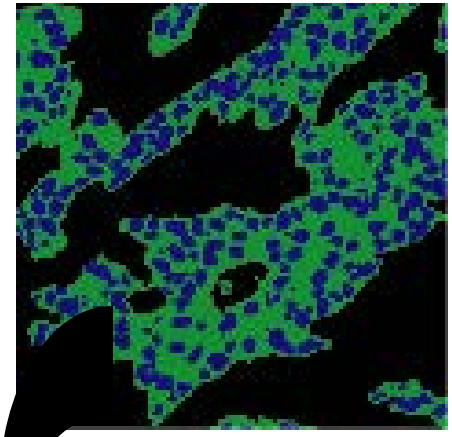


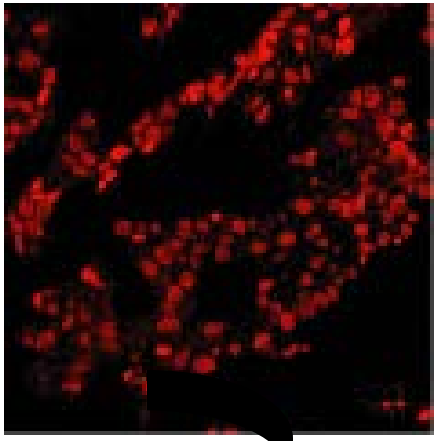
TMA-Tissue Microarray
WTS-Whole Tissue Section



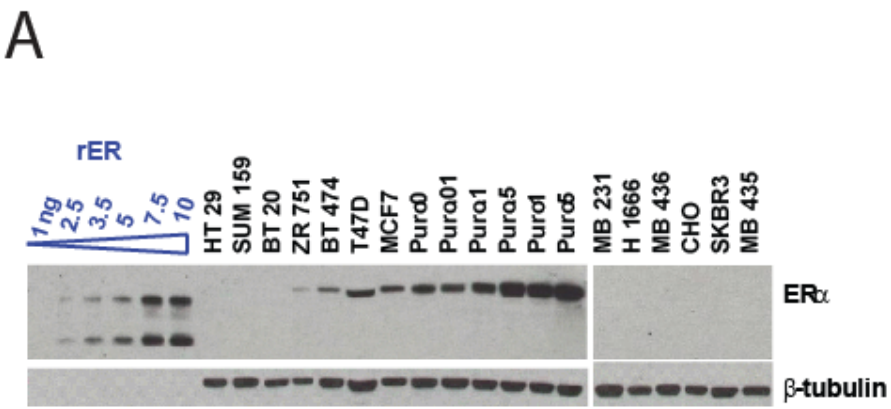
Combine DAPI image and cytokeratin image then cluster to assign each pixel to a subcellular compartment

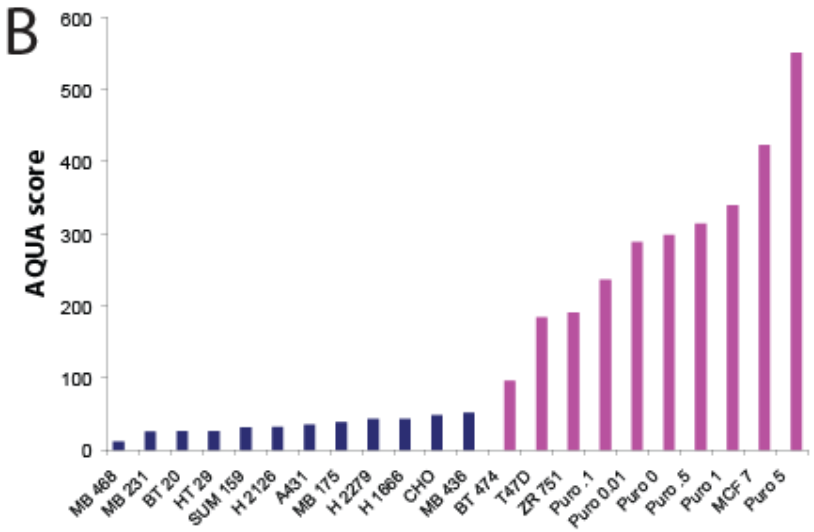
Estrogen Receptor

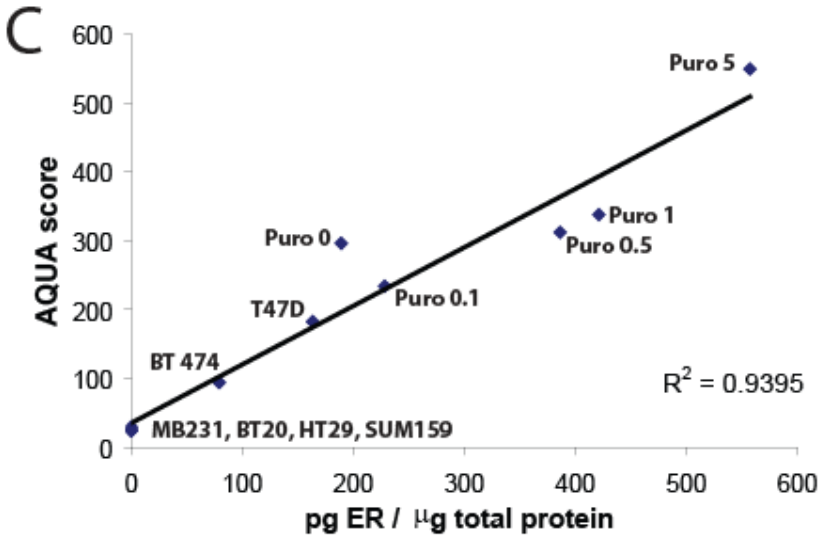


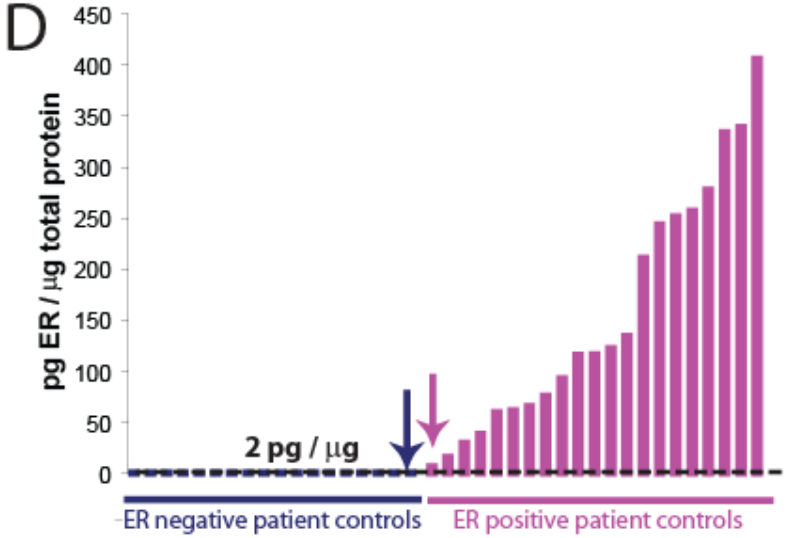


$$\frac{\Sigma \text{ target intensity in compartment pixels}}{\Sigma \text{ compartment pixel area}} = \text{AQUA score}$$







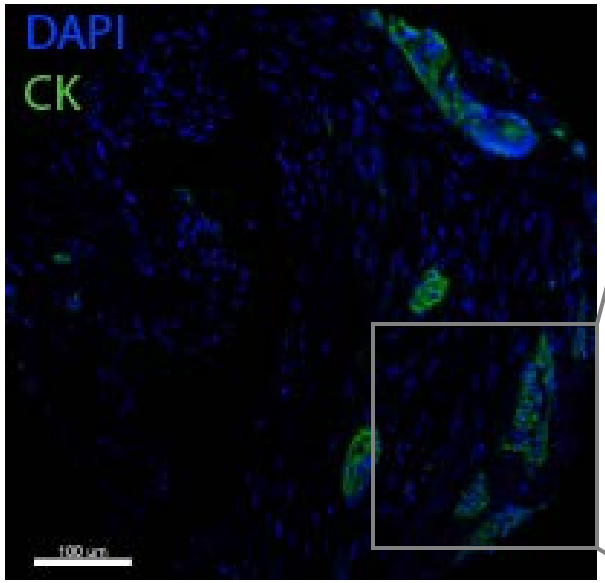


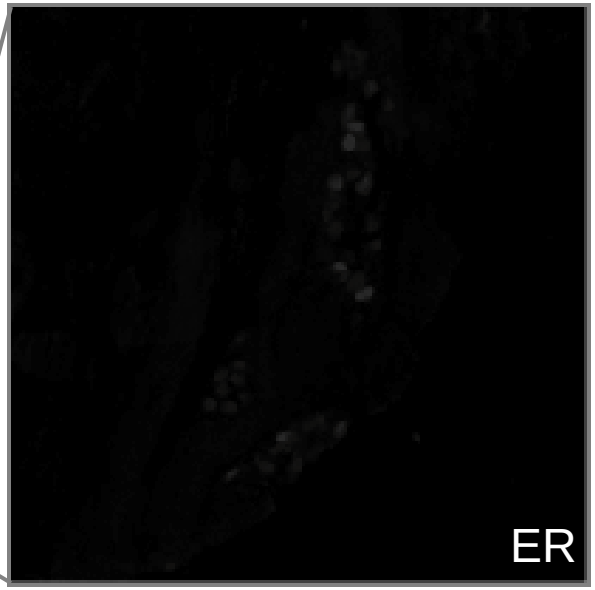
Standardized Index Array

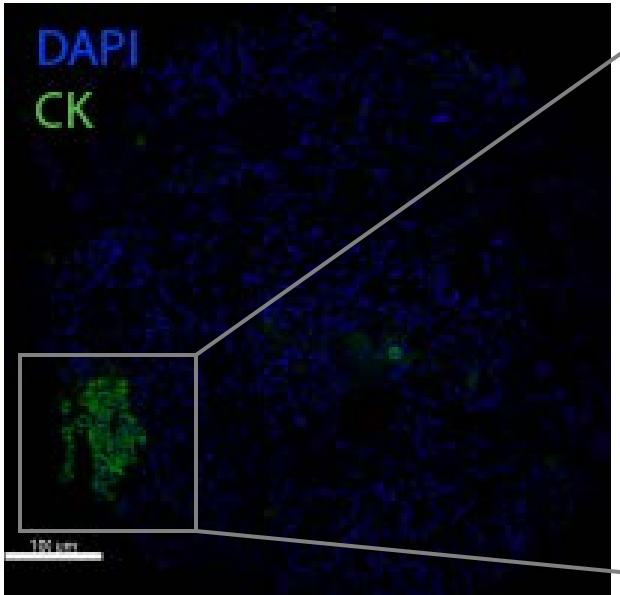
ER antibody used is 1D5

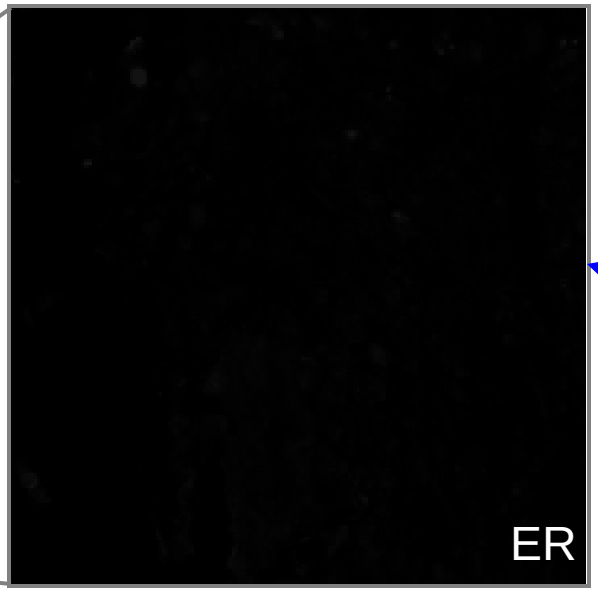
Alley Welsh

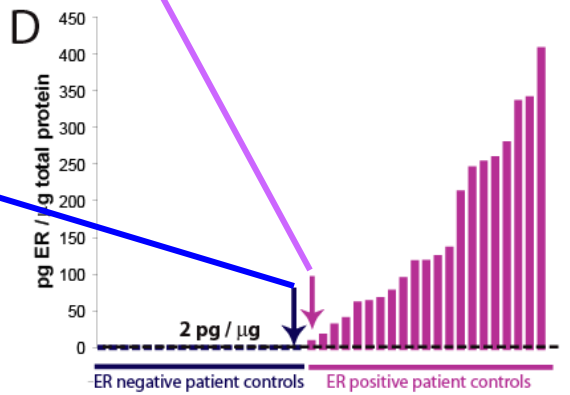
Lowest positive vs. highest negative







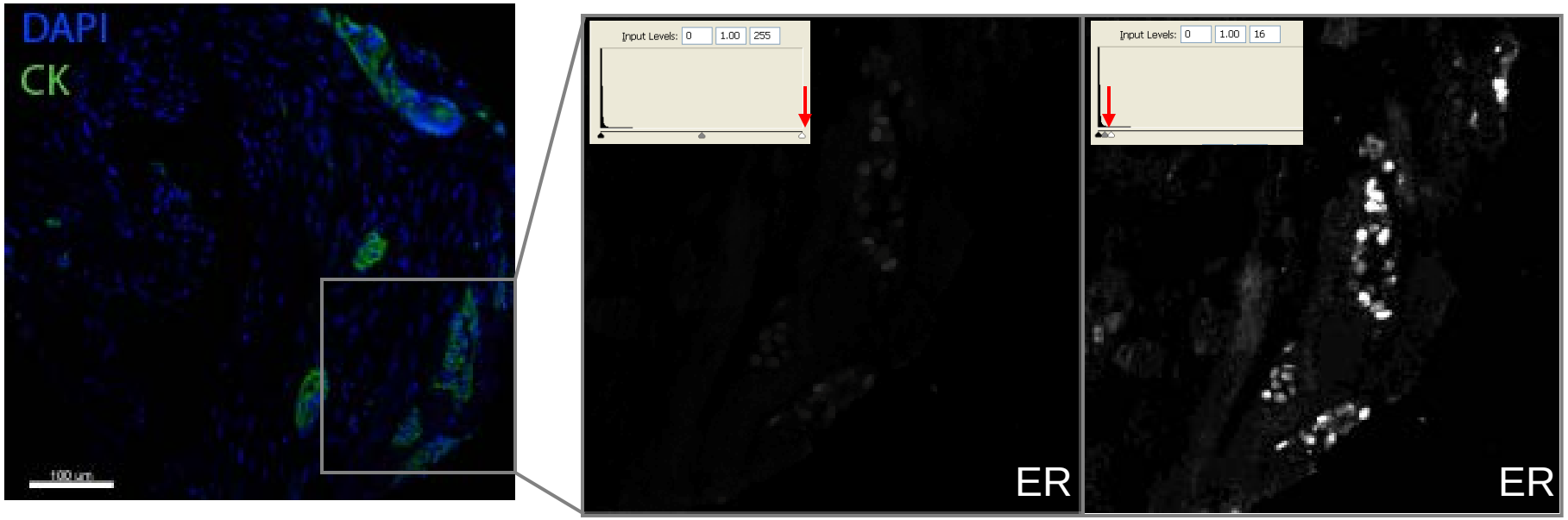


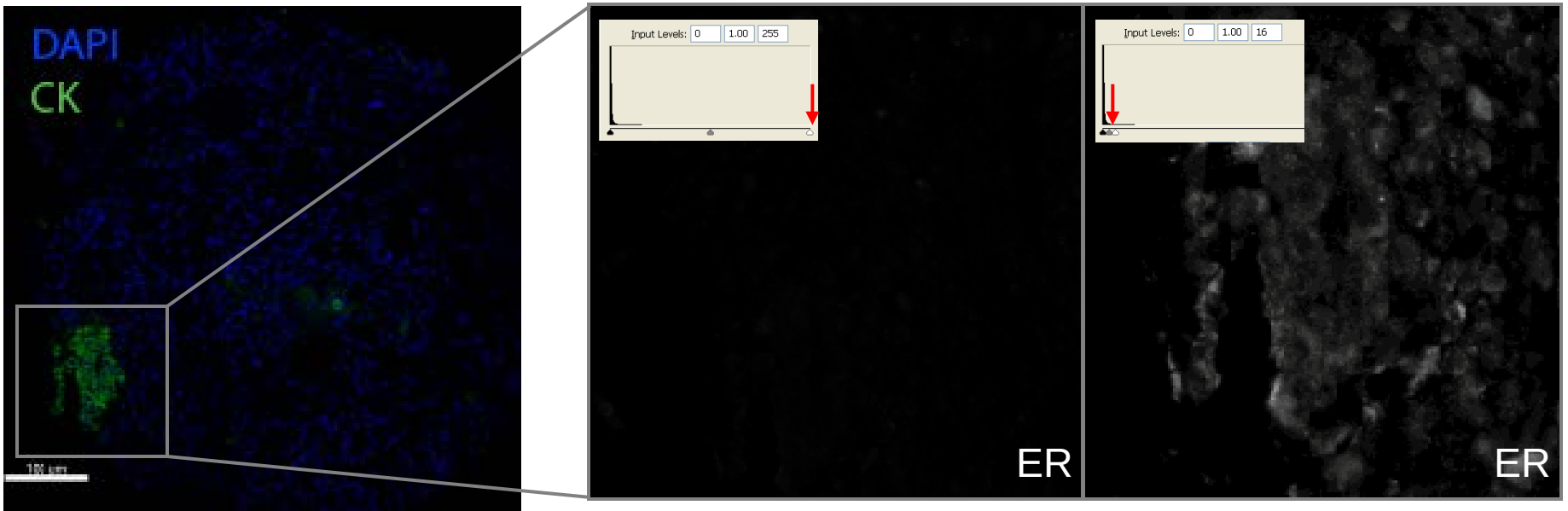


Alley Welsh

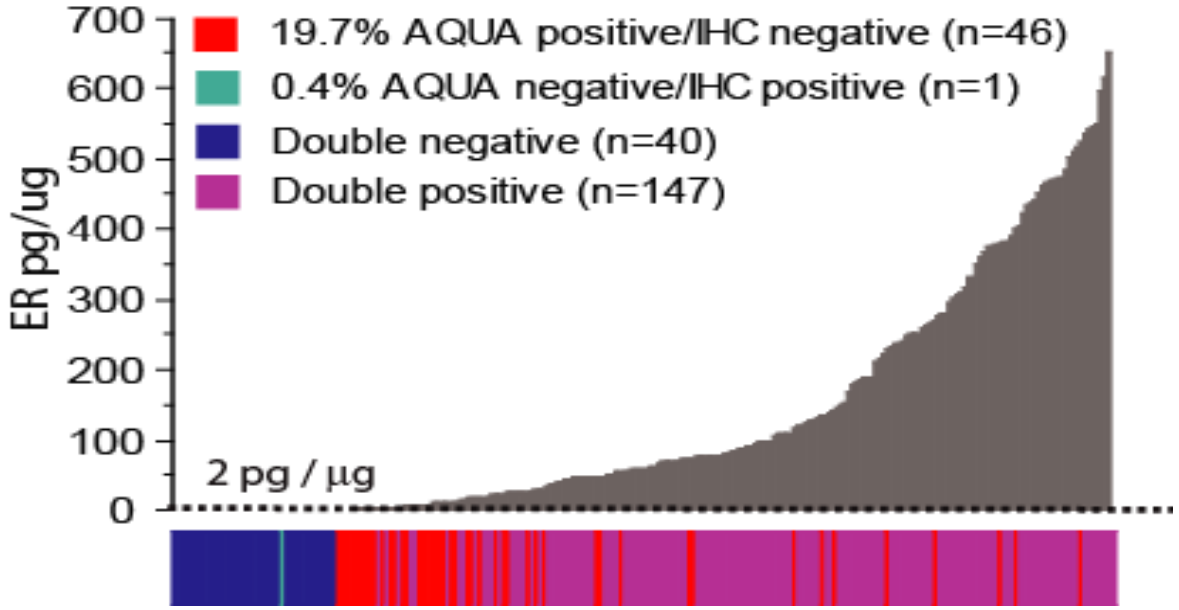
Expanded “levels” to visualize threshold

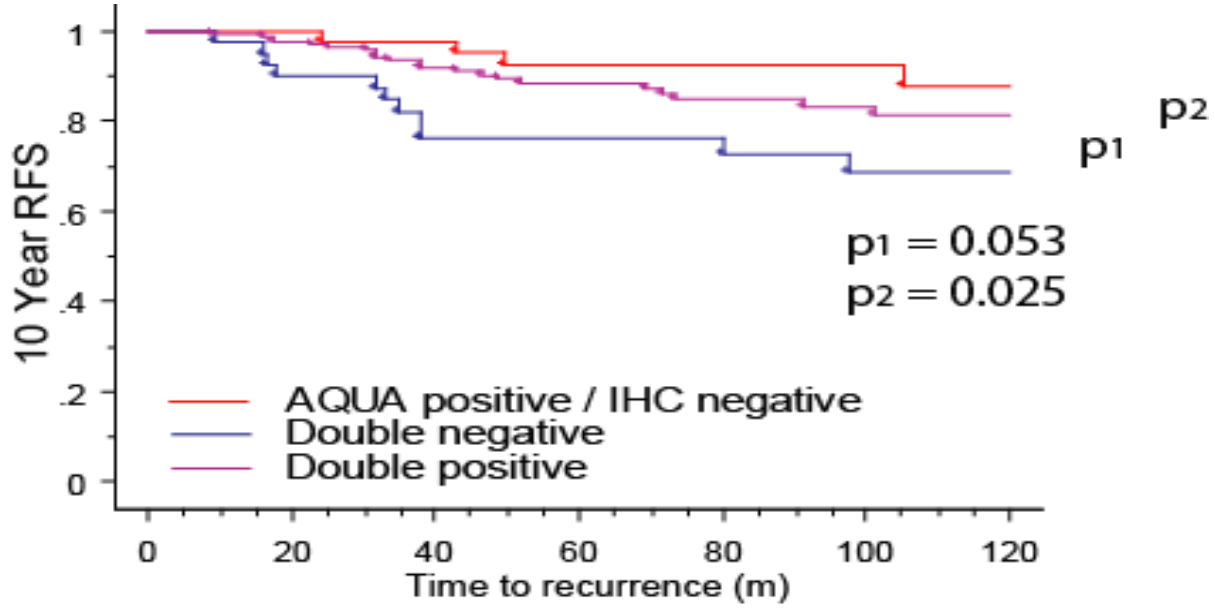
contracted dynamic range of grayscale (max RGB input level 255→16)





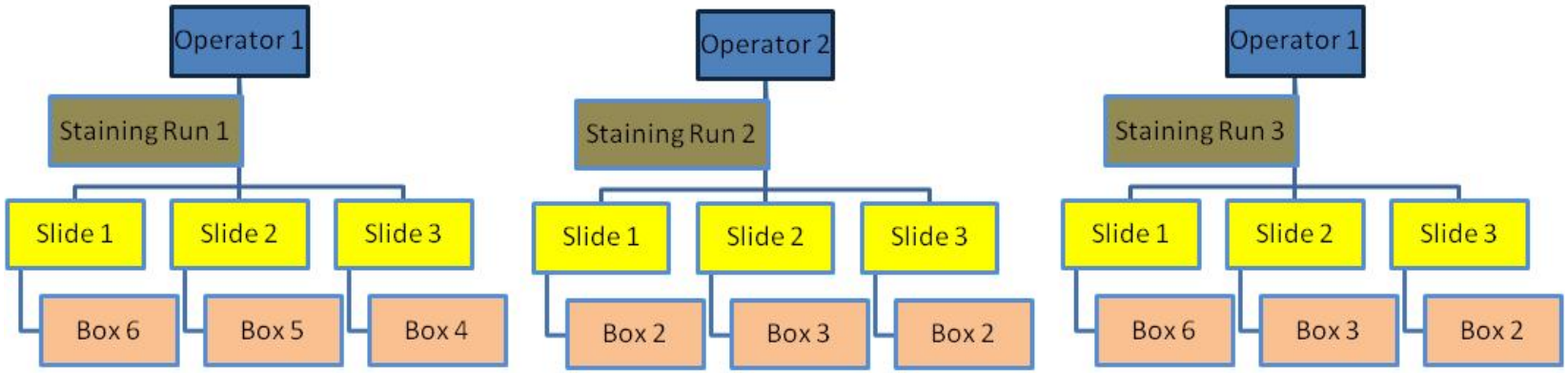
Discordant classification of ER status in YTMA 130 cohort





Welsh et al, JCO 2011

Precision Results (ER-alpha)

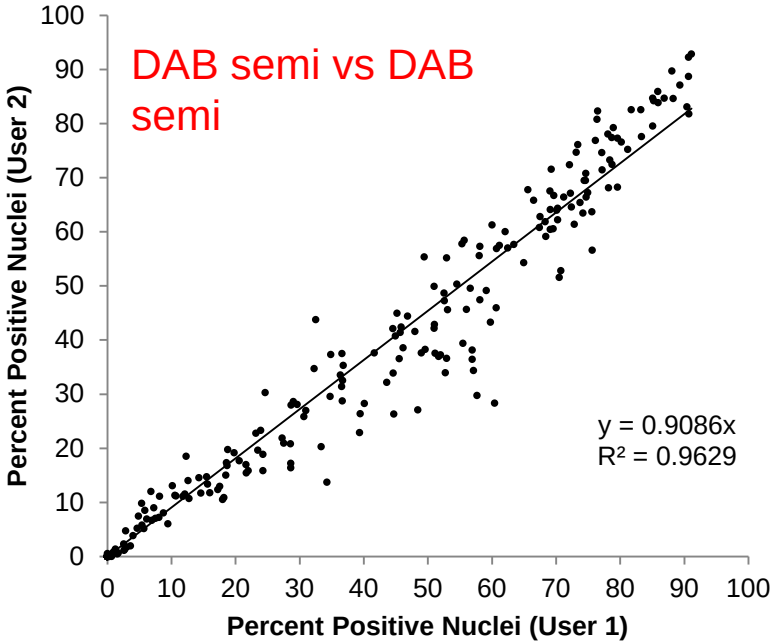


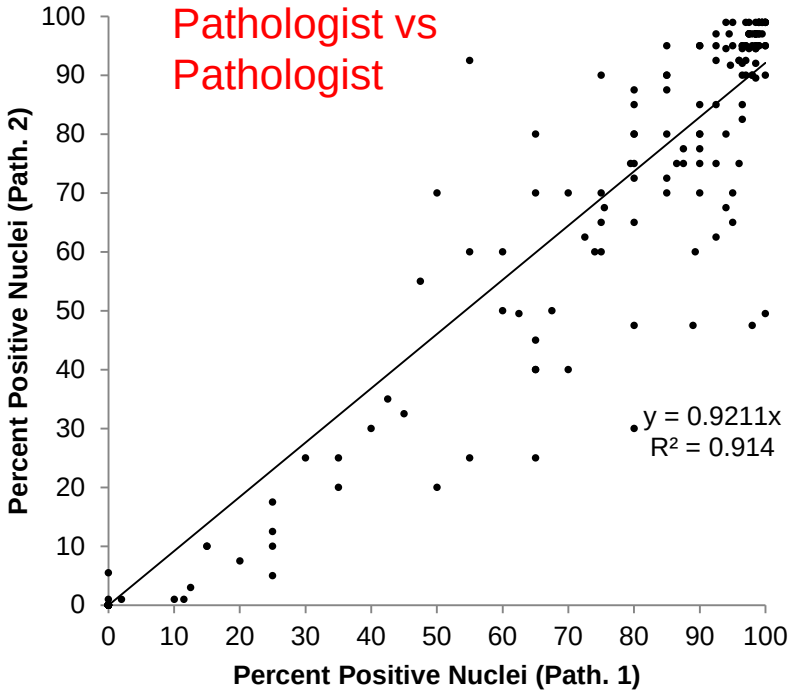
	Pearson R	Slope
Day 1 v. Day 2	**.97**	**.97**
Day 1 v. Day 3	**.97**	**1.01**
Day 2 v. Day 3	**.98**	**1.04**

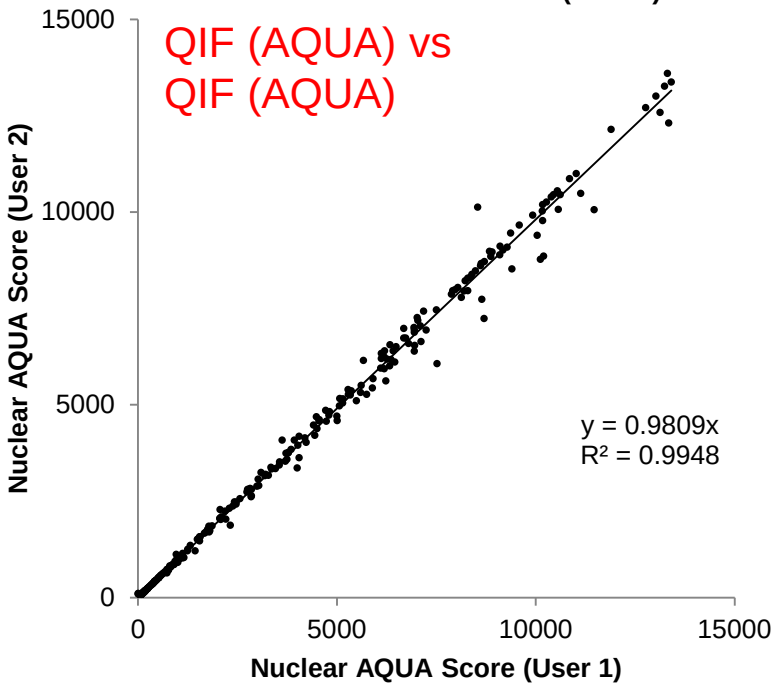
%CV = 4.2

Mark Gustavson and
Jason Christiansen

Comparison Between Methods (reproducibility)

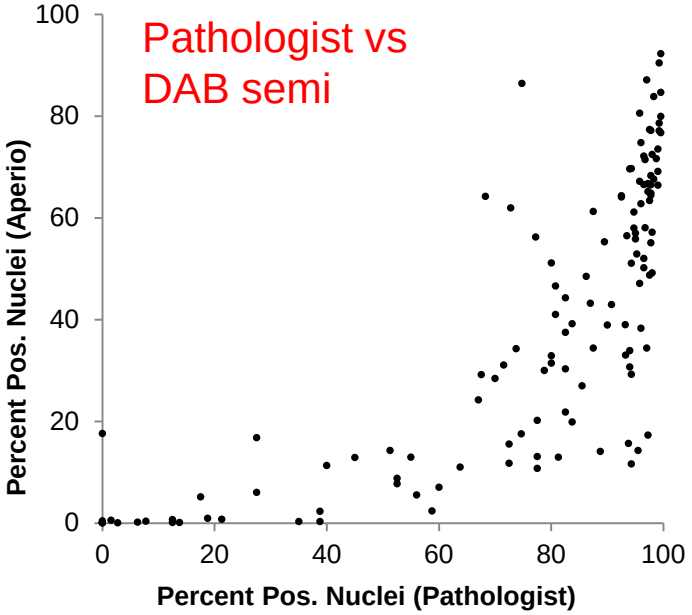


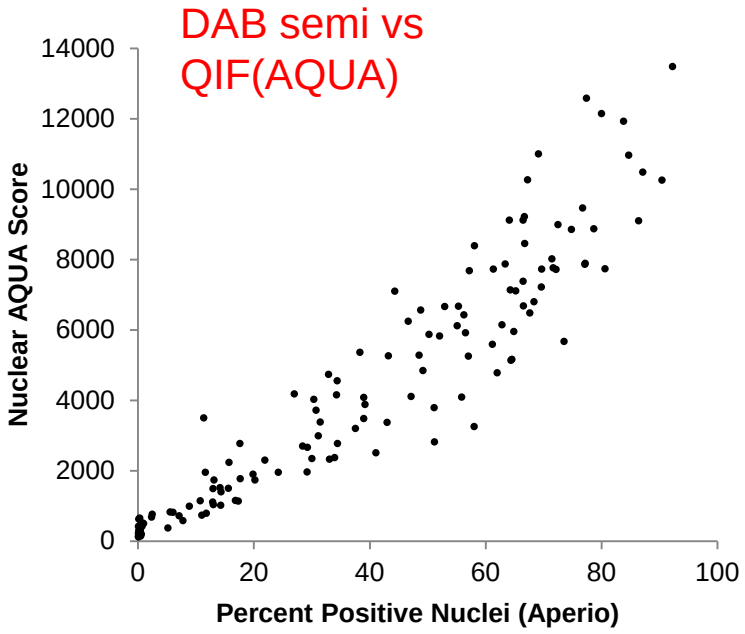


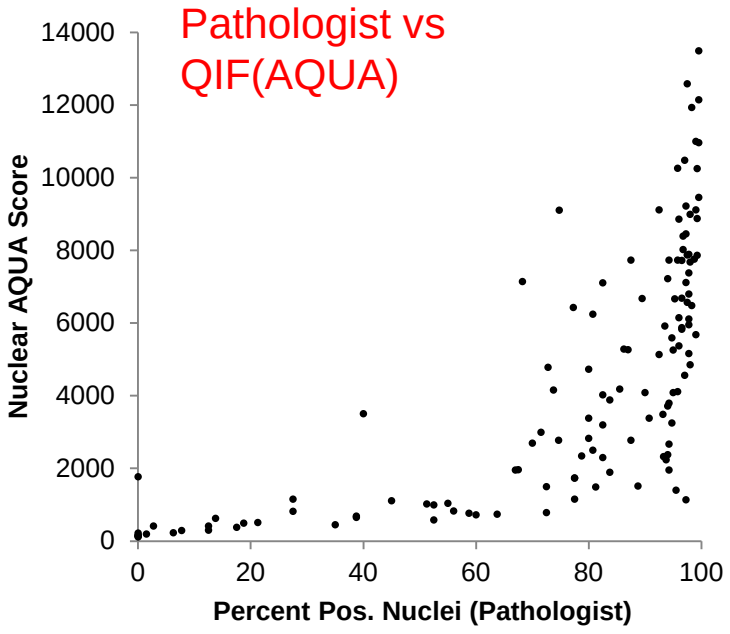


Elizabeth Zarrella and Veronique Neumeister

Comparison Between Methods



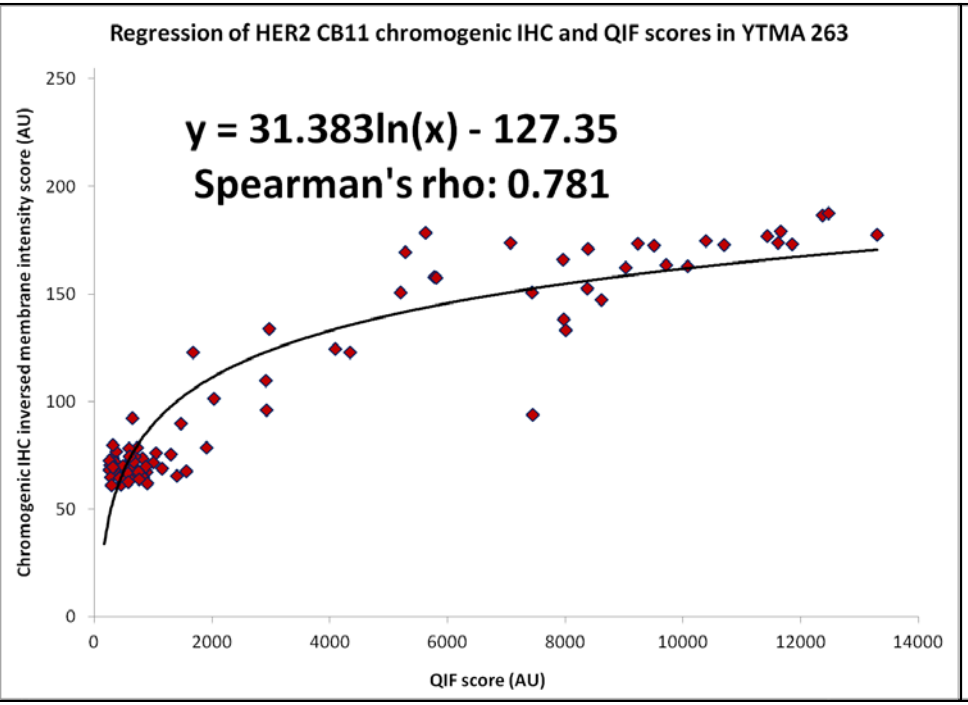




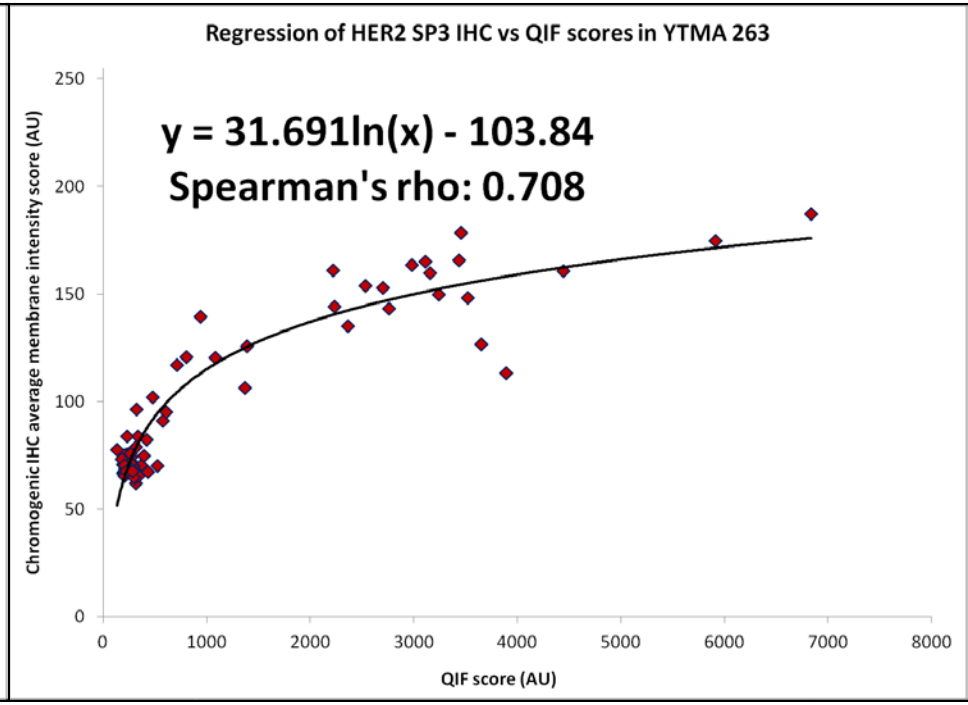
Elizabeth Zarrella and Veronique Neumeister

Regression of IHC vs QIF scores for CB11 and SP3 in YTMA 263

CB11



SP3



Daniel Carvajal

Outline for Presentation

- Using immunohistochemistry (IHC) or quantitative immunofluorescence (QIF) to measure protein on slides
- The challenge of a continuous biomarker – the HER2 example
- The challenge of defining the threshold of detection – The ER example

FDA Cleared Companion Dx antibodies:

Drug	Antigen	Company	Antibody
Trastuzumab	HER2	Ventana	4B5
Trastuzumab	HER2	Leica Biogenex	CB11
Trastuzumab	HER2	Dako	A0485
Endocrine Rx	Estrogen Receptor	Ventana	SP1
Endocrine Rx	Estrogen Receptor	Dako	1D5

What are we using for IHC?

Antibody	N	Percent
4B5	568	55%
A0485	54	5%
CB11	24	2%
Herceptest	296	29%
SP3	52	5%
other	37	4%
Total	1031	

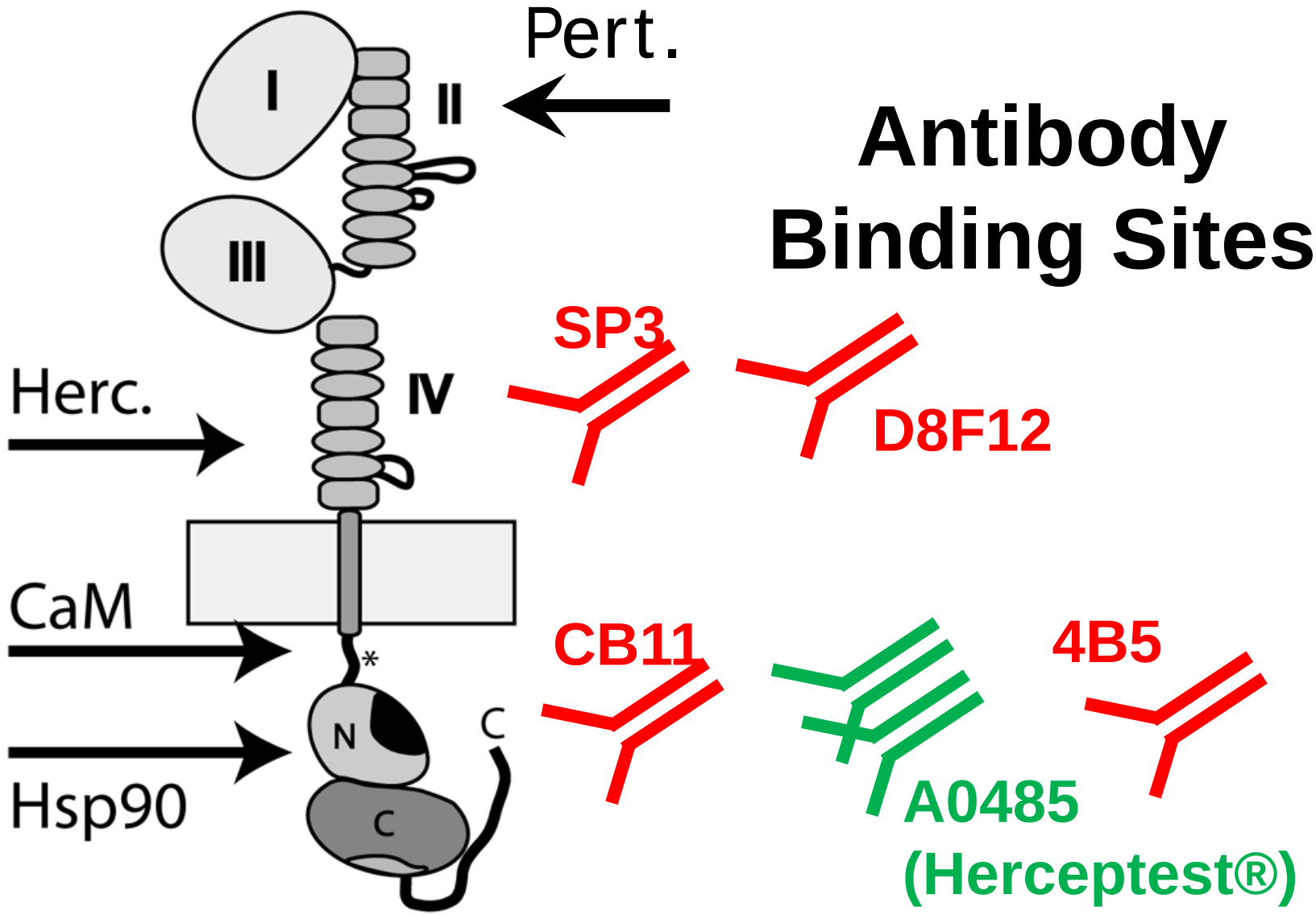
VOLUME 31 · NUMBER 31 · NOVEMBER 1 2013

JOURNAL OF CLINICAL ONCOLOGY

ASCO SPECIAL ARTICLE

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

Antonio C. Wolff, M. Elizabeth H. Hammond,* David G. Hicks,* Mitch Dowsett,* Lisa M. McShane,* Kimberly H. Allison, Donald C. Allred, John M.S. Bartlett, Michael Bilous, Patrick Fitzgibbons, Wedad Hanna, Robert B. Jenkins, Pamela B. Mangu, Soonmyung Paik, Edith A. Perez, Michael F. Press, Patricia A. Spears, Gail H. Vance, Giuseppe Viale, and Daniel F. Hayes**


Pert.
I
II
III
IV
Herc.
CaM
Hsp90
N
C
C
*
Antibody
Binding Sites
SP3
D8F12
CB11
4B5
A0485
(Herceptest®)

Dekker *et al. Breast Cancer Research* 2012, **14**:R93
http://breast-cancer-research.com/content/14/3/R93

Breast Cancer RESEARCH

RESEARCH ARTICLE Open Access

Determining sensitivity and specificity of HER2 testing in breast cancer using a tissue micro-array approach

Tim JA Dekker[1,2], Susan Ter Borg[3], Gerrit KJ Hooijer[3], Sybren L Meijer[3], Jelle Wesseling[4], James E Boers[5], Ed Schuuring[6], Jos Bart[6], Joost van Gorp[7], Wilma E Mesker[2], Judith R Kroep[1], Vincent THBM Smit[8] and Marc J van de Vijver[3*]

Studies done "by eye" cannot see discordance

Table 3 SP3, 4B5 and Herceptest.

		SP3					4B5					Herceptest		
		0,1+	2+	3+			0,1+	2+	3+			0,1+	2+	3+
4B5	0,1+	911	15	0	4B5	0,1+	920	7	0	SP3	0,1+	930	9	0
	2+	30	26	1		2+	42	15	0		2+	27	18	2
	3+	1	7	72		3+	7	8	68		3+	6	3	66
Total		942	48	73	Total		969	30	68	Total		963	30	68

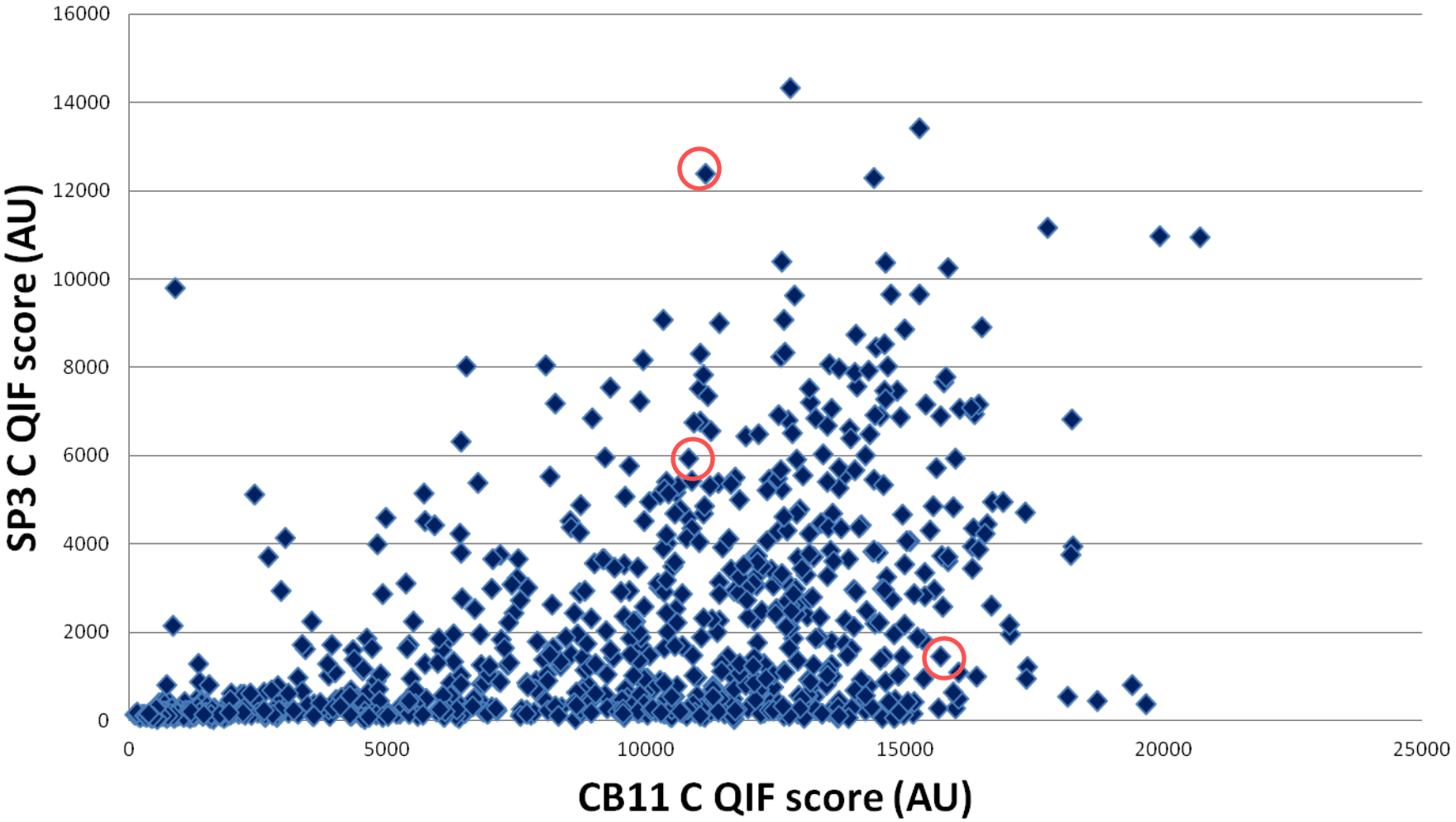
Table 4 HER2 status determined on TMA (SP3, 4B5 and mono color SISH)

Total complete results	1,020 (84.3%)
Negative	932 (91.4%)
Positive	80 (7.8%)
Discordant results	8 (0.8%)
Incomplete results	190 (15.7%)
Total number of cases	1210 (100%)

HER2, human epidermal growth factor receptor 2; SISH, silver *in situ* hybridization; TMA, tissue micro-array.

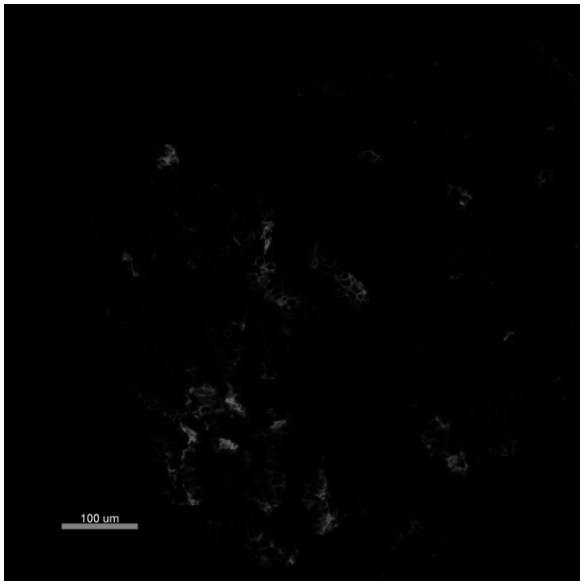
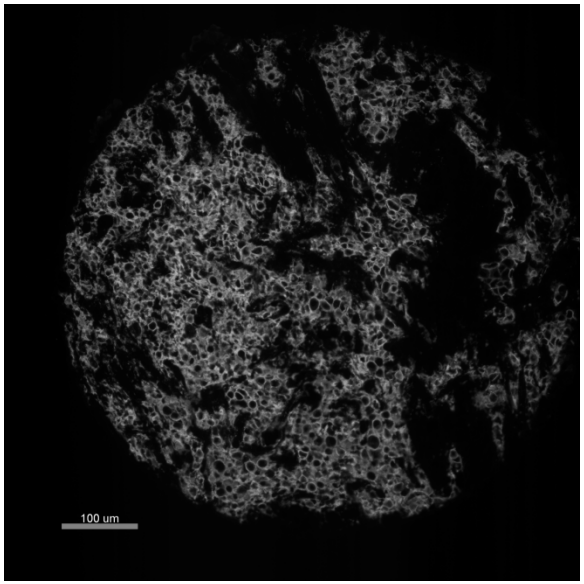
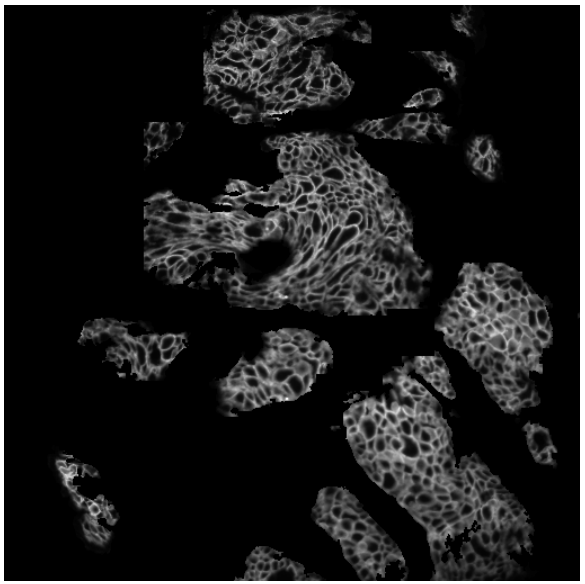
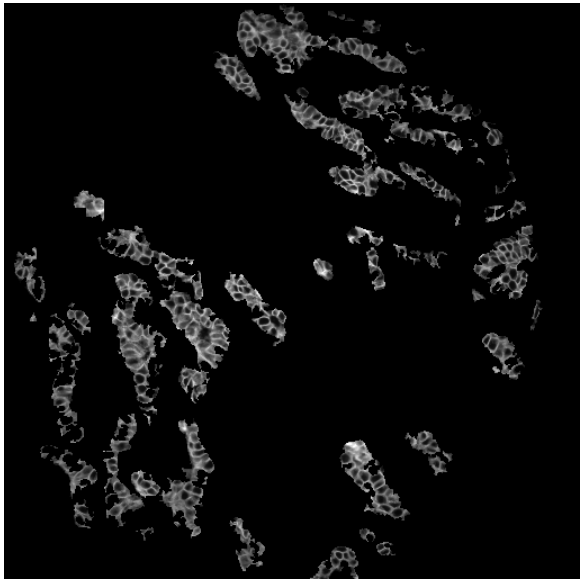
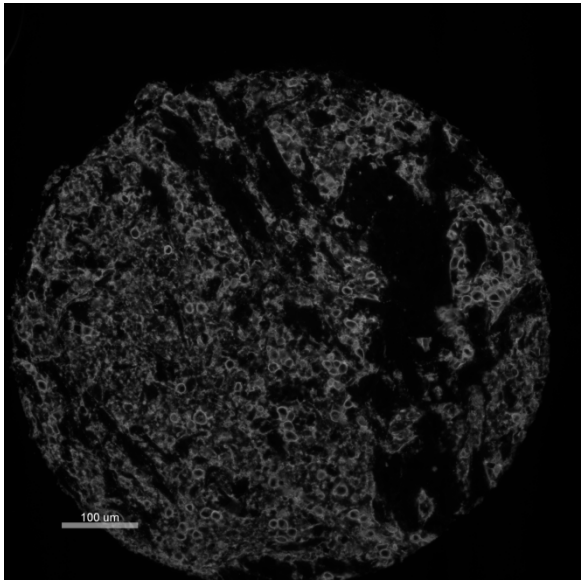
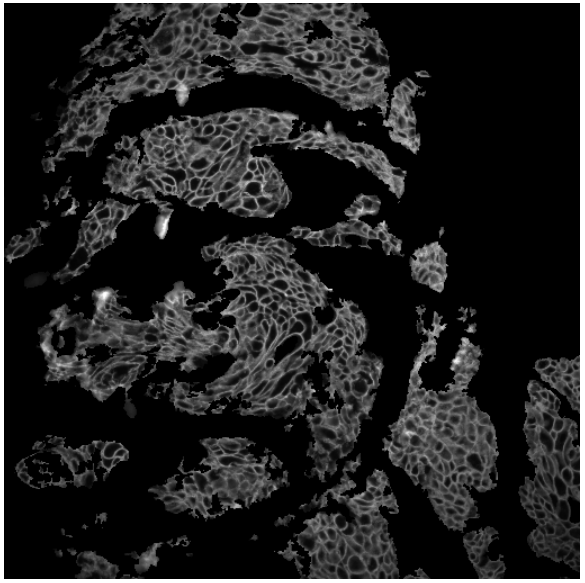
Less than 1% discordance between cytoplasmic and extracellular domain antibodies

SP3 C vs CB11 C (NCCTG 9831 Arm C)



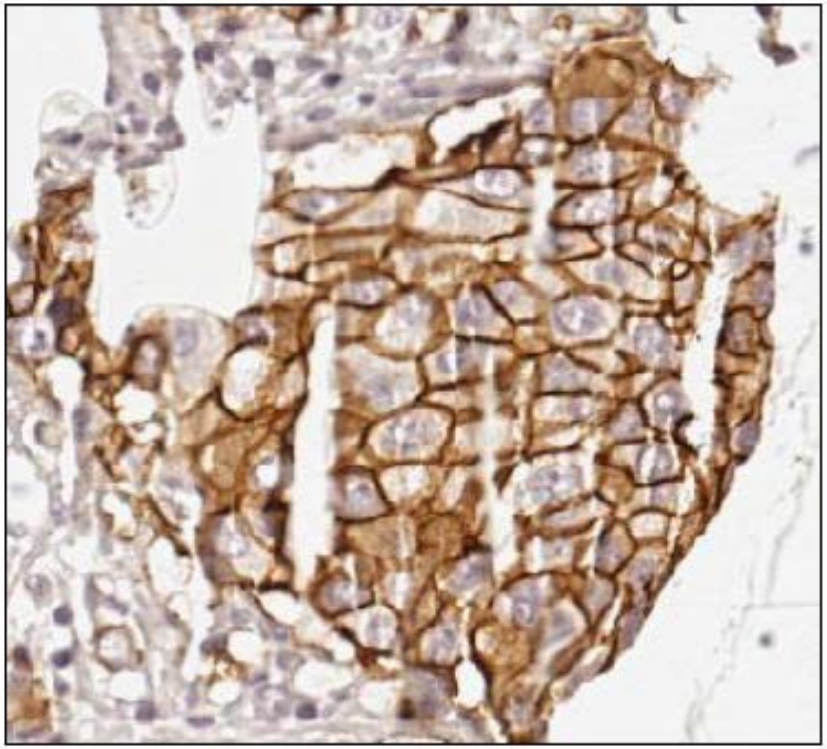
Data from NCCTG9831 from ongoing collaboration with Edith Perez and Karla Ballman

Images

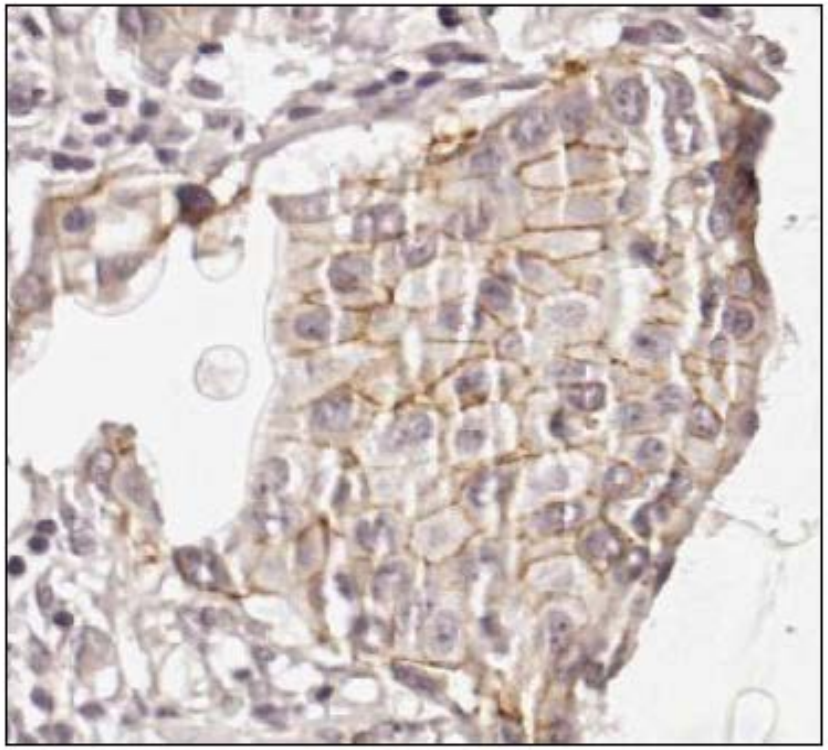
	21	262	78
SP3			
CB11			

DAB shows same effect, but more subtle

CB11 (ICD)

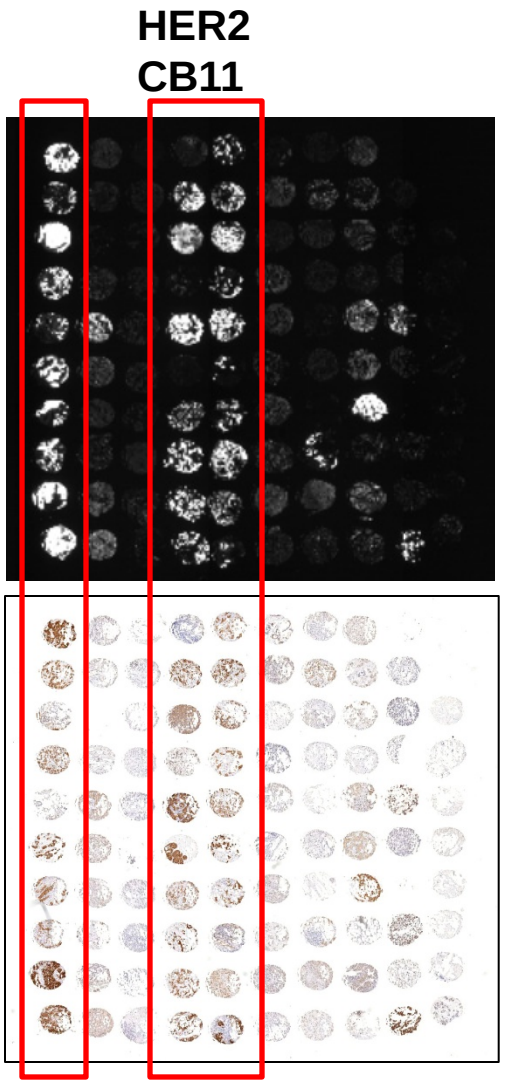


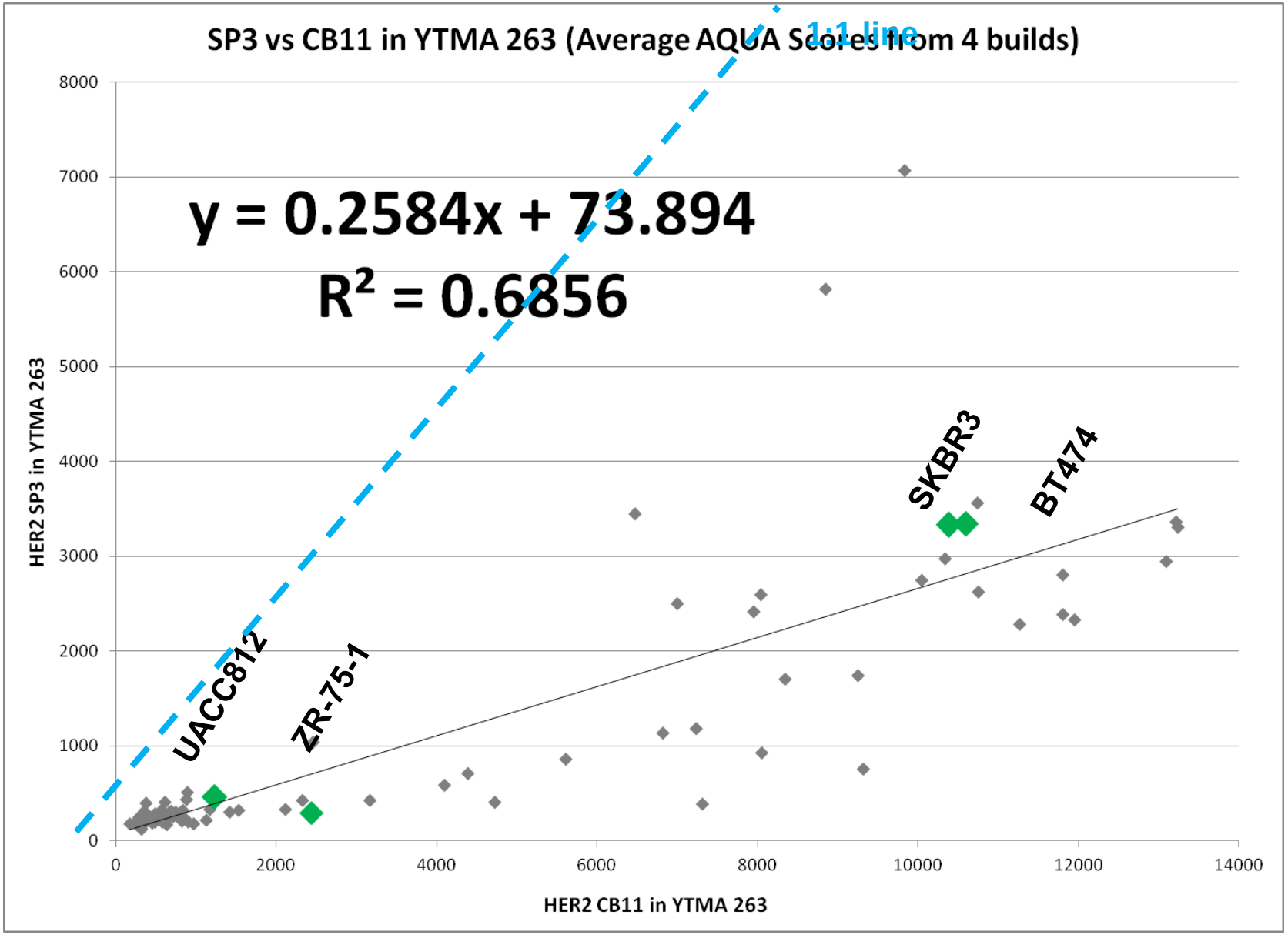
SP3 (ECD)



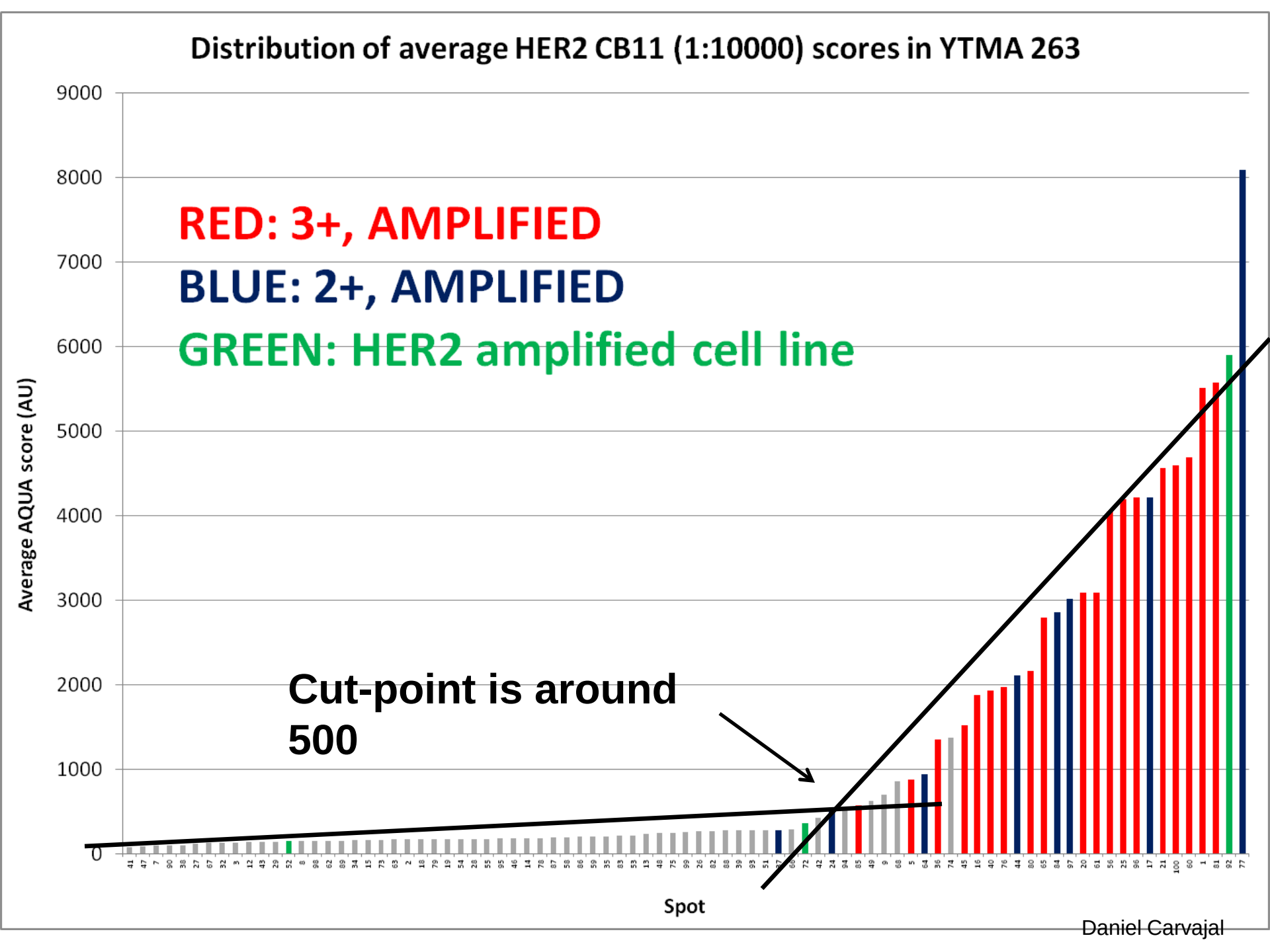
HER2 Standardization Array (YTMA 263)

3+, AMP	1+, NOT AMP	0, NOT AMP	2+, AMP	3+, AMP	1+, NOT AMP	0, NOT AMP	2+, NOT AMP	CELL LINES	NORMAL BREAST TISSUE (NBT)
1	2	3	4	5	6	7	8	9	10
2008	2011	2011	2010	2005	2009	2004	2011	MB453	NBT
2007	2011	2011	2007	2004	2009	2004	2011	BT20	NBT
2011	2011	2011	2006	2004	2008	2009	2011	MCF7	NBT
2011	2011	2011	1998	2003	2007	2009	2011	MB468	NBT
2011	2011	2010	2004	2003	2007	2009	2011	ZR751	NBT
2010	2011	2010	2004	2002	2006	2007	2011	MB361	NBT
2009	2011	2007	2003	2002	2011	2007	2011	SKBR3	NBT
2009	2011	2007	2003	2002	2011	2011	2011	UACC812	NBT
2008	2010	2006	2011	2002	2011	2011	2007	MDA231	NBT
2008	2010	2006	2010	2002	2011	2011	2007	BT474	NBT

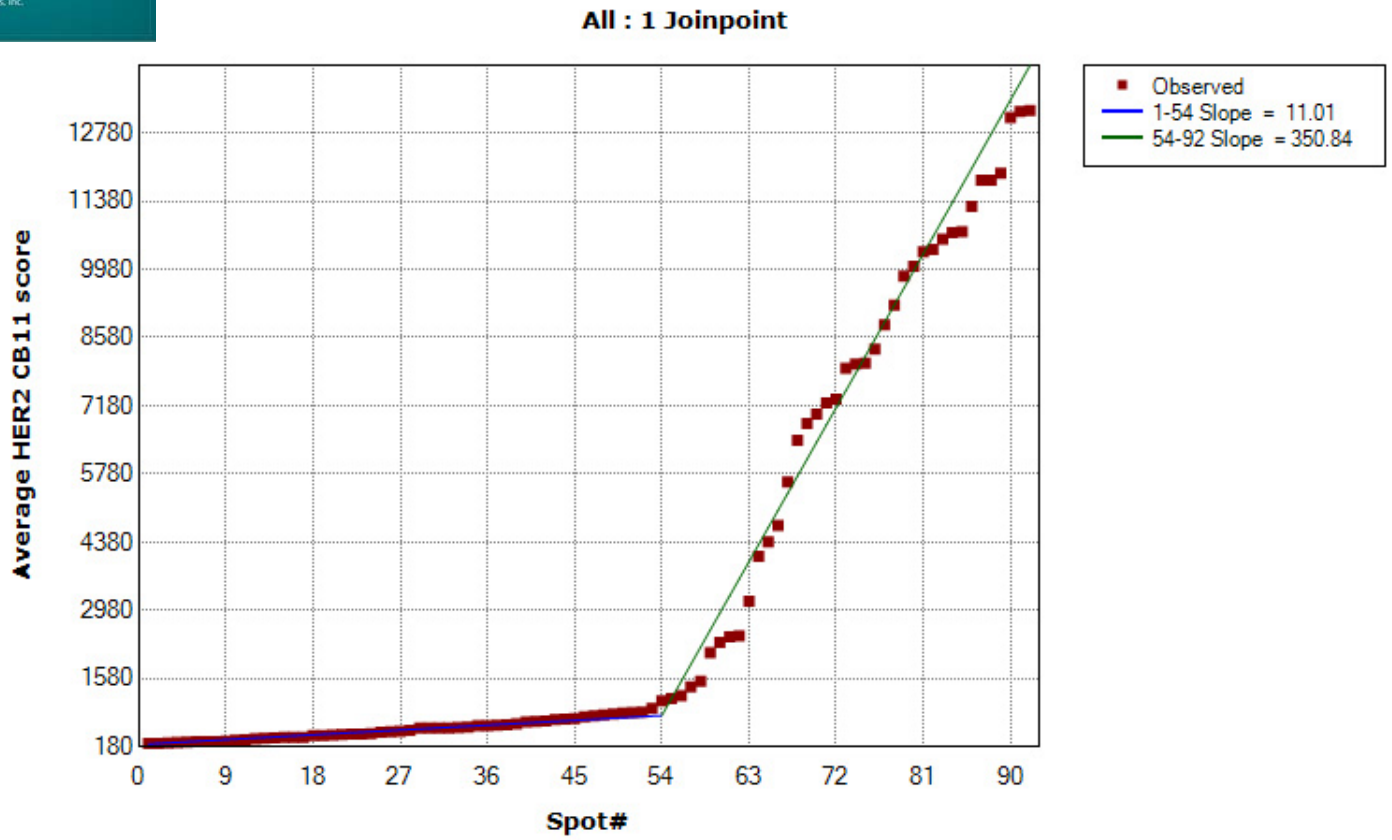




Daniel Carvajal


Distribution of average HER2 CB11 (1:10000) scores in YTMA 263
9000
8000
7000
6000
5000
4000
3000
2000
1000
0
Average AQUA score (AU)
RED: 3+, AMPLIFIED
BLUE: 2+, AMPLIFIED
GREEN: HER2 amplified cell line
Cut-point is around 500
Spot
Daniel Carvajal

HER2 CB11 analysis



Test For Number of Joinpoints

Test Number	Null Hypothesis	Alternate Hypothesis	Numerator Degrees of Freedom	Denominator Degrees of Freedom	Number of Permutations	P-Value	Significance Level~
#1	0 Joinpoint(s)	1 Joinpoint(s) *	2	88	4500	0.0002222	0.0500000

Final Selected Model 1 Joinpoint(s)

Daniel Carvajal

Using Joinpoint to Assess Ab Sensitivity and Specificity to Predict FISH amplification in YTMA 263

HER2 antibody	Joinpoint cut-point (AU)	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)
CB11	1299.3	94.12	85.48	78.05	96.36
A0485	1495.9	90.62	85.71	78.38	94.12
SP3	525.8	88.24	**98.41**	96.77	93.94
D8F12	767.6	90.91	**93.88**	90.91	93.88

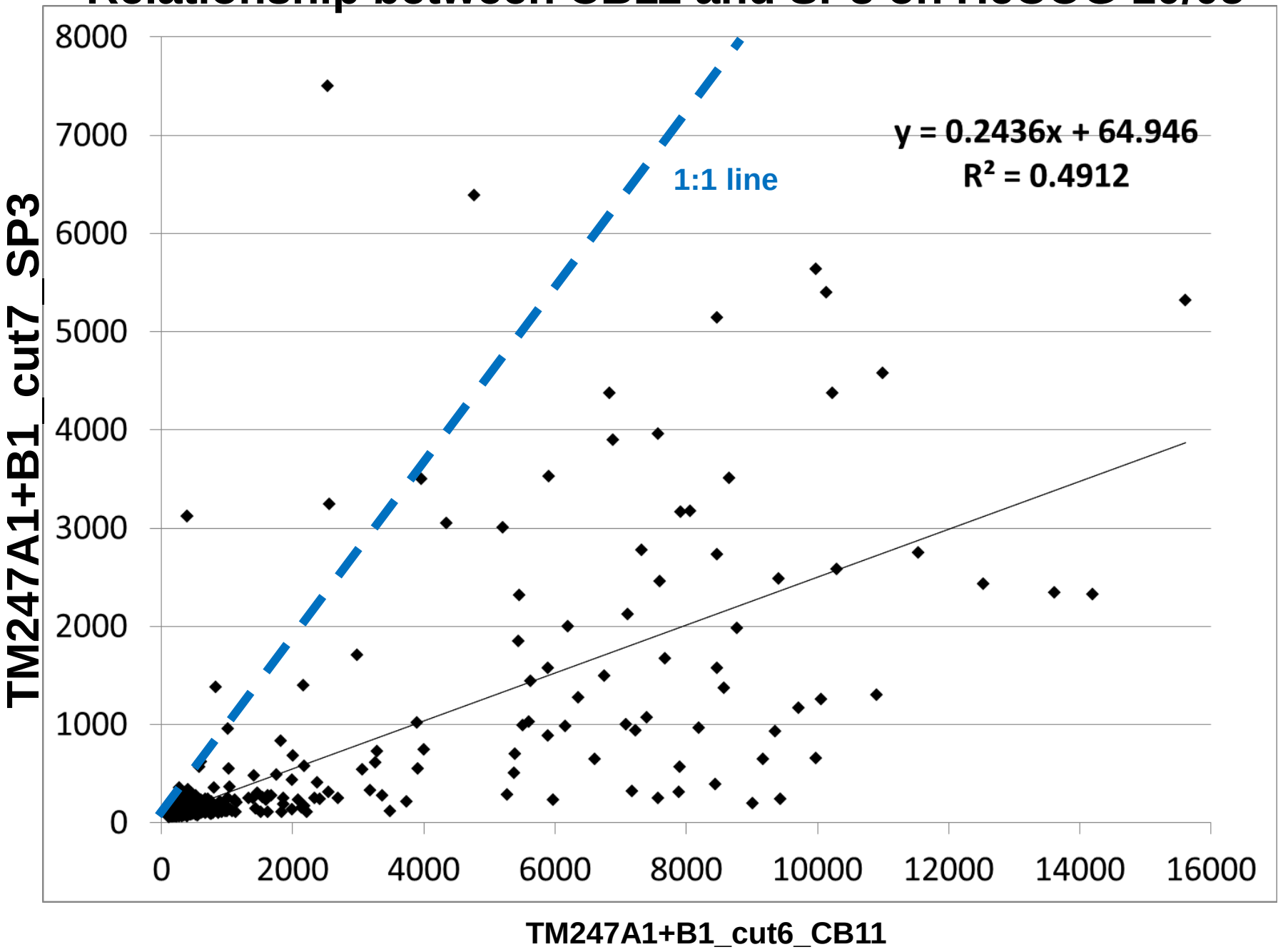
Daniel Carvajal

HeCOG 10/05 patient subgroups

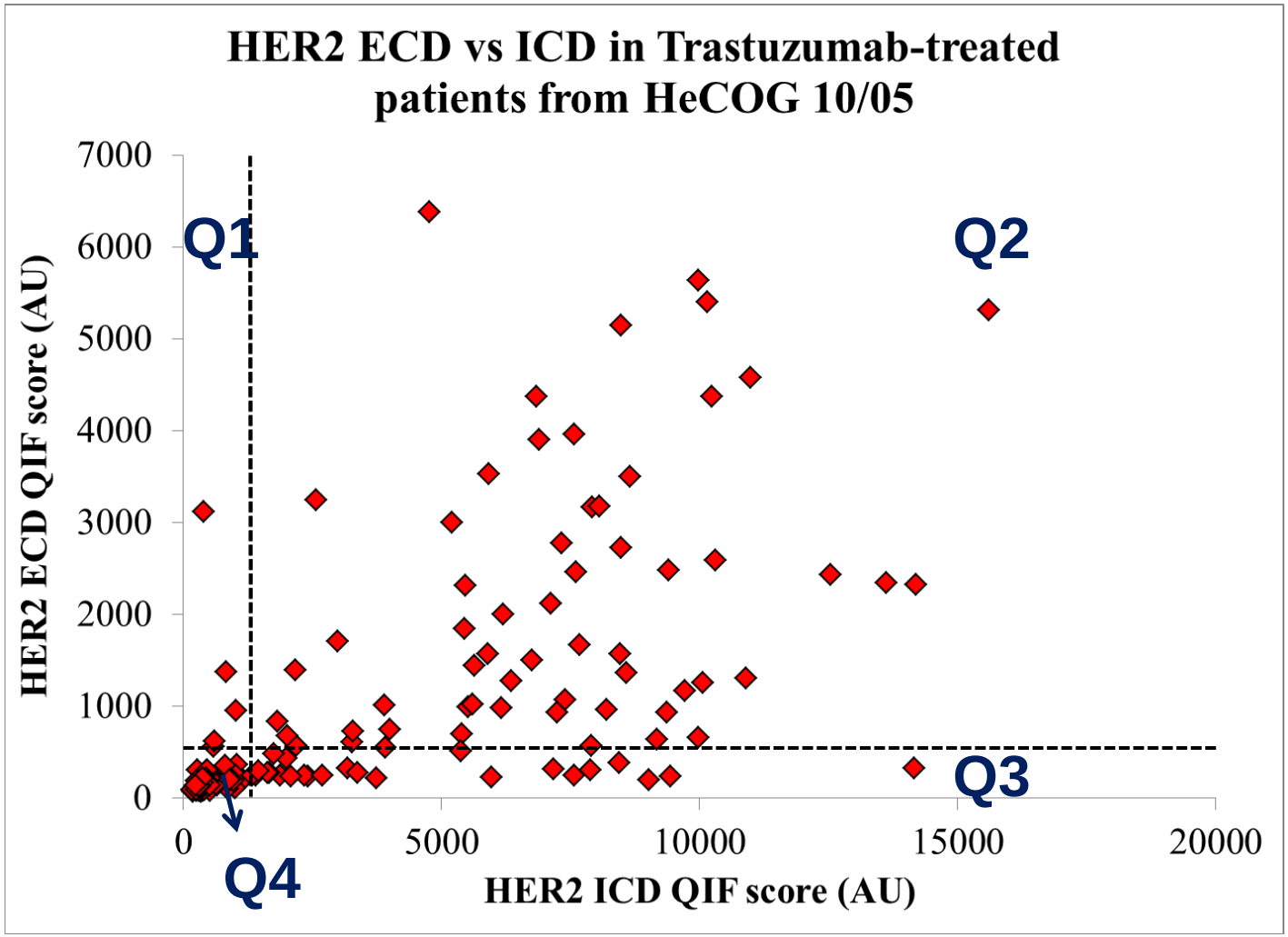
Table 2. Biological subgroups	
Subgroup	**N (%)**
ER(+)/PgR(+)/HER2(-)	366 (51.9)
ER(+)/PgR(+)/HER2(+)	90 (12.8)
ER(+)/PgR(-)/HER2(-)	66 (9.4)
ER(+)/PgR(-)/HER2(+)	23 (3.3)
ER(-)/PgR(-)/HER2(-) (TNBC)	68 (9.6)
ER(-)/PgR(-)/HER2(+)	69 (9.8)
ER(-)/PgR(+)/HER2(-)	15 (2.1)
ER(-)/PgR(+)/HER2(+)	7 (0.9)
Not determined	1 (0.1)

Table 3. Treatment subgroups	
Treatment group	**N (%)**
ER(+) with hormonotherapy	527 (96.7 of ER[+])
ER(+) without hormonotherapy	18 (3.3 of ER[+])
HER2(+) with trastuzumab	186 (98.4 of HER2[+])
HER2(+) without trastuzumab	3 (1.6 of HER2[+])
Hormonotherapy and trastuzumab	111 (15.7 of total)
ER(-) with hormonotherapy	0 (0)
HER2 (-) with trastuzumab	0 (0)

Relationship between CB11 and SP3 on HeCOG 10/05

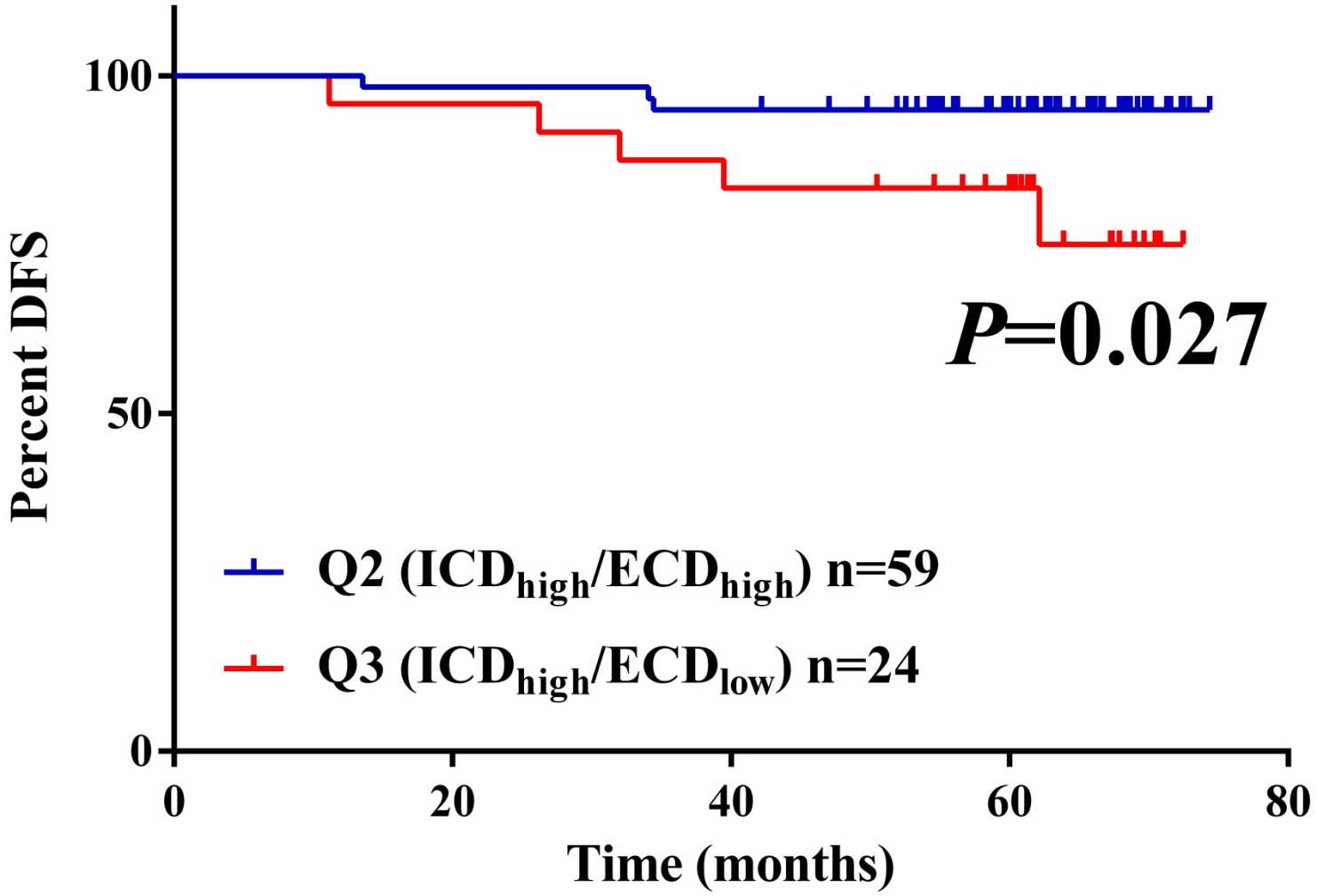


HER2 ICD/ECD protein expression by quadrants



- Cut-points obtained using YTMA 263 slides stained alongside HeCOG and analyzed using Joinpoint software.
- **Quadrants:**
  - **Q1 (ICD_{low}/ECD_{high}):** 5/159 patients **(3.1%)**.
  - **Q2 (ICD_{high}/ECD_{high}):** 59/159 patients **(37.1%)**.
  - **Q3 (ICD_{high}/ECD_{low}):** 24/159 patients **(15.1%)**.
  - **Q4 (ICD_{low}/ECD_{low}):** 71/159 patients **(44.7%)**.

Disease-free survival in HER2-positive, Trastuzumab-treated patients from HeCOG 10/05: Q2 vs Q3



Log rank *P*=0.027; HR=0.23 (95% C.I.: 0.037 to 0.815)

Outline for Presentation

- Using immunohistochemistry (IHC) or quantitative immunofluorescence (QIF) to measure protein on slides
- The challenge of a continuous biomarker – the HER2 example
- The challenge of defining the threshold of detection – The ER example

FDA Cleared Companion Dx antibodies:

Drug	Antigen	Company	Antibody
Trastuzumab	HER2	Ventana	4B5
Trastuzumab	HER2	Leica Biogenex	CB11
Trastuzumab	HER2	Dako	A0485
Endocrine Rx	Estrogen Receptor	Ventana	SP1
Endocrine Rx	Estrogen Receptor	Dako	1D5

Conflicting Papers

VOLUME 24 · NUMBER 36 · DECEMBER 20 2006

JOURNAL OF CLINICAL ONCOLOGY

ORIGINAL REPORT

Immunohistochemical Detection Using the New Rabbit Monoclonal Antibody SP1 of Estrogen Receptor in Breast Cancer Is Superior to Mouse Monoclonal Antibody 1D5 in Predicting Survival

Maggie C.U. Cheang, Diana O. Treaba, Caroline H. Speers, Ivo A. Olivotto, Chris D. Bajdik, Stephen K. Chia, Lynn C. Goldstein, Karen A. Gelmon, David Huntsman, C. Blake Gilks, Torsten O. Nielsen, and Allen M. Gown

From the Genetic Pathology Evaluation Centre at Vancouver General Hospital,

ABSTRACT

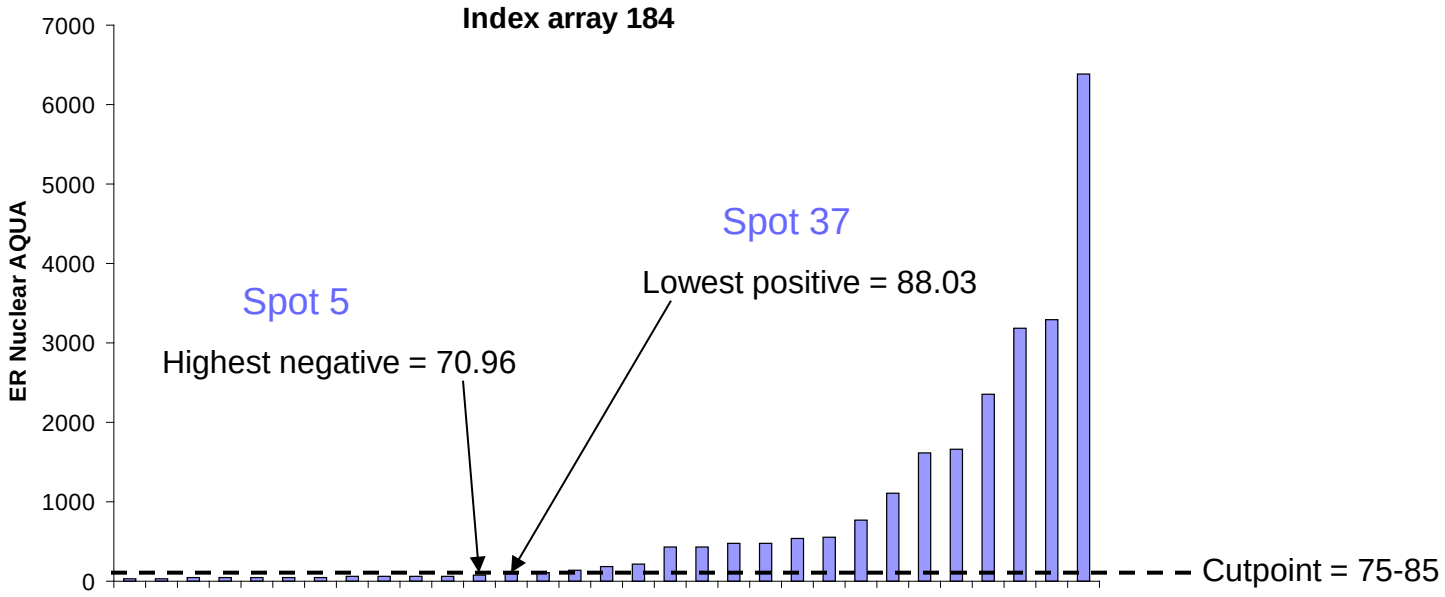
A Comparison of Estrogen Receptor SP1 and 1D5 Monoclonal Antibodies in Routine Clinical Use Reveals Similar Staining Results

Jane E. Brock, MBBS, PhD, Jason L. Hornick, MD, PhD, Andrea L. Richardson, MD, PhD, Deborah A. Dillon, MD, and Susan C. Lester, MD, PhD

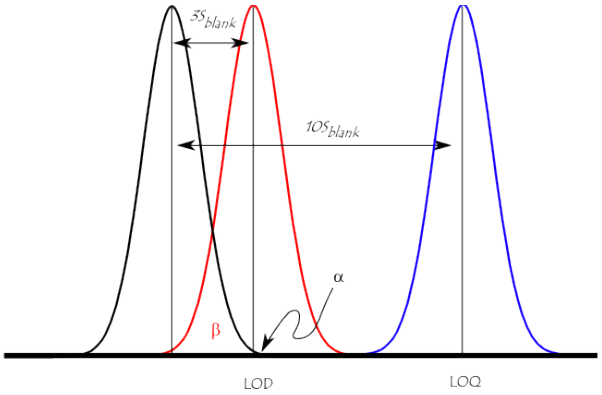
Key Words: Estrogen receptor; SP1; 1D5; Monoclonal antibodies; Breast carcinoma

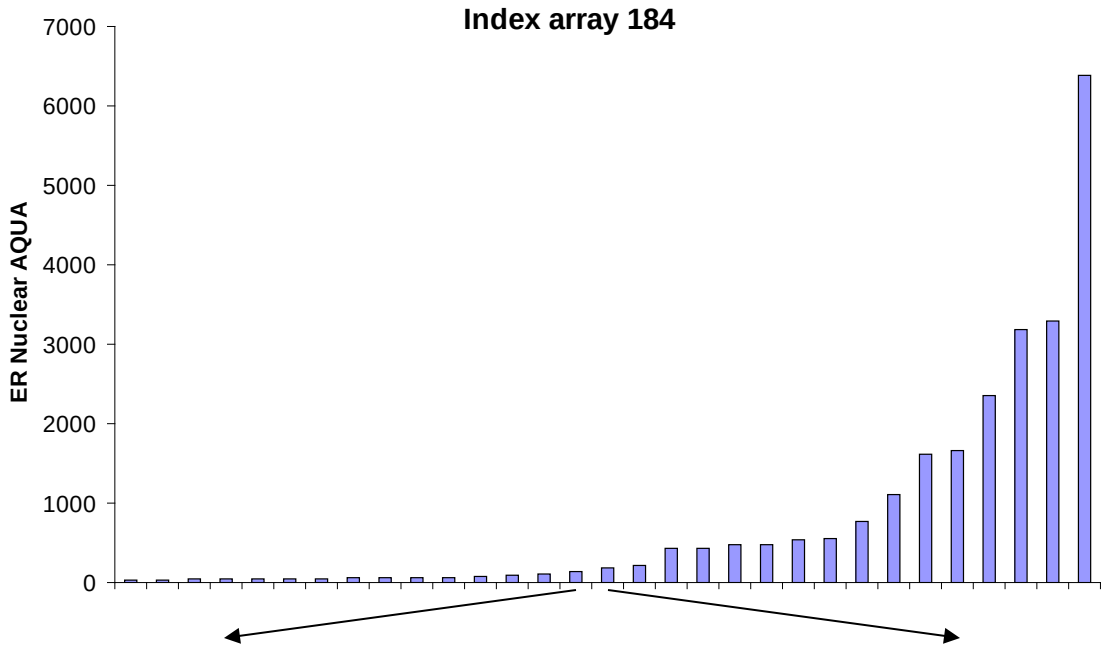
DOI: 10.1309/AJCPSKFWOLPPMEU9

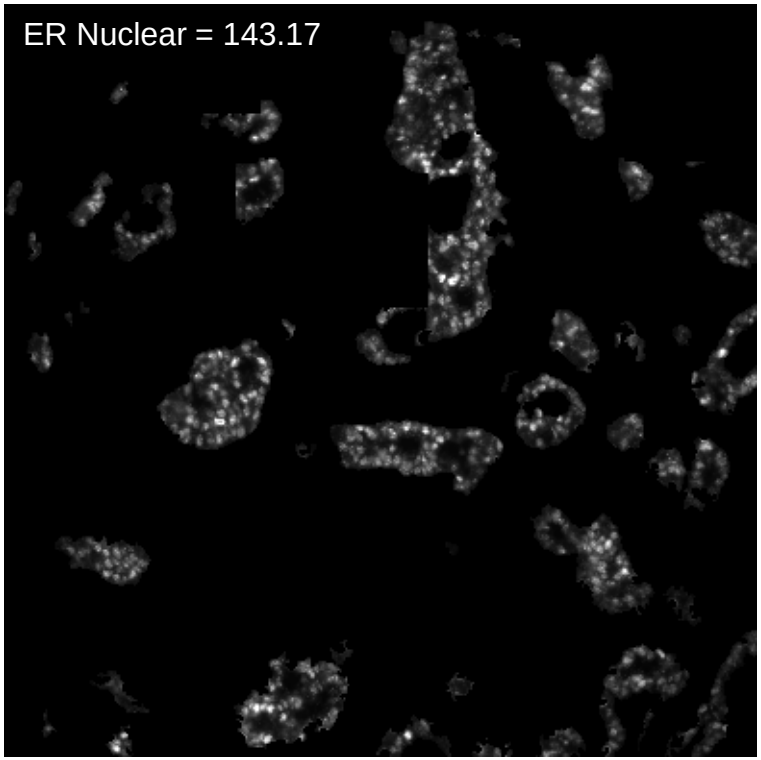
Using an index TMA to define the Threshold of Detection

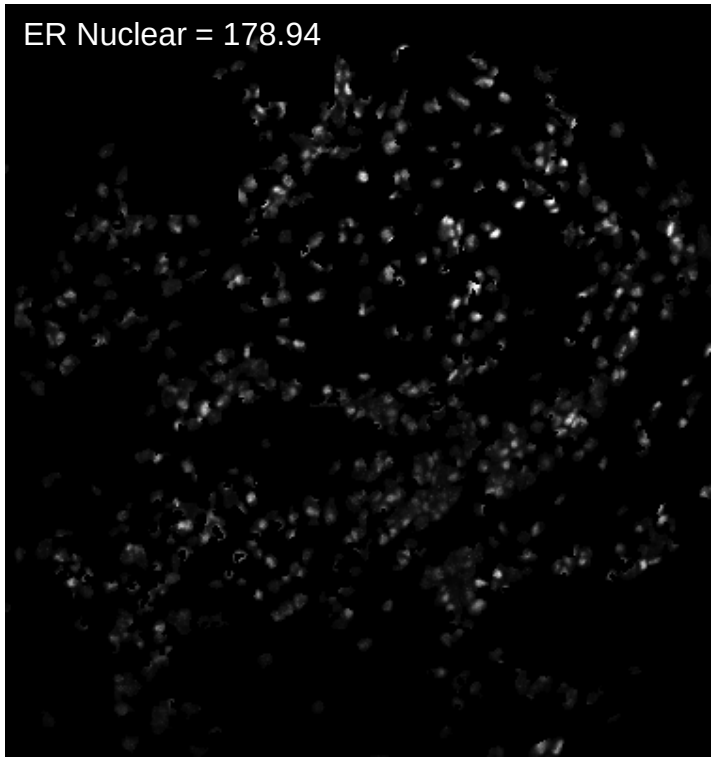


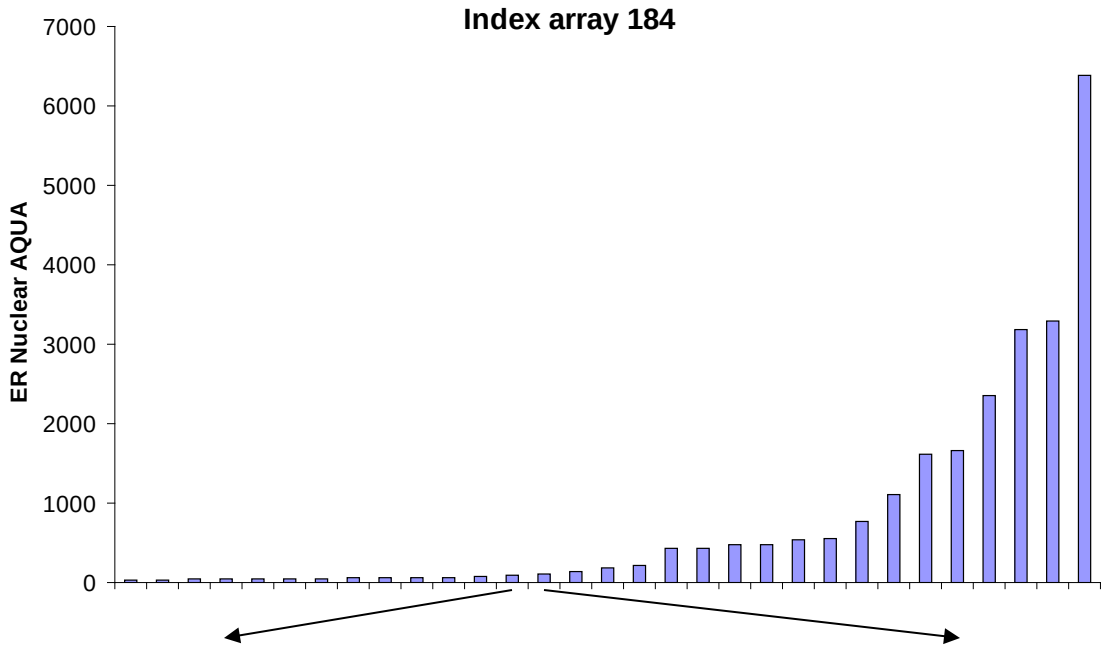
Should we be following the rules for limits of detection and quantification of analytical chemistry?

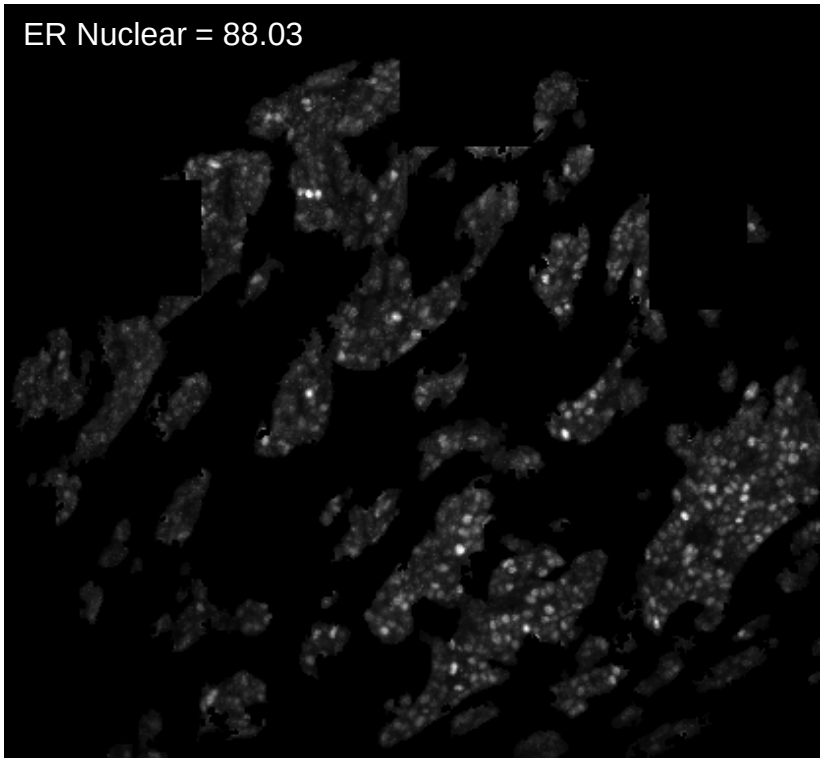


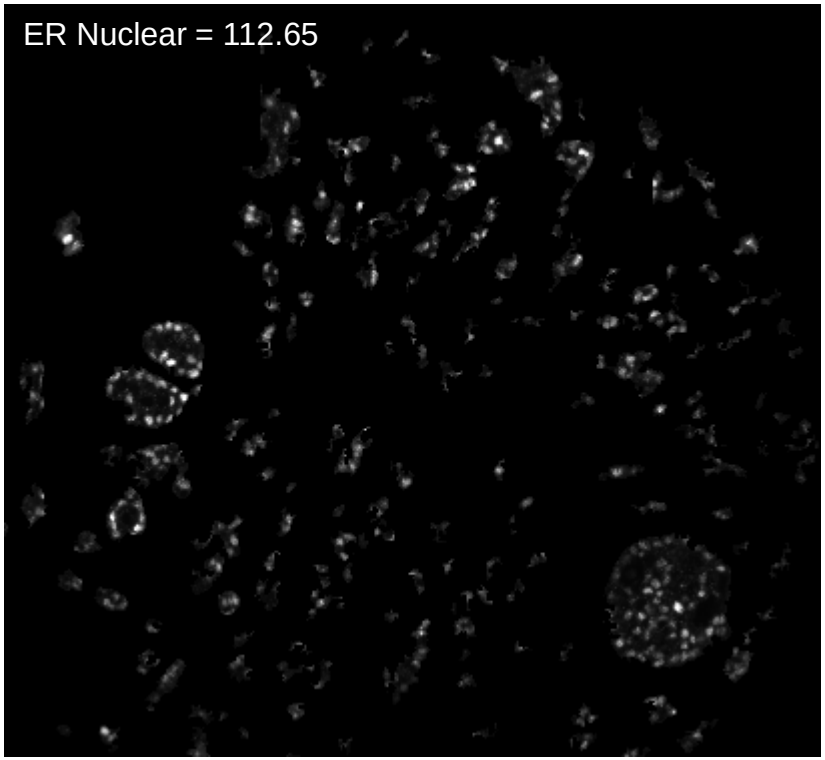

Index array 184
ER Nuclear AQUA
7000
6000
5000
4000
3000
2000
1000
0

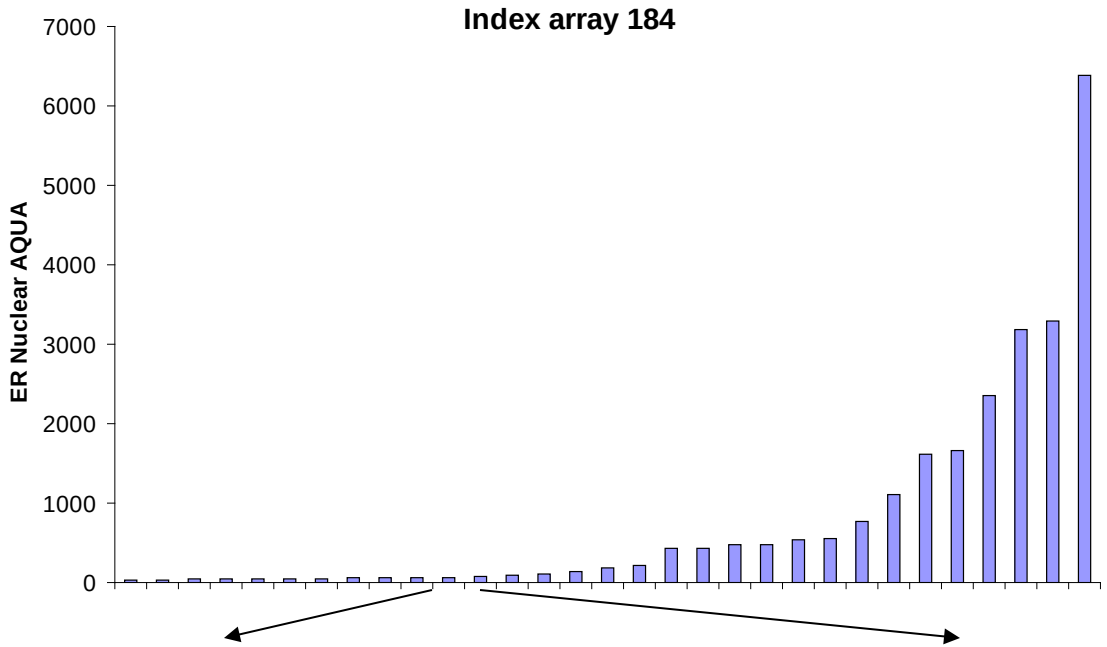

ER Nuclear = 143.17

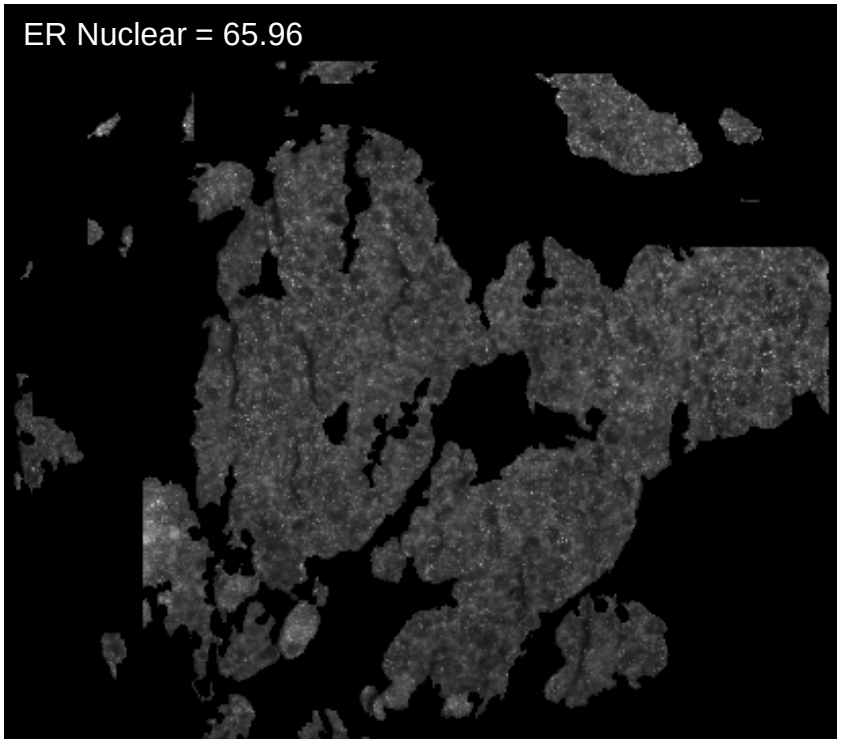

ER Nuclear = 178.94

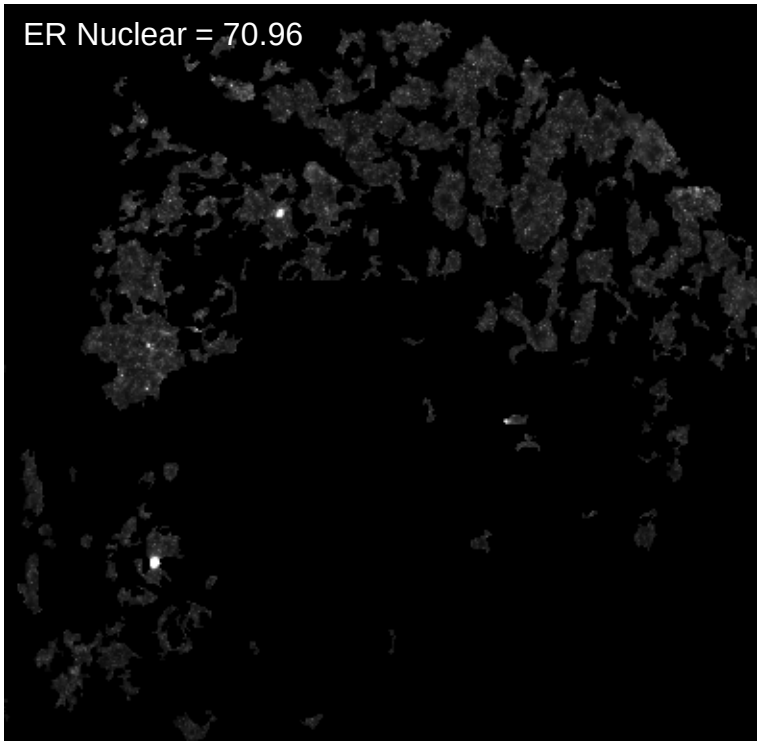

Index array 184
ER Nuclear AQUA
7000
6000
5000
4000
3000
2000
1000
0


ER Nuclear = 88.03

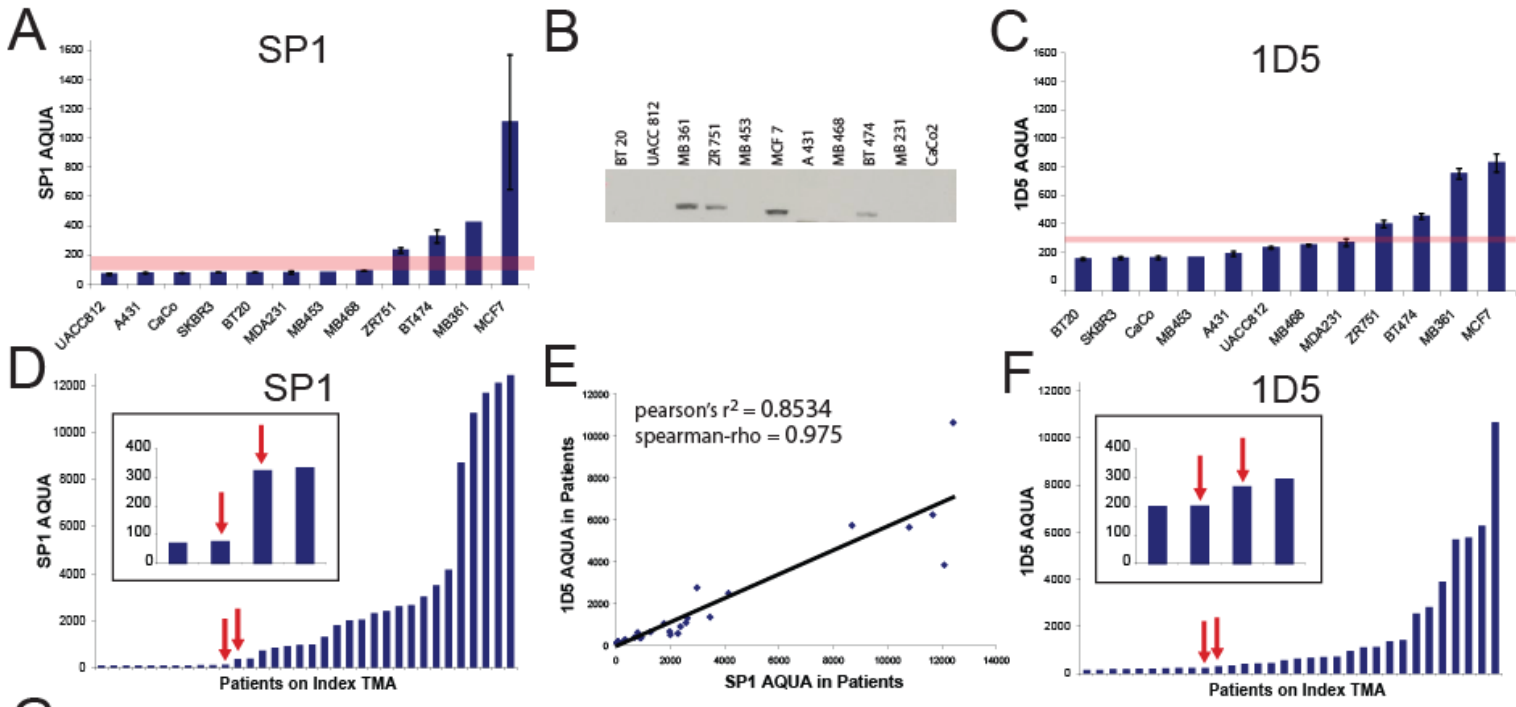

ER Nuclear = 112.65

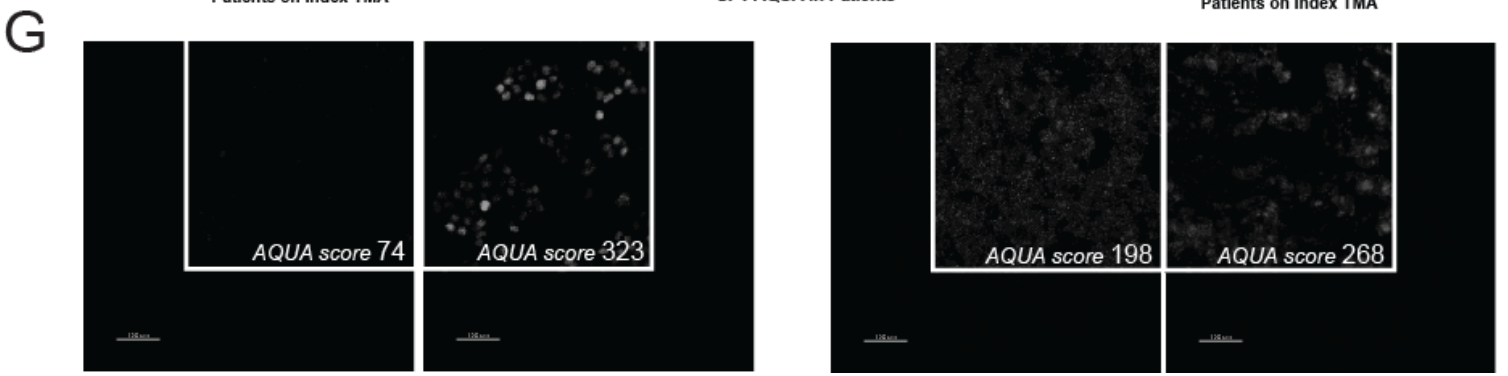

Index array 184
ER Nuclear AQUA
7000
6000
5000
4000
3000
2000
1000
0


ER Nuclear = 65.96


ER Nuclear = 70.96

Comparison of SP1 to 1D5 using QIF (AQUA) on a panel of cell line and patient controls – Higher Affinity leads to better Signal to Noise

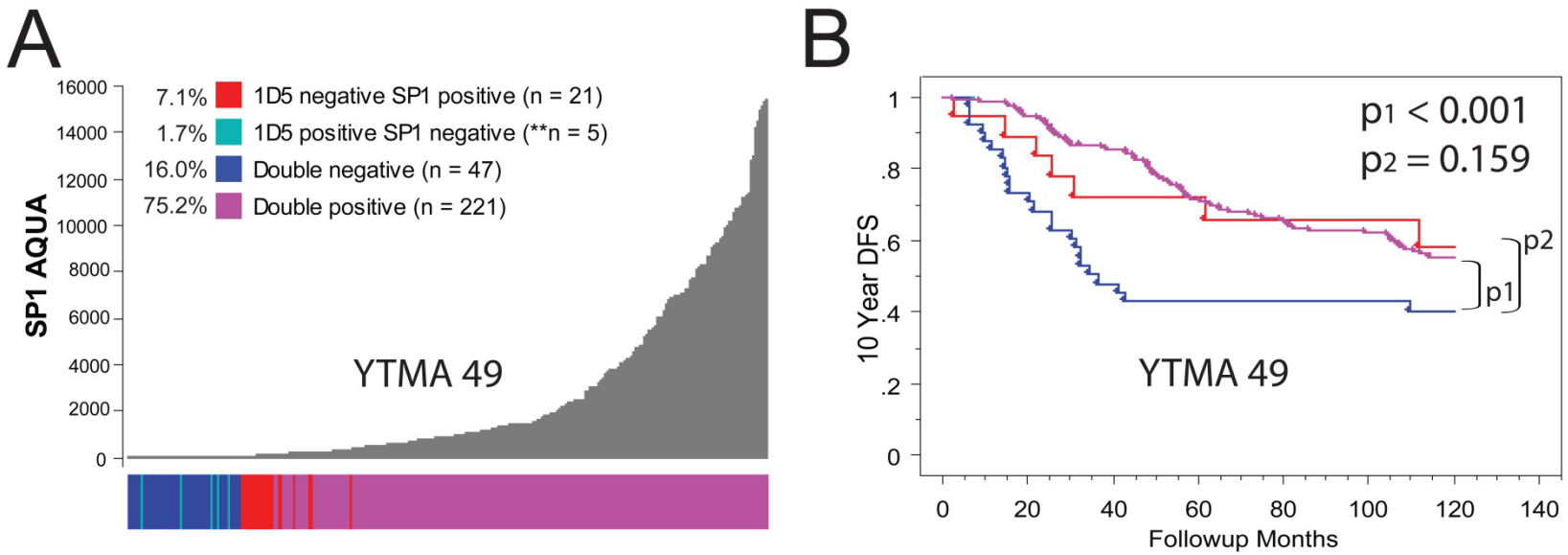




Input levels adjusted to max RGB 17 in inset

Input levels adjusted to max RGB 17 in inset

Comparison of antibodies by QIF assay suggests SP1 is more sensitive than 1D5



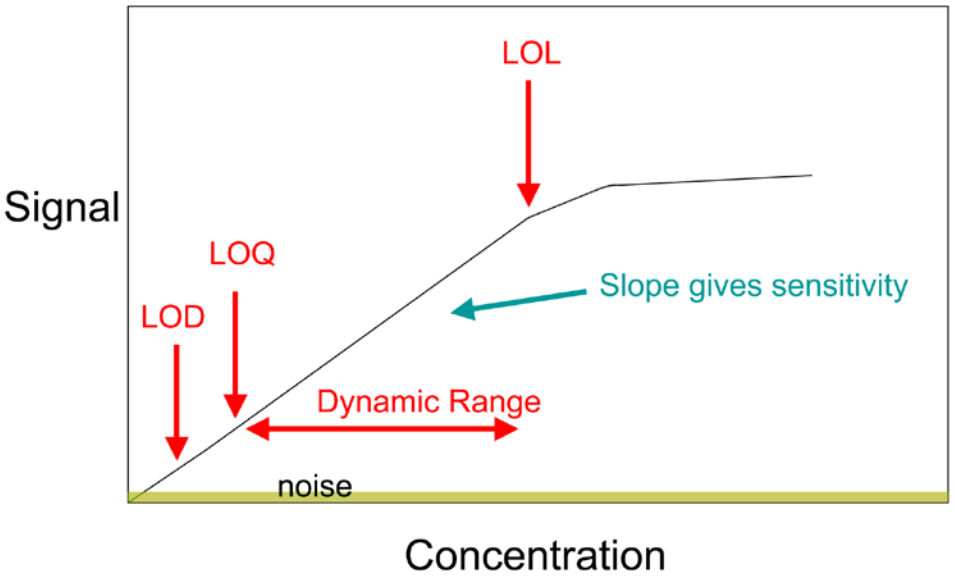
(*Appl Immunohistochem Mol Morphol* 2013;21:139–147)

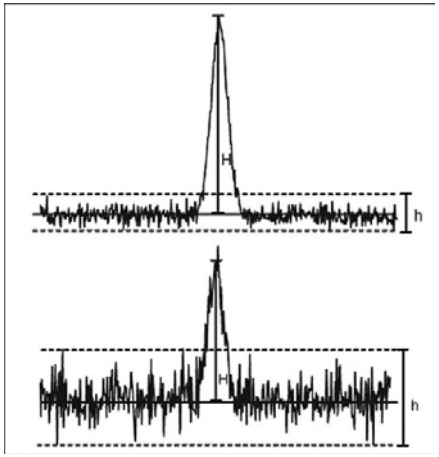
RESEARCH ARTICLE

Quantitative Analysis of Estrogen Receptor Expression Shows SP1 Antibody Is More Sensitive Than 1D5

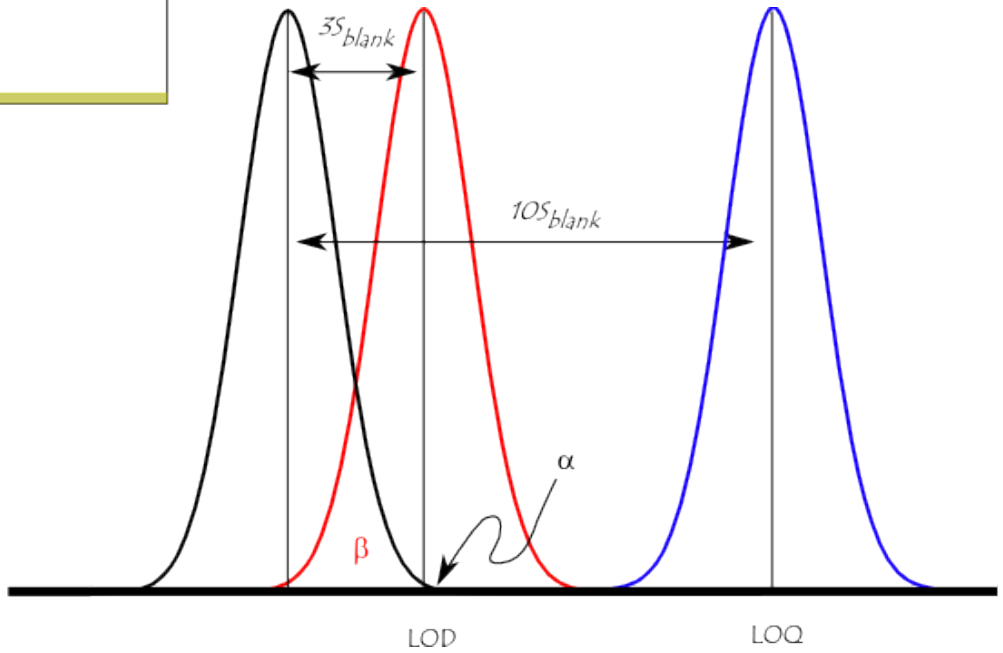
Allison W. Welsh, PhD, Malini Harigopal, MD, Hallie Wimberly, BS, Manju Prasad, MD, and David L. Rimm, MD, PhD

The Goal: To Measure Protein on a slide with the accuracy of Analytical Chemistry





LOD= limit of detection
LOQ = limit of quantification
LOL = limit of linearity



“Lessons Learned”

- Standardization is critical for reproducibility
- Cell lines can help establish cut-points or thresholds
- The human eye is great for many things, but not assessing subtle differences in intensity
- The binding domain and the affinity of the antibody are critical variables

Thanks to:

Rimm Group:
Kurt Schalper
Daniel Carvajal
Jason Brown
Susan Combs
James Smithy
Lauren Moore
Joe McLaughlin
Mehmet Altan
Nemanja Rodic
John McGuire
Joanna Hu
Vasiliki Pelakanou
Mengyao Feng
Nikita Mani
Yan Song
Maria Toki

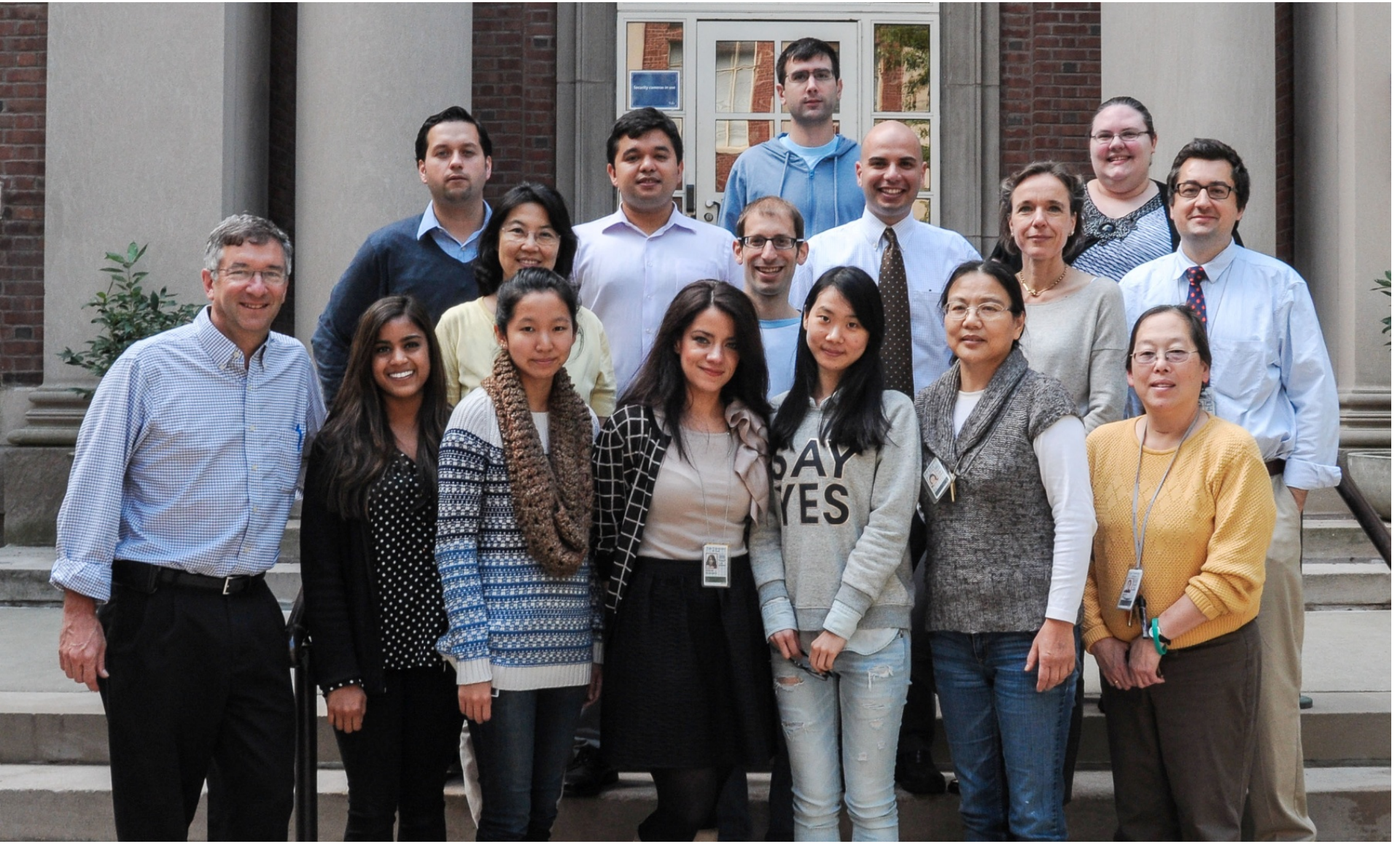
Rimm Lab Alums:
Vamsi Velchetti
Elsa Anagnostou
Anastasios Dimou
Alley Welsh
Robert Camp
Maria Vassilikapoulou
Huan Cheng
Jennifer Bordeaux
Vamsi Velchetti
Elizabeth Zarrella
Hallie Wimberly

Yale Pathology Tissue Services
Lori Charette
Sudha Kumar
Veronique Neumeister
Yalai Bai
(Google YPTS)

Yale Collaborators
Lajos Pusztai
Yuval Kluger
Leiping Chen

Outside Yale Collaborators
David Hicks (U Rochester)
Edith Perez and NCCTG
George Fountzilas – HeCOG
Amanda Psyrri – HeCOG
Konstantine Kalogeras -HeCOG

Work supported by grants from the NCI, DOD, BCRF, and the Susan G Komen Foundation for the Cure



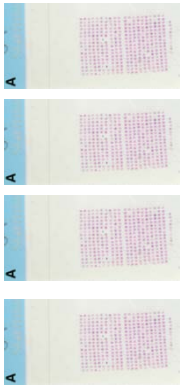
Rimm Lab, Fall 2014

www.tissuearray.org

Methods

Tissue Microarray slide submitted for staining in 4 different labs
Lab1 = Research Histology
Labs 2-4 = Anonymous CLIA labs

4 cuts from Yale TMA 49-9: A cohort of nearly 600 cases from 1962-1982, less than 2% treated with endocrine therapy, node negative treated with surgery alone



1) Lab-to-lab discordance

2) 1%-to10% discordance

Each slide scanned with Bioimagene scanner
read by 2 pathologists (MH and DLR) and 1 grad student (AW)

Lab-to-Lab Discordance

Lab3

Lab4

18.3%

Lab2

	ER pos 1%	ER neg 1%	Totals
ER pos 1%	301	83	384
ER neg 1%	7	111	118
Totals	308	194	502

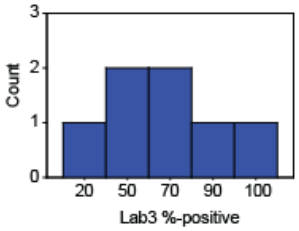
32.2%

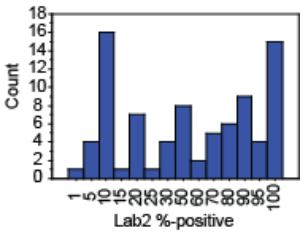
	ER pos 1%	ER neg 1%	Totals
ER pos 1%	234	140	374
ER neg 1%	7	101	108
Totals	241	241	482

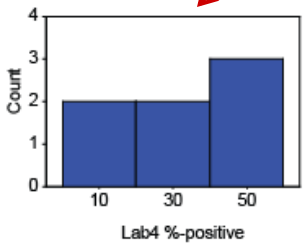
18.7%

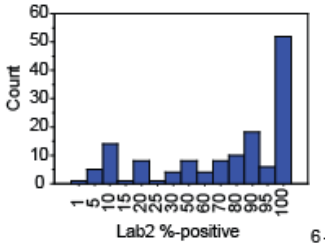
Lab3

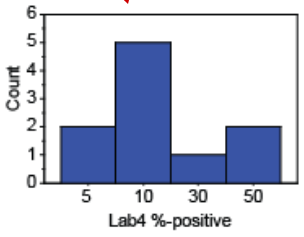
	ER pos 1%	ER neg 1%	Totals
ER pos 1%	231	72	303
ER neg 1%	10	178	188
Totals	241	250	491

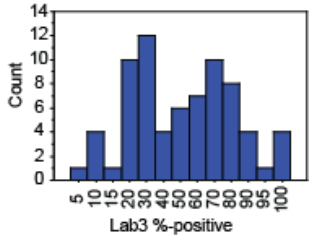






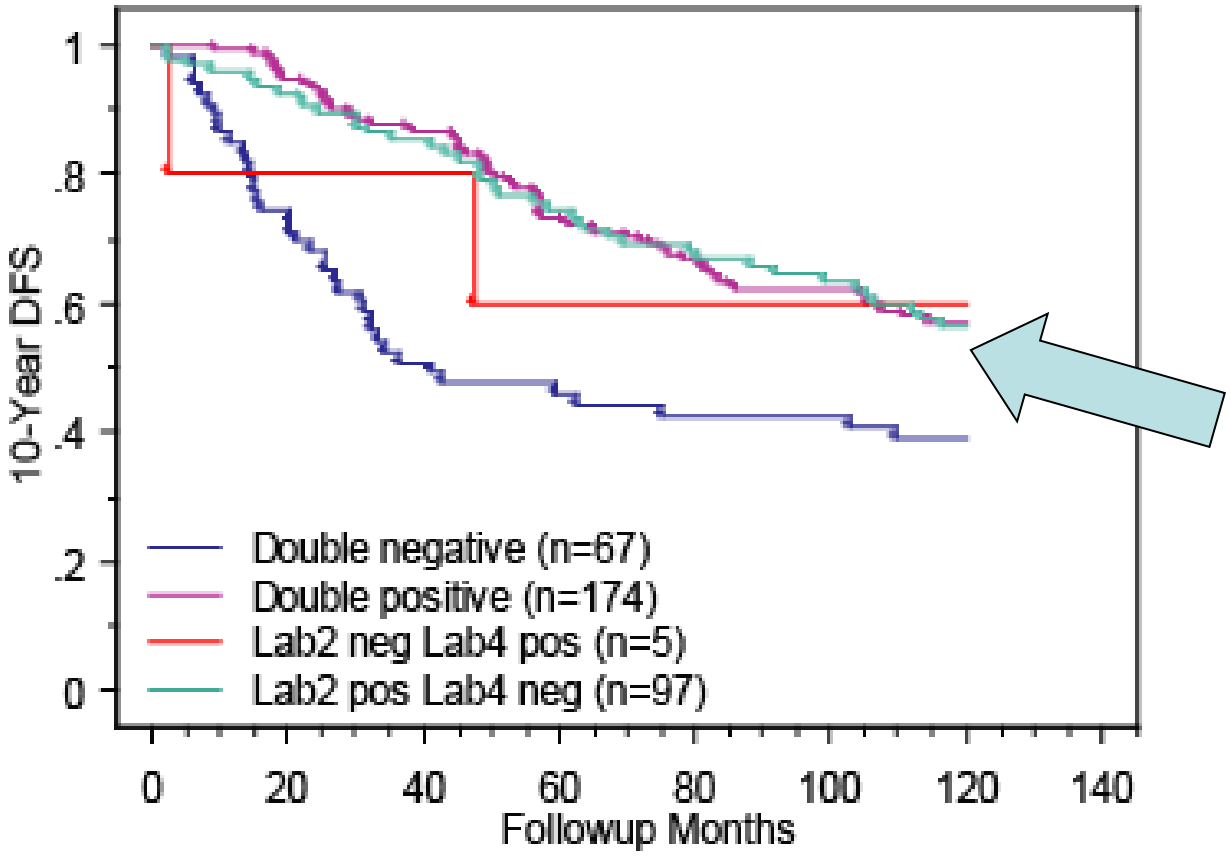






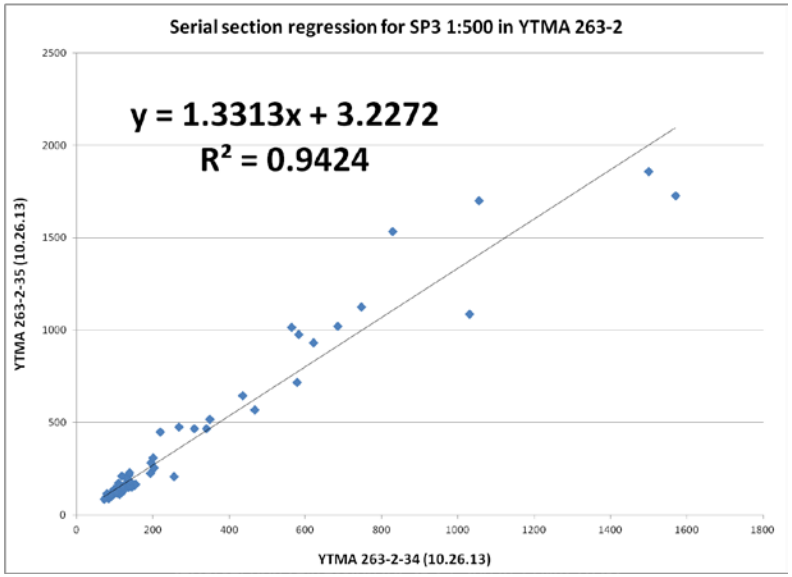
distribution of %-positive in misclassified cases

Some labs show “false negative” results compared to other labs

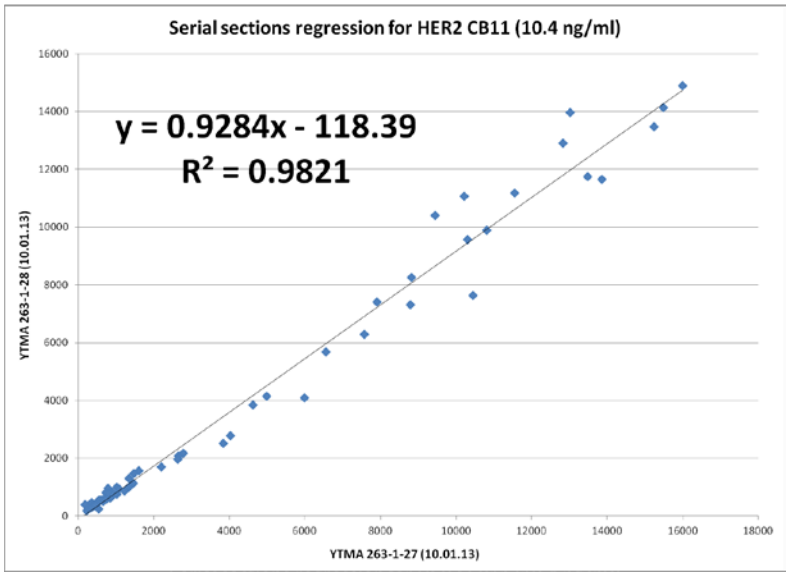


30% of cases were called positive from lab 2 but negative from lab 4 show outcome similar to double positives

Serial section Reproducibility for Domain Specific antibodies



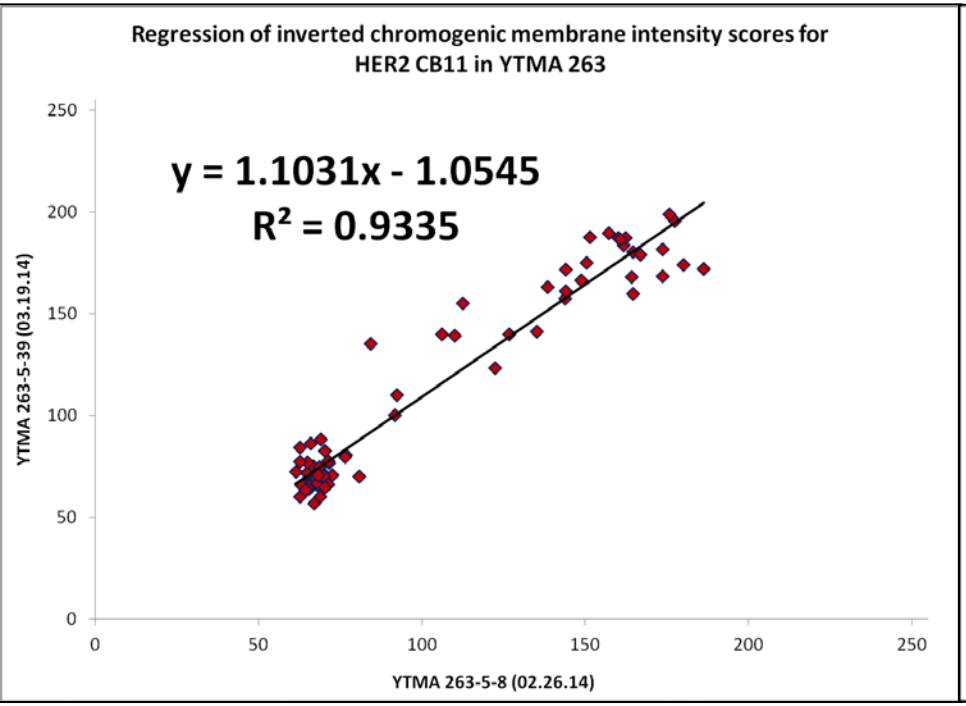
Extra-Cellular Domain (SP3)

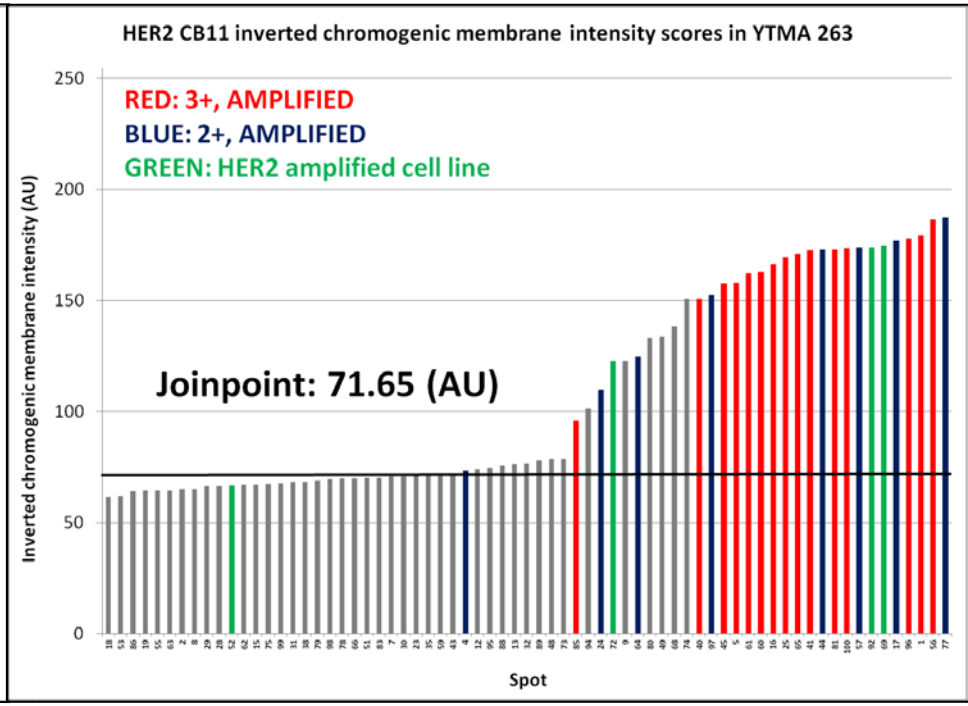


Intra-Cellular Domain (CB11)

Daniel Carvajal

Using Spectrum/Aperio scores from Membrane/Total pixel count for HER2 CB11





Daniel Carvajal

Using Joinpoint to Assess Ab Sensitivity and Specificity to Predict FISH amplification in YTMA 263 (Chromogenic IHC)

HER2 antibody	Joinpoint cut-point (AU)	Sensitivity (%)	Specificity (%)	Positive predictive value (%)	Negative predictive value (%)
CB11	71.65	96.3	64.29	63.41	96.43
SP3	75.75	89.66	82.61	76.47	92.68

Daniel Carvajal