

**DEVELOPING A DATASET
TO VALIDATE COMPUTATIONAL MODELS
THAT ANALYZE DIGITAL PATHOLOGY IMAGES
TO ASSESS TUMOR-INFILTRATING LYMPHOCYTES (TILs) IN
BREAST CANCER**

Brandon D. Gallas

Division of Imaging, Diagnostics, Software Reliability

Office of Science and Engineering Laboratories

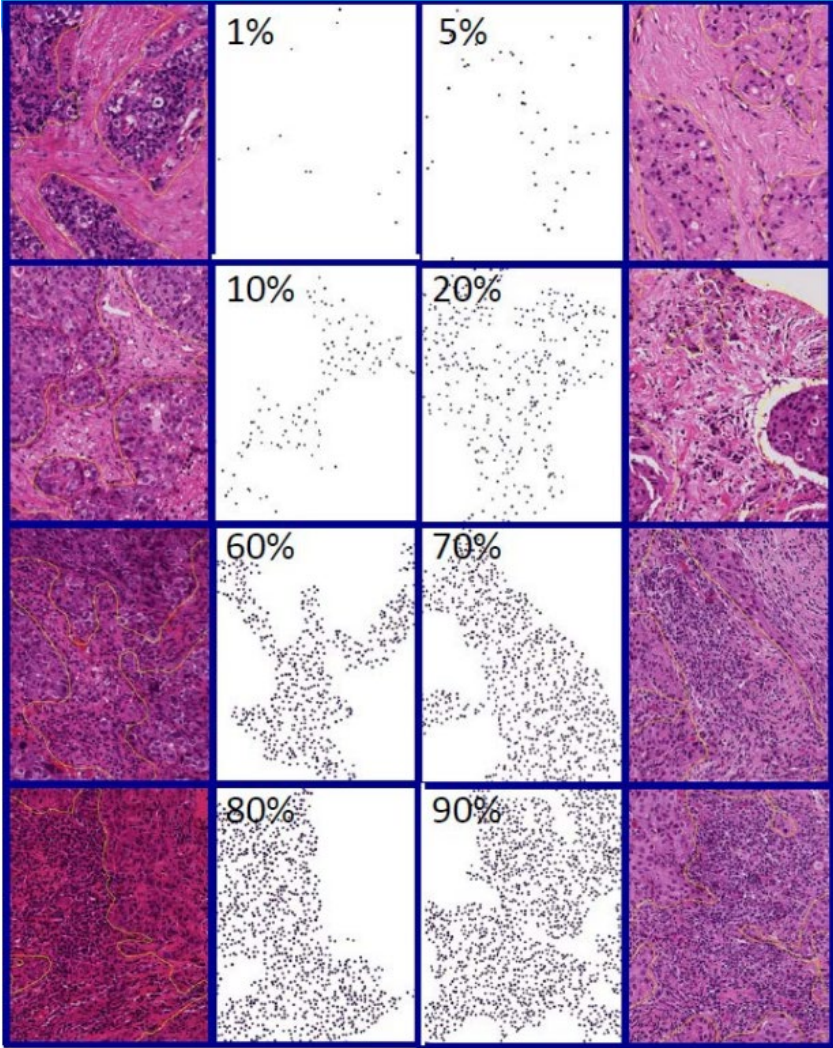
Center for Devices and Radiological Health

U.S. Food and Drug Administration

Outline

- Overview
- Pilot Study
 - Pathologist variability
- Pilot Study Deep Dive
 - Expert Panel Sessions
- Pathologist Training Materials
 - Knowledge-Based
 - Interactive
- Performance == Agreement
 - With experts

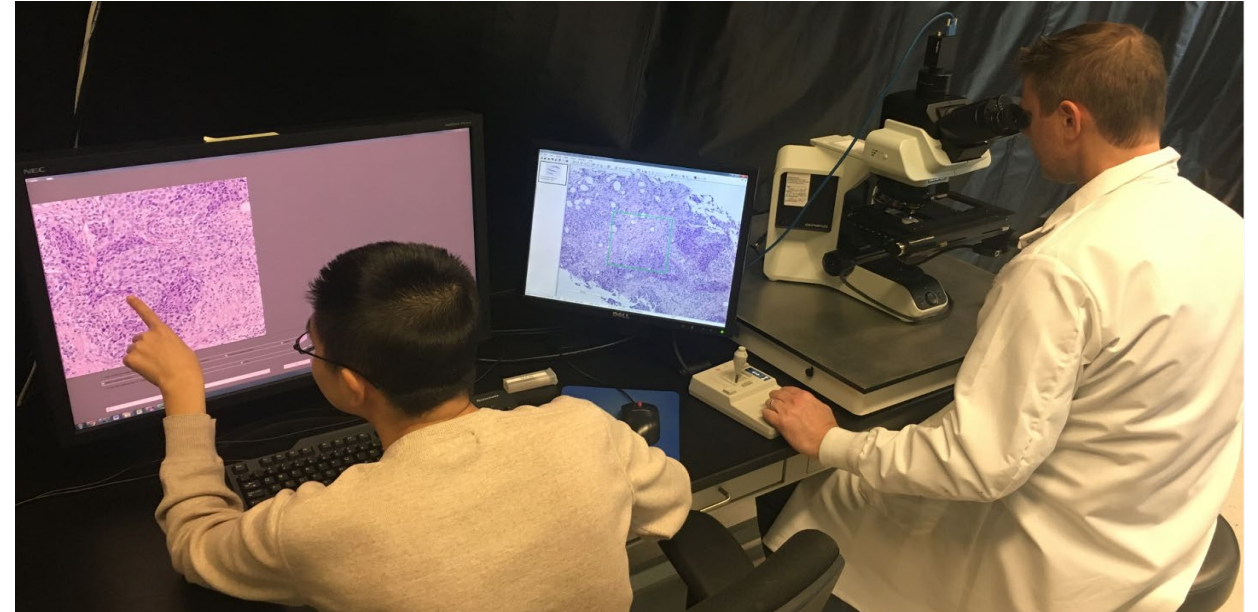
Quantitative Biomarker
TILs: Tumor Infiltrating Lymphocytes



High-Throughput Truthing (HTT) Project



- Clinical context:
 - Breast cancer
 - Quantitative Pathology Biomarker: Stromal Tumor Infiltrating Lymphocytes (sTILs)
- Clinical relevance of sTILs:
 - Prognostic for survival
 - Expected to inform patient management
 - Expected to reduce use of toxic chemotherapies
- Biomarker Evaluation by an Algorithm
 - Reduce burden on pathologist
 - Reproducible
 - Quantitative



- Deliverables/Tools
 - Reference standard data set from pathologists
 - Data-collection methods and platforms
 - Methods to validate a quantitative algorithm



Collaborators

Pathologists, Academics,
Industry, International
Volunteers

- **Mohamed Amgad, MD**
 - Department of Pathology, Northwestern University
- **Kim Blenman, PhD**
 - Yale School of Medicine
- **Weijie Chen, PhD**
 - FDA/CDRH/OSEL/DIDSR
- **Sarah Dudgeon, MPH**
 - CORE Center for Computational Health Yale-New Haven Hospital
- **Kate Elfer, MPH**
 - FDA/CDRH/OSEL/DIDSR
- **Anna Ehinger**
 - Lund University
- **Victor Garcia, MD**
 - FDA/CDRH/OSEL/DIDSR
- **Rajarsi Gupta, MD/PhD**
 - Stony Brook Medicine Dept of Biomedical Informatics
- **Matthew Hanna, MD**
 - Memorial Sloan Kettering Cancer Center
- **Steven Hart, PhD**
 - Department of Health Sciences Research, Mayo Clinic
- **Evangelos Hytopoulos, PhD**
 - iRhythm Technologies Inc
- **Denis Larsimont, MD**
 - Department of Pathology, Institut Jules Bordet
- **Xiaoxian Li, MD/PhD**
 - Emory University School of Medicine
- **Amy Ly, MD**
 - Massachusetts General Hospital
- **Anant Madabhushi, PhD**
 - Case Western Reserve University
- **Hetal Marble, PhD**
 - Massachusetts General Hospital/Harvard Medical School
- **Dieter Pieters**
 - Sint-Maarten Hospital; University of Antwerp; CellCarta
- **Roberto Salgado, PhD**
 - Division of Research, Peter Mac Callum Cancer Centre, Melbourne, Australia; Department of Pathology, GZA-ZNA Hospitals
- **Joel Saltz, MD/PhD**
 - Stony Brook Medicine Dept of Biomedical Informatics
- **Manasi Sheth, PhD**
 - FDA/CDRH/OPQE/Division of Biostatistics
- **Rajendra Singh, MD**
 - Northwell health and Zucker School of Medicine
- **Evan Szu, PhD**
 - Arrive Bio
- **Darick Tong, MS**
 - Arrive Bio
- **Si Wen, PhD**
 - FDA/CDRH/OSEL/DIDSR
- **Bruce Werness, MD**
 - Arrive Bio

High-Throughput Truthing (HTT) Project



- In Transition ...
Preparing Pivotal Study
- Project presentations and publications
- Pathologist training materials
- Access to data-collection Platforms

The screenshot shows a web browser window with the URL <https://ncihub.org/groups/eedapstudies>. The page header includes the NIH logo and the text "NATIONAL CANCER INSTITUTE Center for Biomedical Informatics & Information Technology". The main heading is "eeDAP studies" with a sub-heading "Overview". Below this is the "eeDAP Studies Group Page" title and a descriptive paragraph: "A home for collaborative studies to create tools (methods, data, and code) that advance regulatory science in the area of digital pathology imaging and related artificial intelligence software as a medical device." A sidebar on the left contains a "LOGIN" button and a menu with items: "Overview", "What is HTT?", "HTT Data Collection Training", "Start Data Collection", "Publications & Presentations", and "Contact Us". The main content area features six interactive tiles: "Wiki Home" (links to other project pages), "Evaluation Environment for Digital and Analog Pathology (eeDAP)", "Device Advice: for medical device sponsors submitting to the FDA", "What is HTT?", "HTT Data Collection Training", and "Start Data Collection".

Data-Collection Platforms: Digital

caMicroscope: Open Source

<https://github.com/camicroscope/caMicroscope>

The screenshot shows the caMicroscope interface. On the left is a control panel with a top navigation bar containing icons for home, back, ROI selection, analysis, and help. Below this, the 'ROI Type:' section has four radio button options: 'Intra-Tumoral Stroma' (selected), 'Tumor with No Intervening Stroma', 'Invasive Margin', and 'Other Regions'. Two sliders are visible: the first is labeled '% Tumor-Associated Stroma' with a value of 68%, and the second is labeled 'Please Assess TIL Density' with a value of 12%. A blue 'Save & Next' button is at the bottom of the panel. The main area displays a histology image with a green rectangular ROI box. A small inset image in the bottom right shows a wider view of the tissue with a red box indicating the ROI location.

The screenshot shows the PathPresenter interface. The top toolbar includes zoom levels (0.25x, 2x, 5x, 10x, 20x, 40x, 60x, 100x), a zoom level of 20, an area of 0.25 mm, and a grid option. The main image is a histology slide with a red rectangular ROI box. On the right is a data entry panel. The 'ROI Label:' dropdown is set to 'Intra tumoral stroma'. The 'Description:' field contains 'Test Description'. The '% Tumor-Associated Stroma:' field has a value of 23, with a slider below it also set to 23. The 'TILs:' field has a value of 21, with a slider below it also set to 21. The panel has 'Cancel' and 'Save' buttons at the bottom.

PathPresenter:

<https://pathpresenter.net/about>

Data-Collection Platforms: Microscope

Registers stage coordinates with whole slide image via camera

Allows replication of the digital-mode study design

Computer drives the stage from ROI to ROI

Annotations are independent of the scanner and viewer



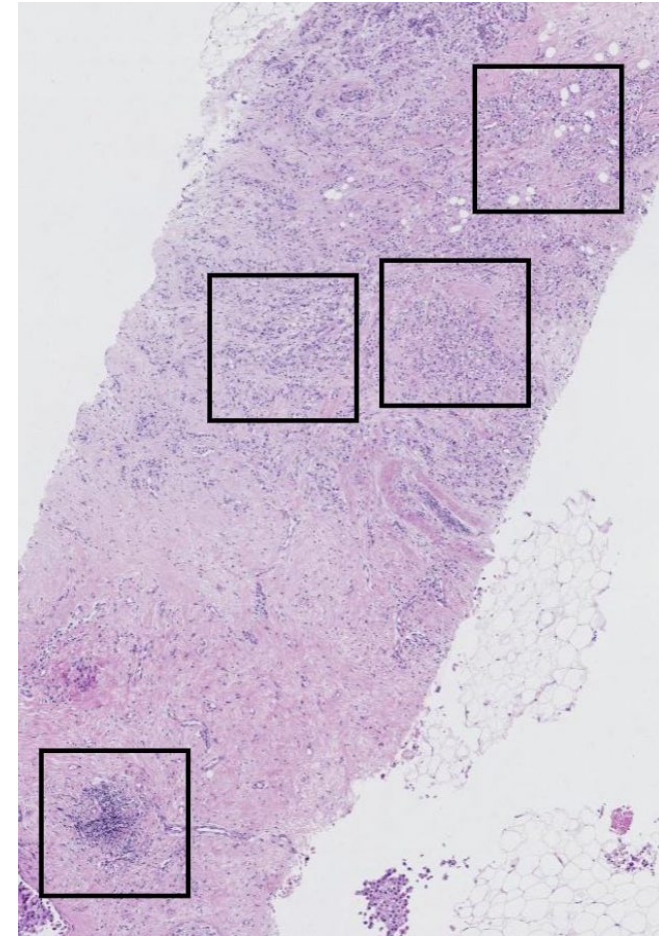
Pilot Study

- Cases:
 - 64 H&E Slides
 - 10 Regions of Interest (ROIs) per Slide
 - Some ROIs are not appropriate for sTIL evaluation

- Evaluation Platforms:
 - 2 digital and 1 microscope

- Readers:
 - 37 readers
 - 7 crowd readers with complete data
 - 7 expert readers are on the collaboration team

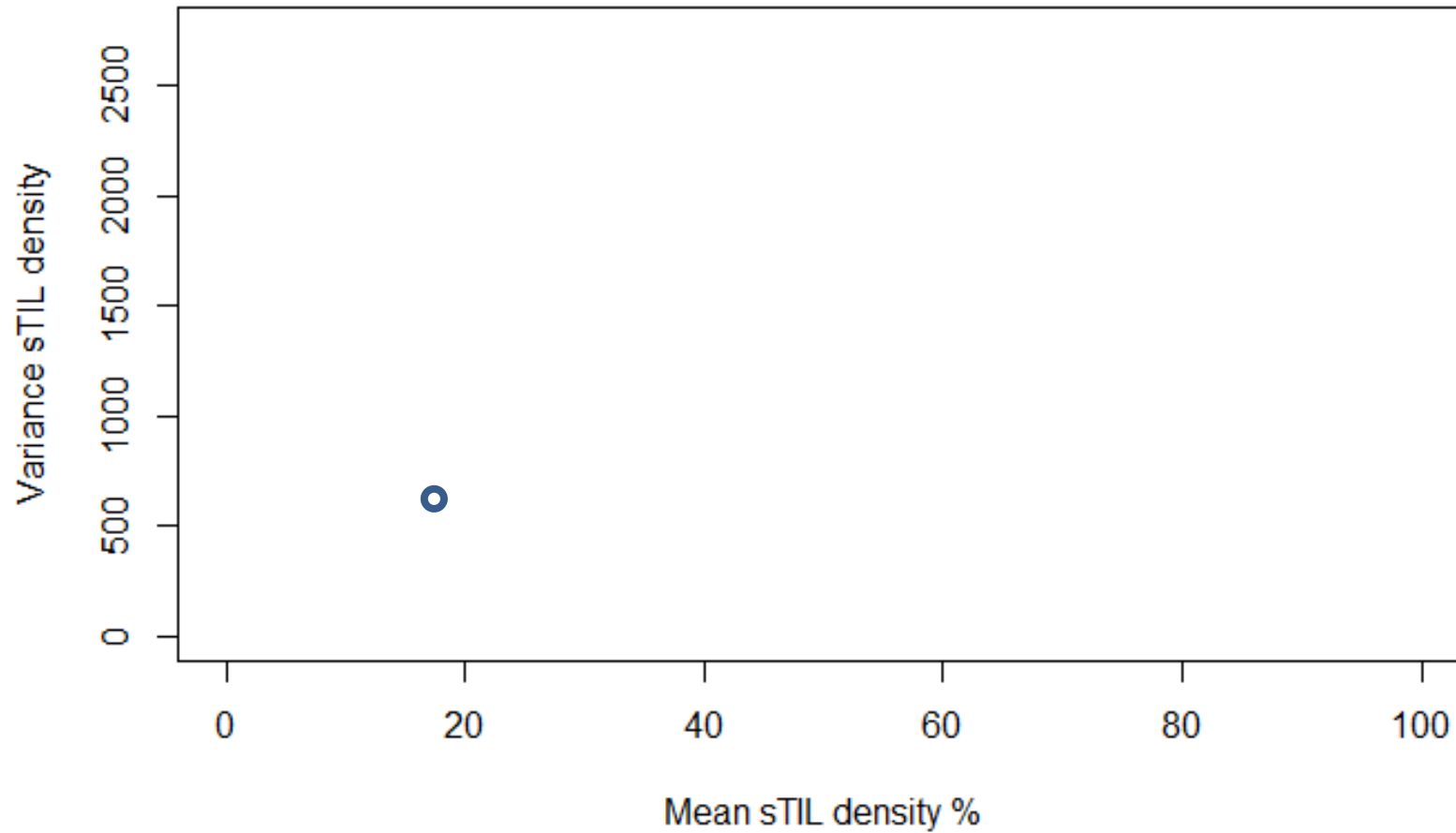
- 7,898 Observations



[R Data Package](https://github.com/DIDSR/HTT)
<https://github.com/DIDSR/HTT>

Initial Analysis of Pilot Study

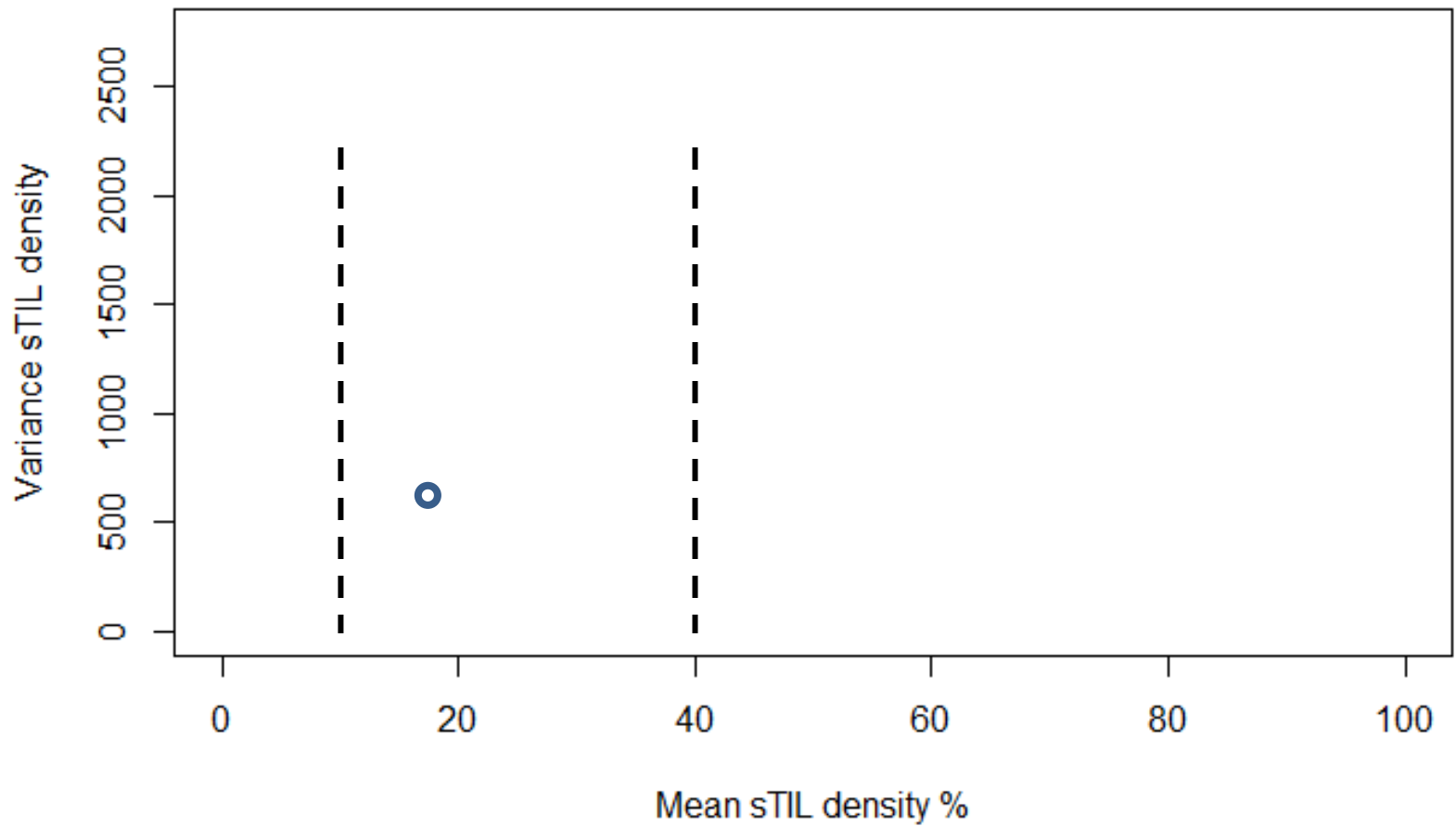
Variance of Pilot Study



- Mean and Variance are averages over all readers

Initial Analysis of Pilot Study

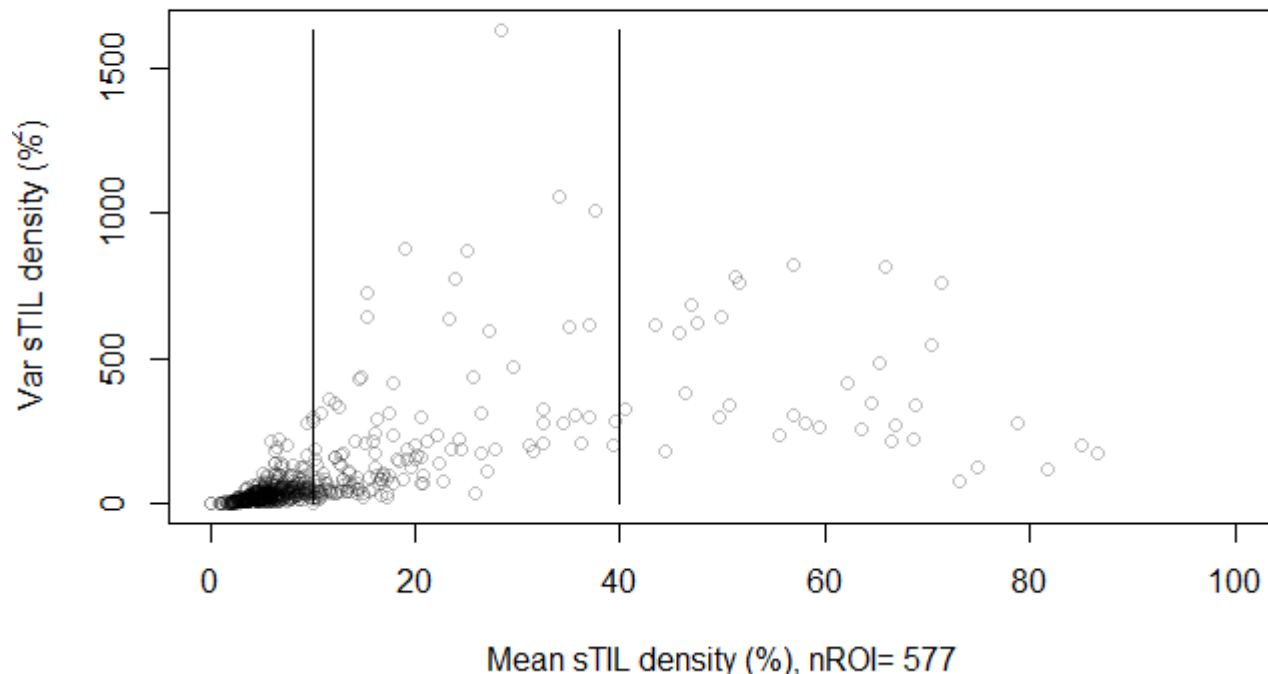
Variance of Pilot Study



- Mean and Variance are averages over all readers
- Vertical dashed lines represent clinical bins
 - low ($\leq 10\%$)
 - medium ($>10\% \ \& \ \leq 40\%$)
 - high ($>40\%$)

Initial Analysis of Pilot Study

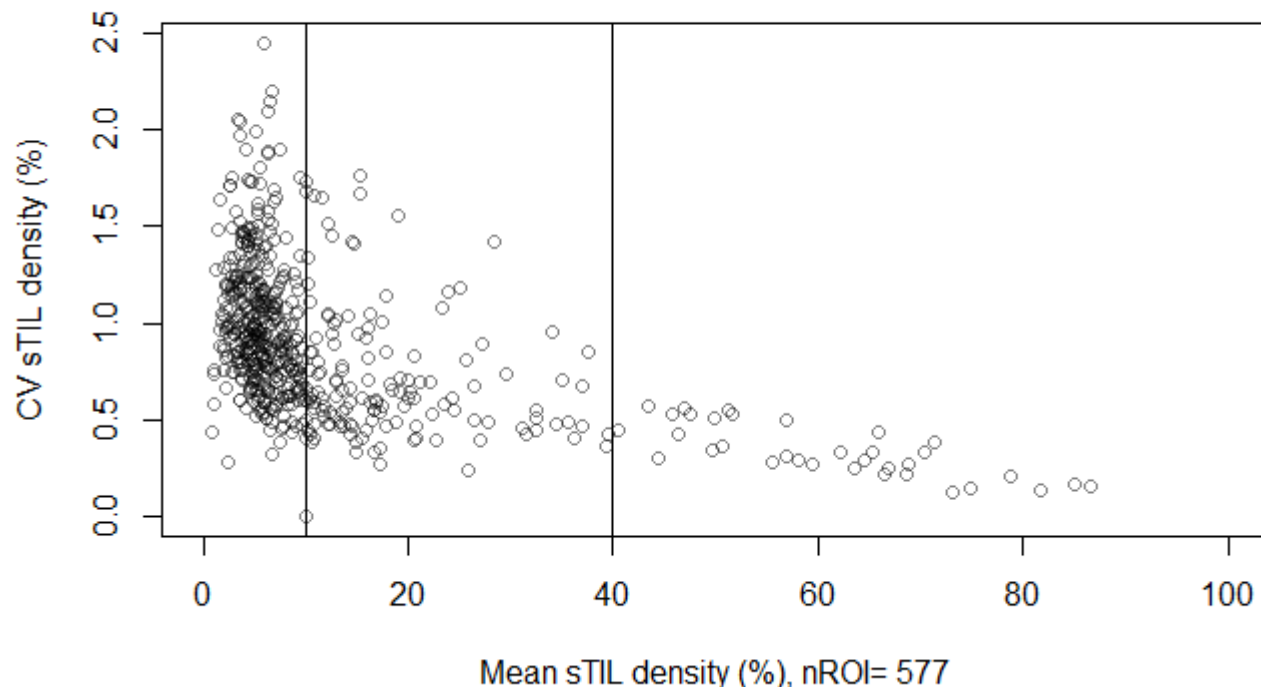
All Pilot Data: Pathologist Variance for each ROI



- Means and Variances are averages over all readers
- Vertical lines represent clinical bins
 - low ($\leq 10\%$)
 - medium ($>10\% \ \& \ \leq 40\%$)
 - high ($>40\%$)
- Variance is increasing with the mean

Initial Analysis of Pilot Study

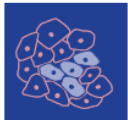
All Pilot Data: Pathologist CV for each ROI



- Means and Variances are averages over all readers
- Vertical dashed lines represent clinical bins
 - low ($\leq 10\%$)
 - medium ($>10\% \ \& \ \leq 40\%$)
 - high ($>40\%$)
- The variance does not increase with mean in a standard way

Pilot Study Deep Dive: Expert Panel Sessions

- Primary purpose
 - Understand pathologist variability
 - Improve instructions to reduce variability
- Subsequent Opportunities
 - Clinical practice training materials
 - Reference standard for pilot study
 - Explore analysis methods








cancers

Garcia et al. 2022, *Cancers*, “...Training Materials...”



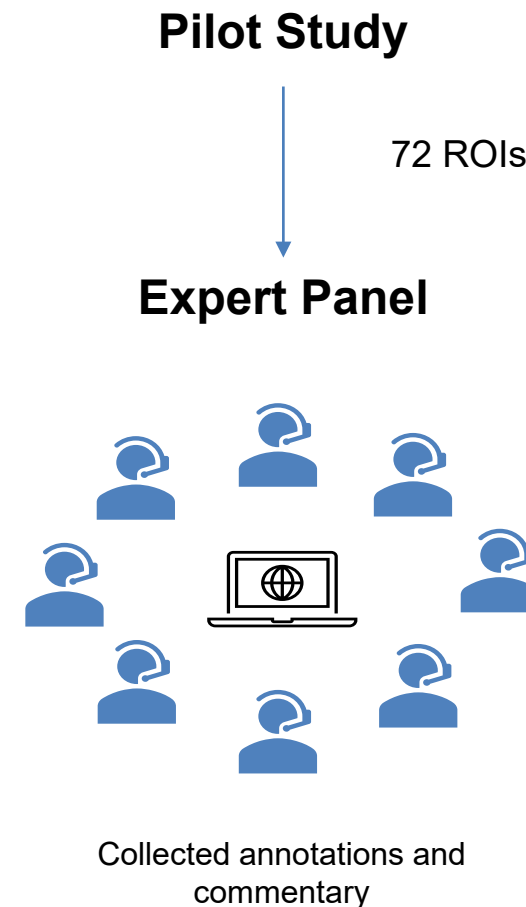
Article

Development of Training Materials for Pathologists to Provide Machine Learning Validation Data of Tumor-Infiltrating Lymphocytes in Breast Cancer

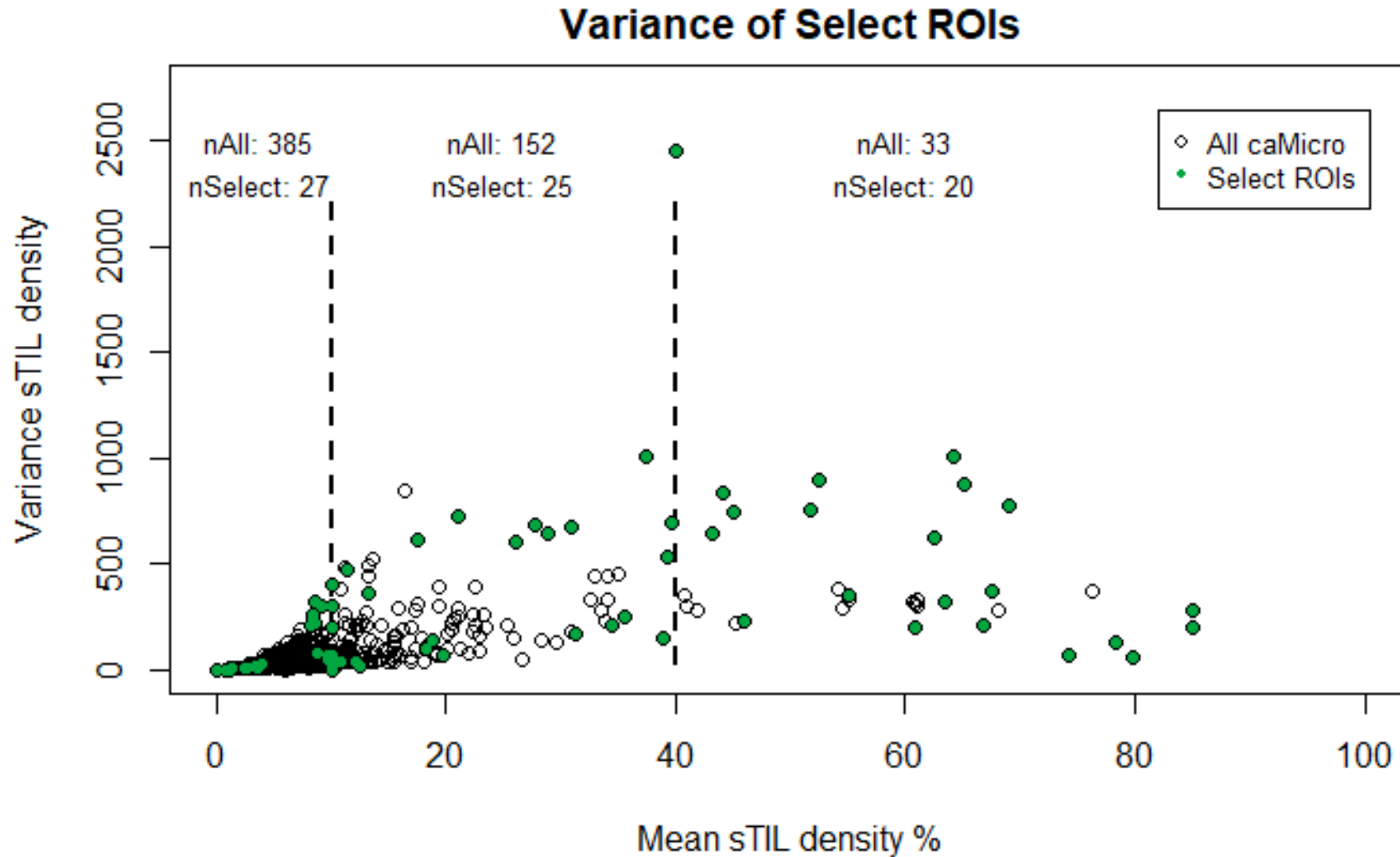
Victor Garcia ^{1,*}, Katherine Elfer ^{1,2}, Dieter J. E. Peeters ^{3,4,5}, Anna Ehinger ⁶, Bruce Werness ^{7,8}, Amy Ly ⁹, Xiaoxian Li ¹⁰, Matthew G. Hanna ¹¹, Kim R. M. Blenman ^{12,13}, Roberto Salgado ^{14,15} and Brandon D. Gallas ¹

Pilot Study Deep Dive: Expert Panel Sessions

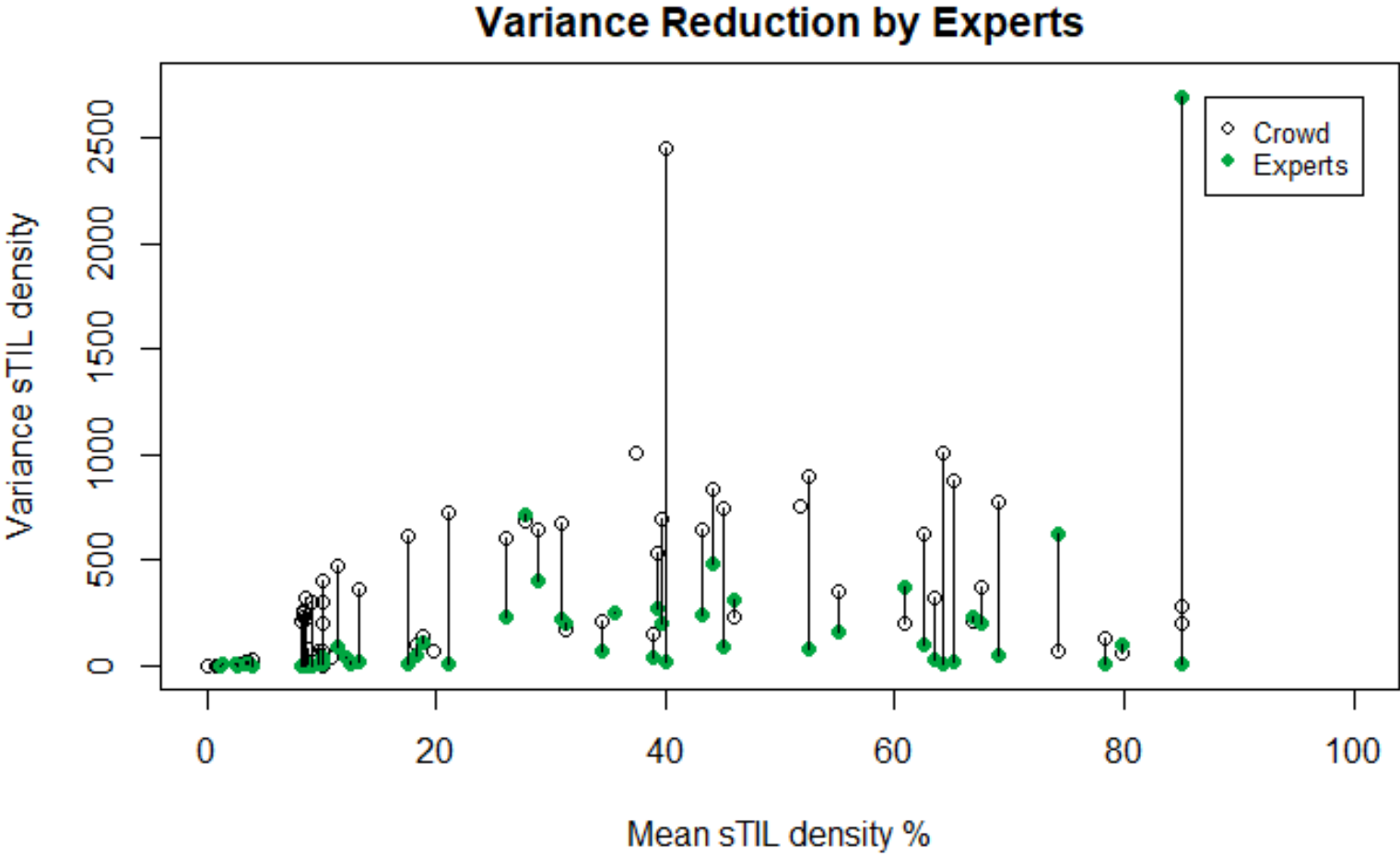
- 72 “Select” ROIs
 - 2:1 mix of high and low pathologist variability cases
 - Stratified sampling on mean density of sTILs (3 bins)
- 8-member expert panel
- 8 recorded, one-hour virtual sessions
- Collect annotations independently
- Digital mode: caMicroscope



Select Data for Expert Panel Sessions



Expert Panel Annotations: Reduced Pathologist Variability



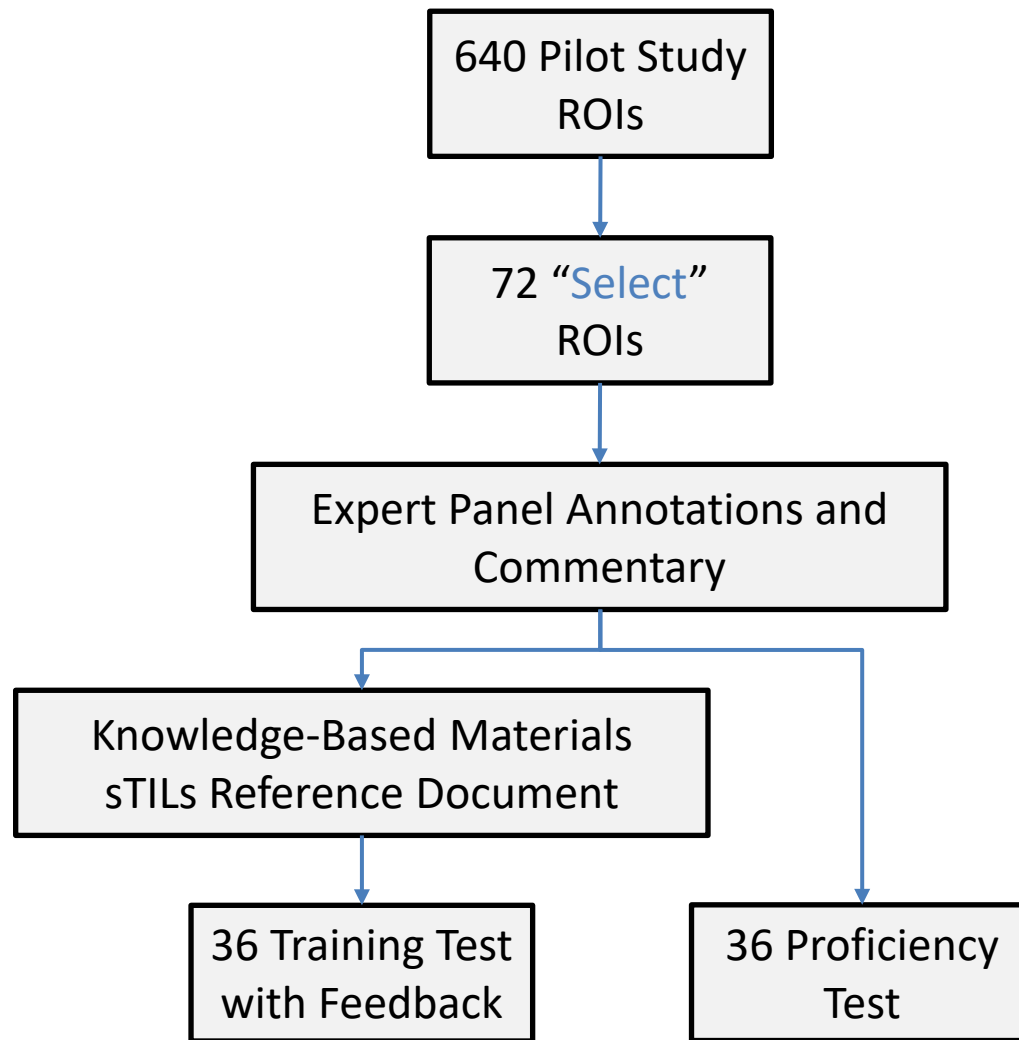
Expert Panel Annotations: Reduced Pathologist Variability



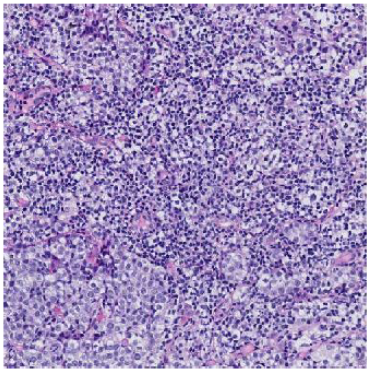
Legend:
Median [IQR]

	All Densities	≤ 10%	10% < % ≤ 40%	> 40%
Crowd - Select	212.24 [39.33 - 549.50]	44.67 [4.05 - 225.28]	246.80 [67.58 - 646.18]	358.75 [210.17 - 762.73]
Experts - Select	14.17 [4.23 - 178.67]	3.07 [0.98 - 4.32]	70.00 [14.17 - 224.17]	96.67 [39.42 - 275.03]

Pathologist Training Materials



sTILs Reference Document



caseID: HTT-TILS-001-04B.ndpi_x24343.2190_y11775.2190

Expert Panel Annotations

ROI Type	Percent Tumor-Associated Stroma	sTILs Density
Evaluable	30	90
Evaluable	60	95
Evaluable	50	92
Evaluable	50	75
Evaluable	60	90
Evaluable	60	90

Mean Percent Tumor-Associated Stroma: 51.7

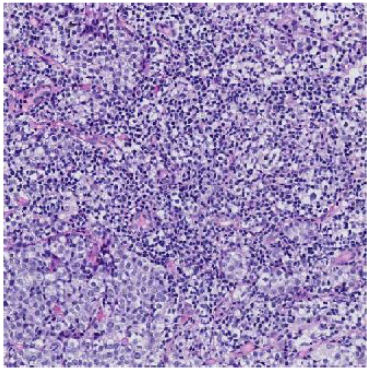
Mean sTILs Density: 88.7

Comments: A challenging case. The high density of lymphocytes results in difficulty determining whether the lymphocytes are located in stroma, or whether they infiltrate tumor cell nests. The presence of small blood vessels and small gaps between lymphocytes suggest the lymphocytes reside within stroma. Occasional tumor cells with small nuclei (possibly degenerating) may be confused for lymphocytes.

Pitfalls: In regions where the sTILs density is very high, the underlying stroma may be obscured. Non-lymphocytes with small nuclei may be confused for lymphocytes.

- ROI
- Expert annotations
- General comments
- Pitfalls, if any

sTILs Reference Document => Pitfalls

caseID: HTT-TILS-001-04B.ndpi_x24343.2190_y11775.2190

Expert Panel Annotations

ROI Type	Percent Tumor-Associated Stroma	sTILs Density
Evaluable	30	90
Evaluable	60	95
Evaluable	50	92
Evaluable	50	75
Evaluable	60	90
Evaluable	60	90

Mean Percent Tumor-Associated Stroma: 51.7
Mean sTILs Density: 88.7


Comments: A challenging case. The high density of lymphocytes results in difficulty determining whether the lymphocytes are located in stroma, or whether they infiltrate tumor cell nests. The presence of small blood vessels and small gaps between lymphocytes suggest the lymphocytes reside within stroma. Occasional tumor cells with small nuclei (possibly degenerating) may be confused for lymphocytes.

Pitfalls: In regions where the sTILs density is very high, the underlying stroma may be obscured. Non-lymphocytes with small nuclei may be confused for lymphocytes.


2



Garcia et al. 2022, *Cancers*, “...Training Materials...”



cancers



Article

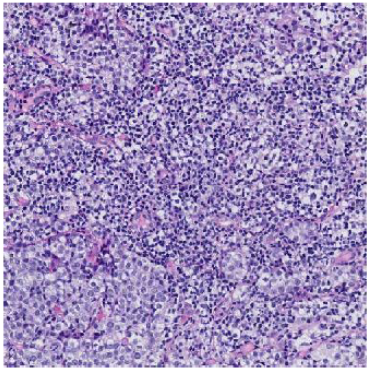
Development of Training Materials for Pathologists to Provide Machine Learning Validation Data of Tumor-Infiltrating Lymphocytes in Breast Cancer

Victor Garcia ^{1,*}, Katherine Elfer ^{1,2}, Dieter J. E. Peeters ^{3,4,5}, Anna Ehinger ⁶, Bruce Werness ^{7,8}, Amy Ly ⁹, Xiaoxian Li ¹⁰, Matthew G. Hanna ¹¹, Kim R. M. Blenman ^{12,13}, Roberto Salgado ^{14,15} and Brandon D. Gallas ¹⁶

Table 6. Summary of pitfalls encountered during the stromal tumor-infiltrating lymphocytes (sTILs) assessment grouped by pitfall type. Region of interest is abbreviated as “ROI”.

Pitfall Type	Pitfall Summary
Percent of Tumor-Associated Stroma	Exclude thick-walled vessels, benign glandular elements, adipocytes, carcinoma in situ, and necrosis from the area of tumor-associated stroma Calculate with respect to the entire ROI area Variations in tumor cell morphology can make it difficult to distinguish stroma from tumor Cells with small/pyknotic nuclei and/or perinuclear clearing can be difficult to categorize
sTILs Density Score	Non-lymphoid cells may be confused for lymphocytes Error in the percent tumor-associated stroma can affect the sTILs density Sparsely distributed tumor cells may be more challenging to quantitate

sTILs Reference Document => Pitfalls Video



caseID: HTT-TILS-001-04B.ndpi_x24343.2190_y11775.2190

Expert Panel Annotations

ROI Type	Percent Tumor-Associated Stroma	sTILs Density
Evaluable	30	90
Evaluable	60	95
Evaluable	50	92
Evaluable	50	75
Evaluable	60	90
Evaluable	60	90

Mean Percent Tumor-Associated Stroma: 51.7
Mean sTILs Density: 88.7

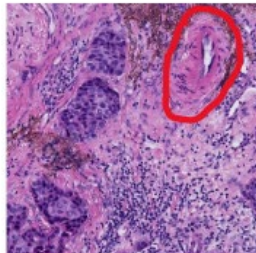
Comments: A challenging case. The high density of lymphocytes results in difficulty determining whether the lymphocytes are located in stroma, or whether they infiltrate tumor cell nests. The presence of small blood vessels and small gaps between lymphocytes suggest the lymphocytes reside within stroma. Occasional tumor cells with small nuclei (possibly degenerating) may be confused for lymphocytes.

Pitfalls: In regions where the sTILs density is very high, the underlying stroma may be obscured. Non-lymphocytes with small nuclei may be confused for lymphocytes.

2

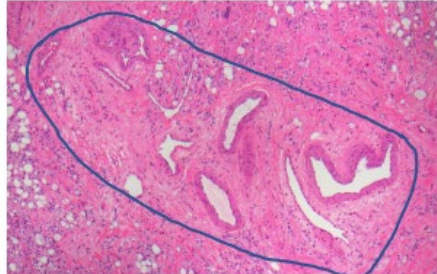


Thick-walled vessels are not considered stroma

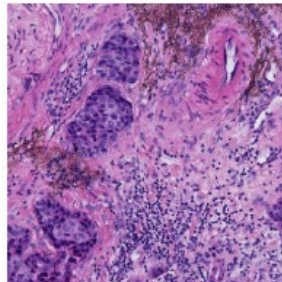


Area of tumoral stroma occupied by mononuclear inflammation X 100
Entire area of tumoral stroma

Thick-walled vessels are not considered stroma



How much tumor-associated stroma is present?

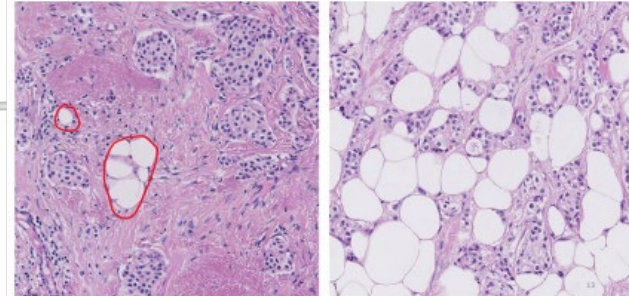


ROI Type	Percent Tumor-Associated Stroma	sTILs Density
Evaluable	75	30
Evaluable	35	60
Evaluable	86	15
Evaluable	75	30
Evaluable	70	25
Evaluable	70	20

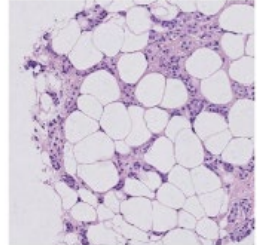
Mean Percent Tumor-Associated Stroma: 68.5
Mean sTILs Density: 30

Example Slides

Adipose tissue is not considered stroma



How much tumor associated stroma is present?



ROI Type	Percent Tumor-Associated Stroma	sTILs Density
Evaluable	10	0
Evaluable	5	1
Evaluable	14	4
Evaluable	20	0
Evaluable	40	0
Evaluable	50	2

Mean Percent Tumor-Associated Stroma: 23.2
Mean sTILs Density: 1.2

Complete Course:

Assessment of Stromal Tumor-Infiltrating Lymphocytes



Objectives

- Describe the **significance** of stromal tumor-infiltrating lymphocytes in triple negative breast cancer.
- Demonstrate knowledge of the **approach** to determining the density of stromal tumor-infiltrating lymphocytes.

Faculty

- Victor Garcia, MD
- Amy Ly, MD
- Matthew Hanna, MD
- Dieter Peeters, MD, PhD
- Roberto Salgado, MD, PhD
- Xiaoxian Li, MD, PhD
- Kim Blenman, PhD, MS
- Katherine Elfer, PhD, MPH
- Bruce Werness, MD
- Anna Ehinger, MD
- Brandon Gallas, PhD

Complete Course:

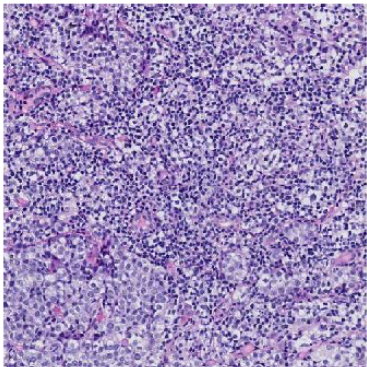
Assessment of Stromal Tumor-Infiltrating Lymphocytes



Content

- Introduction to stromal tumor-infiltrating lymphocytes (**significance**)
 - 12-minute video, not yet published
- TILs Education: What are TILs and their Assessment (**approach**)
 - 8-minute video
 - Created by the International Immuno-Oncology Biomarker Working Group, <https://youtu.be/aPa-pXIBBIU>
- Pitfalls in the sTILs Assessment (**approach**)
 - 27-minute video, not yet published
- The evaluation of tumor-infiltrating lymphocytes (TILs) in breast cancer: recommendations by an International TILs Working Group 2014 (**approach**)
 - Manuscript read time 30-60 minutes
 - Salgado2015_Ann-Oncol_v26p259, <https://www.doi.org/10.1093/annonc/mdu450>

sTILs Reference Document => Interactive Test with Feedback

caseID: HTT-TILS-001-04B.ndpi_x24343.2190_y11775.2190

Expert Panel Annotations

ROI Type	Percent Tumor-Associated Stroma	sTILs Density
Evaluable	30	90
Evaluable	60	95
Evaluable	50	92
Evaluable	50	75
Evaluable	60	90
Evaluable	60	90

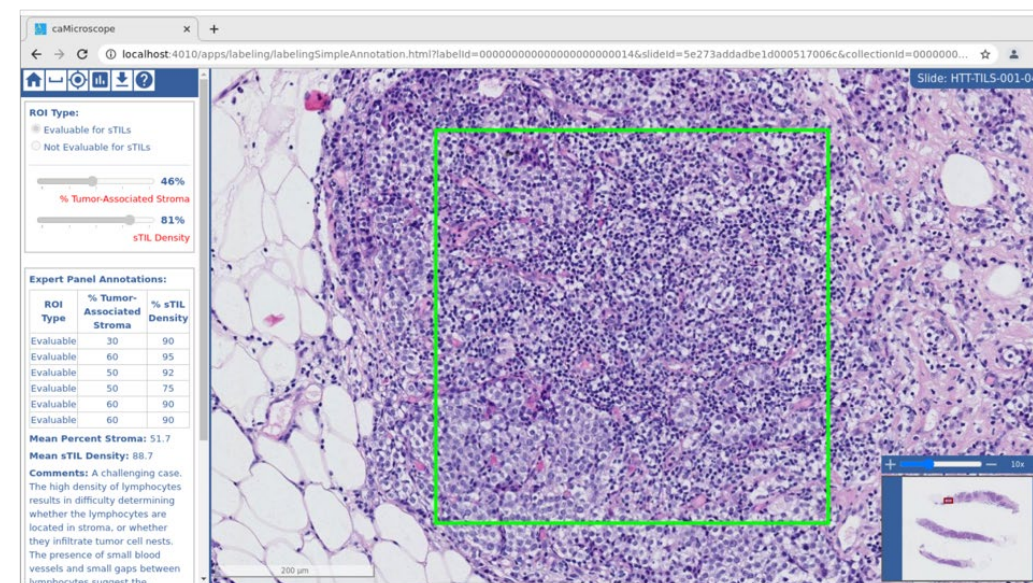
Mean Percent Tumor-Associated Stroma: 51.7
Mean sTILs Density: 88.7

Comments: A challenging case. The high density of lymphocytes results in difficulty determining whether the lymphocytes are located in stroma, or whether they infiltrate tumor cell nests. The presence of small blood vessels and small gaps between lymphocytes suggest the lymphocytes reside within stroma. Occasional tumor cells with small nuclei (possibly degenerating) may be confused for lymphocytes.

Pitfalls: In regions where the sTILs density is very high, the underlying stroma may be obscured. Non-lymphocytes with small nuclei may be confused for lymphocytes.

2

36 cases

caMicroscope

localhost:4010/apps/labeling/labelingSimpleAnnotation.html?labelId=000000000000000000000000146&slideId=5e273addadbe1d000517006c&collectionId=0000000...

Slide: HTT-TILS-001-04B

ROI Type:
 Evaluable for sTILs
 Not Evaluable for sTILs

46%
 % Tumor-Associated Stroma

81%
 sTIL Density

Expert Panel Annotations:

ROI Type	% Tumor-Associated Stroma	% sTIL Density
Evaluable	30	90
Evaluable	60	95
Evaluable	50	92
Evaluable	50	75
Evaluable	60	90
Evaluable	60	90

Mean Percent Stroma: 51.7
 Mean sTIL Density: 88.7

Comments: A challenging case. The high density of lymphocytes results in difficulty determining whether the lymphocytes are located in stroma, or whether they infiltrate tumor cell nests. The presence of small blood vessels and small gaps between lymphocytes suggest the lymphocytes reside within stroma. Occasional tumor cells with small nuclei (possibly degenerating) may be confused for lymphocytes.

Interactive test with feedback



ROI Type:

Evaluable for sTILs
 Not Evaluable for sTILs

46%
 % Tumor-Associated Stroma

81%
 sTIL Density

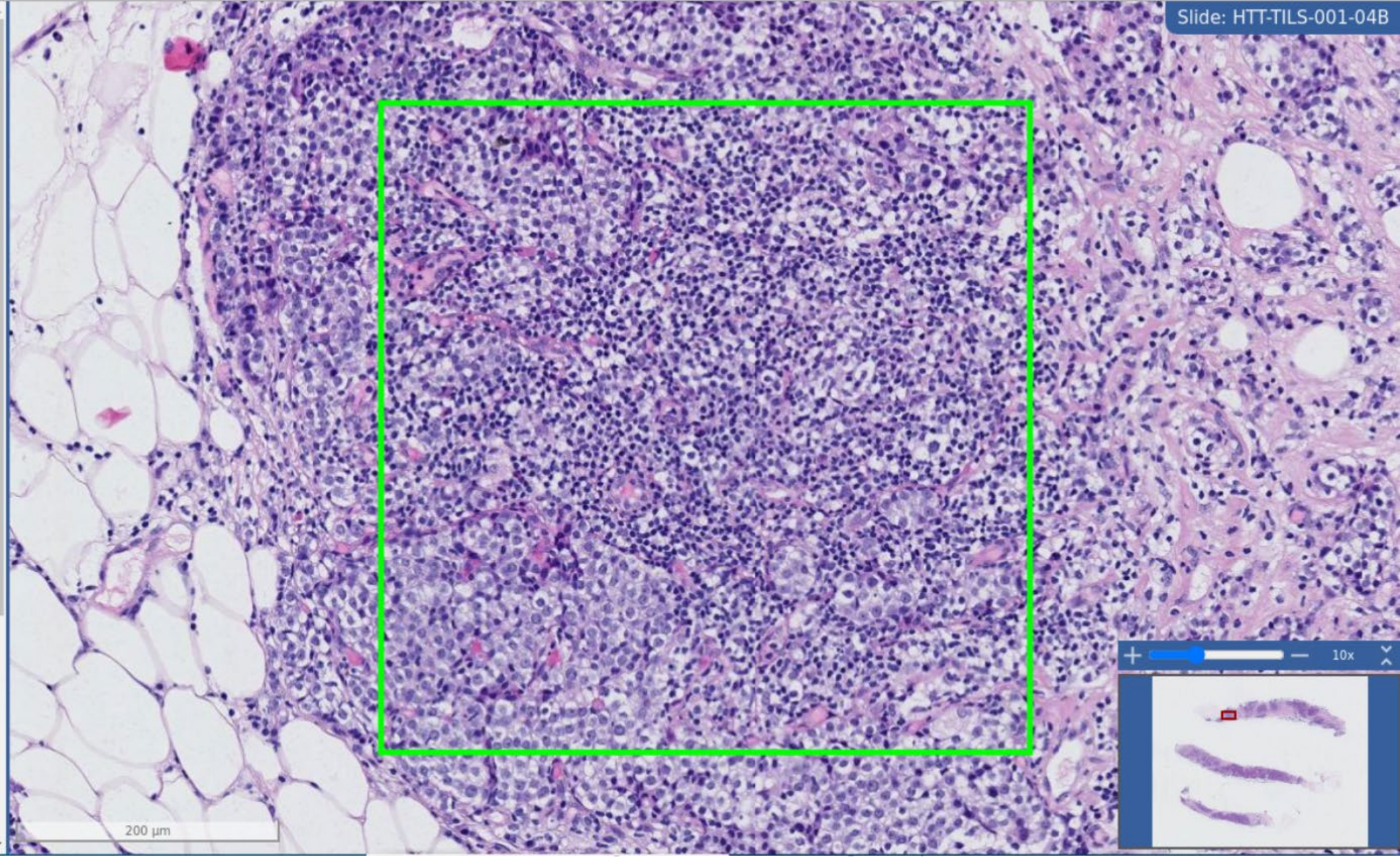
Expert Panel Annotations:

ROI Type	% Tumor-Associated Stroma	% sTIL Density
Evaluable	30	90
Evaluable	60	95
Evaluable	50	92
Evaluable	50	75
Evaluable	60	90
Evaluable	60	90

Mean Percent Stroma: 51.7

Mean sTIL Density: 88.7

Comments: A challenging case. The high density of lymphocytes results in difficulty determining whether the lymphocytes are located in stroma, or whether they infiltrate tumor cell nests. The presence of small blood vessels and small gaps between lymphocytes suggest the





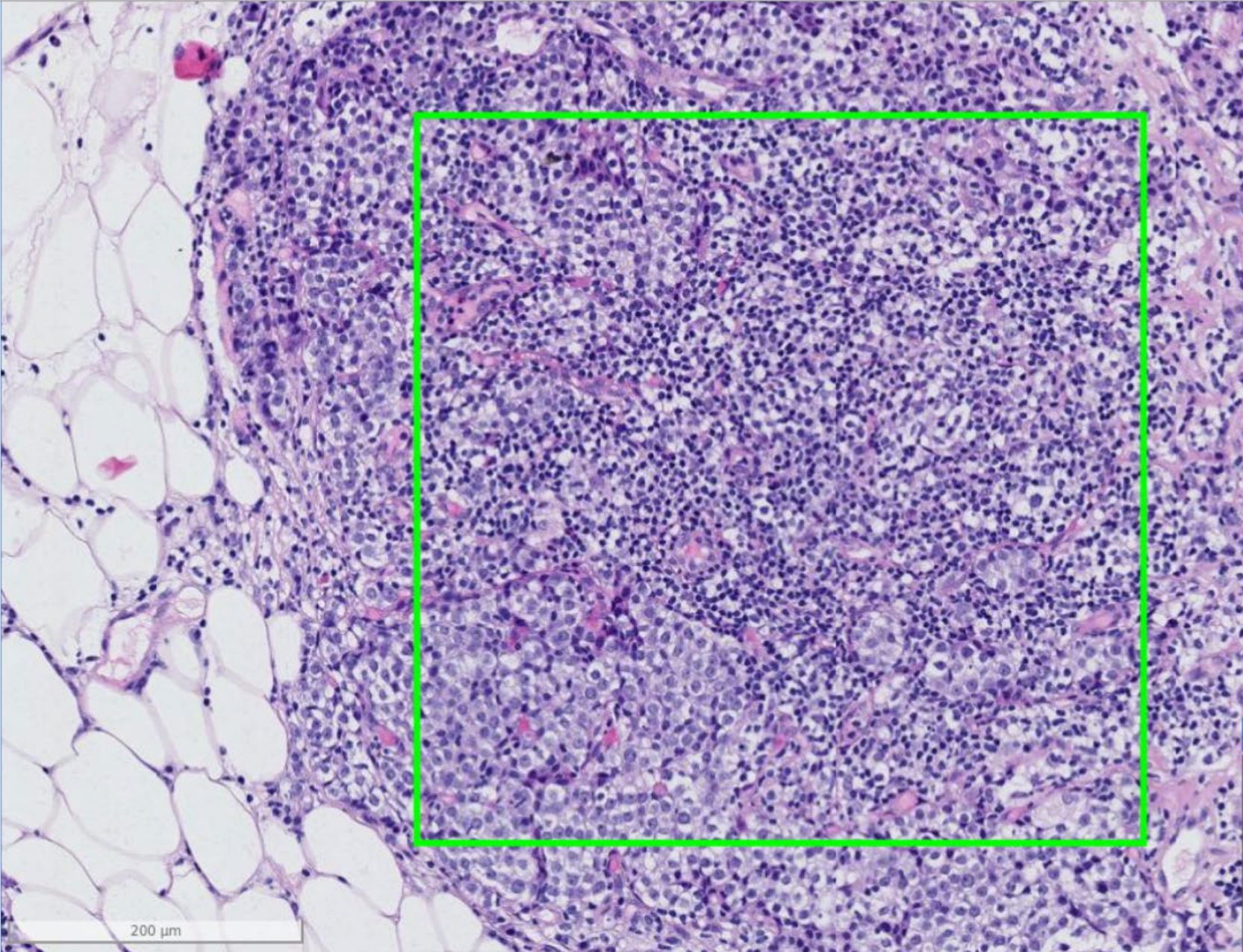
ROI Type	% Tumor-Associated Stroma	% sTIL Density
Evaluable	30	90
Evaluable	60	95
Evaluable	50	92
Evaluable	50	75
Evaluable	60	90
Evaluable	60	90

Mean Percent Stroma: 51.7
Mean sTIL Density: 88.7

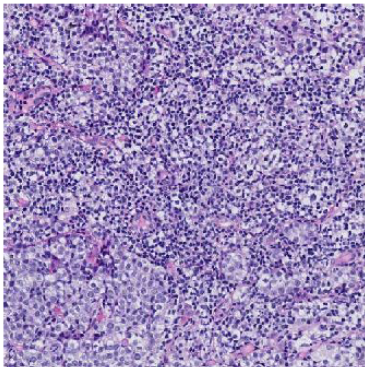
Comments: A challenging case. The high density of lymphocytes results in difficulty determining whether the lymphocytes are located in stroma, or whether they infiltrate tumor cell nests. The presence of small blood vessels and small gaps between lymphocytes suggest the lymphocytes reside within stroma. Occasional tumor cells with small nuclei (possibly degenerating) may be confused for lymphocytes.

Pitfalls: In regions where the sTILs density is very high, the underlying stroma may be obscured. Non-lymphocytes with small nuclei may be confused for lymphocytes.

Next



sTILs Reference Document => Proficiency Test



caseID: HTT-TILS-001-04B.ndpi_x24343.2190_y11775.2190

Expert Panel Annotations

ROI Type	Percent Tumor-Associated Stroma	sTILs Density
Evaluable	30	90
Evaluable	60	95
Evaluable	50	92
Evaluable	50	75
Evaluable	60	90
Evaluable	60	90

Mean Percent Tumor-Associated Stroma: 51.7
Mean sTILs Density: 88.7

Comments: A challenging case. The high density of lymphocytes results in difficulty determining whether the lymphocytes are located in stroma, or whether they infiltrate tumor cell nests. The presence of small blood vessels and small gaps between lymphocytes suggest the lymphocytes reside within stroma. Occasional tumor cells with small nuclei (possibly degenerating) may be confused for lymphocytes.

Pitfalls: In regions where the sTILs density is very high, the underlying stroma may be obscured. Non-lymphocytes with small nuclei may be confused for lymphocytes.

2

36 cases



- No feedback
- Pathologists will be scored against the experts
- HTT participants must pass the proficiency test

How should we determine ...



- If a crowd pathologist is an expert?
- If an AI/ML model is good enough?

- First thought

- Bland-Altman Plots
- Limits of Agreement (LOA)

- How do we incorporate multiple readers ... multiple experts?

STATISTICS IN BIOPHARMACEUTICAL RESEARCH
2022, VOL. 00, NO. 0, 1–10
<https://doi.org/10.1080/19466315.2021.2063169>



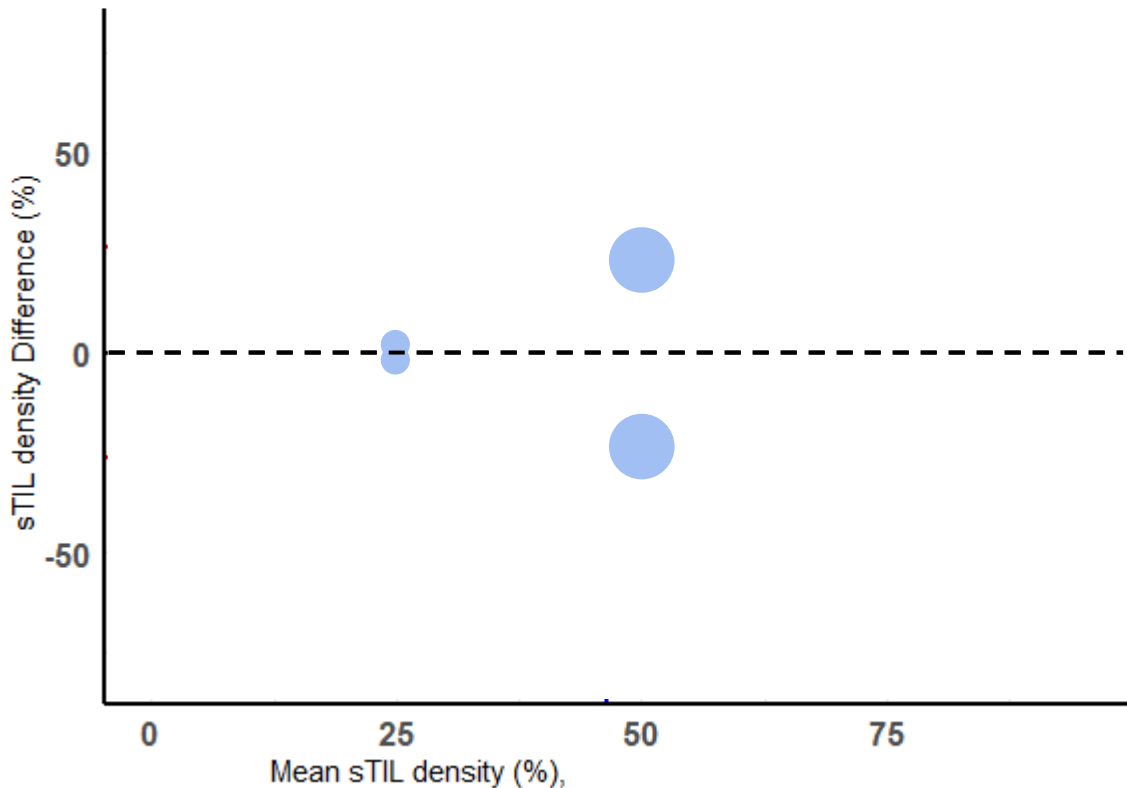
Check for updates

Three-Way Mixed Effect ANOVA to Estimate MRMC Limits of Agreement

Si Wen and Brandon D. Gallas

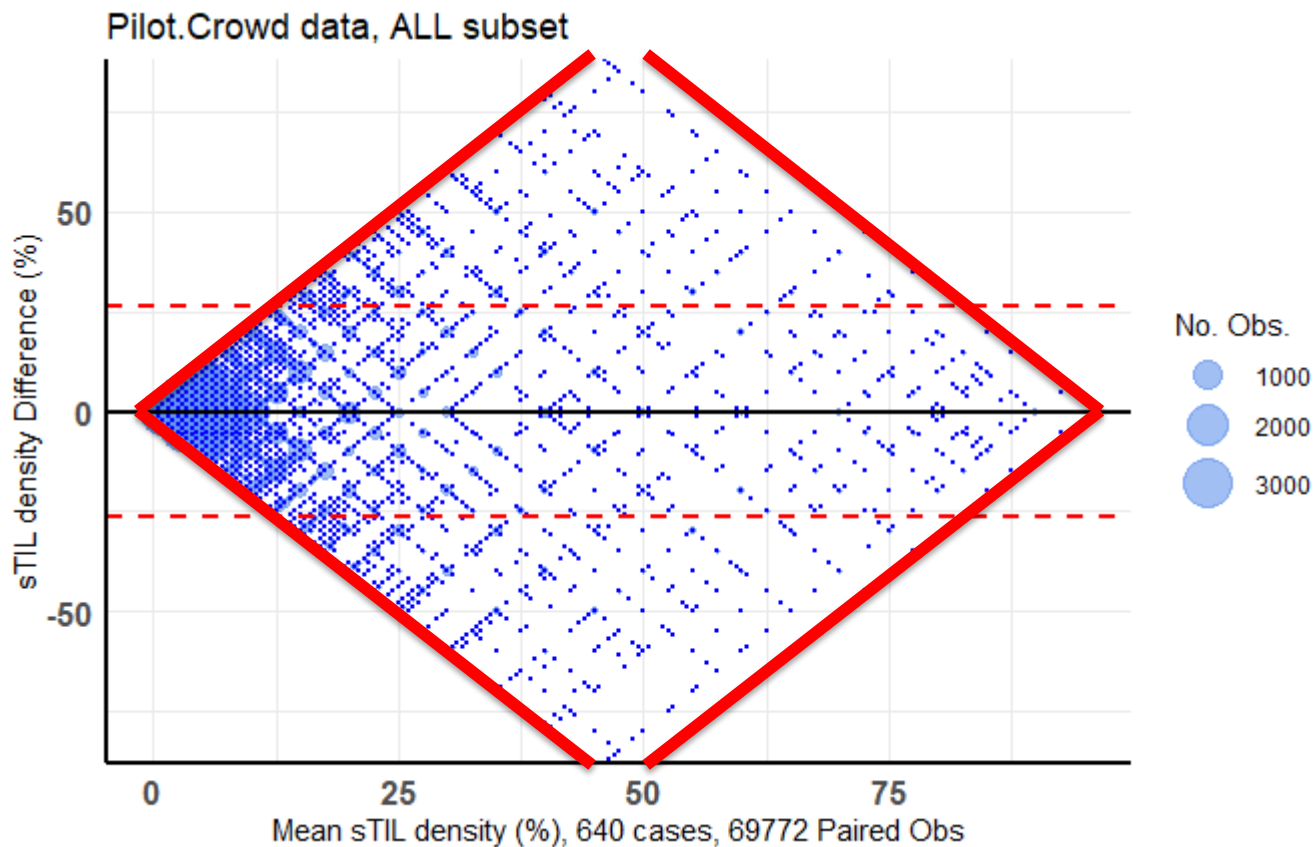
CDRH/OSEL Division of Imaging, Diagnostics, and Software Reliability, U.S. FDA, Silver Spring, MD

Mean Difference (Bland-Altman) Plots



- One point represents scores
 - Two readers
 - One ROI
- X-axis is the mean
- Y-axis is the difference
- Size of point scales with duplicates
 - Multiple readers or multiple ROIs
- Vertically symmetric by construction
 - Assume readers are equivalent
 - Include Difference: Reader 1 – Reader 2
 - Include Difference: Reader 2 – Reader 1

Mean Difference (Bland-Altman) Plots for **seven** pathologists with complete pilot data



- Differences not independent
 - Multiple readers, Multiple Cases
 - Fully-crossed data
- Differences not identically distributed
 - Differences increase with the mean and then decrease
- Differences not normally distributed
 - Cone of maximum possible difference

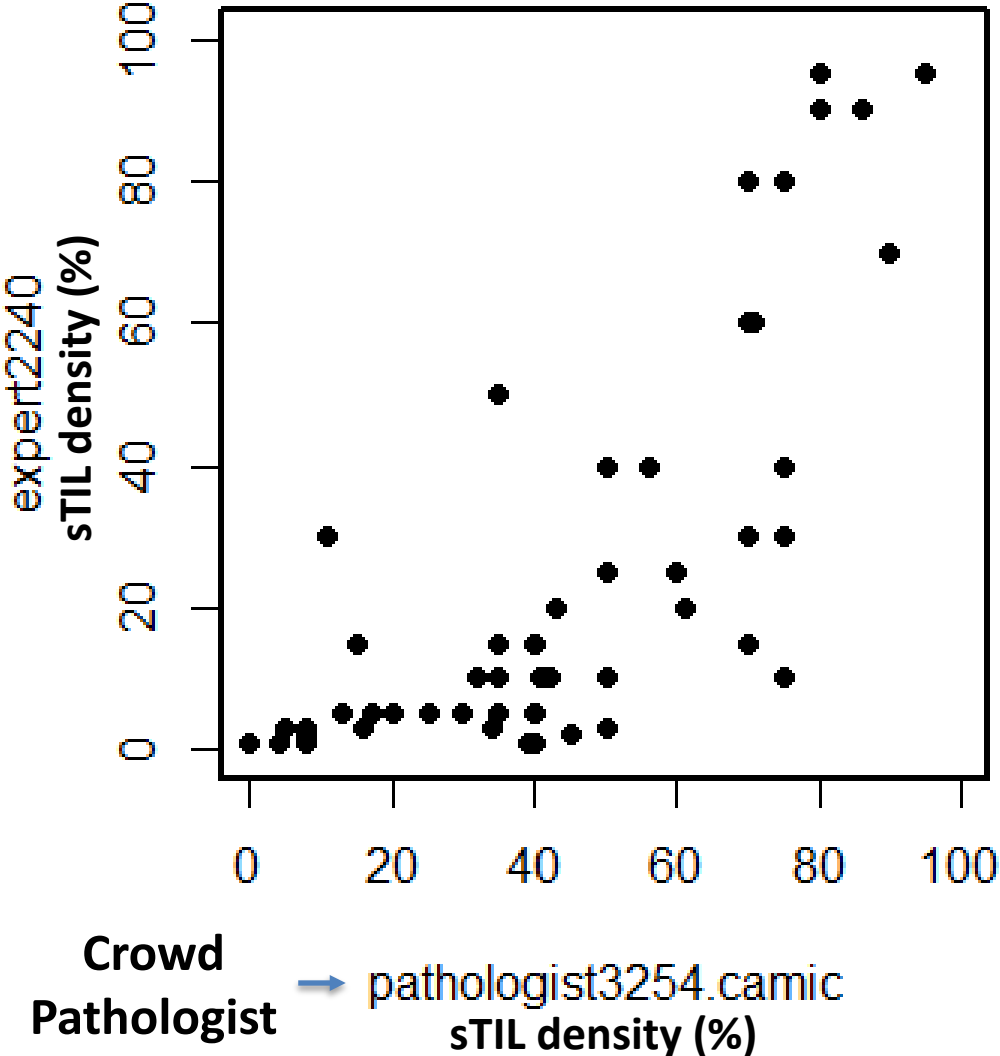
How should we determine ...

- If a pathologist is an expert?
- If an AI/ML model is good enough?

- First thought
 - Bland-Altman Plots
 - Limits of Agreement (LOA)

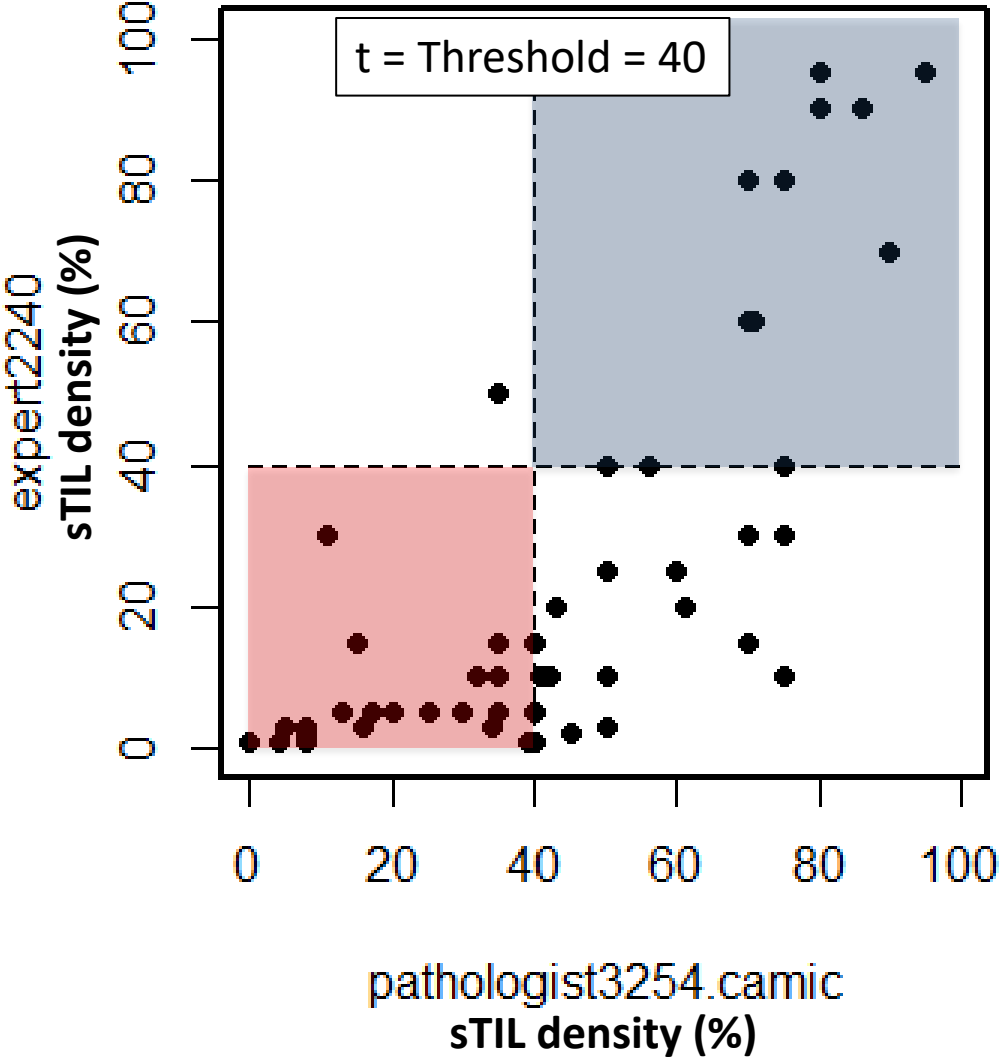
- Assumptions not satisfied ... Good for exploratory analysis
 - What next?

Crowd Pathologist vs. Expert, nObs = 59

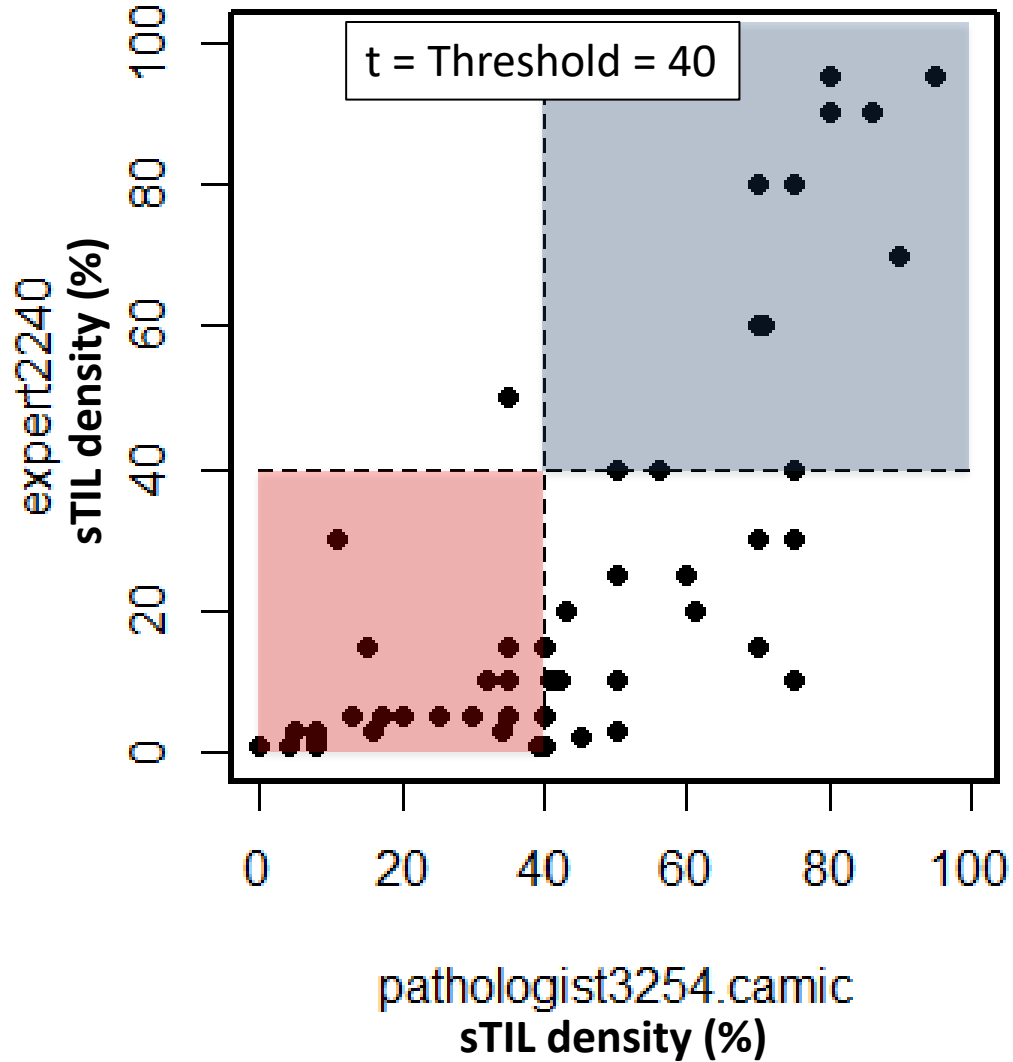


- Crowd pathologist
 - Typical data
 - Substitute “AI Model”
- SELECT data
 - 72 cases, some labeled not evaluable
- Not clustered around diagonal
- Not normally distributed

Crowd Pathologist vs. Expert, nObs = 59



Crowd Pathologist vs. Expert, nObs = 59



		crowd	
		≤ t	> t
Expert	> t	1	10
	≤ t	30	18

Crowd-Expert Agreement

threshold	expert	crowd
40	expert2240	pathologist3254.camic

		crowd		Row Total	Fraction Agree	Standard Error
		$\leq t$	$> t$			
Expert	$> t$	1	10	11	0.91	0.0867
	$\leq t$	30	18	48	0.63	0.0699

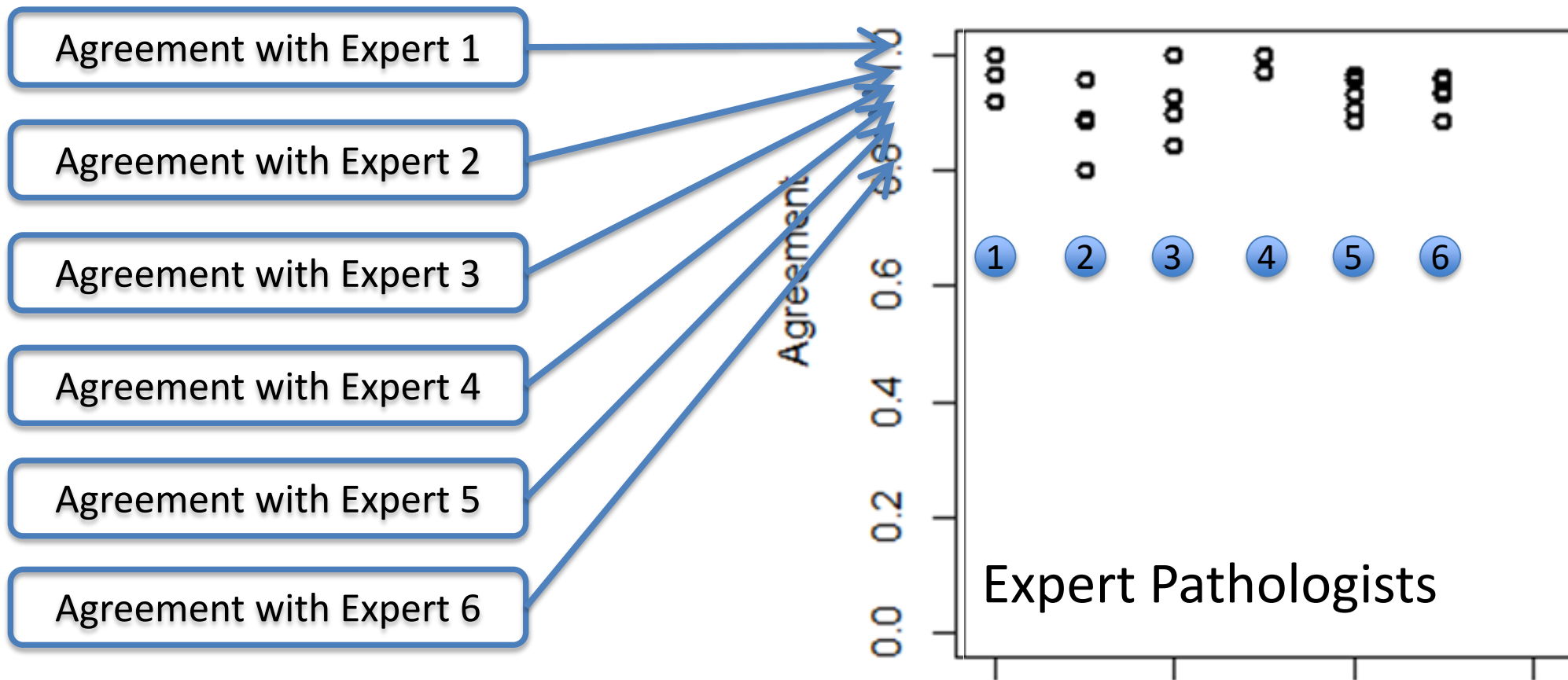
Crowd-Expert Agreement

threshold	expert	crowd
40	expert2240	pathologist3254.camic

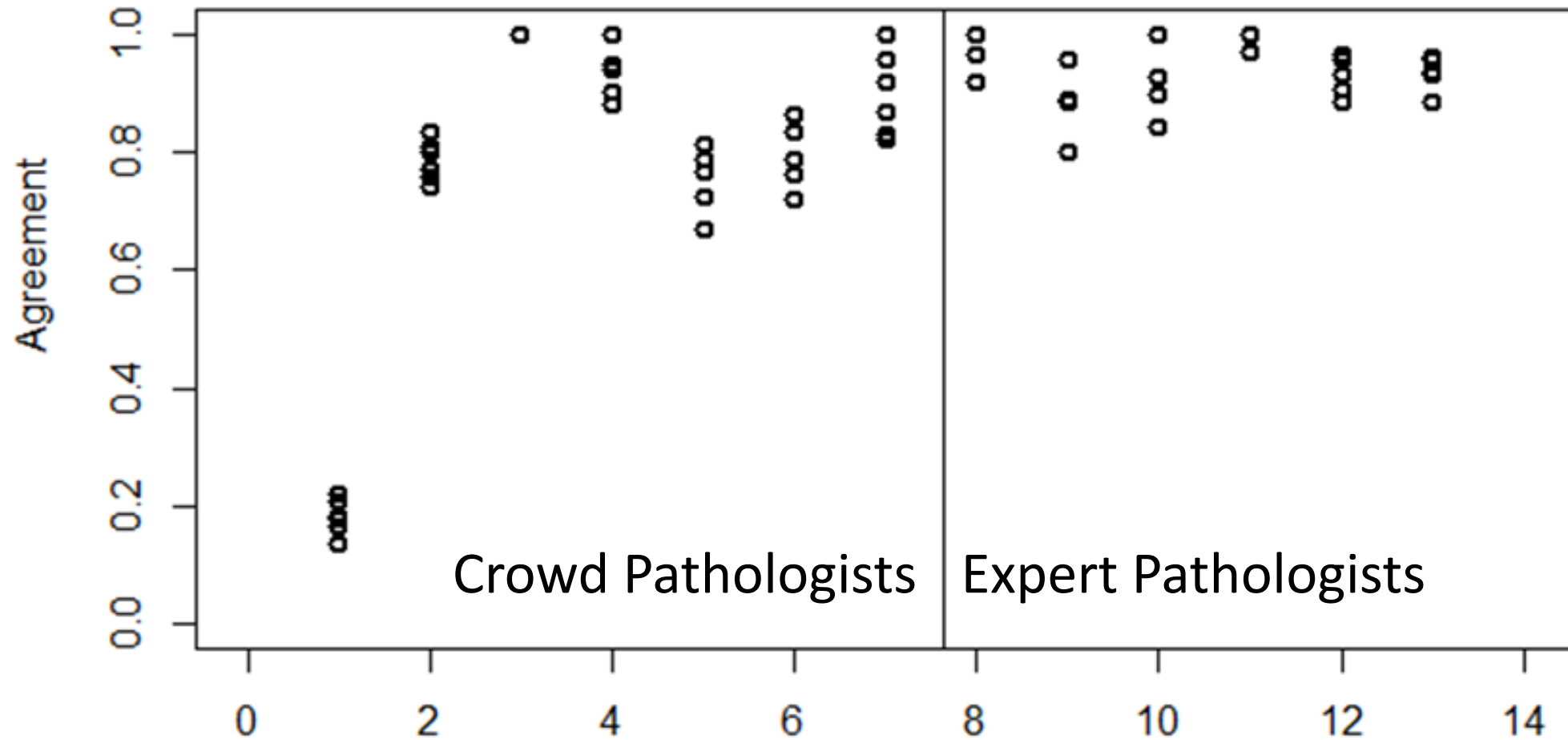
		crowd		Row Total	Fraction Agree	Standard Error
		$\leq t$	$> t$			
Expert	$> t$	1	10	11	0.91	0.0867
	$\leq t$	30	18	48	0.63	0.0699

- TPF = Fraction Agree “ $> t$ ”
- FPF = Fraction Agree “ $\leq t$ ”
- TPF and FPF
understood to be
- Crowd-Expert Agreement
- Compare Crowd to all Experts

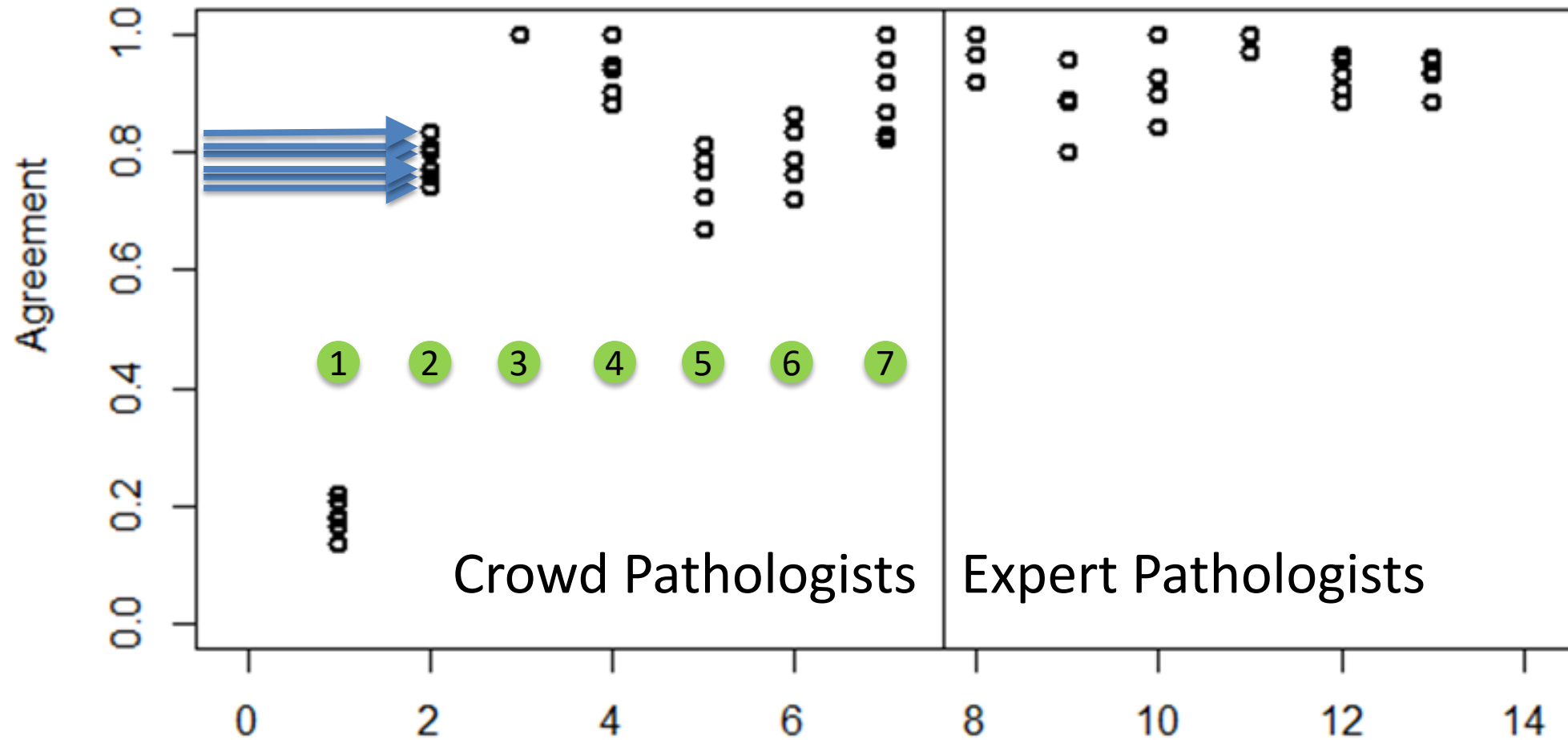
Expert vs. Expert, sTIL density ≤ 10



Agreement with Experts: sTIL density ≤ 10



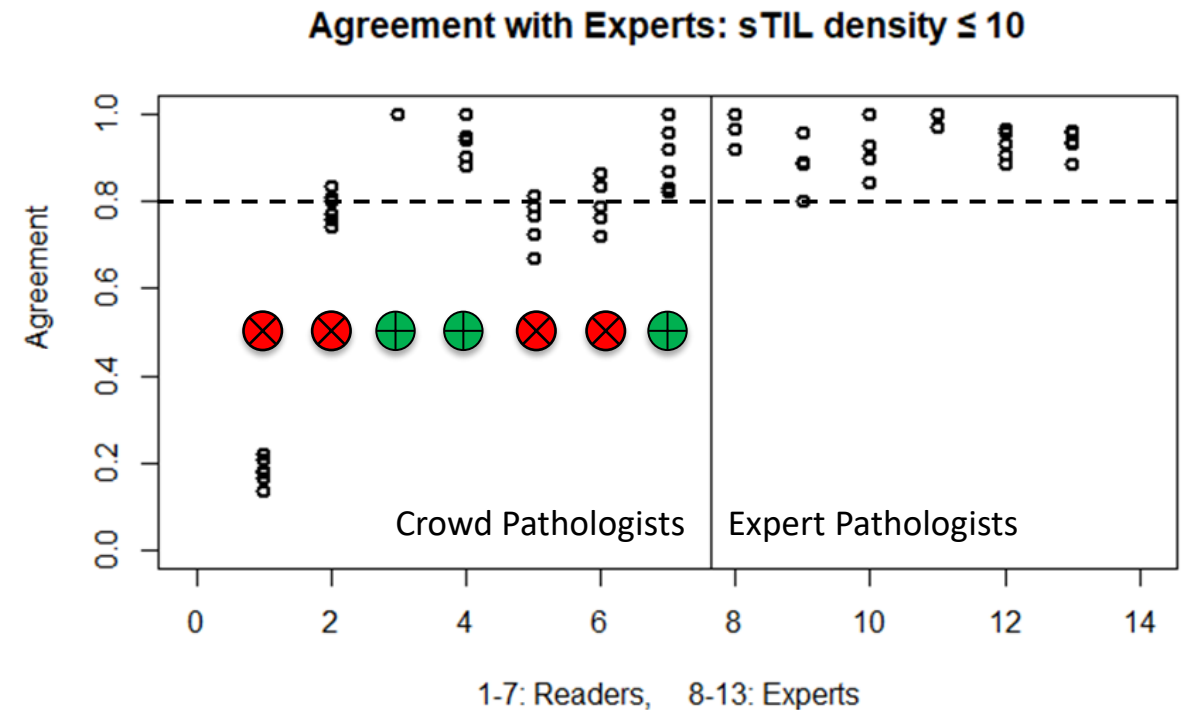
Agreement with Experts: sTIL density ≤ 10



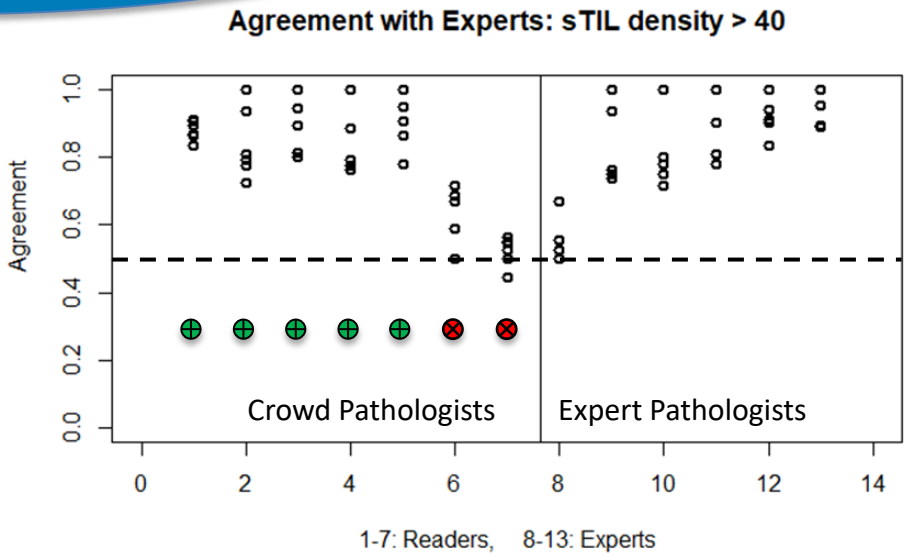
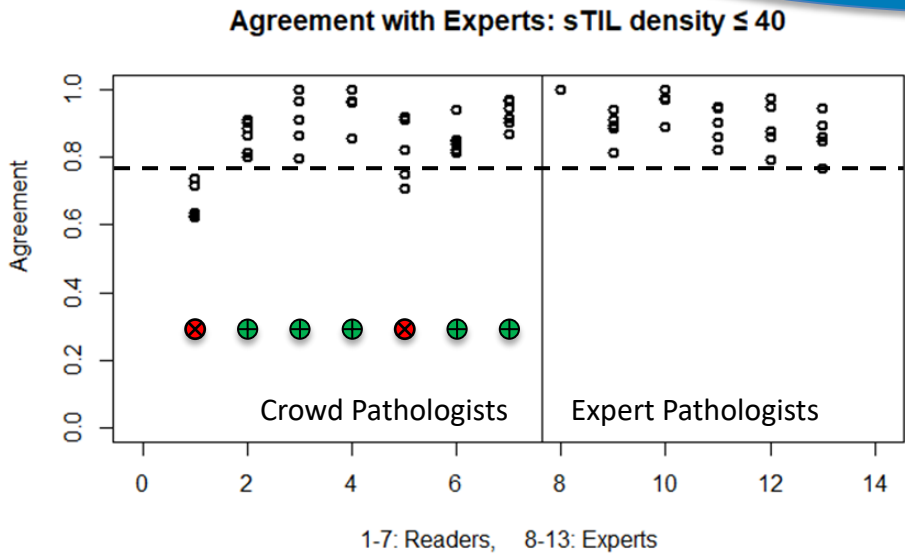
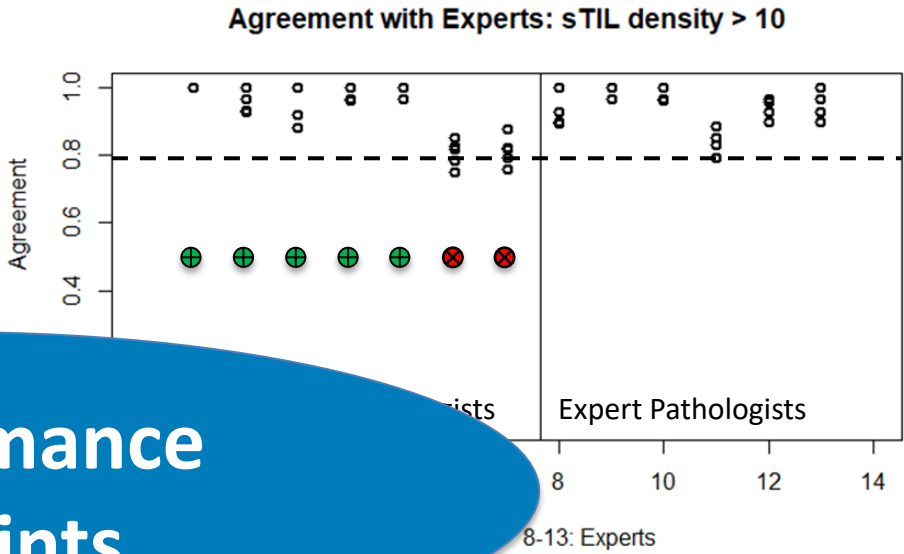
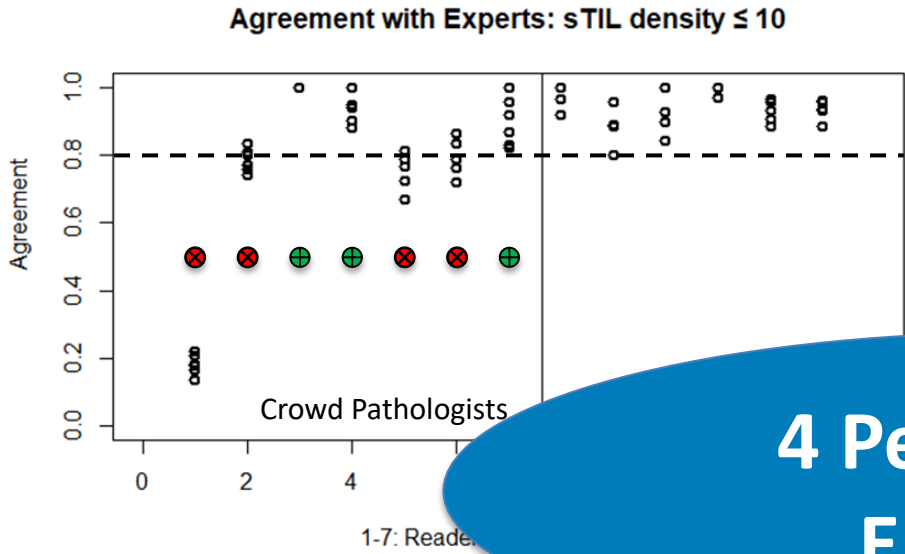
How should we determine ...



- If a pathologist is an expert?
- Current strategy for proficiency test
 - Crowd pathologist agreement with each expert must be greater than
 - Lowest expert-expert agreement for all pairs of experts
- Multiple performance endpoints
 - Add agreement above the threshold
 - Add agreement at additional thresholds
- This strategy does not immediately generalize beyond proficiency test.
 - As experts are added, lowest agreement decreases.

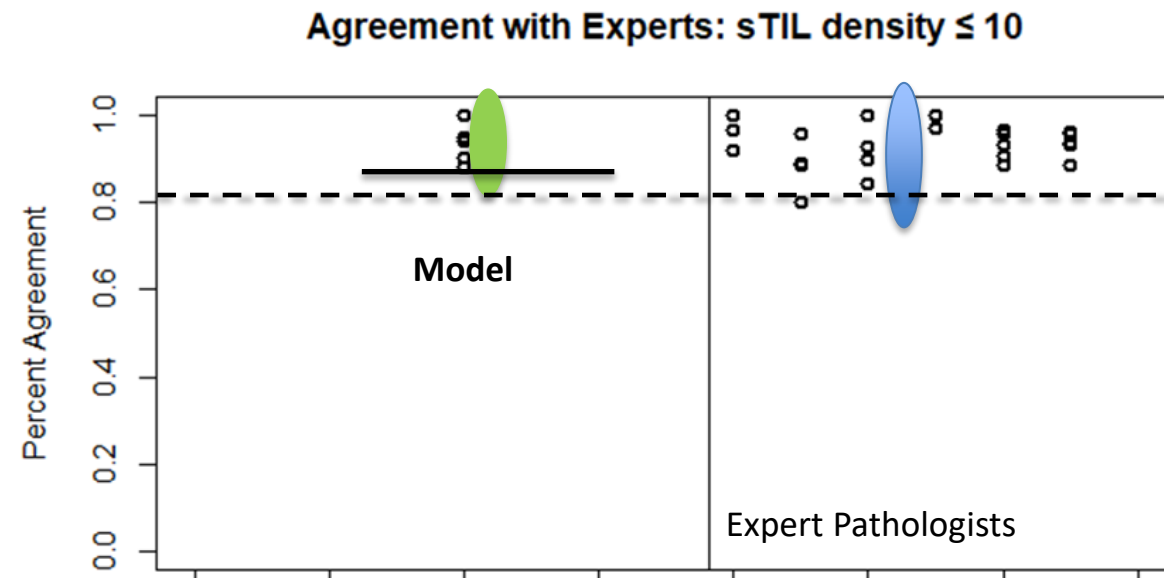


4 Performance Endpoints



How should we determine ...

- If an AI/ML model is good enough?
- *Under development*
- Multi-Expert Multi-Case (MEMC) analysis method
 - For each threshold and agreement above and below (multiple hypotheses) ...
 - Study result is the lower 97.5 percentile of model-to-expert agreement
 - Comparator is the lower 97.5 percentile of expert-expert agreement
- Need to account for expert and case variability and correlations



Summary and Thoughts

- A lot has been done. A lot still to do.
- Lessons learned from pilot study (and deep dive)
 - Pathologist variability can be significant
 - Pathologist variability can be reduced
 - Pathologist variability is not well behaved
 - Need tools to account for pathologist variability
- Tools (deliverables) from pilot study (and deep dive)
 - Pathologist training materials
 - Data to explore and model
 - Data-collection tools
- Lessons and tools broadly support
 - Biomarker validation
 - AI/ML model validation
 - Community is hungry for this research

Amplifying Tools (Deliverables)



Medical Device Development Tools



- Reduce burden to sponsors
 - Skip the design of the clinical trial
 - Know performance evaluation methods FDA will accept
 - Replace 40-70 pages of a submission with,
“We used the MDDT dataset and our algorithm performance was ...”
- Reduce burden to FDA
 - Qualify data and analysis methods once to support multiple sponsors

Regulatory Science Tool Catalog



- iMRMC software package
 - Software to do multi-reader multi-case analysis of reader studies

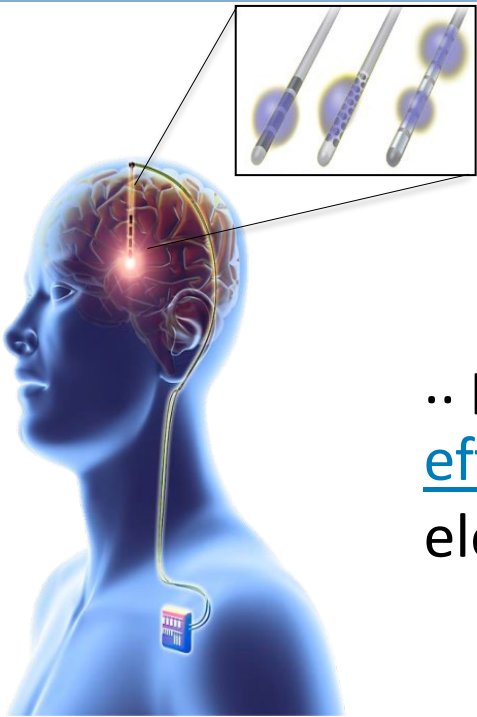


State of the Project

- Pivotal study slide and metadata sourcing
 - Huge effort
 - RCAs with 2 sites, one more in process
 - Received first batch of slides and metadata (n=86)
 - Target n=200
- Patient population not discussed
 - Triple-negative breast cancer biopsies
 - Metadata (demographic data, cancer stage, nuclear grade, ...)
- Statistical analysis plan under development
 - No peeking at pivotal study data
 - Results can impact target
- We are on the home stretch to launch the pivotal study
- ROI selection
 - Targeting pitfalls
 - Stratify by sTILs density
- Finalizing knowledge-based training
- Finalizing interactive training modules
- Moving one digital data-collection platform to precision FDA
 - Control
 - Security, trust

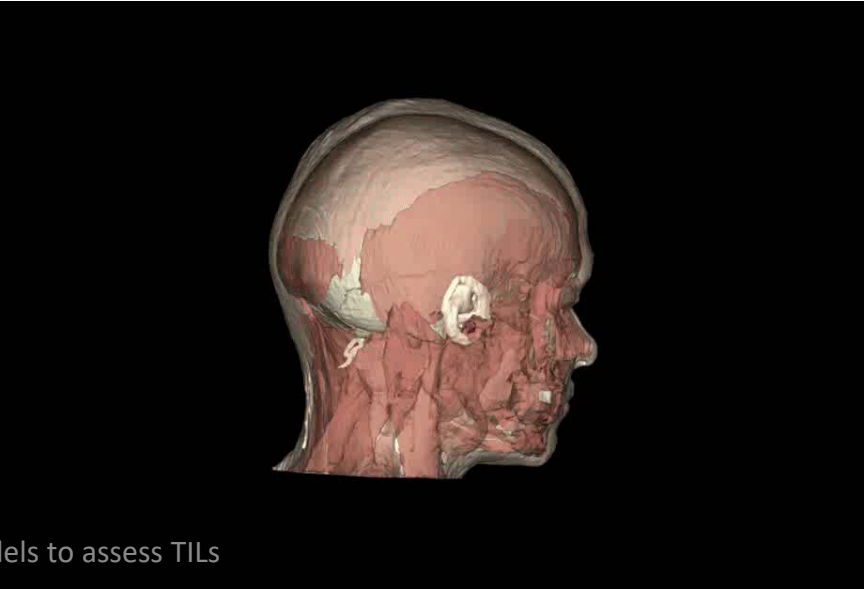
Title and content (black background)

CDRH Mission

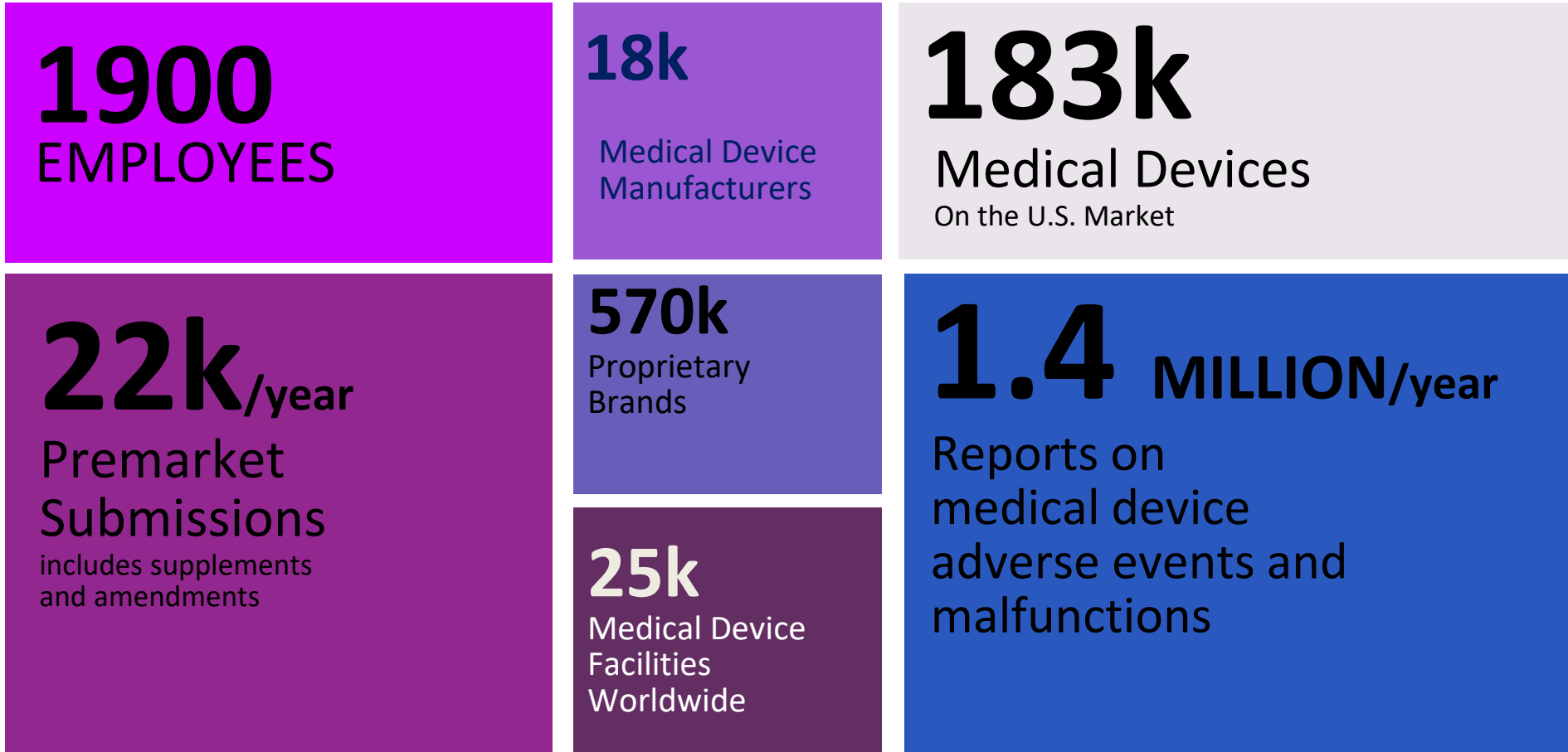


.. protect and promote the health of the public by ensuring the safety and effectiveness of **medical devices** and the safety of radiation-emitting electronic products...

We facilitate medical device innovation by advancing regulatory science, providing industry with predictable, consistent, transparent, and efficient regulatory pathways, and assuring consumer confidence in devices marketed in the U.S.



CDRH in Perspective



Office of Science and Engineering Laboratories (OSEL)

- Conduct laboratory-based regulatory research to facilitate development and innovation of safe and effective medical devices and radiation emitting products
- Provide scientific and engineering expertise, data, and analyses to support regulatory processes
- Collaborate with colleagues in academia, industry, government, and standards development organizations to develop, translate, and disseminate science and engineering-based information regarding regulated products
- <https://www.fda.gov/about-fda/cdrh-offices/office-science-and-engineering-laboratories>

OSEL in Perspective

183
FEDERAL EMPLOYEES
Up to 180 visiting scientists

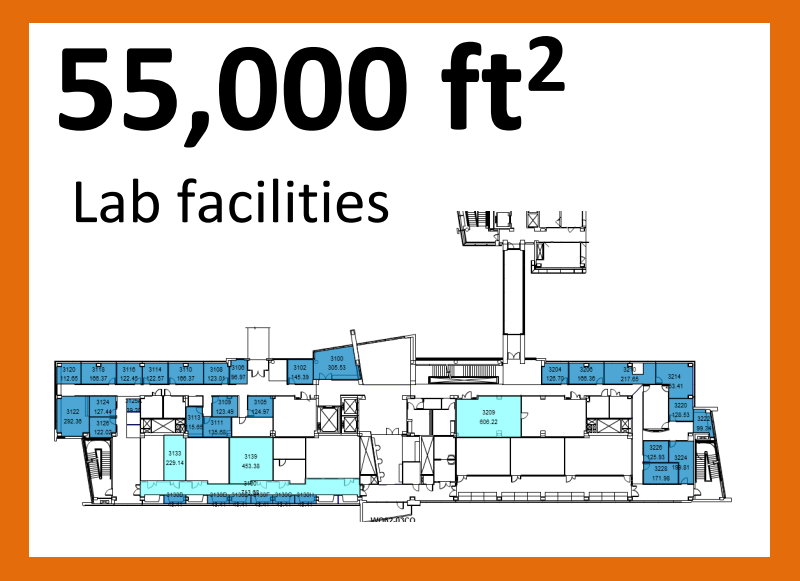
140 Projects
In 27 Laboratories
and Program
Areas

400/year
Peer reviewed presentations,
articles, and other public disclosures

2,500k/year
Premarket
Regulatory consults

75
Standards and
conformity
assessment
committees

70%
Staff with post
graduate degree



Division of Imaging, Diagnostics and Software Reliability (DIDSR)



- Develop least burdensome approaches for regulatory evaluation of imaging and big-data devices
 - Efficient clinical trials accounting for reader variability, simulation tools, in silico phantoms and imaging trials, addressing issues related to imperfect / missing reference standards, and limited data for training/testing of machine classifiers
- Develop measures of technical effectiveness of imaging and big-data technologies
 - Phantoms, laboratory measurements, computational models

DIDSR in Perspective

35

FEDERAL EMPLOYEES
14 Fellows/Students
3 Open Staff Positions

145/year

Peer reviewed articles, code and presentations

4 Program Areas

- AI/ML
- Medical Imaging and Diagnostics
- Digital Pathology
- Mixed Reality (AR/VR/XR)

550/year

Premarket
Regulatory consults

~15,000 ft²

DIDSR Lab and facilities

