Collecting and Annotating Digital Pathology Images to Assess Computational Pathology

Brandon D. Gallas

Division of Imaging, Diagnostics, Software Reliability

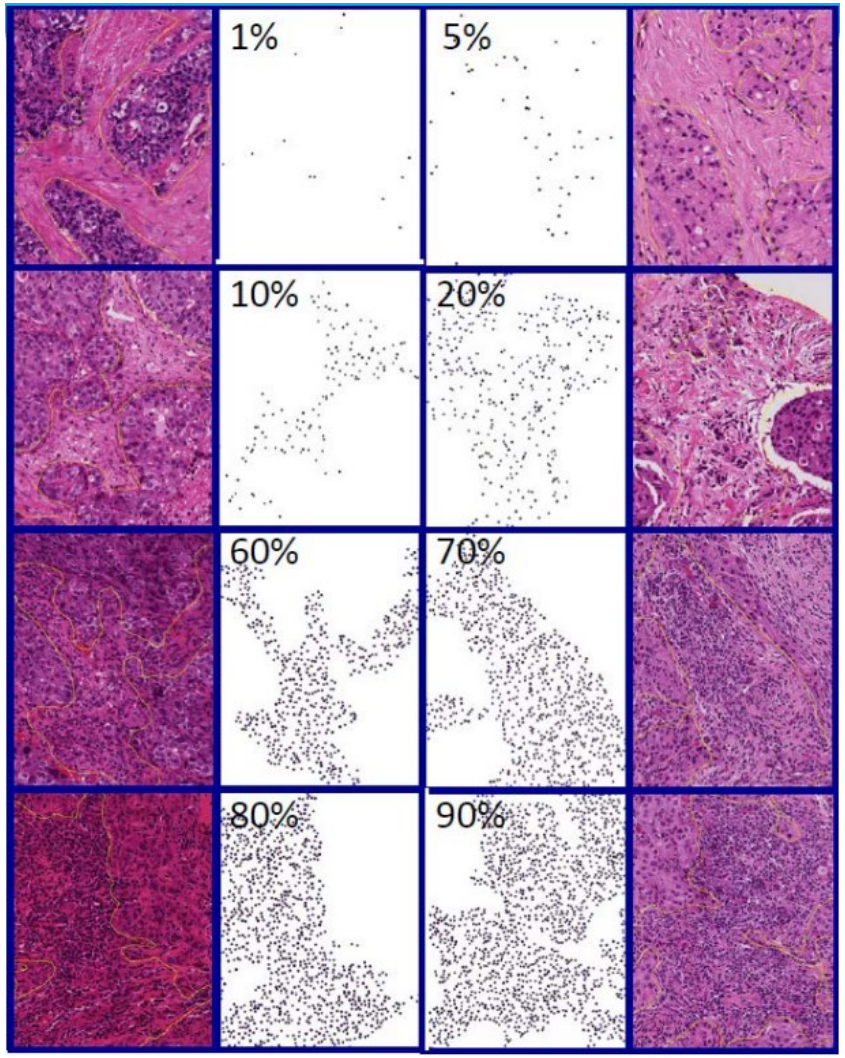
Office of Science and Engineering Laboratories
Center for Devices and Radiological Health
U.S. Food and Drug Administration

From Mammograms to Microwaves
https://www.fda.gov/about-fda/fda-organization/center-devices-and-radiological-health

Outline

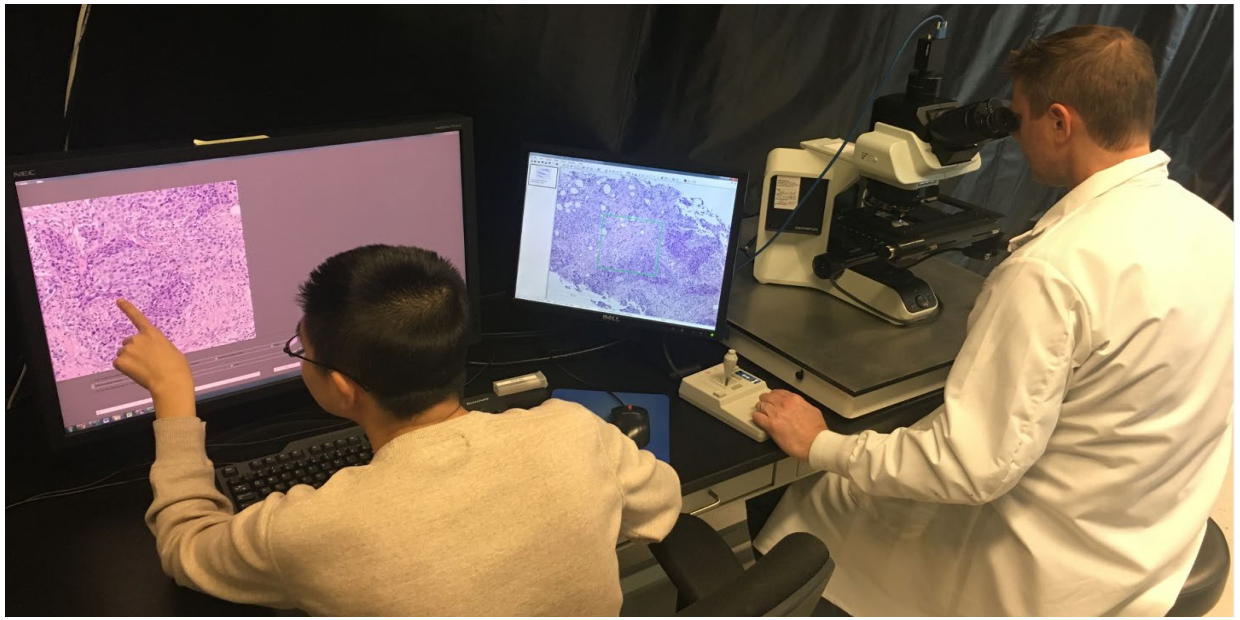
- HTT: High Throughput Truthing Project
- Training Materials
- Website
- Publications
- Related Activities

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



- Tools for AI-enabled Software Devices
  - Reference standard data set from pathologists
  - Data-collection methods and platforms
  - Methods to validate a quantitative algorithm

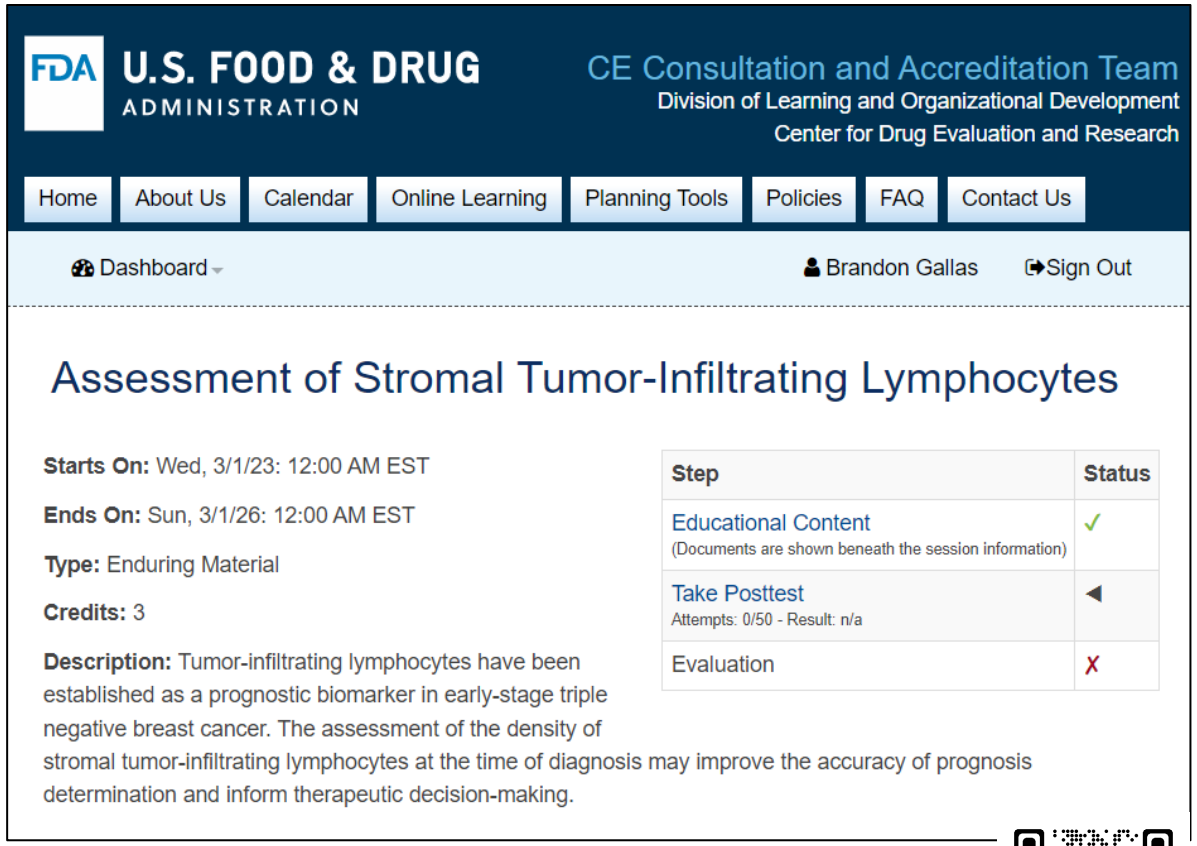
CME Course:
Assessment of Stromal Tumor-Infiltrating Lymphocytes

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

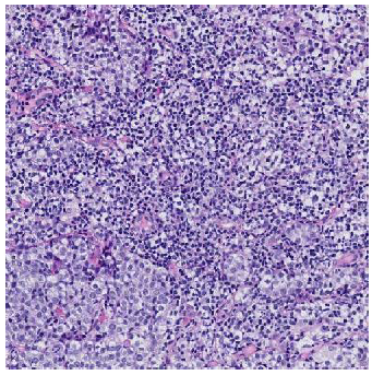
Victor Garcia, MD
Amy Ly, MD
Matthew Hanna, MD
Dieter Peeters, MD, PhD
Roberto Salgado, MD, PhD
Xiaoxian Li, MD, PhD
Kim Blenman, PhD, MS
Katherine Elfer, PhD, MPH
Bruce Werness, MD
Anna Ehinger, MD
Brandon Gallas, PhD



https://ceportal.fda.gov/

192 participants

sTILs Reference Document and 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

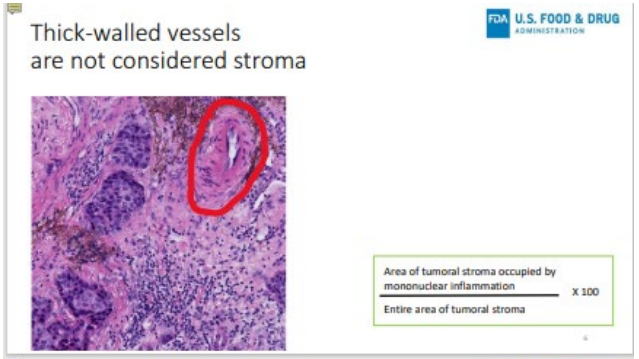
Mean Percent Tumor-Associated Stroma: 51.7

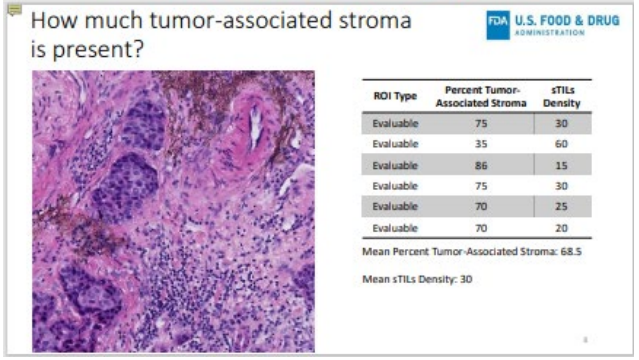
Mean sTILs Density: 88.7

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.

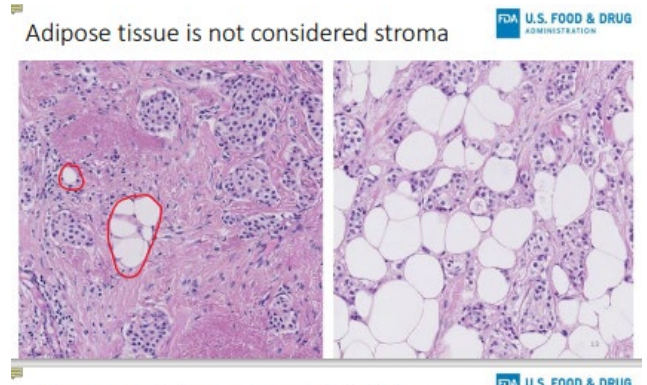
Pitfalls: In regions where the sTILs density is very high, the underlying stroma may be obscured. Non-lymphocytes with small nuclei may be confused for lymphocytes.

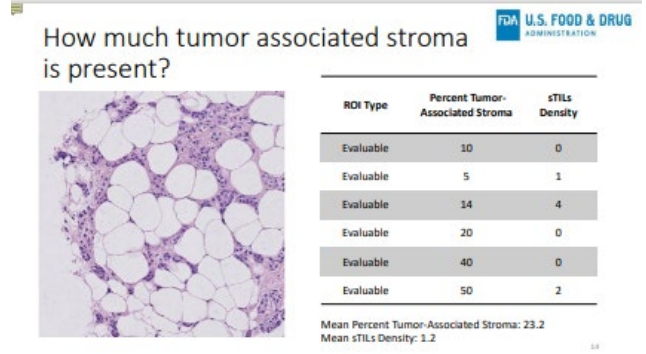
2





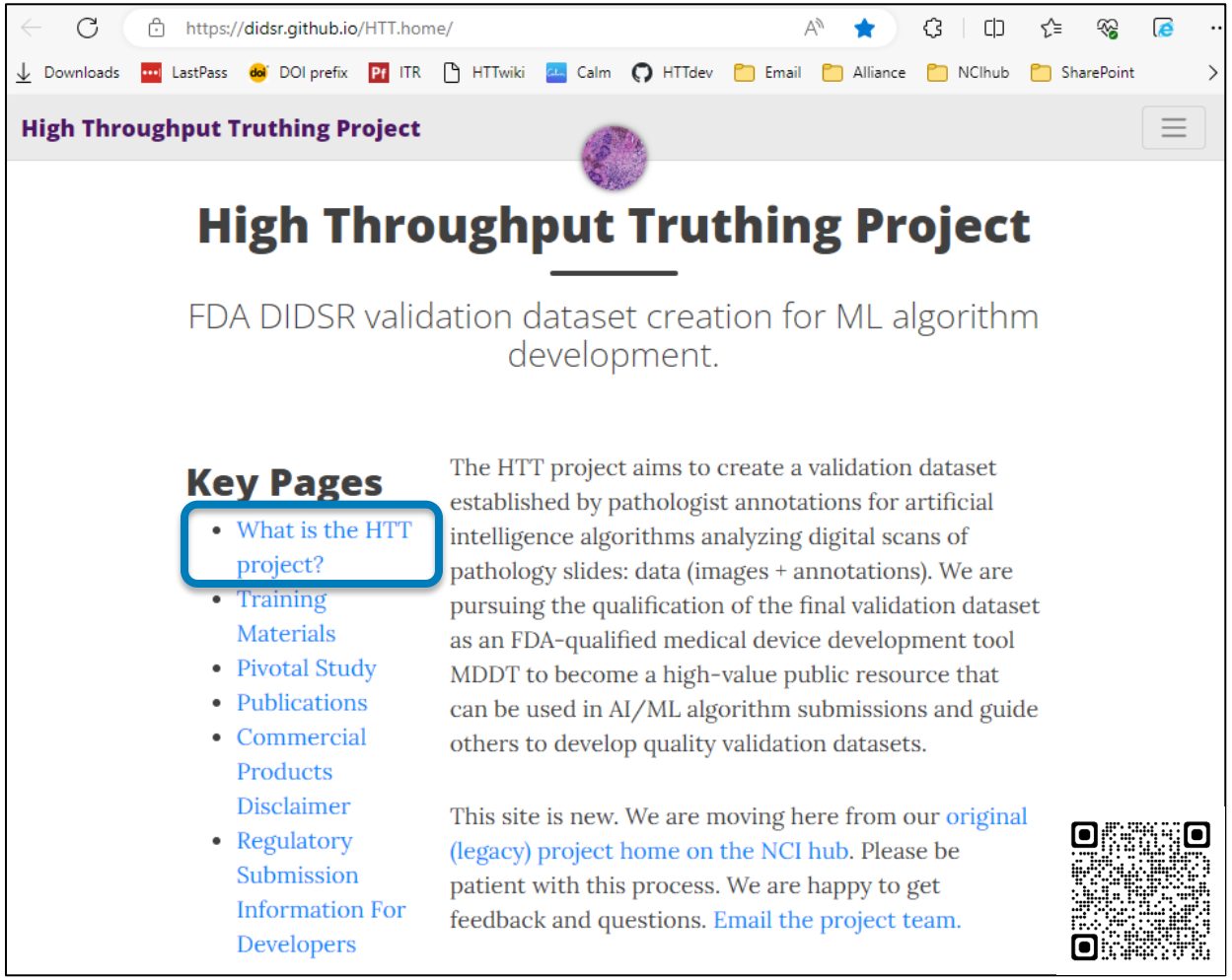
Example Pitfalls





HTT project home

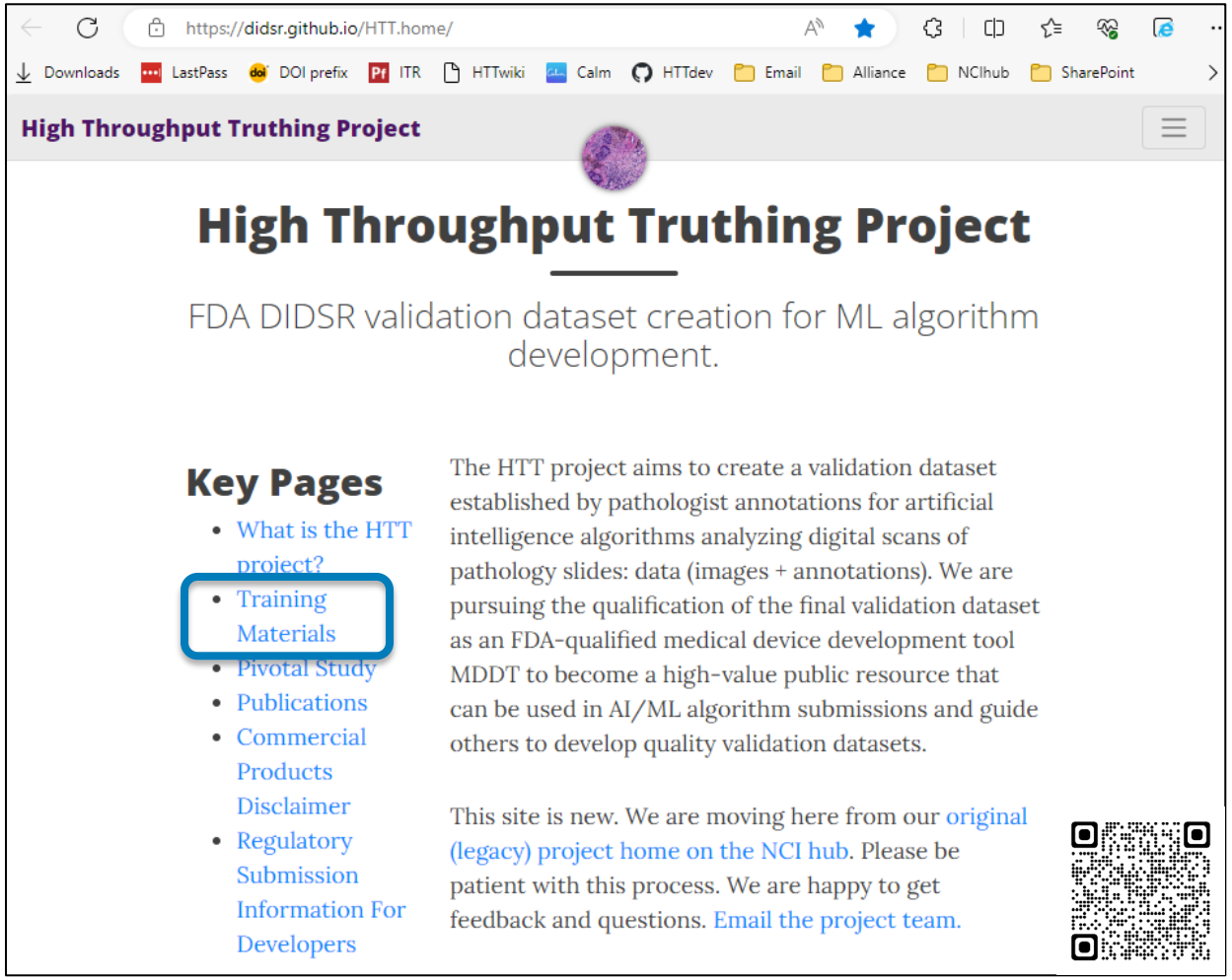
- https://didsr.github.io/HTT.home/



https://didsr.github.io/HTT.home/

HTT project home: Training Materials

- CME Course Info and Mirror
- Interactive Training
  - Test with feedback
  - Proficiency test
  - Example test report
  - Reference document
    (feedback test content)

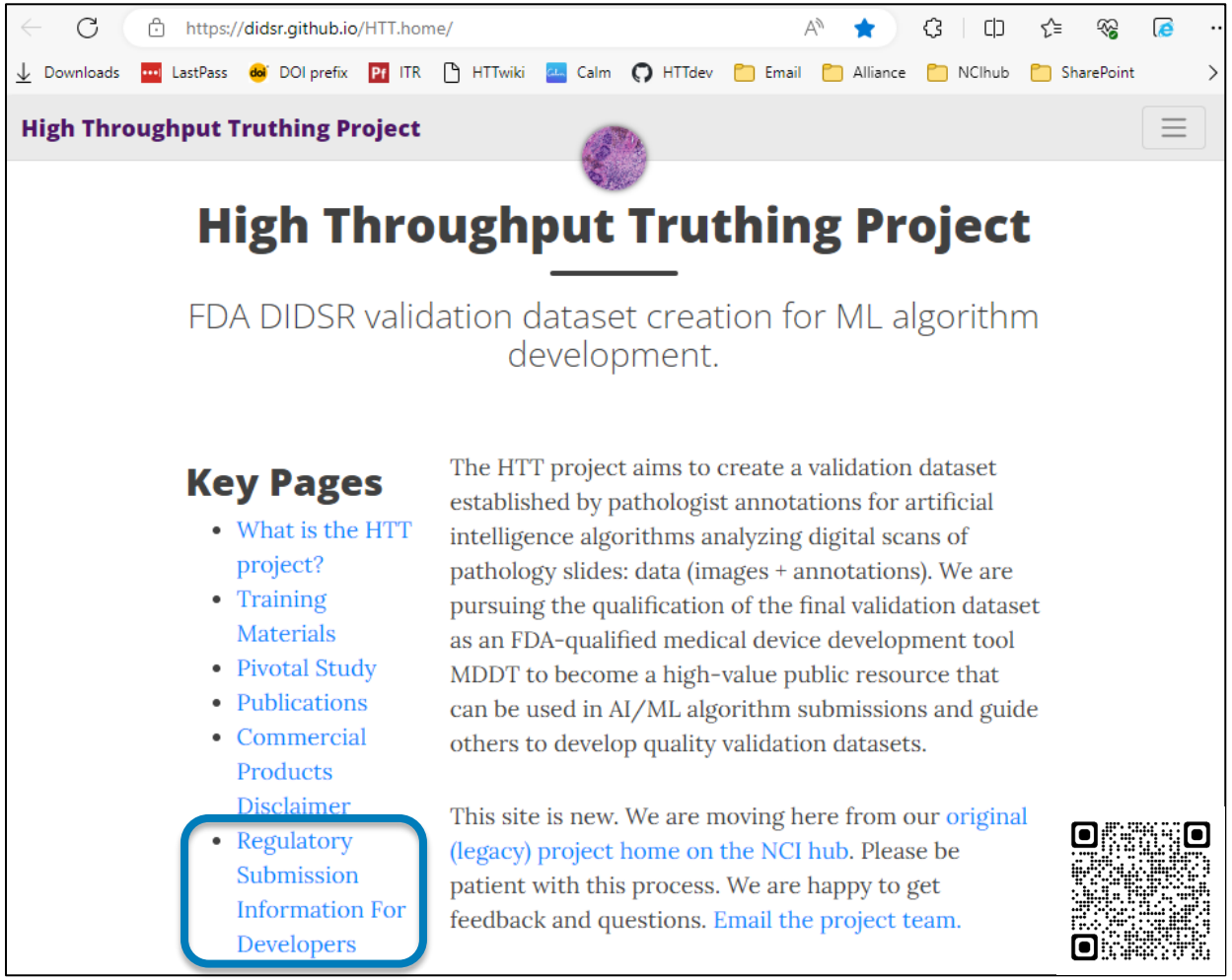


https://didsr.github.io/HTT.home/

HTT project home: Regulatory Submission Information

Links to information for Developers of Medical Imaging Algorithms

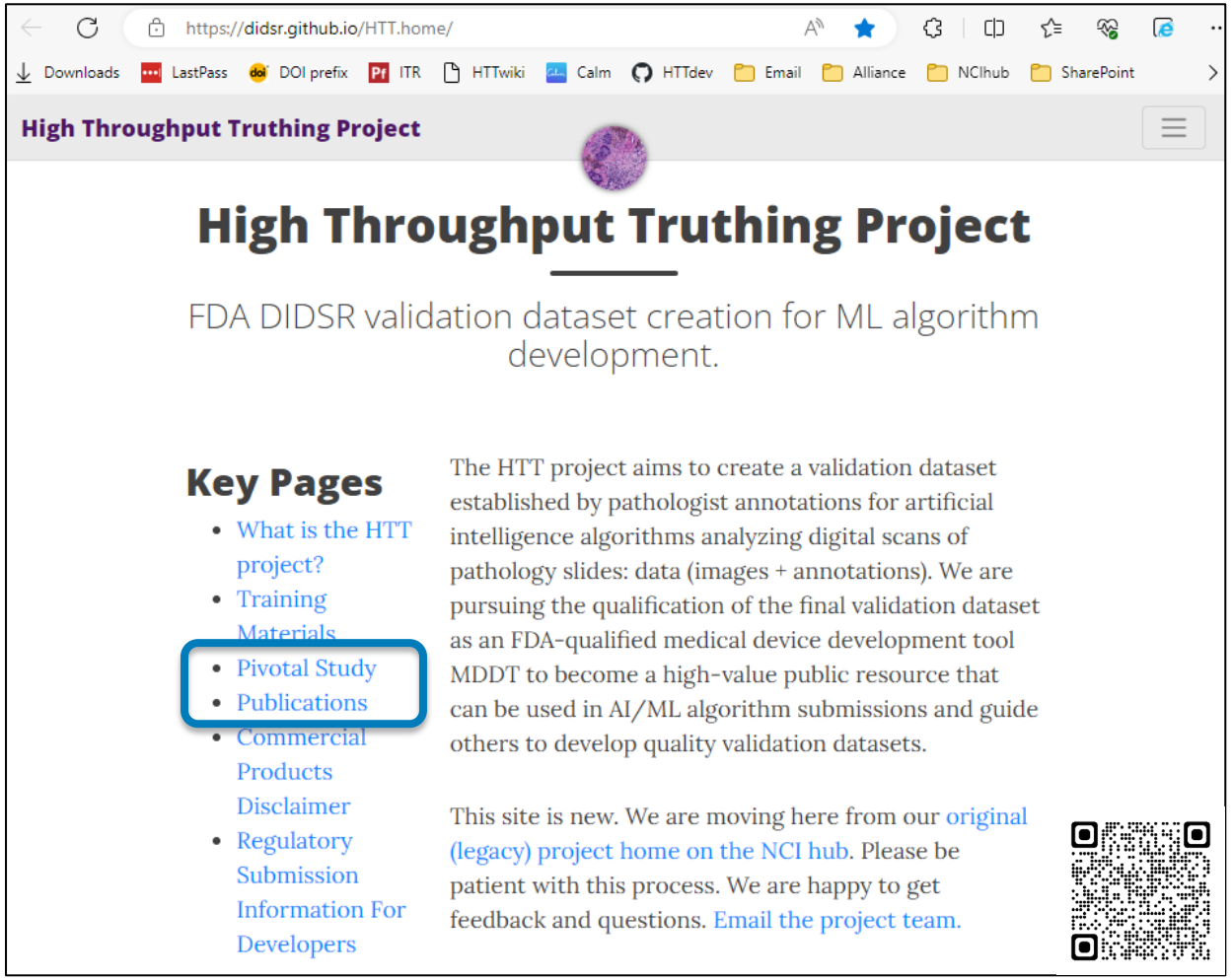
- Basic regulatory pathways
- Related guidance documents
- Decision summaries of related devices
- Presentations by FDA scientists



https://didsr.github.io/HTT.home/

HTT project home: Project Information

- Pivotal study entry point
  - Recruiting pathologists to be the reference standard
- Related publications

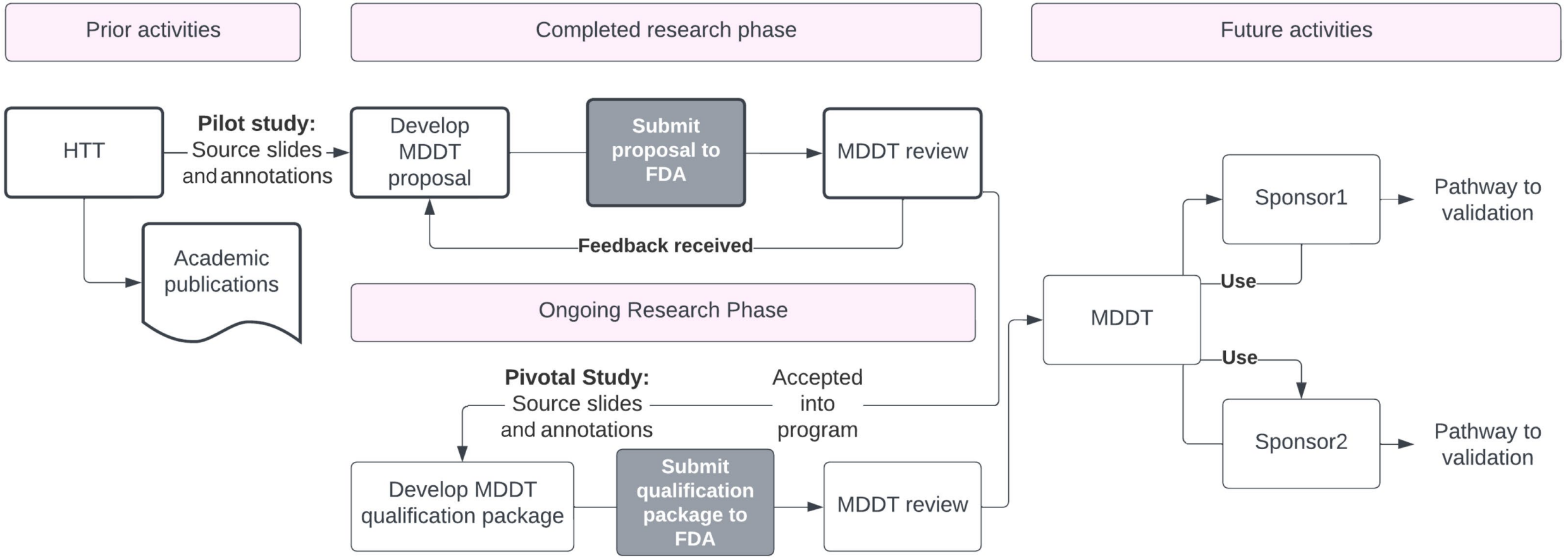


https://didsr.github.io/HTT.home/

"Initial interactions with the FDA on developing a validation dataset as a medical device development tool,"

S. Hart et al. (2023), Journal of Pathology, Vol. 261, p. 378-384

https://doi.org/10.1002/path.6208



"Training pathologists to assess stromal tumor infiltrating lymphocytes in breast cancer synergizes efforts in clinical care and scientific research"

A. Ly et al. (2023), Histopathology

https://doi.org/10.1111/his.15140

ROI Type:

Evaluable for sTILs

Not Evaluable for sTILs

50%
% Tumor-Associated Stroma

5%
sTILs Density

Expert Panel Annotations:

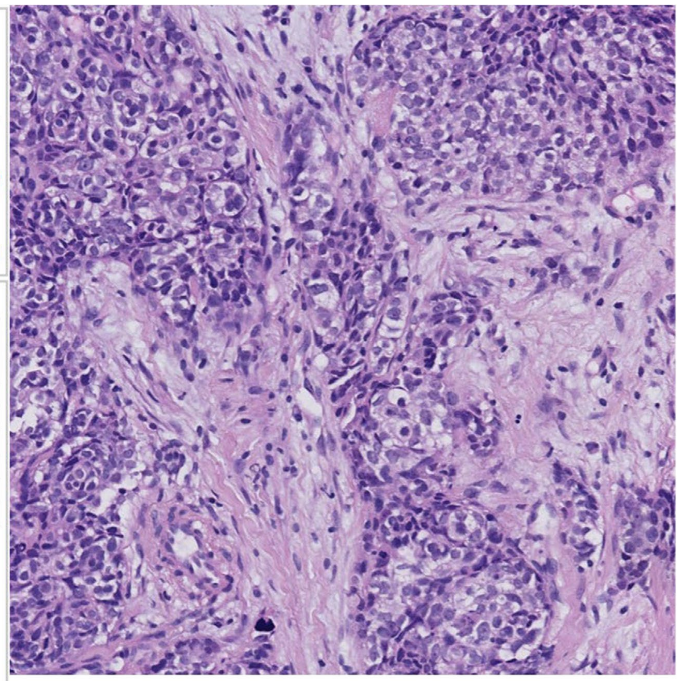
ROI Type	% Tumor-Associated Stroma	% sTIL Density
Evaluable	30	5
Evaluable	40	9
Evaluable	50	7
Evaluable	50	3
Evaluable	40	1
Evaluable	50	5

Mean Percent Stroma: 43.3

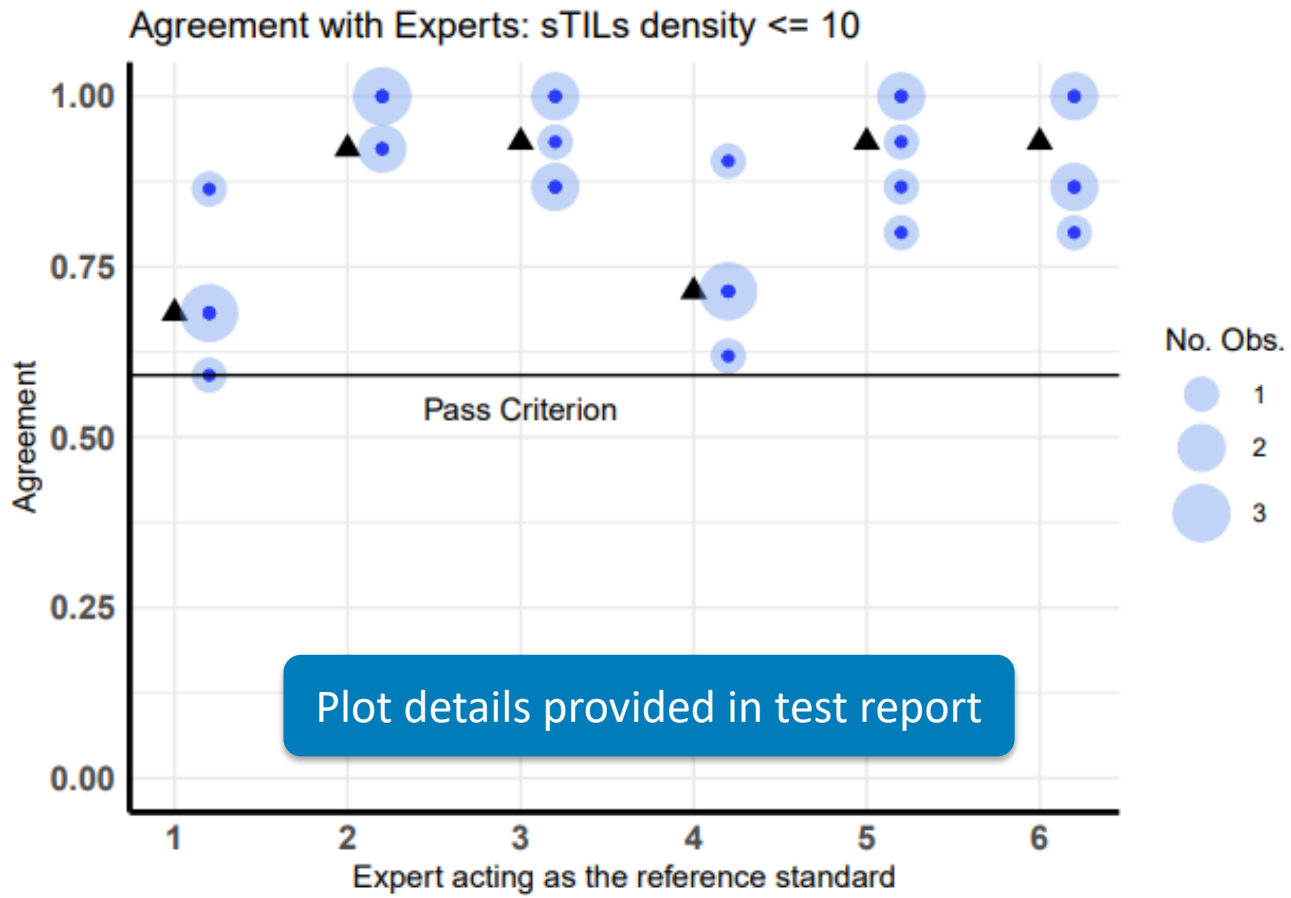
Mean sTILs Density: 5

Comments: It is difficult to distinguish between fibroblasts and sTILs in this case. The cells in the middle of the ROI are a bit wider than the other cells, so they probably are cancer cells that have artifact as a result of tissue processing. Though strong suspicion for a cancer cell, it could be a macrophage, which we see after treatment, and expect that an algorithm will have difficulty making this distinction on H&E stain.

Pitfalls: Non-lymphocytes may be confused for lymphocytes if there is tissue fixation artifact. Axially sectioned fibroblasts may be mistaken for lymphocytes.



- Region of Interest
- Annotation Fields
- Expert Panel Annotations
- Comments
- Pitfalls

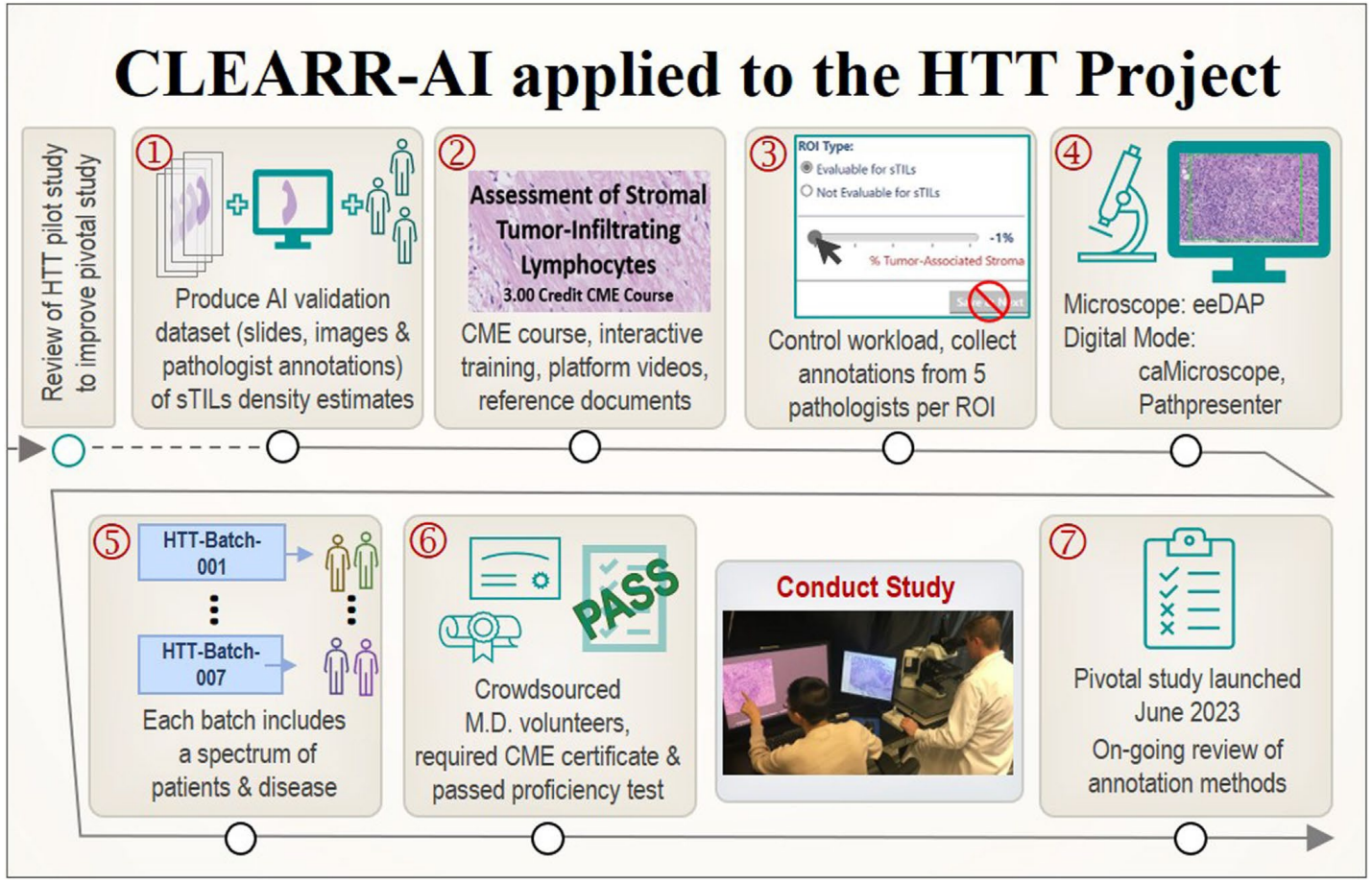


Plot details provided in test report

"Reproducible Reporting of the Collection and Evaluation of Annotations for Artificial Intelligence Models"

K. Elfer et al. (2024), Modern Pathology, Vol. 37, Issue 4, p. 100439

https://doi.org/10.1016/j.modpat.2024.100439



Study Component	Explanation	Origin
1. Objectives	Project objectives, dataset use case (training, testing models), degree of annotation (patient, image, ROI, or feature) and patient population.	Wahab *et al.* 1. Objectives and 2. Diagnostic/ Prognostic Algorithm (modified) and 7. Degree of Annotation
2. Data Dictionary	Training materials and reference documents(s) describing image features, anatomic/biological context, and details on annotations: • Types of annotation: nominal, ordinal, quantitative, or a mixture thereof; • Constructs: arrows, outlines, slider bars, text boxes).	Wahab *et al.* 3. Annotation Data Dictionary (modified), 5. Annotation Levels, and 6. Annotation Constructs
3. Study Design	Specify the study design • Number of annotators, number of cases, number of annotators per case • Methods to randomize cases or viewing conditions • Time limits for annotation tasks • Methods to assist or combine annotations (adjudication)	Wahab *et al.* 9. Workload Distribution (modified)
4. Annotation Methods	Determine how annotators will encounter the data (digital platforms, in-person viewing, or mixed-methods), and what tools will be used to access and view the images.	Wahab *et al.* 4. Selection of Annotation Software, (modified)
5. Image Curation	Specify annotator, patient, and image sampling methods for the entire study and individual sub-groups of a study (e.g. e.g. case sampling methods including enrichment, and stratified sampling): • Annotator: exhaustive in an allotted time-period, random subset, or stratified sampling of subgroups; • Patient: demographic and clinical sub-groups, target features, and variability of features case sampling methods including enrichment, stratified sampling, and ROI selection; • Image: quality and number of images in each subgroup.	Wahab *et al.* 11. Quality Review (modified)
6. Annotators	Define annotators: number of total annotators, number of annotators per case, qualifications (training and requirements), and how they were recruited.	[New]
7. Quality Review	During and after the annotation study, identify, review and discuss adherence to the above components of the template, report the collected data, and report any deviations. Specify whether this was a single study or part of a larger study.	Wahab *et al.* 8. Phase of Annotation and 11. Quality Review (modified)

Status of HTT Pivotal Study

- **6/2023**: Training launched
  (CME & Interactive Training)
- **6/2023**: Pivotal Study launched
- **8/2023**: New Website launched

- **52**: Pathologist Inquiries
- **5**: Active Pathologists
- **4**: Pathologists not passing proficiency
  - Training is a hurdle
- **155** cases
  - Still sourcing
- **113** cases curated
  - First pass ROI selection
- **40** cases batched for pivotal study
  - 5 pathologists per case: nearly complete
- **48** cases imminently batched

Related Activities

Pathology Innovation Collaborative Community

PIcc – "Pie See See"

- FDA participates in PIcc
  - https://pathologyinnovationcc.org/

- Look for Joe!

Joe Lennerz · 1st
Chief Scientific Officer, BostonGene, Ma, USA

- Regulatory Landscape Survey

https://qualtricsxmq9n4cl9pg.qualtrics.com/jfe/form/SV_4Sf41xG9Gm6XQMu

Related Activity: FNIH BC-CSC:

Foundation for NIH Biomarkers Consortium Cancer Steering Committee

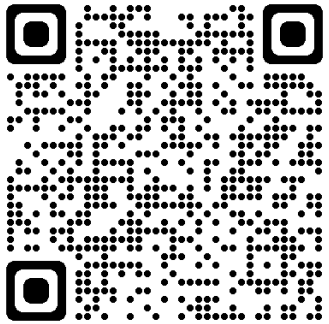
- PIcc coordinating project pitch to FNIH
  - New PIcc members welcome to the effort
- Presented vision to create a pipeline of real-world data for validating AI models
  - FNIH meeting 11/2023
  - Summary: https://fnih.org/our-programs/biomarkers-consortium-csc-scientific-symposium/
- Currently producing a skeleton proposal for 1-1 discussions with FNIH members

Call for Feedback!

ARPA-H FDA/CDRH Medical Imaging Data **Marketplace**

- *A self-sustaining, federated, national marketplace to catalyze transformative medical and health AI innovations*

- Network survey to provide feedback
  - https://investorcatalysthub.org/medical-imaging/

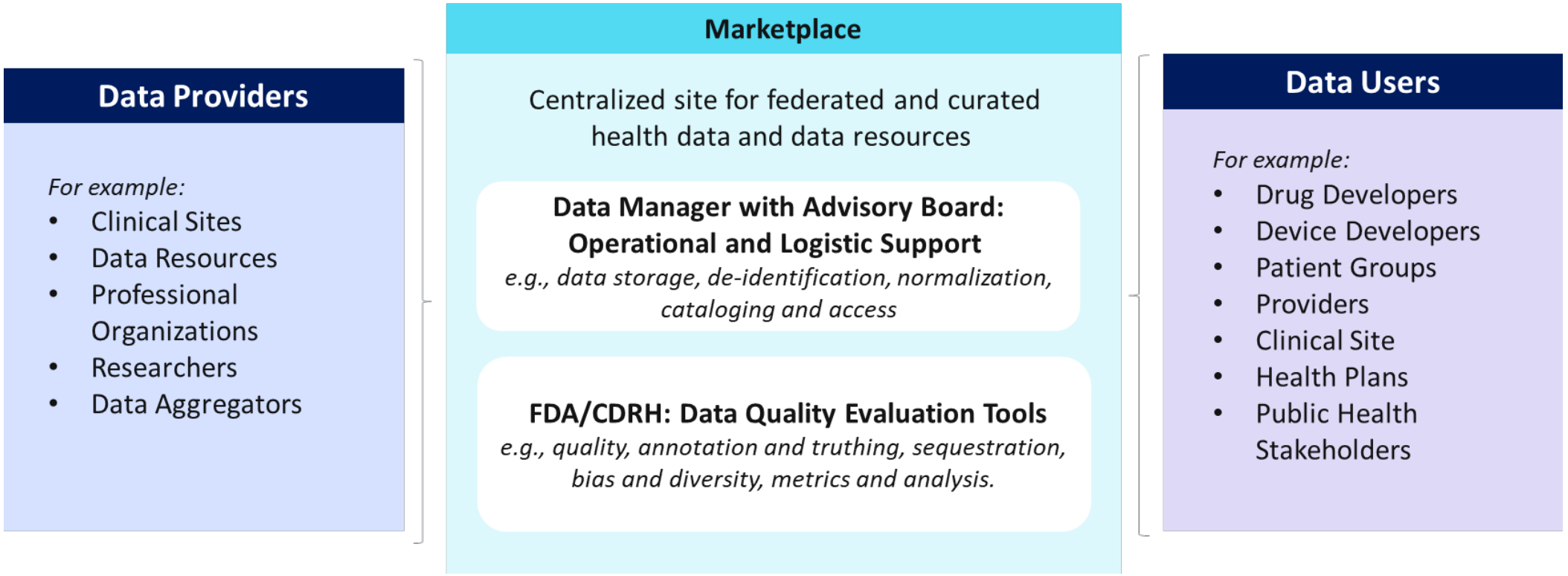


- Email for more Information:
  - midm@arpa-h.gov

Call for Feedback!

ARPA-H FDA/CDRH Medical Imaging Data Marketplace

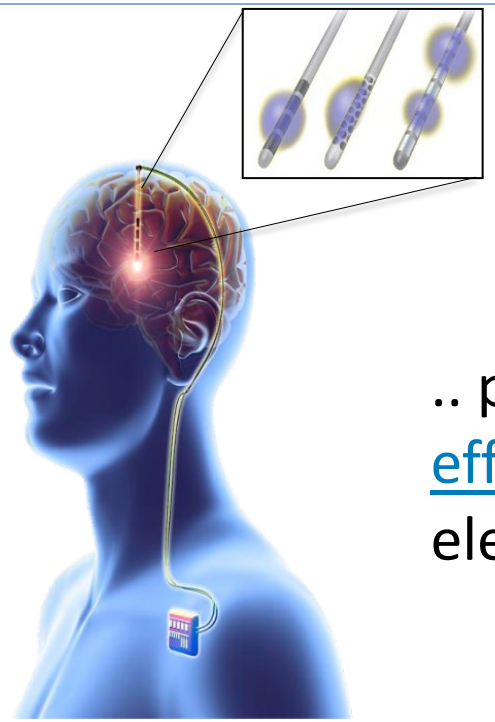
- *A self-sustaining, federated, national marketplace to catalyze transformative medical and health AI innovations*



Summary

- @Pathologists
  - CME and interactive training for the assessment of TILs in breast cancer
  - **Recruiting pathologists** to be the reference standard to validate AI
- @AI-Developers
  - Regulatory information for developers of medical imaging algorithms
  - Materials and lessons learned for creating a validation dataset
  - Approaches to evaluating pathologists against experts
  - Feedback from FDA reviewers
- @All
  - Find out more at: https://didsr.github.io/HTT.home/
  - Related digital and computational collaborations

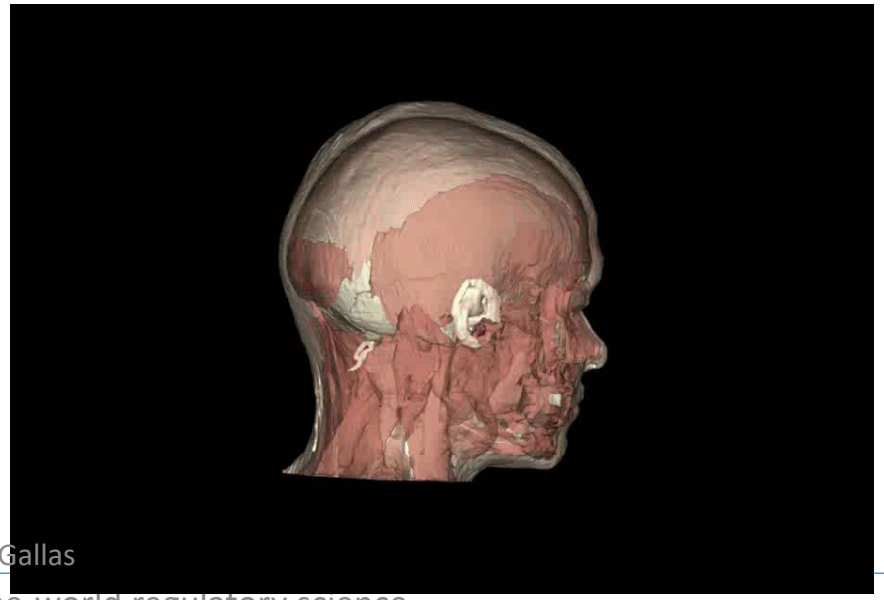
Title and content (black background)



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

1900
EMPLOYEES

18k
Medical Device Manufacturers

183k
Medical Devices
On the U.S. Market

22k/year
Premarket Submissions
includes supplements and amendments

570k
Proprietary Brands

25k
Medical Device Facilities Worldwide

1.4 MILLION/year
Reports on medical device adverse events and malfunctions

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

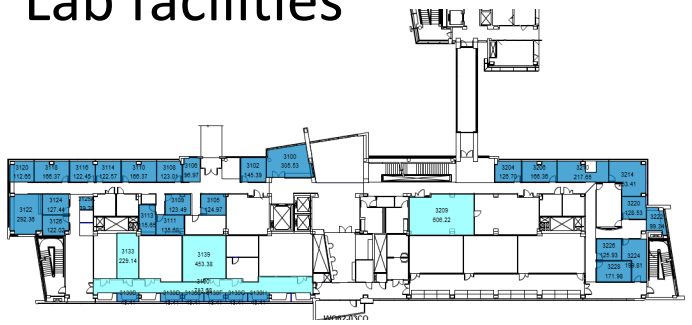
400/year
Peer reviewed presentations, articles, and other public disclosures

2,500k/year
Premarket Regulatory consults

75
Standards and conformity assessment committees

70%
Staff with post graduate degree

55,000 ft²
Lab facilities



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