Recruiting Pathologists to Truth Images

Host Site: Yale School of Medicine (Summer-Early Fall 2021)
Site Lead (Primary Contact): Dr Kim Blenman (kim.blenman@yale.edu)
Host Site: Stony Brook Medicine, (Summer-Early Fall2021)
Site Lead (Primary Contact TBD)
Digital Mode: https://ncihub.org/groups/eedapstudies

Researchers from the U.S. Food and Drug Administration, alongside academic, clinical and industry colleagues, are collecting pathologist annotations and quantifications of tumor infiltrating lymphocytes (TILs) as data for AI/ML algorithm validation. We are asking pathologists to score at least one full batch of 80 Regions of Interest (ROIs) as part of a research study. We anticipate that one batch will take 30 minutes in digital mode and 60 minutes on our microscope platform. We will train contributors prior to collection; training takes approximately 30 minutes and can be done ahead of time. The data are intended to inform the FDA's approach to novel algorithm validation, ensuring high quality commercial products with a faster FDA-pipeline to approval.

Specifically, you will be presented with pre-selected ROIs digitally or on a microscope (Figure 1). For each ROI, you will label the ROI and enter the percent of tumor-associated stroma and the tumor-associated stromal TIL density via a graphical user interface (GUI, Figure 2), which are numbers from 0-100.

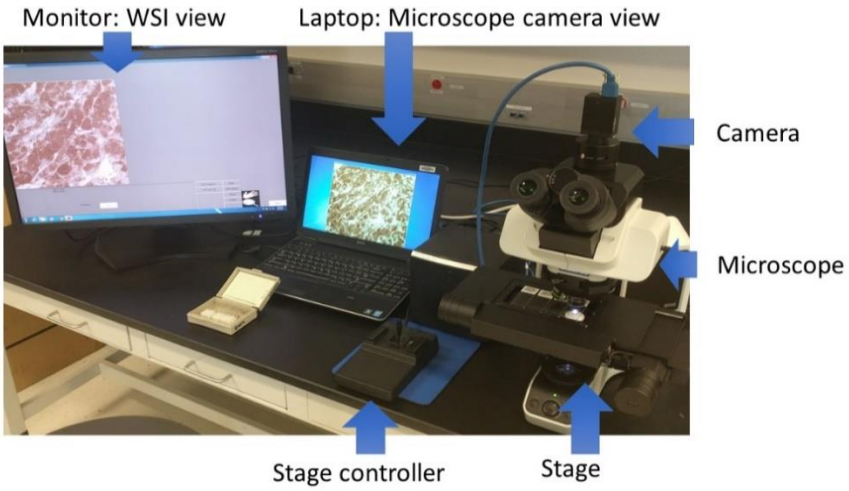

Figure 1: Microscope Setup. Computer controlled stage automatically navigates to next ROI.

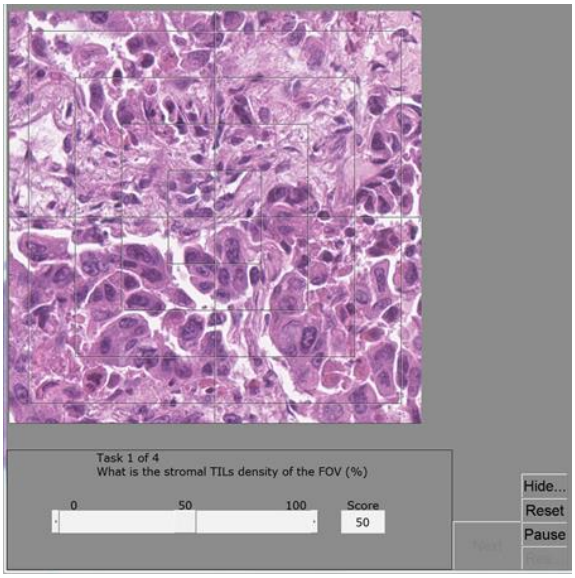

Figure 2: Data capture system for TIL evaluation with slider bar or keyboard data entry.

Please visit this wiki page for information about the project, data-collection training, IRB participation consent, and to begin data collection: https://ncihub.org/groups/eedapstudies.

All the best,

Sarah Dudgeon, Site Coordinator (FDA Intern 2021; Yale PhD Candidate; Sarah.Dudgeon@yale.edu),
Katherine Elfer, Project Coordinator (Katherine.Elfer@fda.hhs.gov) and
Brandon Gallas, Project Lead (Brandon.Gallas@fda.hhs.gov).

HTT Project Update

The HTT project's objective is to a) create a validation dataset fit for a regulatory purpose and b) apply statistical methods to evaluate variance in ground truth for a given dataset. Pursuing this objective is expected to inform regulatory frameworks and be instructive to others to develop their own validation datasets. Specifically, the project deals with the assessment of tumor infiltrating lymphocytes (TILs) in H&E-stained breast cancer slides – board-certified pathologists are asked to quantify TILs by viewing regions of interest (ROIs) on breast biopsy and core H&Es under a standardized microscope setup or in digital mode. All project info can be found here: https://ncihub.org/groups/eedapstudies.

Recent Accomplishments:

- 8/8/2021 (accepted): Presentation at the American Statistical Society Joint Statistical Meetings
  - "Pathologist Agreement from Quantitative Measurements: a Pilot Study"
  - In coordination with this presentation and the corresponding proceedings paper, we intend to make public a GitHub repository with an R data package containing the pathologist annotations.
  - Need: Looking for more perspectives and methods to do the statistical analysis.
- 6/2021: Updated data collection training nearly complete. Next, we will create a test with feedback and a proficiency test. Then we will establish an online CME opportunity that will create the "credential" for study participation.
  - Need: Looking for feedback on the training.
- 6/2021: We have executed a research collaboration agreements (RCA) for slide sourcing at one site and are nearing completion of an RCA with another site.
- 6/2021: Hired six summer research assistants to support the related research.
- 5/2021: Poster in the 2021 FDA Science Forum
  - "High Throughput Truthing (HTT) of pathologist annotations as a reference standard for validating artificial intelligence in digital pathology"
  - Link to poster
- 5/2021: Presentation at the Pathology Informatics Summit.
  - "High Throughput Truthing (HTT): Pathologist Agreement from a Pilot Study"
  - Link to presentation information and slides
- 4/2021: Accepted to 2nd Stage Interviews for the Burroughs Wellcome Fund Innovation in Regulatory Science Award to support slide sourcing and WSI platform development. Unfortunately we were not selected for funding in this year's award cycle.
- 2/2021: Manuscript accepted for publication in the *Journal of Pathology Informatics*. Pre-publication copy available at https://arxiv.org/abs/2010.06995.
  - "A Pathologist-Annotated Dataset for Validating Artificial Intelligence: A Project Description and Pilot Study".
- 2/2021: Awarded internal funding from the FDA/CDRH Office of Women's Health. This funding will support a research assistant and moderate incidentals for slide sourcing and in-person data collection events.
- 1/2021: Target data collection for pilot study achieved: 640 regions of interest were evaluated by at least 5 pathologists.
- Previous update available here: https://ncihub.org/groups/eedapstudies/wiki/HTTupdate20201023.

HELP! To continue our work, we need more pathologist annotations on existing slides, U.S. sites willing to host in-person data collection events, and sharing of physical slides for pathologists to review. Please see below for instructions on how to contribute:

- ***Digital* Data Collection:** We are looking for data collectors to provide annotations on our digital platforms: PathPresenter and caMicroscope.
  - Pathologists or pathology residents: complete the training and registration HERE.

- ***In-Person* Data Collection:** We are recruiting data collectors and host sites for collecting microscope-mode data using the eeDAP platform
  - Summer 2021: We are recruiting data collectors to contribute at either Yale School of Medicine or Stony Brook Medicine
  - Ideal collection: 5 readers per batch, 8 total batches, just over 1 hour per batch (can be broken into half batches of ~30 minutes each)

- **Sourcing Glass Slides:** We are sourcing glass slides of TILs in breast cancer for our pivotal study.
  - We have executed a research collaboration agreements (RCA) for slide sourcing at one site and are nearing completion of an RCA with another site.
  - We have a template research plan for all Research Collaboration Agreements
  - We'd like to identify at least one more U.S. site that can provide slides.

If you want to help, please contact:

Katherine Elfer, Project Coordinator (Katherine.Elfer@fda.hhs.gov)
Brandon Gallas, Project Lead (Brandon.Gallas@fda.hhs.gov)
Kim Blenman, Yale Site Lead (kim.blenman@yale.edu)