CDRH Critical Path Application FY 2018

Submissions will be accepted no later than November 6, 2017

For content and technical questions regarding the application, please contact: Mimi.Nguyen@fda.hhs.gov

Section I: Basic Information

Lead Investigator

Name	Brandon D. Gallas			
Position Title	Research Physicist			
Office	OSEL - Office of Science and Engineering Laboratories			
Division/Lab	Division of Imaging, Diagnostics, and Software Reliability			
Project Role *(350 characters)*	I will lead the development of the collaboration mechanisms (MTAs, CRADAs, IRBs), the experiments, the analyses, and the communications. I will be leveraging a digital pathology working group with over 130 members including all kinds of stakeholders that I and my colleagues have built over the last four years (search "wsi working group").			
Office Phone	+1 (301) 796-2531	E-mail	brandon.gallas@fda.hhs.gov	PI Expected Time Commitment *Expressed as % of total workload*
Supervisor	Kyle Myers			45%, but probably more

Proposal Title *(500 characters)*

High-throughput truthing of microscope slides to validate artificial intelligence algorithms analyzing digital scans of pathology slides: leveraging data collected in international "grand challenges".

Critical Path Status: (•) New Application ( ) Ongoing CP Project

Was a similar proposal also submitted to the FY18 Office of the Chief Scientist Proposal Call? (Full proposals only) ( ) Yes (•) No

Plain Language Summary

Provide a summary of the project in lay language and explain how this proposal helps CDRH achieve its vision. (3100 characters)

The microscope is going digital; glass slides are being digitized by devices called whole slide imaging (WSI) devices. The first WSI device for primary diagnosis was cleared this past spring (4/12/17) and more are coming. Among the potential benefits of this technology is to enable artificial intelligence (AI), computer algorithms. AI promises to reduce the pathologist's burden of searching and enumerating certain cells or cellular features on the slides as the pathologist evaluates a case and produces his or her report; let the computer do it. The regulatory question is then, "How well can the computer algorithms do the tasks at hand?" The most practical ground-truth for evaluating the performance of an algorithm is a pathologist's assessment of the WSI images. The problem with this kind of truth is that clinicians make mistakes and don't agree. Furthermore, there is a loss of information in the scanning process. The scanners have limited spatial and color resolution and currently produce a 2D slice of a 3D specimen. In this work, we plan to investigate a high-throughput algorithm truthing study and aid computer algorithm developers by producing a public resource for use in a regulatory submission.
We have developed a hardware and software evaluation environment for digital and analog pathology (eeDAP). eeDAP allows us to automatically present pre-specified regions of interest or individual cells and cellular features on a microscope for pathologist evaluation. This allows us to compare location-specific computer algorithm results to microscope-based pathologist evaluations. Last week, we installed eeDAP on a multi-head microscope and completed a data-collection session, collecting evaluations from 12 pathologists simultaneously in a single visit to MSKCC. That is high throughput truthing.
In recent years, "grand challenges" have been organized that offer algorithm developers an opportunity to compare computer algorithms on a common set of images in a controlled public setting. We plan to leverage the materials, results, and expertise produced by challenges that are based on WSI images. Two of our collaborators are the challenge organizers of http://tupac.tue-image.nl/ and https://camelyon16.grand-challenge.org/organizers/. Using the glass slides from the challenges, we plan to design, execute, and analyze studies with pathologists that will yield regulatory-grade performance results and a template for the evidence module of an FDA submission. We will do this in the public domain so that the community will benefit along with the challenge participants.
This work is a natural evolution of the PI's efforts to nurture a community for discussing topics related to technical and pathologist performance using WSI images (https://nciphub.org/groups/wsi_working_group) and the Medical Device Development Tool qualification of the enabling technology eeDAP (https://nciphub.org/groups/eedapstudies). The work will speed access to safe and effective state-of-the-art computer algorithms to help pathologists provide patients with better information.

Section II: Proposal Information

Research Strategy: Aims

Please list the specific objectives of the proposed research (3000 characters)

Objective 1: We will establish a CRADA with the collaborators to allow for the sharing of glass slides, digital images, and other resources. During this objective we will review all the materials related to the challenges, determine priorities for collecting truth data, and plan shipping and networking needs for sharing the materials.
Objective 2: We will write a protocol for collecting the truth data. This will begin with background research: analysis of challenge materials, analysis of preliminary studies (see Sect. VII: "Additional Information"), and a literature review. The key result of the background research will be determining the study endpoints and study size. We will present a complete protocol to CDRH reviewers and request feedback in the same way as the Medical Device Innovation Consortium received feedback from the FDA: "Protein-Based Multiplex Assays: Mock Presubmissions to the US Food and Drug Administration". Regnier, et al, Clinical Chemistry 56:2 165–171 (2010). We will finalize the protocol based on the feedback from the CDRH reviewers.
Objective 3: We will organize data collection events, "reader studies" with pathologists providing truth in the form of image annotations, scoring cells or regions. Pathologists will use the microscope for some data collection and digital images for other. Preparing for data collection, we will curate the data and pilot the study. This will require pathologist and algorithm annotations to identify the cases and image content that will be most appropriate and lead to an effective and efficient study. This will also require IRB approvals which must be requested well in advance.
Objective 4: Establish methods to evaluate the performance of artificial intelligence algorithms and produce a template of the evidence module for future submissions to CDRH. The evidence module will contain sections describing the algorithm architecture, the algorithm development process, the images and data used to train the algorithm, the images and data used to test the algorithm, the procedures used to define reference results (the truth), the statistical analyses, and the results. All the challenge participants will be invited to submit the results of their artificial intelligence algorithm. We will present the report to CDRH reviewers for feedback and revise the report to satisfy all stakeholders.

Research Strategy: Approach and Methods

Please describe the study design including the overall strategy, methodology, and analyses to be used to accomplish the specific aims of the research. (3300 characters)

The novel aspect of this project is to classify individual cells (or other features) in pre-specified ROIs during a session on a 14-head microscope. This is a high-throughput study in the sense that we will collect data from 14 pathologists at the same time. We can navigate the microscope to pre-specified ROIs using a system we developed at CDRH called eeDAP: evaluation environment for digital and analog evaluation. eeDAP is a software and hardware platform for designing and executing digital and analog pathology studies where evaluation ROIs in the digital image are registered to the real-time view on the microscope (Gallas2014_J-Med-Img_v1p037501).

The overall strategy for data collection is best discussed by example. We just completed a study using a 14-head microscope at MSKCC (Nov. 3, 2017). We installed our camera and computer controlled stage onto a MSKCC microscope, an Olympus BX45 frame that has 14-heads mounted on it. There were 12 study pathologists who counted mitotic figures (MFs) in 40 ROIs on 4 glass slides. The study pathologists also classified 128 candidate MFs. The counts and classifications were done without discussion under instructions to make decisions independently, and the classifications were confidence scores as opposed to binary yes/no classifications: a 0-100 scale, threshold=50. We believe that there is more information in multi-level scores that can lead to smaller studies or better statistical power. Most of the candidate MFs were identified during a preliminary image curation study where 4 pathologists counted and marked the locations of MFs arising from the same 40 ROIs. We call them "candidate" MFs because only 21 of 92 were unanimously identified. The distribution of agreement results across the candidates reflects a spectrum of candidates from obvious and likely MFs to subtle and improbable MFs. In fact, in order to include candidate MFs with even lower pathologist agreement, we asked one of the image curation pathologists to re-evaluate 12 of the original 40 ROIs. These ROIs had only one candidate or none identified during the original image curation. For this re-evaluation we asked the image curation pathologist to lower his threshold for identifying MFs. He identified 36 more candidates. This spectrum allows an efficient and effective evaluation of an AI algorithm.
The key principle that will guide the analysis methods will be that the results and error bars will account for reader and case variability. The reader is the pathologist and the case is the patient, a patient image, a region of interest within a patient image (ROI), or an individual cell within a patient image. Dr. Gallas has developed, validated, and publicly shared such tools and source code (https://github.com/DIDSR/iMRMC/releases). The current tools treat data where the truth is known. Dr. Gallas has been developing analogous tools for data that does not have truth or defines truth based on an expert panel, as is typically the case in pathology. This project will support that development.

Other Project Details

Please check all that apply to this application

☒ Human Subjects, Sample or Data Use

For more information on RIHSC, please contact Dora Vega.

☒ Animal Specimen/Data Use

For information on IACUC, please visit the IACUC intranet page.

☐ IT Infrastructure Support

Research Timeline, Milestones, and Deliverables

Please provide specific information regarding staff, time spent on project, and projected timeline.
Include a plan for interactive discussions with stakeholders. (4200 characters)

Q1 - Q2 (Objective 1): We will establish needed MTAs, CRADAs, and IRBs with the collaborators to allow for the sharing of glass slides, digital images, and other resources. During this objective we will review all the materials related to the challenges that our listed collaborators have organized (http://tupac.tue-image.nl/ and https://camelyon16.grand-challenge.org/organizers/), determine priorities for collecting truth data, and plan shipping and networking needs for sharing the materials.
Q1 - Q3 (Objective 2): We will begin the development of the evidence module for an FDA submission and we will develop a draft protocol and study design of two reader studies aligned with the challenges described at http://tupac.tue-image.nl/ and https://camelyon16.grand-challenge.org/organizers/. We will base this work on analysis of the challenge materials, and related data sets.
Q3-Q4 (Objective 2): We will present the draft protocol to the WSI working group (by email, webinar, and face-to-face meeting). The WSI WG is a community for discussing topics related to technical and pathologist performance evaluation using WSI images (https://nciphub.org/groups/wsi_working_group). The WSI WG is led by Dr. Gallas and Dr. Gavrielides from CDRH/OSEL Division of Imaging, Diagnostics, and Software Reliability, Dr. Treanor from Leeds Teaching Hospitals NHS Trust, and Dr. Hewitt from Laboratory of Pathology, Center for Cancer Research, National Cancer Institute, National Institutes of Health. There are over 130 members including all kinds of stakeholders from industry, government (including FDA), academia, clinicians, hospital and health provider network department chairs, VPs, and CEOs. The purpose for sharing the draft protocol is to get feedback from the community (consensus-building), recruit additional partners for conducting the reader studies and for participating in the challenges, and to disseminate tools, expertise, and the entire experience. We need this project to be transparent and in the public domain so that FDA scientists may collaborate hand-in-glove with industry to produce regulatory-grade truth data and analyses.
Q4-Q5 (Objective 2): We will update the protocol, share it with CDRH reviewers, and request feedback in the same way as the Medical Device Innovation Consortium received feedback from the FDA: "Protein-Based Multiplex Assays: Mock Presubmissions to the US Food and Drug Administration". Regnier, et al, Clinical Chemistry 56:2 165–171 (2010). We will finalize the protocol based on the feedback from the CDRH reviewers.
Q4-Q6 (Objective 3): We will organize data collection events, "reader studies" with pathologists providing truth in the form of image annotations and scoring of cells or regions. Pathologists will use digital images for some data collection and the microscope for other (the current clinical-practice baseline for performance). Preparing for data collection, we will curate the data and pilot the studies. Data collection will happen at the institutions of our collaborators, we will recruit other partners and sites via the WSI WG, and we will propose an "Interactive Microscopy" session for the 2019 meeting of the United States and Canadian Academy of Pathologists.
Q6-Q8 (Objective 4): We will bring together the reader study data and the algorithm results and produce a template for the main evidence module of an FDA submission. This will involve the core collaborators and the algorithm developers that participated in the challenges. Like the protocol, we will present a draft to the WSI WG to generate feedback for revisions, to invite additional algorithm developers to participate, and to keep that community up to date. After appropriate revisions, we will then share the evidence module with CDRH reviewers for feedback and revise the report to satisfy all stakeholders.

High throughput truthing study at Memorial Sloan Kettering Cancer Center. 12 pathologists marking (on paper) the locations of mitotic figures in 40 regions of interest and then classifying 128 pre-determined candidate mitotic figures on a multi-level confidence scale. All done in 120 minutes.

If you would like to add 1 graphic to this page, please click in the space above and upload. It will adjust to the space provided.

Section III: Regulatory Science Alignment and Public Health Impact

Critical Path Categories
Please select all that apply to your project.

- ☐ 1. Biomarkers
- ☐ 2. Models/Modeling
- ☒ 3. Trial Design & Strategy
- ☐ 4. Manufacturing
- ☒ 5. Bioinformatics

Regulatory Science Priorities
Please identify at most two priorities that apply to your proposal.

Develop methods and tools to improve and streamline clinical trial design

Leverage precision medicine and biomarkers for predicting medical device performance, diagnostics and prognosis

Regulatory Impact/ Alignment with Regulatory Science Priorities
Discuss the regulatory impact of your proposal and how your proposal addresses and/or supports CDRH's regulatory science priorities. (1700 characters)

At the end of our project, we will have a regulatory-grade evidence module for a digital pathology AI algorithm (bioinformatics, software as a device) suitable for submitting to the FDA. The protocol, study design, raw line-level data and analyses with code will have been developed and shared in the public domain, allowing anyone to use the resources and follow the path for a similar submission (an AI algorithm that performs a different task). By collaborating with challenge organizers we are leveraging data and expertise that are not traditionally involved in regulatory decision making, but should be. We are training a whole new group of stakeholders on the regulatory process and best practices in clinical trials. The public leaderboard of challenge participants identify 23 participants, from around the world. By sharing our protocol and data analyses with the WSI WG and then CDRH reviewers in order to get feedback and make improvements, we are leveraging the expertise of many stakeholders. This will be consensus-building in real time to improve and streamline clinical trials. Finally, we are collecting evaluations on pre-specified ROIs and cells on the microscope and correlating them with the specific location on the digital image. This strengthens our ability to evaluate algorithm performance in terms of statistical power, which translates to smaller studies. The high-throughput nature of collecting data on a multi-head microscope should be compared to the burden of sending slides to different pathologists and waiting.

Regulatory Science Metrics
Select the metrics that can be used to measure the success and impact of this proposal. For information on how to complete this section, please visit the CDRH Metrics SharePoint.

Level 4: Research Imperatives (max 3)	Level 3: Internal Outputs (max 3)	Level 2: External Impacts (max 2)
Collaboration	Updated and Streamlined Review Process	CDRH Tools in Use by Industry
Effective Dissemination	Communications Impact	
Public Health Need		

Outcomes for Evaluating Impact
Using the selected metrics, please describe how this work will change CDRH processes and impact the public health. Provide details that demonstrate how the research imparative metrics will be met, specifically addressing each metric chosen. Then provide plans to produce internal outputs and accomplish external impacts. (2100 characters)

Like everything else, the microscope is going digital. The volume of data is staggering. MSKCC alone is scanning 40,000 slides a month. This data is meaningless without pathologist annotations and, ultimately, the clinical outcomes. The pathologist annotations are fueling a huge computational pathology revolution. “Deep learning” is everywhere, and it is coming to the FDA.
Dissemination: The WSI WG is hosted on NCIPhub (National Cancer Informatics Program Hub). NCIPhub is a site for community research and collaboration in cancer research and informatics. Users around the world can share resources, host online communities, and use collaboration tools . Similarly we are using GitHub, an international code hosting platform for version control and collaboration. Both of these hubs also allow private projects where projects can start and mature before going public.
We believe that the CDRH-developed eeDAP system can play an important role in training and validating the algorithms at the cell annotation level. We have been accepted to produce a submission to the MDDT program. There are five institutions that have contributed to this effort, acting as a subgroup of the larger WSI WG and providing regular progress updates to the larger group, all open access (https://nciphub.org/groups/eedapstudies/wiki).
The WSI WG has convened four face-to-face meetings attached to annual meetings of the Association of Pathology Informatics and the Digital Pathology Association. We have also hosted many organizational T-cons and webinars. We use NCIPhub to share all the presentations given and plans being made. The infrastructure and audience are in place to receive updates, archive information, and become active project participants. We plan to host some meetings at the FDA to enable the feedback from CDRH reviewers on the protocol and the ultimate evidence module/report.

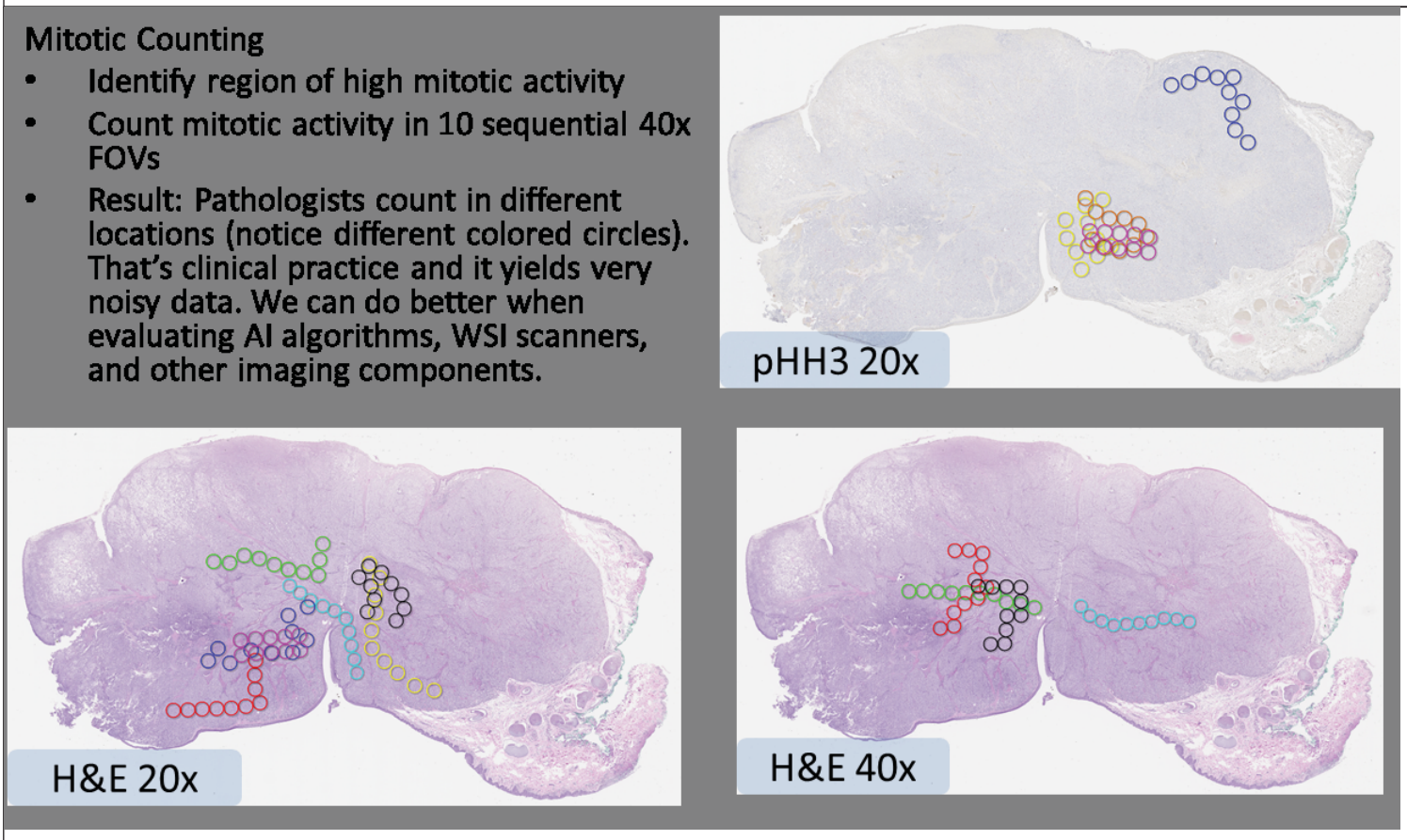
Public Health Impact

Discuss the public health need this proposal addresses (disease area, device type, etc) and how it addresses the Center's vision. (3300 characters)

Digital pathology systems have the potential to help pathologists be more effective, more efficient, and have more impact. Increased effectiveness may be possible with the use of image analysis software (enhancement filtering, quantitation, and AI algorithms), facilitated consulting, and better education. Increased efficiencies may be possible because of high-throughput whole-slide scanning, automated image analysis performing laborious tasks, and integration with the patient digital record (archiving, annotation, indexing, and retrieval). Impact can most definitely be increased with telepathology for underserved populations and the use of AI algorithms when there are no pathologists available.

The proposed project will shape the regulatory process with the engagement of a broad spectrum of stakeholders. In a community setting, we will discuss what data needs to be collected to support a regulatory claim for software as medical device, and we will develop an outline of how that data can be packaged for a regulatory submission.

The statistical analysis tools mentioned at the end of the "Research Strategy: Approach and Methods" appear as a small element to this project, but they are perhaps the most generalizable. They allow for the evaluation of the AI algorithms, the WSI scanning device, or any of the other components in the imaging chain, like the monitor. The analysis tools and the study designs are also not specific to digital pathology. They can be utilized in any imaging study where there is no ground truth (reference standard) and performance can only be given in terms of within and between (expert) reader agreement. Furthermore, the reader need not be a human. It may, in fact, be an AI algorithm, or a population of AI algorithms built on different architectures or training data.



If you would like to add 1 graphic to this page, please click in the space above and upload. It will adjust to the space provided.

Section IV: Collaborator Information

Participating CDRH Collaborators

Name Dr. Cheng Cui

Office/Division OMPT/CDRH/OIR/DMGP/MPCB **Supervisor** Eunice Lee

Role Advisor **FTE % Time Commitment** 10%

Contribution
Describe the expertise of this collaborator and how this contributes to the proposal. (500)

Dr. Cui is a senior reviewer with expertise in pathology and cytology. He will receive quarterly updates on the project progress and provide informal, high-level feedback regarding study design and analyses. He will also help coordinate interactions between the project collaborators and his group and educate the collaborators on the regulatory submission processes, including the pre-submission program.

Name Marios Gavrielides

Office/Division OMPT/CDRH/OSEL/DIDSR **Supervisor** Kyle Myers

Role Collaborator/Working Group **FTE % Time Commitment** 10%

Contribution
Describe the expertise of this collaborator and how this contributes to the proposal. (500)

Dr. Gavrielides has experience developing artificial intelligence algorithms and conducting reader studies. He also has expertise in the IRB process at the FDA. He will help design and execute reader studies, including the development and organization of study materials (slides, images, training documents, IRBs).

Non-CDRH Collaborators

List all, if known

Name Dr. Yukako Yagi, PhD

Organization Memorial Sloan Kettering Cancer Center

Category Academic/University **Collaborative Mechanism** To be determined

Contribution
Describe the expertise of this collaborator and how this contributes to the proposal. (500)

Dr. Yagi is a Digital Pathology Engineer and the Director of Pathology Digital Imaging at MSKCC. She collaborates with several pathologists at MSKCC that are intimately involved in studying the value of eeDAP and have artificial intelligence interests and expertise. She has access to slides and scanners. She helped arrange an MTA with FDA that brought an eeDAP system to MSKCC (equipment was shared by FDA, Dr. Gallas). She also arranged for the successful high throughput truthing study.

Name	Dr. Partha Mitra, PhD
Organization	Cold Spring Harbor Laboratory
Category	Academic/University
Collaborative Mechanism	CRADA
Contribution *Describe the expertise of this collaborator and how this contributes to the proposal. (500)*	Dr. Mitra is a physicist doing research in AI. He is also leading a neuroscience lab involved in scanning brains (mouse now, human future) and understanding brain architecture. He has access and expertise in slide prep and scanning. He has collaborations with experts in psychophyics and algorithms at CSHL. He also has collaborations with researchers in India, a severely underserved country. He also has collaborations with Northwell health, the largest health provider in New York state with more than 60 pathologists. Please see his letter of support.

Name	Dr. Jeroen van der Laak, PhD
Organization	Radboud University Medical Center
Category	Academic/University
Collaborative Mechanism	To be determined
Contribution *Describe the expertise of this collaborator and how this contributes to the proposal. (500)*	Dr. van der Laak has expertise on the development of artificial intelligence algorithms with pathology images and the collection of truth data from pathologists. He was a lead organizer in two "grand challenges" in which computational researchers were invited to work on a diagnostically relevant problem in histopathology. He has offered the glass slides from the challenges and his time to the execution of this project. Please see his letter of support.

Name	Dr. Mitko Veta, PhD
Organization	Technische Universiteit (University of Technology), Eindhoven
Category	Academic/University
Collaborative Mechanism	To be determined
Contribution *Describe the expertise of this collaborator and how this contributes to the proposal. (500)*	Dr. Veta has expertise on the development of artificial intelligence algorithms with pathology images and the collection of truth data from pathologists. He was a lead organizer in three "grand challenges" in which computational researchers were invited to work on a diagnostically relevant problem in histopathology. He has offered his time to the execution of this project. Please see his letter of support.

Name	
Organization	
Category	
Collaborative Mechanism	
Contribution *Describe the expertise of this collaborator and how this contributes to the proposal. (500)*	

Section V: Financial Information

Estimated FY17 CP Budget

List the expected project expenses to be covered by ONE year of CP funding.

Budget Category	Item Description	Justification	Cost
ORISE	Masters level bioengineer	Maintain hardware, prepare study materials (slides and images), coordinate logistics.	
ORISE	Masters level statistical programmer	0.5 FTE: Starting with CDRH-developed tools for sizing and analyzing reader studies, expand them to analyze data when truth is not known (under direction of senior statistician).	
ORISE	Travel costs	The bioengineer will travel to data collection sites to conduct data collection events, transporting equipment.	
		Total Requested Budget	

Out Year Costs

Project the expected expenses for continuing the project AFTER ONE year of CP funding.

Budget Category	Item Description	Justification	Cost
ORISE	Masters level bioengineer	Maintain hardware, prepare study materials (slides and images), coordinate logistics.	
ORISE	Masters level statistical programmer	0.5 FTE: Starting with CDRH-developed tools for sizing and analyzing reader studies, expand them to analyze data when truth is not known (under direction of senior statistician).	
ORISE	Travel costs	The bioengineer will travel to data collection sites to conduct data collection events, transporting equipment.	
		Total Budget	

Section VI: Reviewers

Technical Reviewers

Provide up to three (3) FDA employees who could be technical reviewers for your project. These reviewers must be familiar with the research need/area but not involved with this specific proposal.

Name and Title	Center/Office	E-mail

Section VII: Additional Information for Reviewers/Leadership

Is there any additional information you would like to share with the reviewers and leadership? You may upload a maximum of 5 documents to this proposal but it is not guaranteed that all files will be reviewed by reviewrs. This can include but is not limited to letters of support from external collaborators/stakeholders, CVs/Resumes, and recent publications.

	Name of File	Description
Attach	letterOfSupport-MitraCSHL.pdf	Letter of support from Dr. Partha Mitra
Attach	letterOfSupport-VanDerLaakRadboud.pdf	Letter of support from Dr. Jeroen van der Laak
Attach	letterOfSupport-VetaTUE.JPEG	Letter of support from Dr. Mitko Veta
Attach	letterOfSupport-YagiMSK.pdf	Letter of support from Dr. Yukako Yagi
Attach		

References

Please use this space to list any references relevant to your proposal. (2600 characters)

Data from previous studies available to inform project on designing, executing, and analyzing future studies:

A study of 12 pathologists at four sites comparing two stains (Hematoxylin and Eosin vs. Phospho-histone h3) and three viewing modalities (the microscope, 40x WSI digital images, 20x WSI digital images). There were 113 cases: 1 case = 1 slide from a canine oral melanoma patient. Censored survival data was available on 67 patients. The pathologists were asked to count mitotic figures (MFs) following clinical practice: Starting in the tumor region on the slide, count MFs in 10 sequential 40x fields of view. There were 29,130 counts collected in total. Drs. Mark Simpson DVM/PhD and Charles Halsey DVM from the NIH Center for Cancer Research, Laboratory of Cancer biology and Genetics, Molecular Pathology Unit, were the PIs of the study with Dr. Gallas advising on study design and data analysis.

A study of 5 pathologists using slides from the study described above. The pathologists provided counts and locations of MFs on a common set of regions of interest (40 ROIs). This is the image curation study described in the methods section that yielded 128 candidate MFs for the high-throughput truthing study just conducted on the 14-head microscope at MSKCC.

The high-throughput truthing study just completed and described in the methods section. There were 12 pathologists that marked (and counted) the locations of MFs in 40 ROIs. For each ROI, they also classified the candidate MFs contained within. There were 128 of these candidates in total.

For questions, please contact: Mimi.Nguyen@fda.hhs.gov