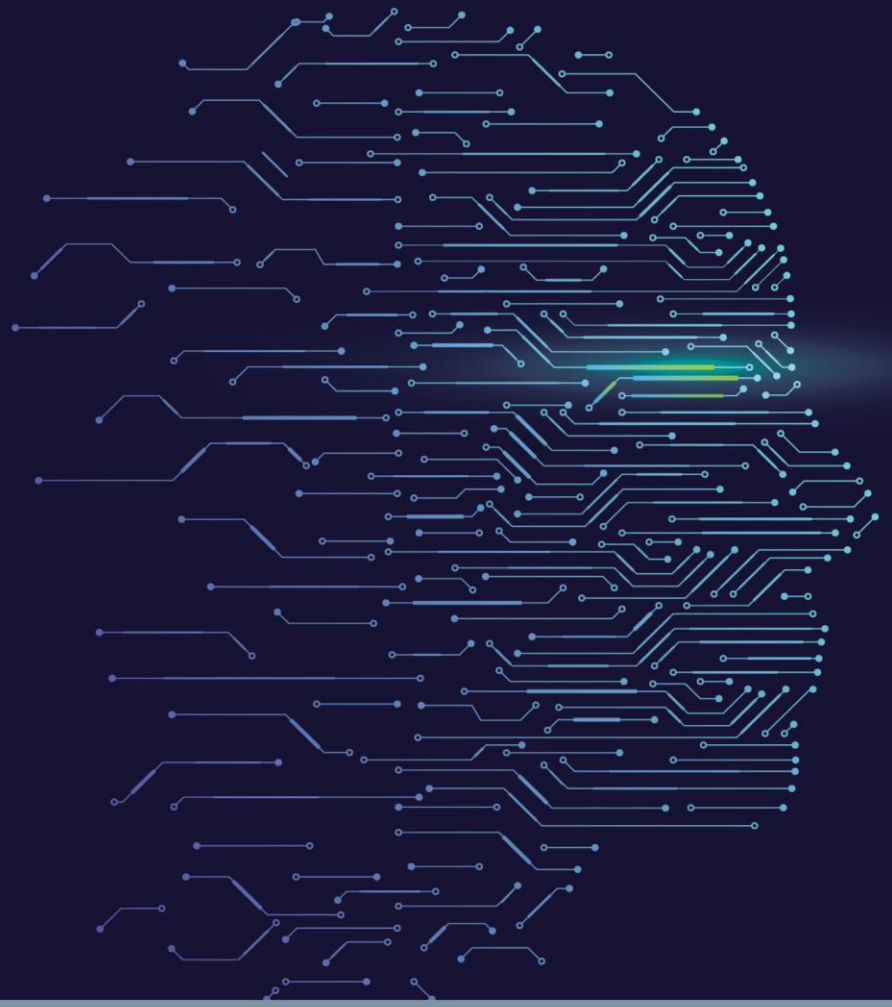
2019 IMAGING INFORMATICS SUMMIT



Evaluating Artificial Intelligence Devices at the FDA and Related Collaborations and Initiatives

Brandon Gallas, PhD

Research Physicist and Mathematician

Division of Imaging, Diagnostics, and Software Reliability

OSEL, CDRH, FDA

ACR® American College of Radiology

WASHINGTON

- Dr. Gallas has no conflicts to report.

Attendees will …

- Learn where to get information and help
- Understand that it may be less burdensome to start small and grow when it comes to submissions of algorithms to CDRH
- Be able to outline the core content of submissions of algorithms to CDRH
- Be able to distinguish a stand-alone study from a clinical study
- Be able to discuss FDA-led initiatives and other collaborations

Outline

- General Info for Submissions to CDRH: Imaging and AI
- Some History of DIDSR:
  Division of Imaging, Diagnostics, and Software Reliability
  - Contributions to the field of radiology
  - Guidance and Consensus Building: Image Quality Evaluation
- Recent DIDSR research
- Forming a collaboration Alliance

This talk is based on FDA's Current Thinking

Our current thinking changes over time just like science!

Useful Advice: Start with a narrow IFU for CAD

- Tie IFU to one imaging system
- Expand indications over time
  - Other imaging systems & protocols
  - Algorithm updates/improvements
- Possibly less burdensome
  - FDA knows device and performance

<u>Less burdensome methods</u>

- Technical arguments
  - Phantoms, Simulation
- Reuse cases (rescan film, slides)
  - New reader study
- Studies with fewer cases or fewer readers
- Stand-alone performance only
- No statistical hypothesis test

Useful Advice: Start with a narrow IFU for CAD

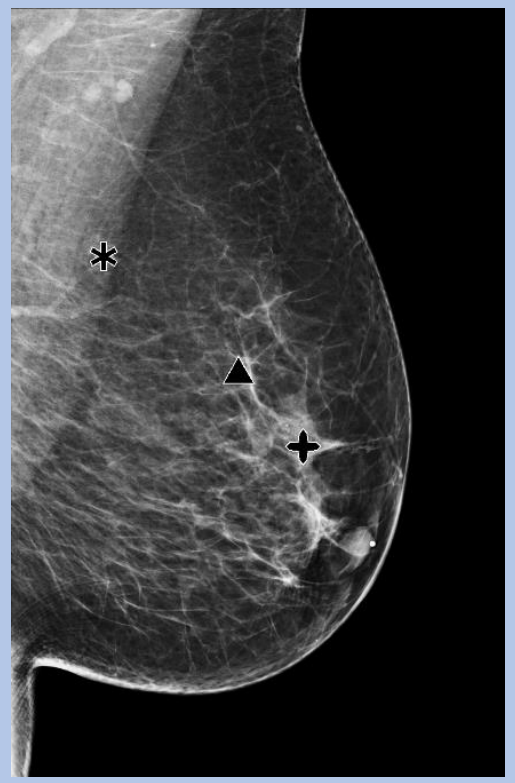
Example CADe

- R2 ImageChecker (P970058)
  - The ImageChecker M1000 is a computer system intended to identify and mark regions of interest on routine screening mammograms to bring them to the attention of the radiologist after initial reading has been completed. Thus, the system assists the radiologist in minimizing observational oversights by identifying areas on the original mammogram that may warrant a second review.

Useful Advice: Start with a narrow IFU for CAD

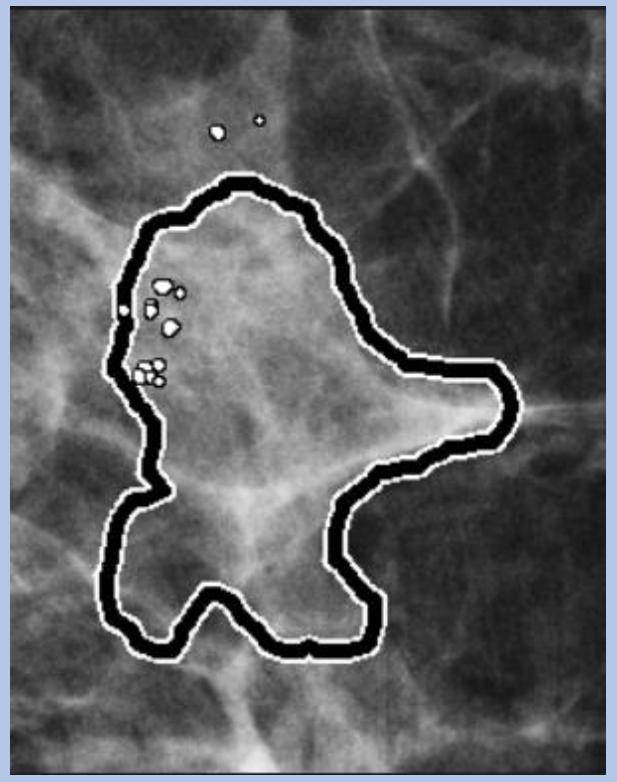
- Original device included a film digitizer!

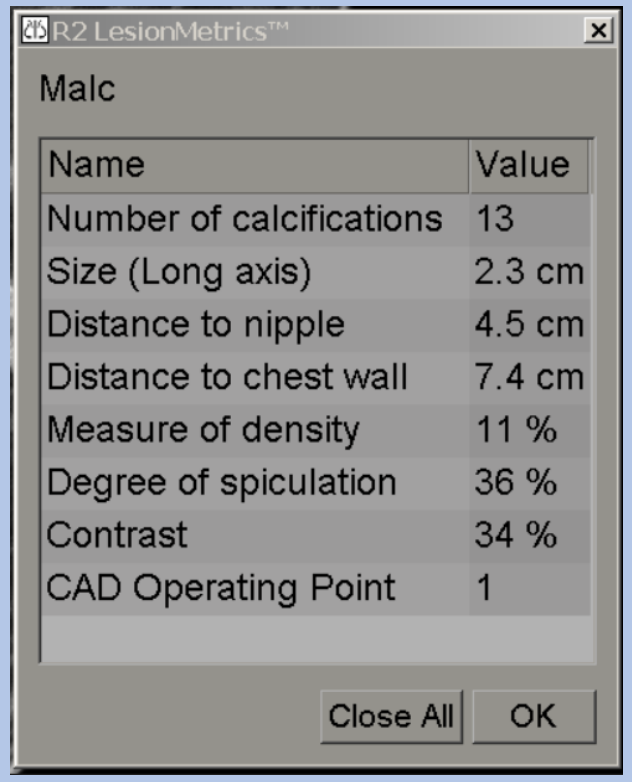
- “For each of the films, the video monitors display the corresponding low resolution images and markers…”

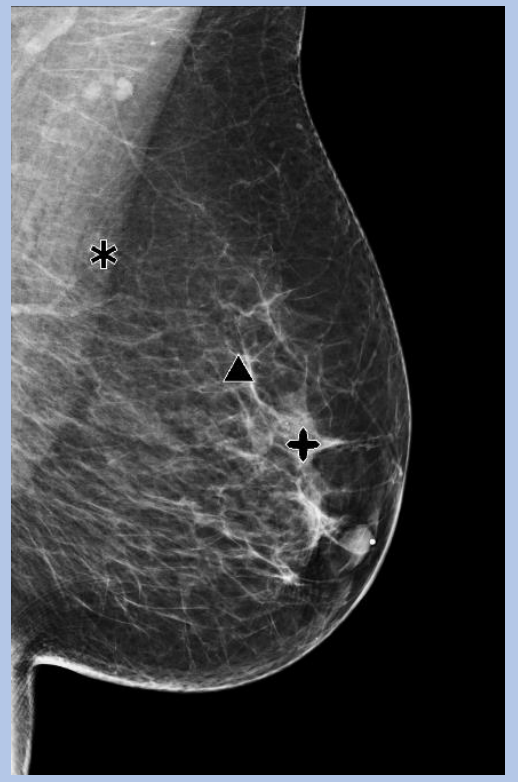


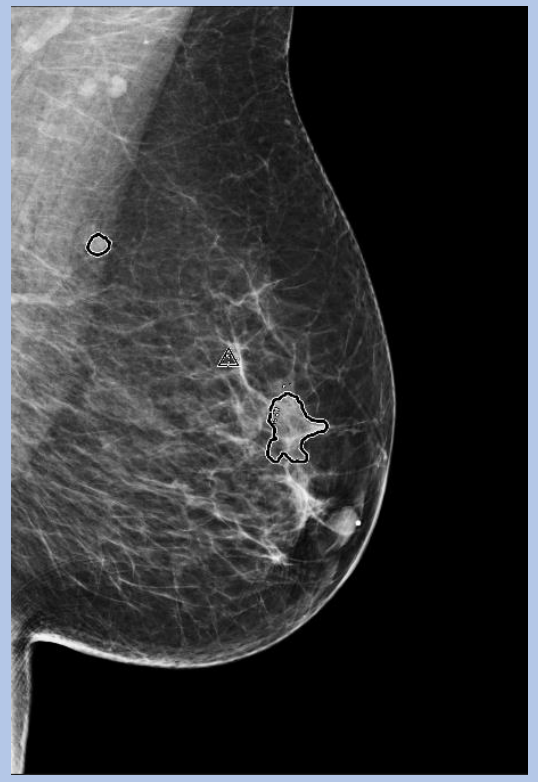
http://www.hologic.ca/image-analytics#overlay-context=closeup-peerview-cad

ImageChecker® Analytics









http://www.hologic.ca/image-analytics#overlay-context=closeup-peerview-cad

ImageChecker Submission History

1998

Approval of Original submission

1. Hardware changes and minor bugs and enhancements

1999

2. Performance change

3. Post approval study protocol

4. New marker (correlated masses)

5. Alternative film digitizer

2000

6. Performance change

7. Label change with respect to efficacy

8. New marker (subtle vs. obvious masses)

2001

9. New marker (subtle vs. obvious calcifications)

10. Indications expanded from screening to diagnostics

11. Indications expanded to digital images (GE Senographe 2000)

2002

12. Label change with respect to efficacy

13. Transparent marker (see image under marker)

14. Label change

2 decades
27 updates

2003

15. New Manufacturing facility

16. Choice of new operating points (high and low sensitivity), operates on analog and GE FFDM images, operates on GE FFDM images "formatted for presentation", reduces false-negatives of oversized malignant calcification clusters

17. Alternative film digitizer

18. Indications expanded to Fischer Senoscan FFDM

2003

19. Indications expanded to Hologic Selenia FFDM

2005

20. Indications expanded to include Siemens Novation FFDM

21. More operating points

2006

22. Change label to include specificity (previously it was sensitivity and false marks per image)

2007

24. New manufacturing facility

20012

25. Algorithm updates and indications expanded to GE Senograph Essential

2014

26. Indications expanded to C-view images Hologic Selenia Dimensions (Tomosynthesis) system

2016

27. New manufacturing facility

ImageChecker Submission History

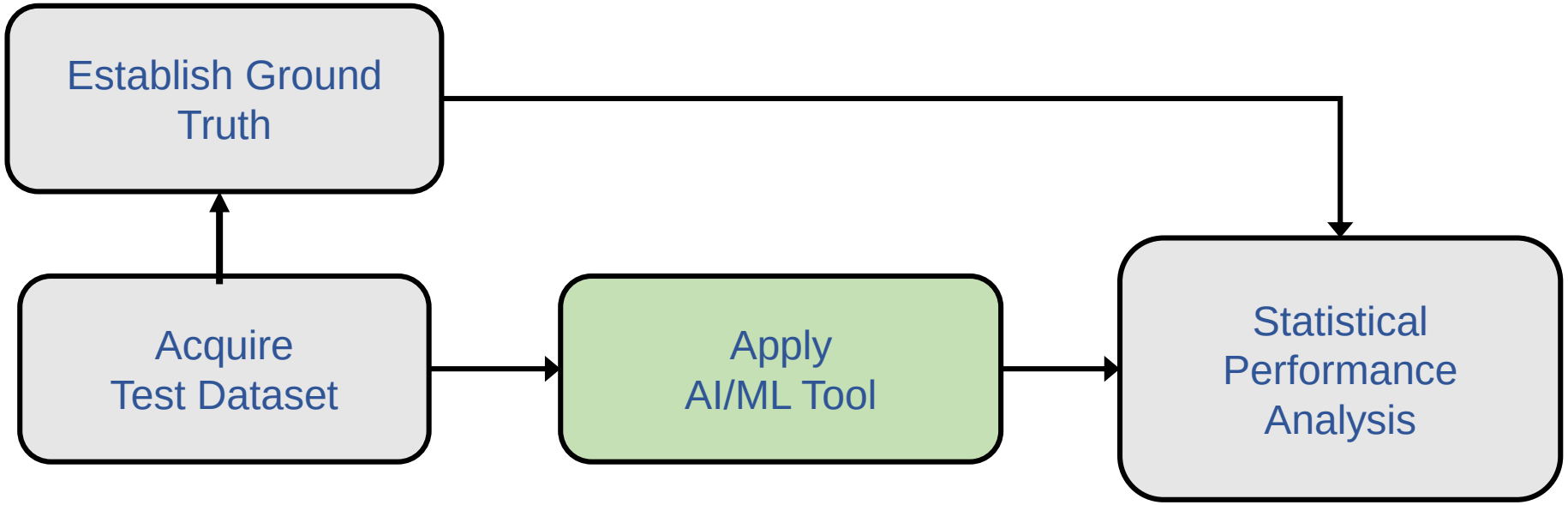
- Performance changes
- Screening to diagnostics
- Expand hardware
  - Alternative film digitizer **x2**
  - Digital imagers **x4**
- New marker
  - Correlated masses
  - subtle vs. obvious masses
- New operating points
  - Add higher and lower sensitivity
  - More operating points
- Indications expanded to C-view images Hologic Selenia Dimensions (Tomosynthesis) system

Core Content of Submissions for computer aids in Radiology

- Find a predicate
- Description
  - Indications for use
  - Clinical context, clinical workflow
  - Patient and clinician population
  - Imaging system and protocols
- Technological Characteristics
  - Algorithm design and function
  - Processing steps
  - Features
  - Models and classifiers
  - Training paradigm
- Imaging modality
  - Manufacturer and Model
  - Imaging parameters and techniques
- Databases: Training and Testing
  - Document data use
  - Sites, dates, collection protocols, patient characteristics
  - Training and testing sets must be Independent
- Reference standard
- Assessment
  - Depends on algorithm type: Aid vs. Automatic
  - Stand-alone performance study: No human in the loop
  - Clinical Performance: human in-the-loop

Stand-alone performance study: No human in the loop

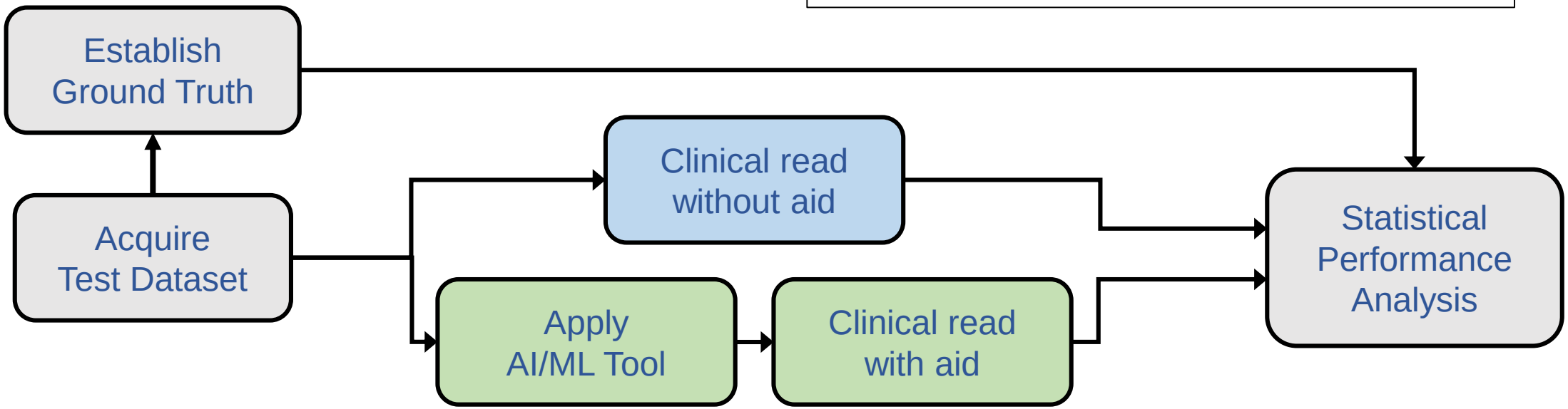
- Performance of algorithm by itself, independent of any interaction with user
  - Intrinsic functionality of device



Clinical Performance: human in-the-loop

- Assessment of clinicians' performance utilizing the device
  - Many possible study designs
    - Prospective/retrospective
    - Multi-reader multi-case designs

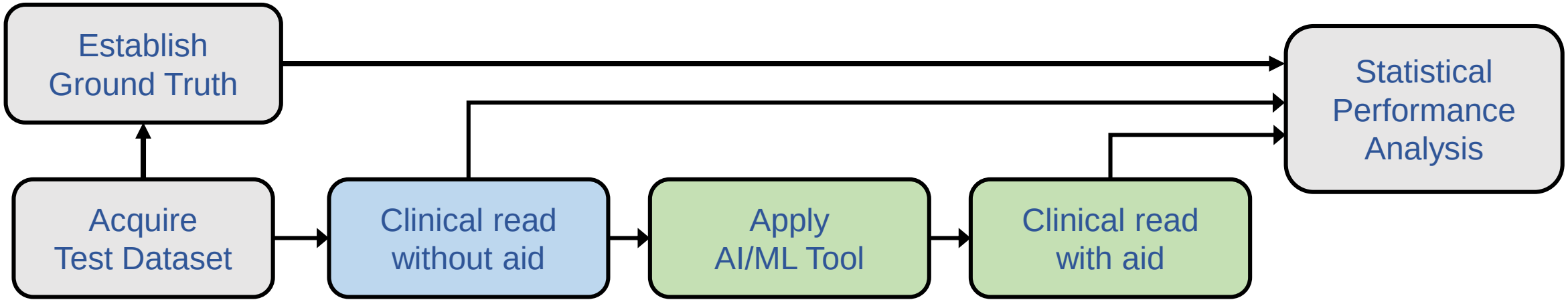
Independent crossover design
Need to balance reading order
Need washout



Clinical Performance: human in-the-loop

- Assessment of clinicians' performance utilizing the device
  - Many possible study designs
    - Prospective/retrospective
    - Multi-reader multi-case designs

Sequential design
Build in useful correlation
No washout needed



Actual submission feedback

- A device review is not unlike a manuscript or grant review
  - Clarity, Conciseness
  - Good science
  - Reproducible research

- **STARD 2015**: List of Essential Items for Reporting Diagnostic Accuracy Studies
  - Required by *Radiology*, the journal.
  - Not required by FDA.

Actual submission feedback: Study Analysis Plans

- Please provide a primary endpoint with clinical meaning as well as justification for expected performance in terms of a hypothesis test.
- Please provide a sample size calculation for both readers and images included in the study based on the proposed endpoint and hypothesis test.
- Analyses should account for the uncertainty from multiple readers and multiple cases and the correlations that arise from the study design (multiple readers reading the same cases).
- Please use statistical and mathematical equations and descriptions in addition to words.
- We welcome a simulation study to describe how you plan to do the analyses.

Actual submission feedback: Hardware

- Please provide details of how the imaging data were/are to be collected (e.g., make and model of the imaging device and imaging protocol).

- In your premarket submission, you should demonstrate that your algorithm is robust to variability across device manufacturers.

Actual submission feedback: Hardware Modeling

- It may be acceptable to supplement analyses of clinical study data by incorporating models of the performance characteristics of the range of devices or by other arguments with appropriate justification.

- If you plan to conduct such modeling, we recommend discussing the specifics with us prior to conducting this type of analyses.

- Need to see the protocol to provide guidance; however, the FDA is open to phantom based validation if the protocol is appropriate.

Actual submission feedback: Generalizability

- Please evaluate the clinical accuracy of your device across the range of intended imaging devices, multiple operators, and multiple sites.

- In a random splitting, the test set is expected to have the same characteristics as the training set. Thus, your proposed study design may lead to overestimation of the performance of your algorithm in the test set and may not be generalizable

- We recommend that you conduct your external clinical validation study using a unique data set, separated by time and site from your training data set to avoid biasing your study results.

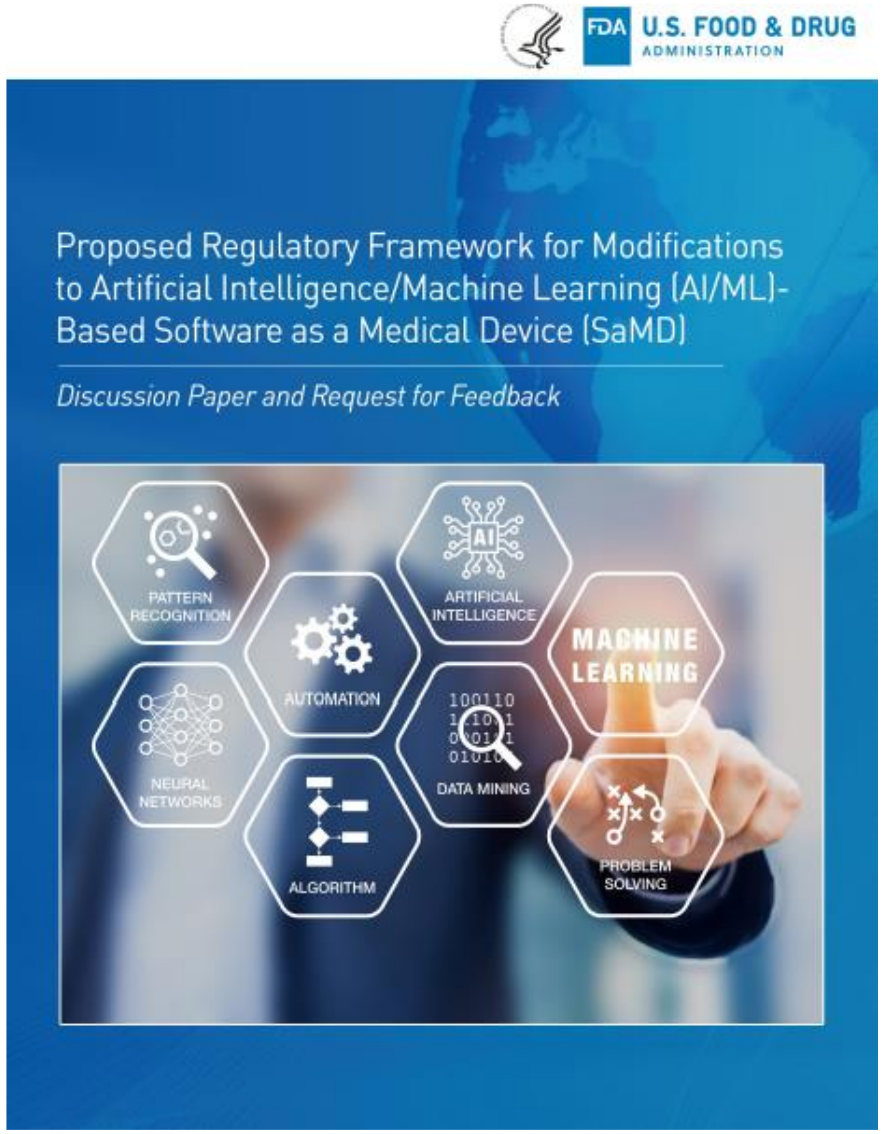
Actual submission feedback: Locking the algorithm

- Your device's algorithm, including any clinical cutoff(s), should be locked down before the start of the analytical and clinical studies to validate (i.e., test) the performance.

- To mitigate the bias discussed, it is important to pre-specify and finalize the cutoff and all other aspects pertaining to model selection and development of the software before examining any of the data that will be used for validating the software.

We are open to working with a sponsor toward clearance of an adaptive algorithm.

We have yet to clear/approve an adaptive algorithm … in Radiology … that we know of.



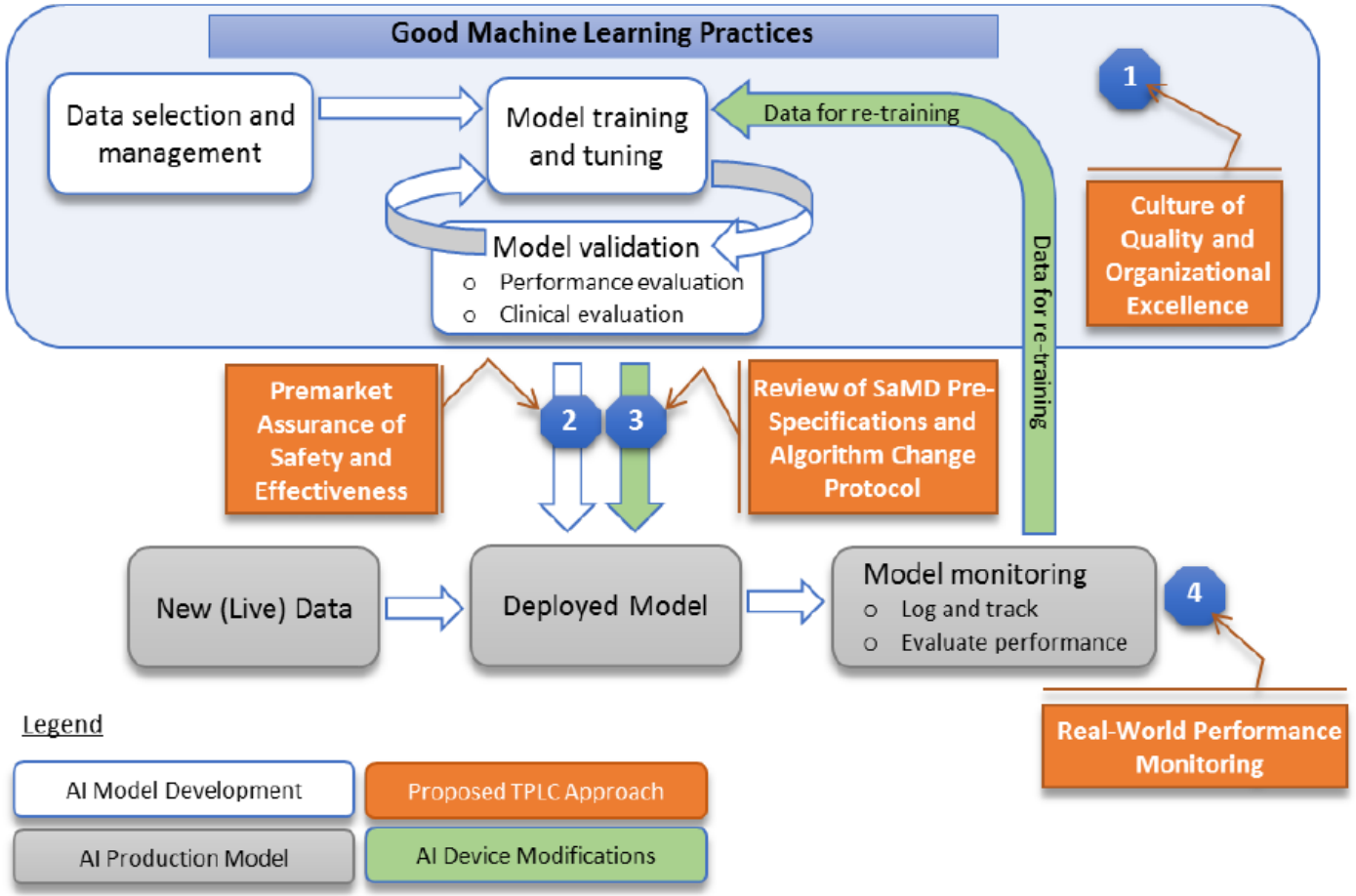
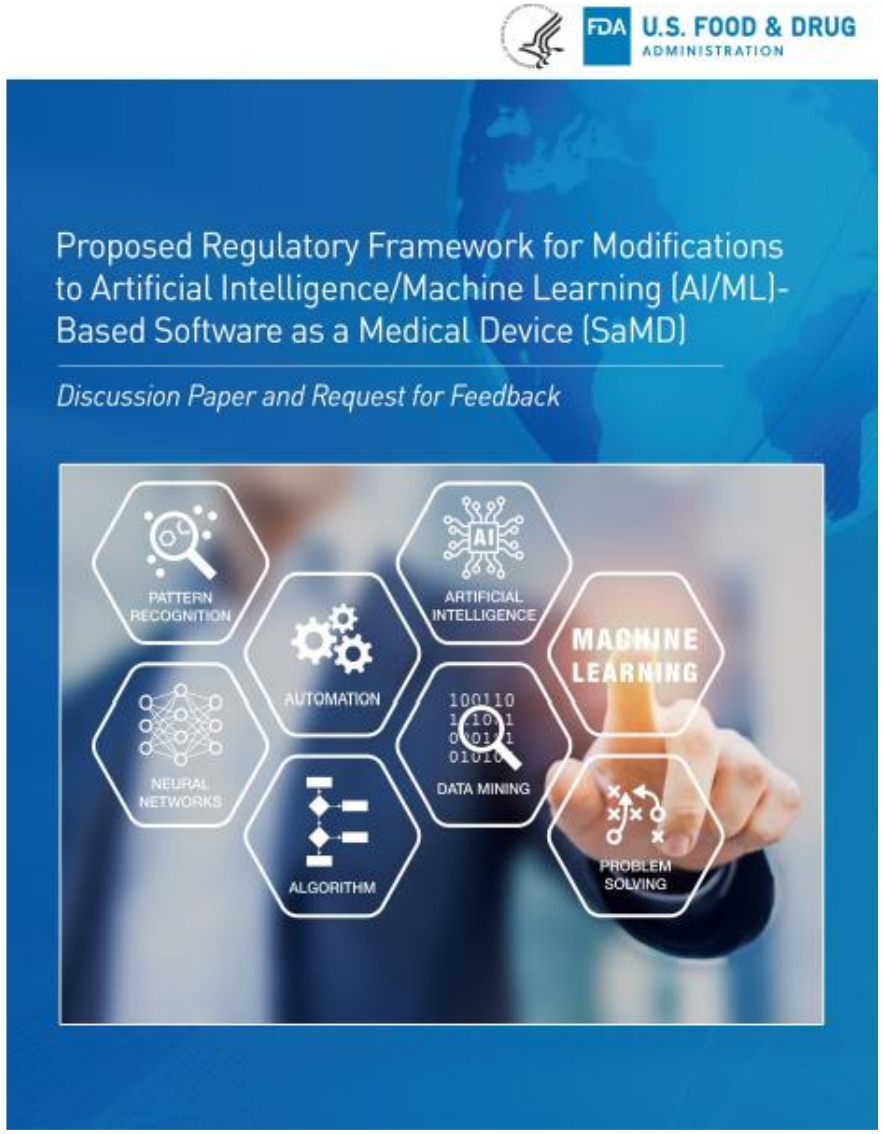


Figure 2: Overlay of FDA's TPLC approach on AI/ML workflow



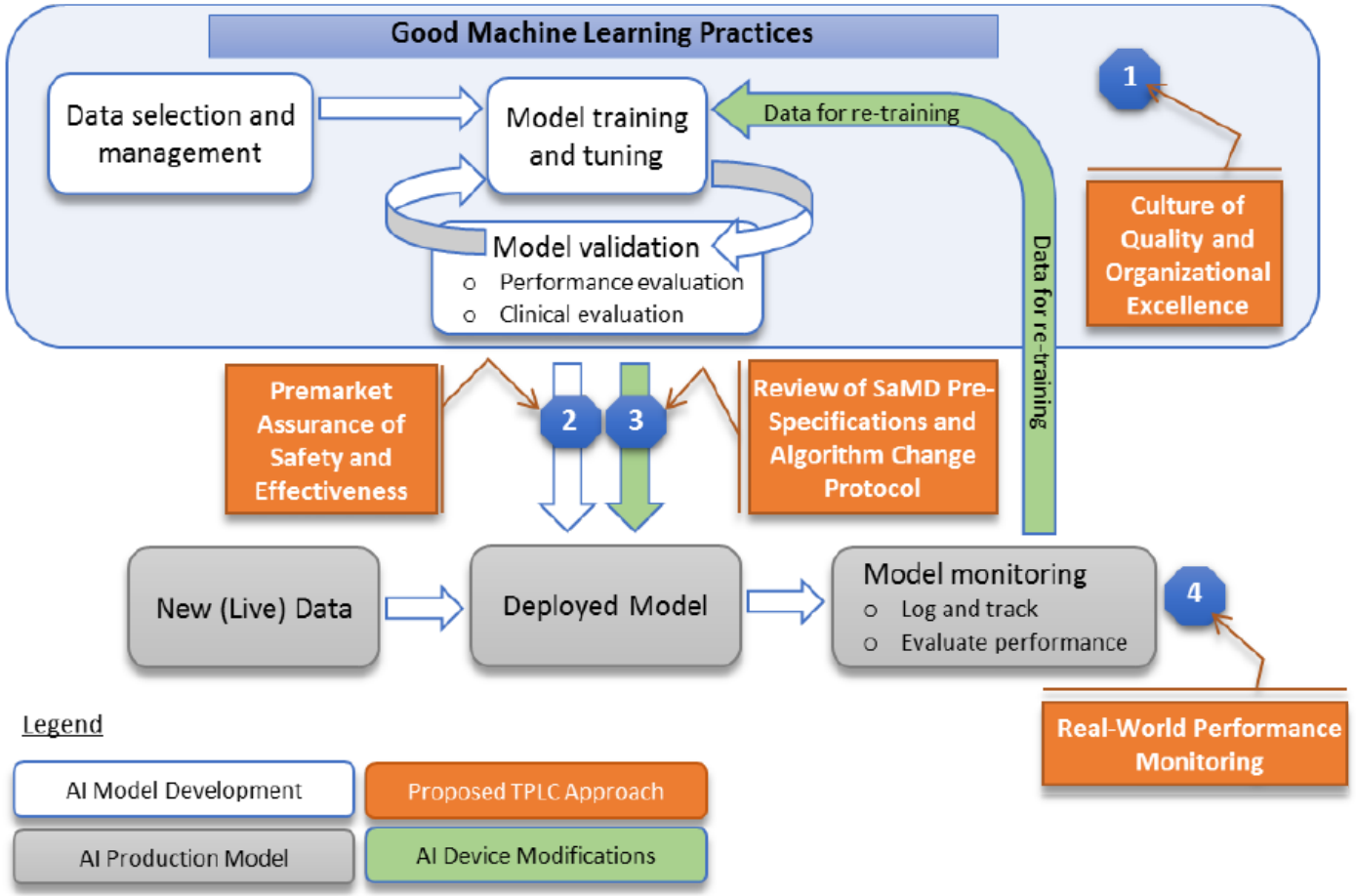
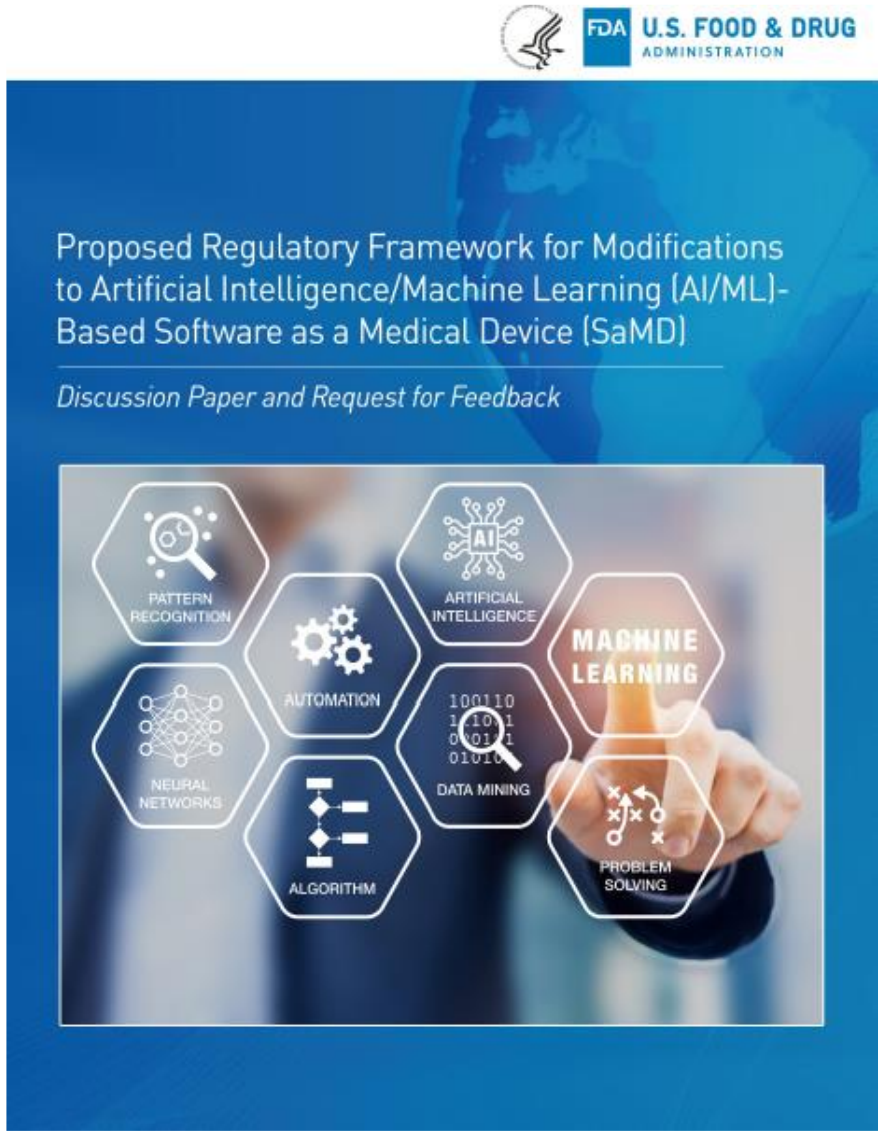


Figure 2: Overlay of FDA's TPLC approach on AI/ML workflow



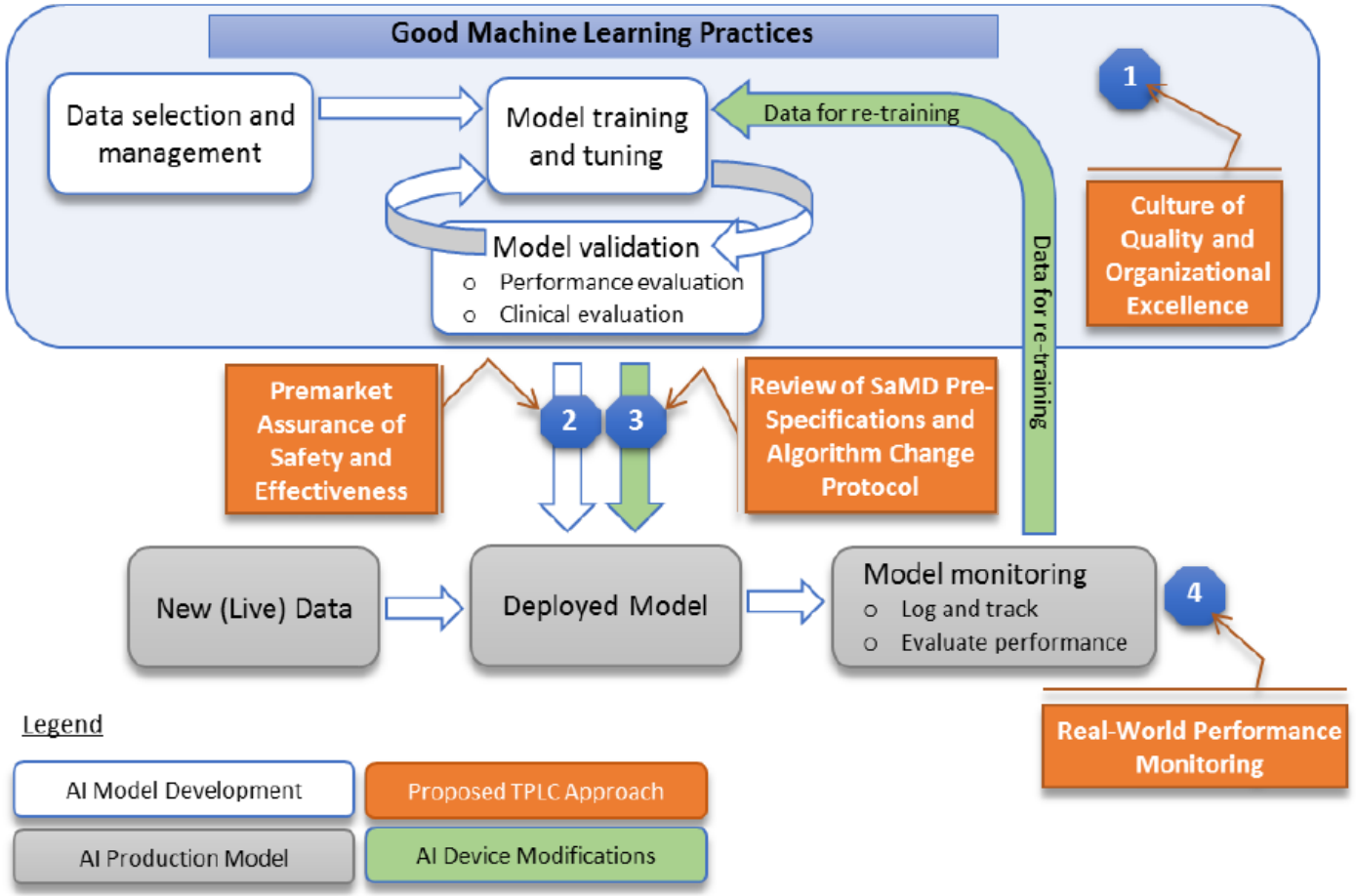


Figure 2: Overlay of FDA's TPLC approach on AI/ML workflow

Data Management	➢ For new training & test data: • Collection protocols • Quality assurance • Reference standard determination ➢ Auditing and sequestration of training and test sets
Re-training	➢ Re-training objectives ➢ Changes related to: • ML methods, including architecture and parameters • Data pre-processing ➢ Criteria to initiate performance evaluation
Performance Evaluation	➢ Assessment metrics ➢ Statistical analysis plans ➢ Frequency and triggers for evaluation ➢ Performance targets ➢ Methods for testing with "clinicians in the loop" when necessary
Update Procedures	➢ Software verification and validation ➢ When and how updates will be implemented ➢ Plans for global and local updates ➢ Communication and transparency to users

Figure 4: Algorithm Change Protocol components

Some History of DIDSR:

Division of Imaging, Diagnostics, and Software Reliability

Bureau of Radiological Health (BRH) → DIDSR

- **1971**, Executive Order: BRH staff reassigned to FDA.
- **1972**: DIDSR's founders helped organize SPIE's first "Medical Imaging" meeting.
- **1982**, Organizational units at the FDA that regulated medical devices and radiation-emitting products merged to form the Center for Devices and Radiological Health (CDRH).

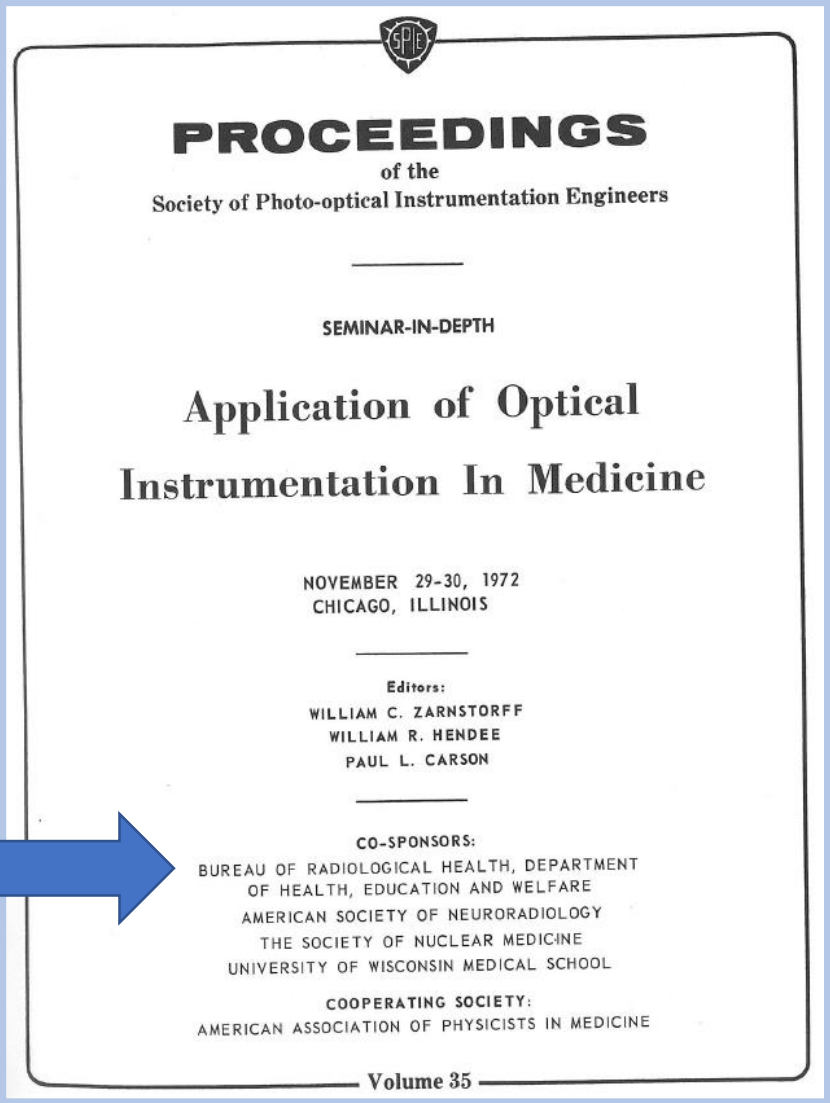
- David Brown, Tom Fewell, Pam Clatterbuck, Roger Schneider, Mal Bruce, Mary Pastel, Ralph Shuping, Robert Jennings, Robert Wagner

Picture circa 1974

Some History of DIDSR:
Division of Imaging, Diagnostics, and Software Reliability

Bureau of Radiological Health (BRH) → DIDSR

- **1971**, Executive Order: BRH staff reassigned to FDA.
- **1972**: DIDSR's founders helped organize SPIE's first "Medical Imaging" meeting.
- **1982**, Organizational units at the FDA that regulated medical devices and radiation-emitting products merged to form the Center for Devices and Radiological Health (CDRH).



PROCEEDINGS
of the
Society of Photo-optical Instrumentation Engineers

SEMINAR-IN-DEPTH

Application of Optical Instrumentation In Medicine

NOVEMBER 29-30, 1972
CHICAGO, ILLINOIS

Editors:
WILLIAM C. ZARNSTORFF
WILLIAM R. HENDEE
PAUL L. CARSON

CO-SPONSORS:
BUREAU OF RADIOLOGICAL HEALTH, DEPARTMENT OF HEALTH, EDUCATION AND WELFARE
AMERICAN SOCIETY OF NEURORADIOLOGY
THE SOCIETY OF NUCLEAR MEDICINE
UNIVERSITY OF WISCONSIN MEDICAL SCHOOL

COOPERATING SOCIETY:
AMERICAN ASSOCIATION OF PHYSICISTS IN MEDICINE

Volume 35

Some History of DIDSR:

Division of Imaging, Diagnostics, and Software Reliability

Bureau of Radiological Health (BRH) → DIDSR

- **1971**, Executive Order: BRH staff reassigned to FDA.
- **1972**: DIDSR's founders helped organize SPIE's first "Medical Imaging" meeting.
- **1982**, Organizational units at the FDA that regulated medical devices and radiation-emitting products merged to form the Center for Devices and Radiological Health (CDRH).

Work of one young DIDSR Founder, Bob Wagner

Image Quality Indices
MTF, NPS
ROC

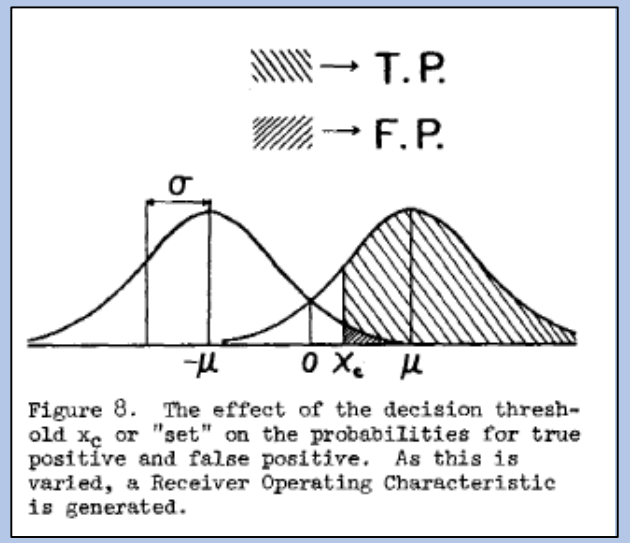


Figure 8. The effect of the decision threshold x_c or "set" on the probabilities for true positive and false positive. As this is varied, a Receiver Operating Characteristic is generated.

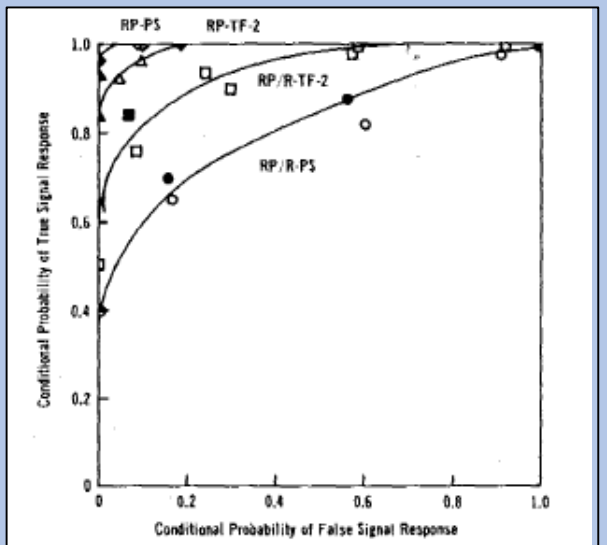


Figure 9. R.O.C. curves obtained by Goodenough for an observer detecting small low contrast object radiographed with various screen-film combinations. The variation of the threshold is achieved by a rating procedure described by Goodenough (Ref. 46; courtesy of David J. Goodenough).

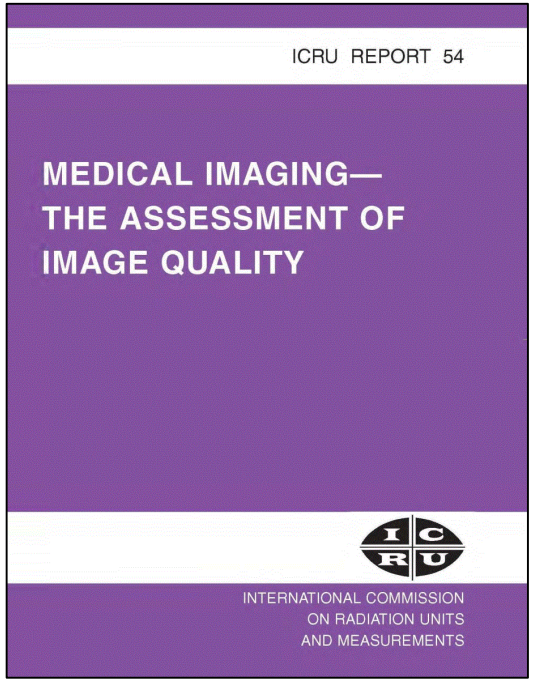
Some History of DIDSR:
Guidance and Consensus Building

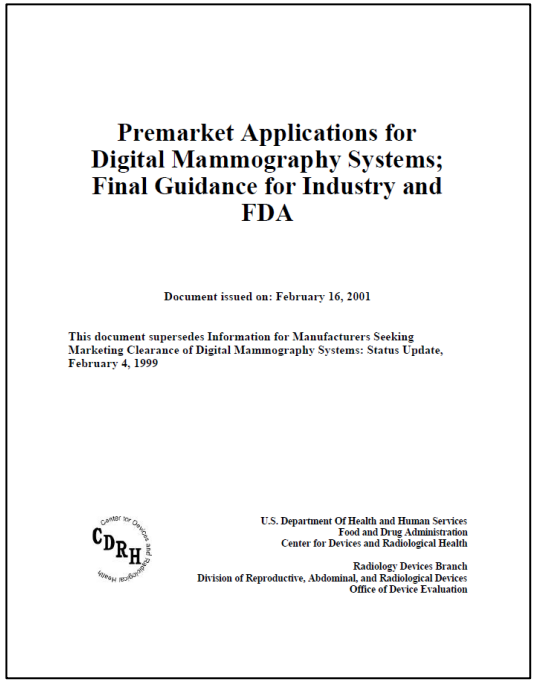
- 1996: ICRU Report 54
- 2001: Guidance on FFDM
- 2008: ICRU Report 79

Image Quality Indices: MTF, NPS, DQE

ROC studies and MRMC analysis
Enriched reader studies

Modeling ideal decision maker
Modeling human decision maker



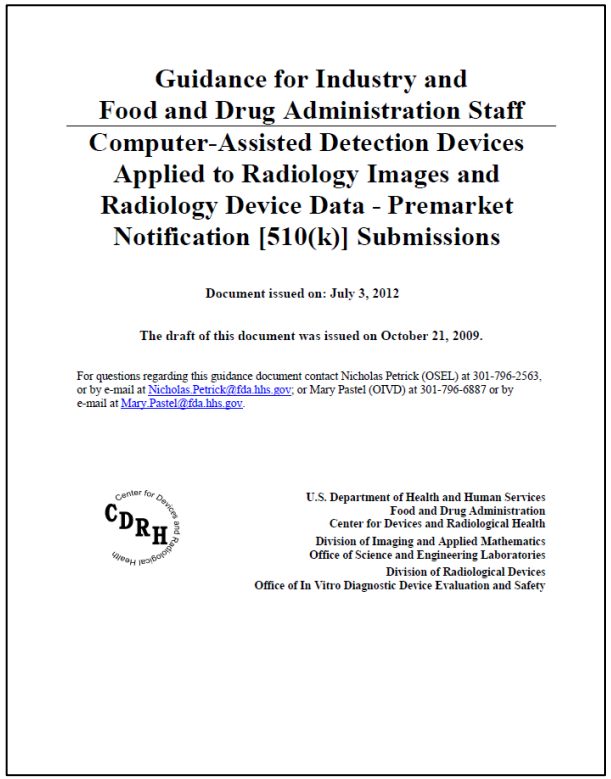


Some History of DIDSR:

Guidance and Consensus Building: Image Quality Evaluation

- 1996: ICRU 54
- 2001: Guidance on FFDM
- 2008: ICRU 79
- 2012: Guidance on CADe

Non-clinical = Stand-alone performance study
No human in the loop



Guidance for Industry and Food and Drug Administration Staff

Computer-Assisted Detection Devices Applied to Radiology Images and Radiology Device Data - Premarket Notification [510(k)] Submissions

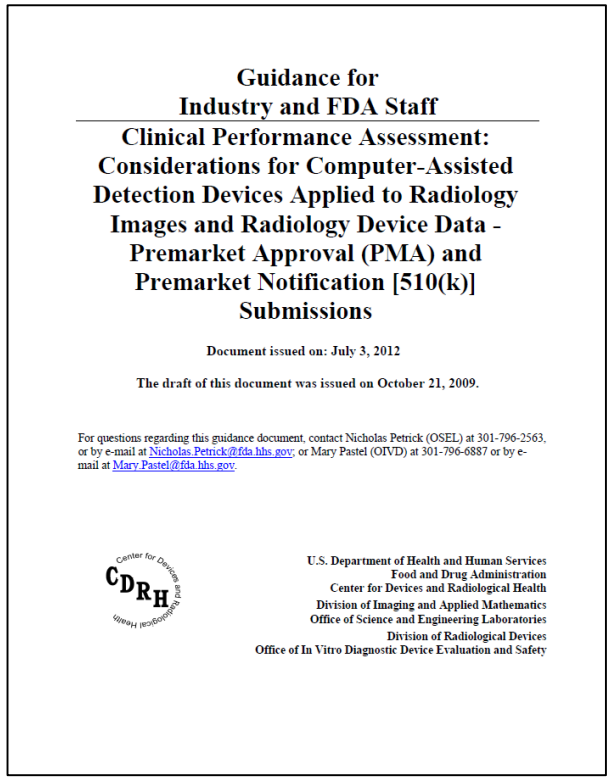
Document issued on: July 3, 2012

The draft of this document was issued on October 21, 2009.

For questions regarding this guidance document contact Nicholas Petrick (OSEL) at 301-796-2563, or by e-mail at Nicholas.Petrick@fda.hhs.gov; or Mary Pastel (OIVD) at 301-796-6887 or by e-mail at Mary.Pastel@fda.hhs.gov.

U.S. Department of Health and Human Services
Food and Drug Administration
Center for Devices and Radiological Health
Division of Imaging and Applied Mathematics
Office of Science and Engineering Laboratories
Division of Radiological Devices
Office of In Vitro Diagnostic Device Evaluation and Safety

Clinical = Reader Study
Human in the loop



Guidance for Industry and FDA Staff

Clinical Performance Assessment: Considerations for Computer-Assisted Detection Devices Applied to Radiology Images and Radiology Device Data - Premarket Approval (PMA) and Premarket Notification [510(k)] Submissions

Document issued on: July 3, 2012

The draft of this document was issued on October 21, 2009.

For questions regarding this guidance document, contact Nicholas Petrick (OSEL) at 301-796-2563, or by e-mail at Nicholas.Petrick@fda.hhs.gov; or Mary Pastel (OIVD) at 301-796-6887 or by e-mail at Mary.Pastel@fda.hhs.gov.

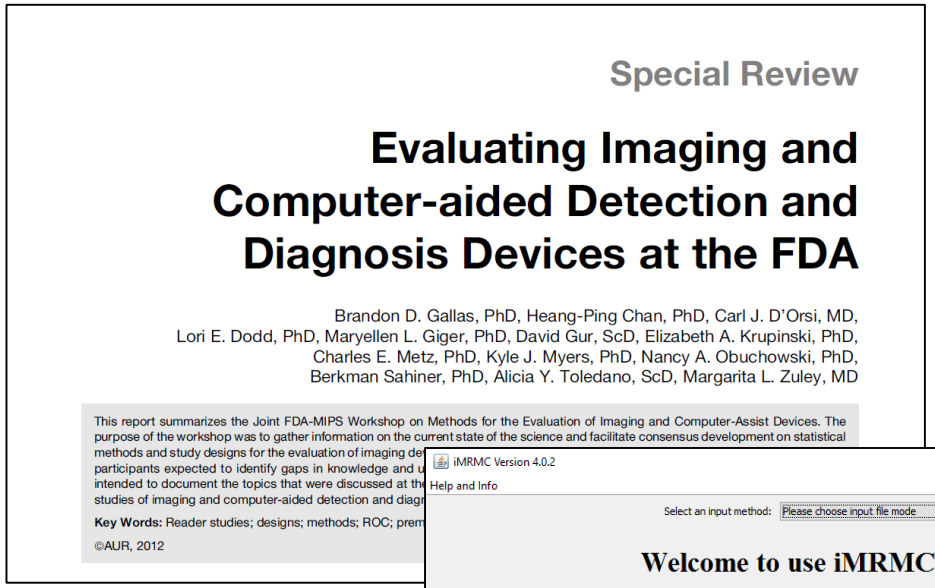
U.S. Department of Health and Human Services
Food and Drug Administration
Center for Devices and Radiological Health
Division of Imaging and Applied Mathematics
Office of Science and Engineering Laboratories
Division of Radiological Devices
Office of In Vitro Diagnostic Device Evaluation and Safety

https://www.fda.gov/regulatory-information/search-fda-guidance-documents/clinical-performance-assessment-considerations-computer-assisted-detection-devices-applied-radiology
https://www.fda.gov/regulatory-information/search-fda-guidance-documents/computer-assisted-detection-devices-applied-radiology-images-and-radiology-device-data-premarket-0

Some History of DIDSR:

Guidance and Consensus Building: Image Quality Evaluation

- 1996: ICRU 54
- 2001: Guidance on FFDM
- 2008: ICRU 79
- 2012: Guidance on CAD
- 2012: Whitepaper on reader studies
- 2013: Software for MRMC analysis of reader studies



Special Review

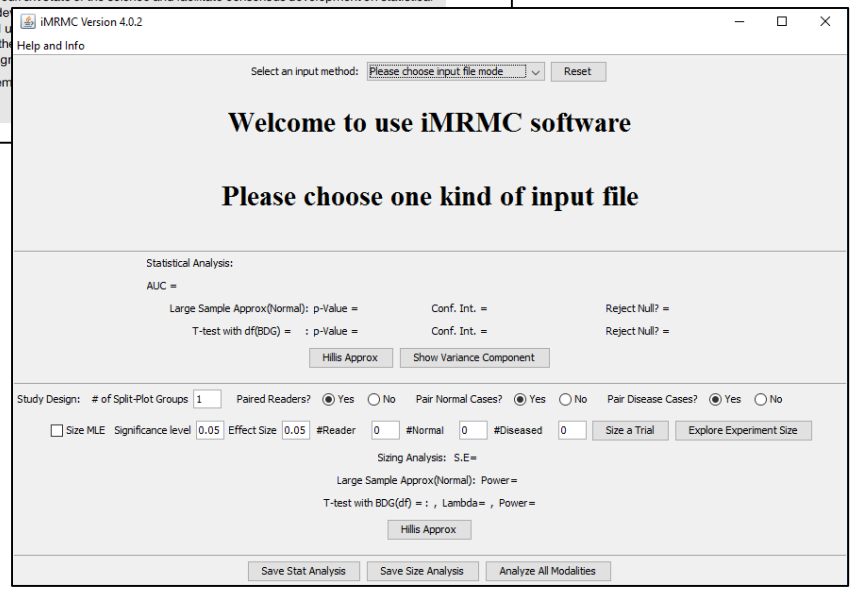
Evaluating Imaging and Computer-aided Detection and Diagnosis Devices at the FDA

Brandon D. Gallas, PhD, Heang-Ping Chan, PhD, Carl J. D'Orsi, MD, Lori E. Dodd, PhD, Maryellen L. Giger, PhD, David Gur, ScD, Elizabeth A. Krupinski, PhD, Charles E. Metz, PhD, Kyle J. Myers, PhD, Nancy A. Obuchowski, PhD, Berkman Sahiner, PhD, Alicia Y. Toledano, ScD, Margarita L. Zuley, MD

This report summarizes the Joint FDA-MIPS Workshop on Methods for the Evaluation of Imaging and Computer-Assist Devices. The purpose of the workshop was to gather information on the current state of the science and facilitate consensus development on statistical methods and study designs for the evaluation of imaging de participants expected to identify gaps in knowledge and u intended to document the topics that were discussed at th studies of imaging and computer-aided detection and diagr

Key Words: Reader studies; designs; methods; ROC; prem

©AUR, 2012



Recent DIDSR Research:

"Impact of prevalence and case distribution in lab-based diagnostic imaging studies"

Full-field digital mammography
vs. screen-film mammography

- 5 sub-studies
- 20 radiologists/study
- 60-175 cases per study
- 20,382 total observations

Study design to reduce regulatory burden.
Examples to follow.
Tools to use.

- Demonstrate:
  - Split-plot study design
  - MRMC analysis tools
  - Prevalence effect on Sensitivity/Specificity
  - ROC curves invariant to prevalence

- **All data, functions, and scripts online:**
  - https://didsr.github.io/viperData/
  - MRMC sample size analysis
  - Electronic case report form
  - Instructions for reporting ROC scores

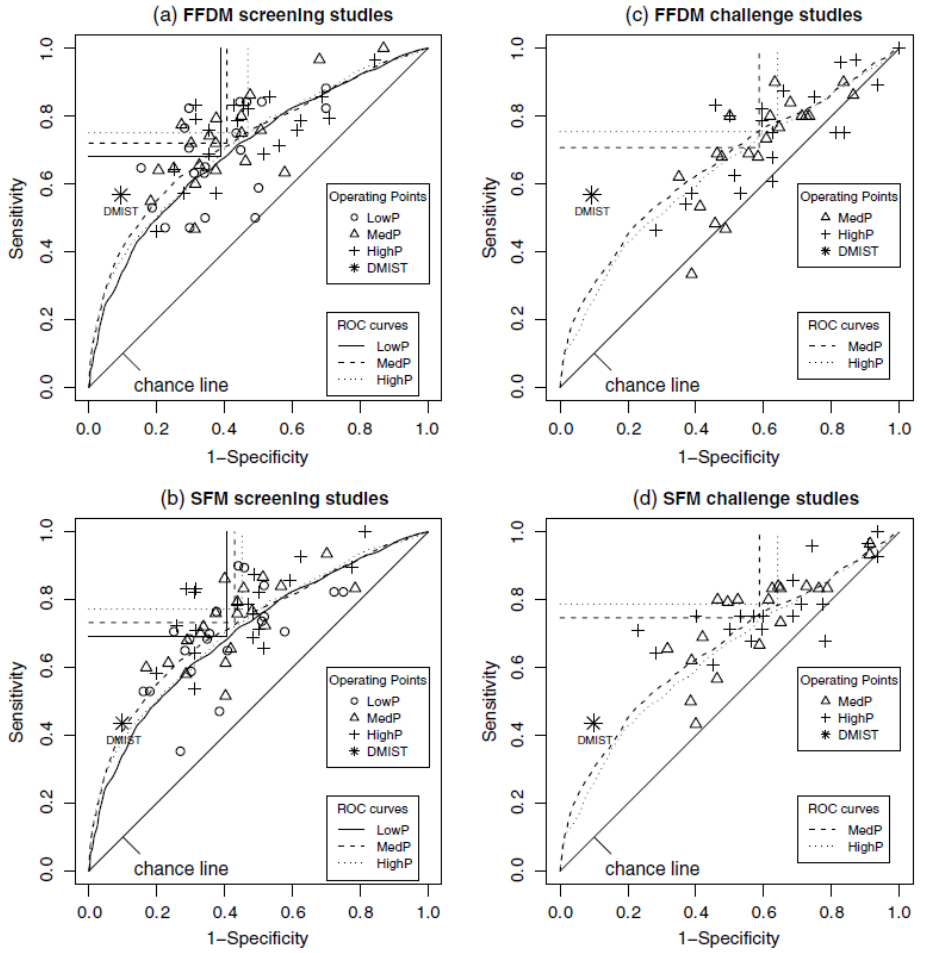


Fig. 3 Plots of reader-averaged ROC curves, reader-averaged (1-Spec., Sens.) operating points (the vertical and horizontal crossings), and reader-specific operating points (denoted by the symbols). Study populations are restricted to women with dense breasts (heterogeneously dense and extremely dense). Reader-averaged ROC curves of different prevalences are very close. Reader-averaged operating points move up and to the right as prevalence increases. (a) FFDM screening studies. (b) FFDM challenge studies. (c) SFM screening studies. (d) SFM challenge studies.

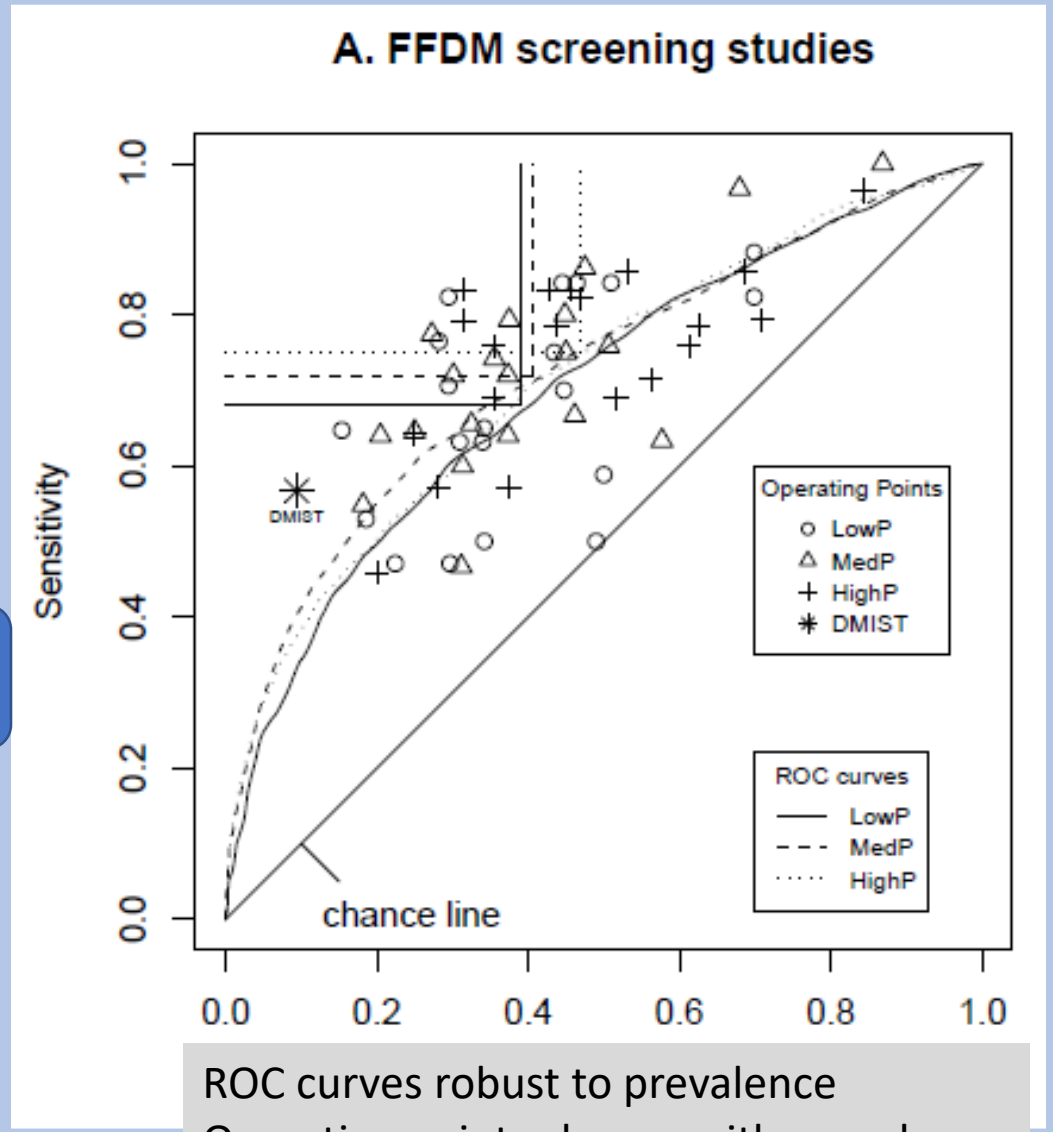
Recent DIDSR Research: VIPER

- "Impact of prevalence and case distribution in lab-based diagnostic imaging studies"
  - Full-field digital mammography vs. screen-film mammography
  - 5 sub-studies
  - 20 radiologists/study
  - 60-175 cases per study
  - 20,382 total observations

Reproducible science.

- Demonstrate:
  - Split-plot study design
  - MRMC analysis tools

- **All** data, functions, and scripts online:
  - https://didsr.github.io/viperData/
  - MRMC sample size analysis
  - Electronic case report form
  - Instructions for reporting ROC scores

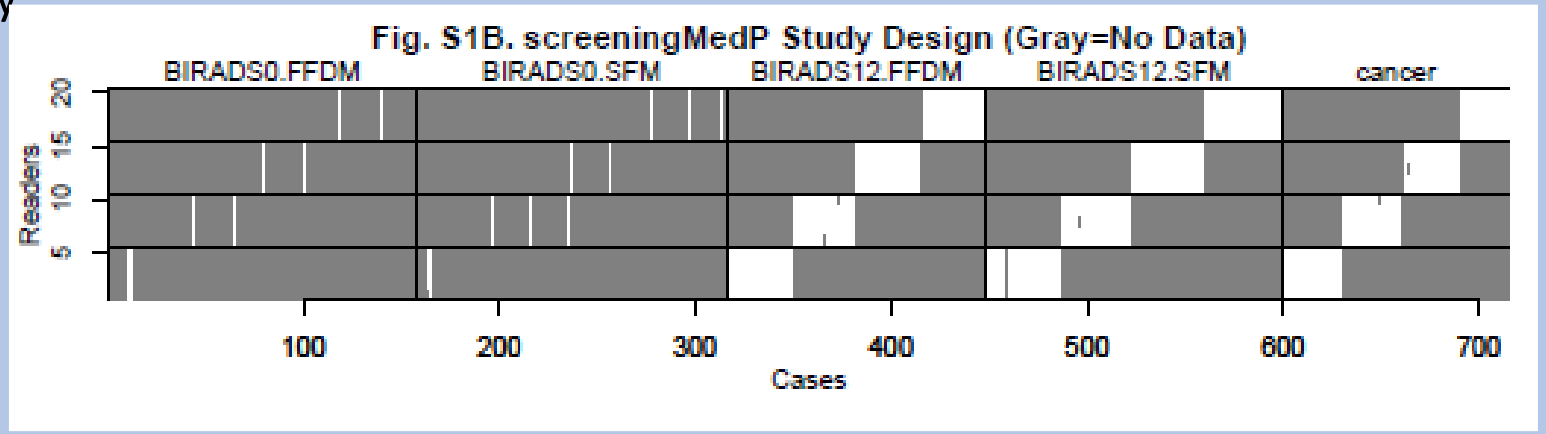


ROC curves robust to prevalence
Operating points change with prevalence
Study operating points don't represent clinical operating point (* DMIST)

Recent DIDSR Research: VIPER

Study design to reduce regulatory burden.

- "Impact of prevalence and case distribution in lab-based diagnostic imaging studies"
  - Full-field digital mammography vs. screen-film mammography
  - 5 sub-studies
  - 20 radiologists/study
  - 60-175 cases per study
  - 20,382 total observations



- Demonstrate:
  - Split-plot study design
  - MRMC analysis tools

- **All** data, functions, and scripts online:
  - https://didsr.github.io/viperData/
  - MRMC sample size analysis
  - Electronic case report form
  - Instructions for reporting ROC scores

Split-plot design

Efficient use of cases, reader workload, and total observations.

Each case is read by multiple readers, reducing the noise from one observation

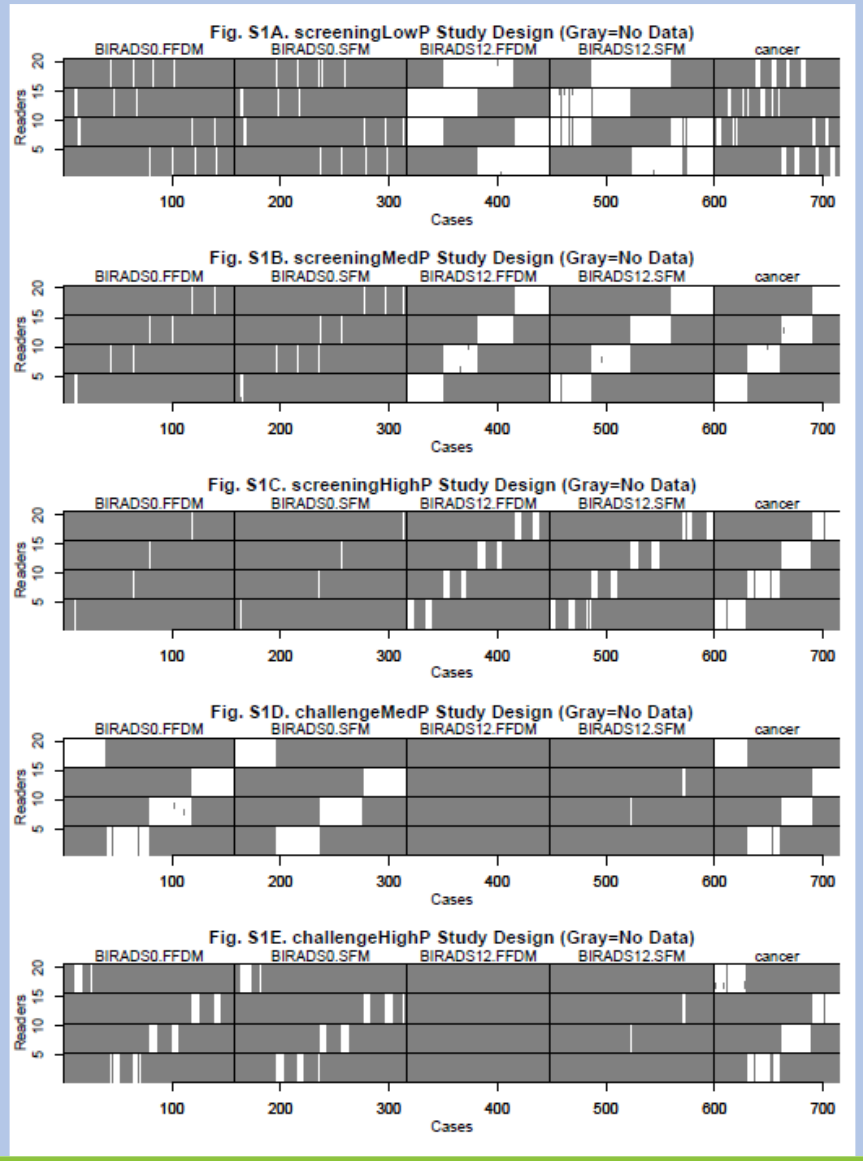
Each case is not read by all readers, avoiding diminishing returns.

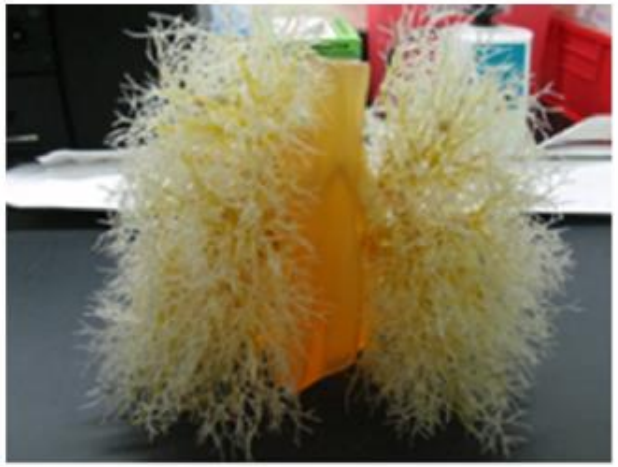
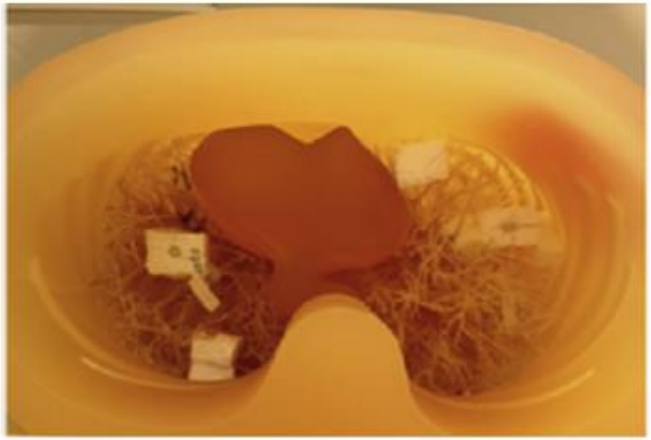
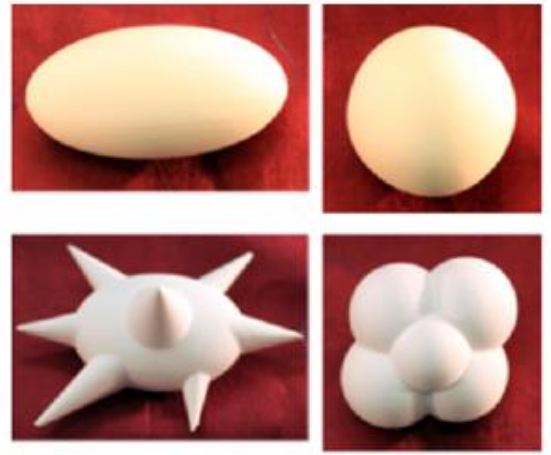
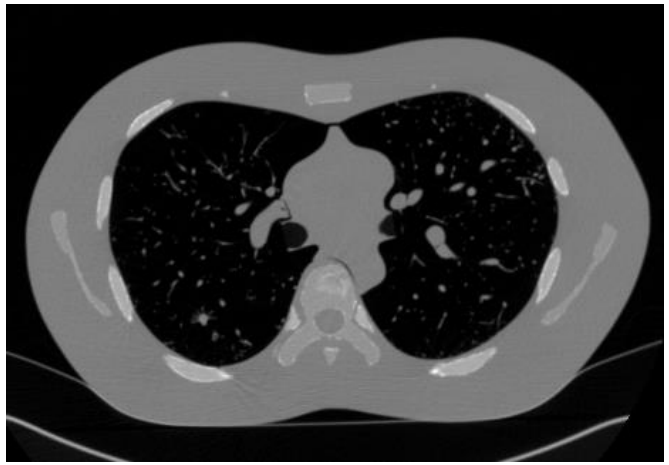
Recent DIDSR Research: VIPER

- "Impact of prevalence and case distribution in lab-based diagnostic imaging studies"
  - Full-field digital mammography vs. screen-film mammography
  - 5 sub-studies
  - 20 radiologists/study
  - 60-175 cases per study
  - 20,382 total observations

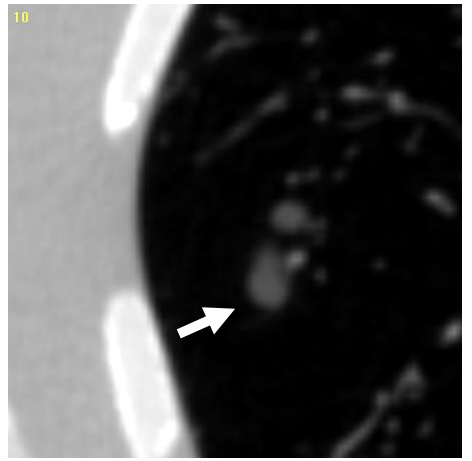
- Demonstrate:
  - Split-plot study design
  - MRMC analysis tools

- **All** data, functions, and scripts online:
  - https://didsr.github.io/viperData/
  - MRMC sample size analysis
  - Electronic case report form
  - Instructions for reporting ROC scores



- Thorax “lung” phantom
- Embedded synthetic nodules
  - Variable shapes, sizes, densities
  - Used in pre-market submissions



Gavrielides et al., “**A resource for the assessment of lung nodule size estimation methods: database of thoracic CT scans of an anthropomorphic phantom**”, *Optics Express*, vol. 18, n.14, pp. 15244-15255, 2010.

4433 scans total. 738 series downloaded per week (avg.) (LINK to Lung Phantom data at TCIA)
TCIA: The Cancer Imaging Archive

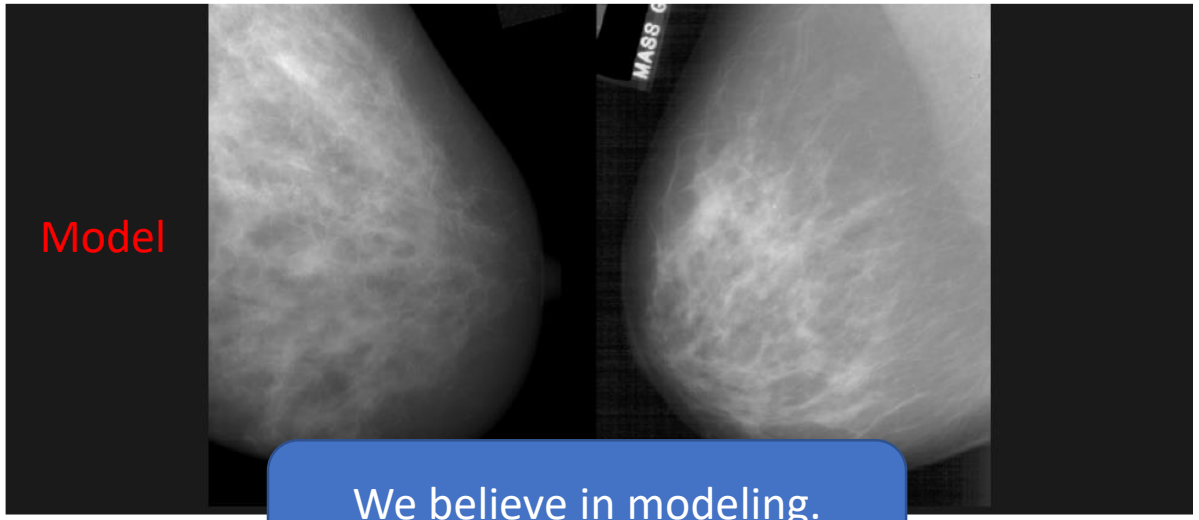
The VICTRE Project: The First All-In-Silico Imaging Clinical Trial

A Badano, A Badal, S Glick, C Graff, F Samuelson,
D Sharma, R Zeng, and K Myers

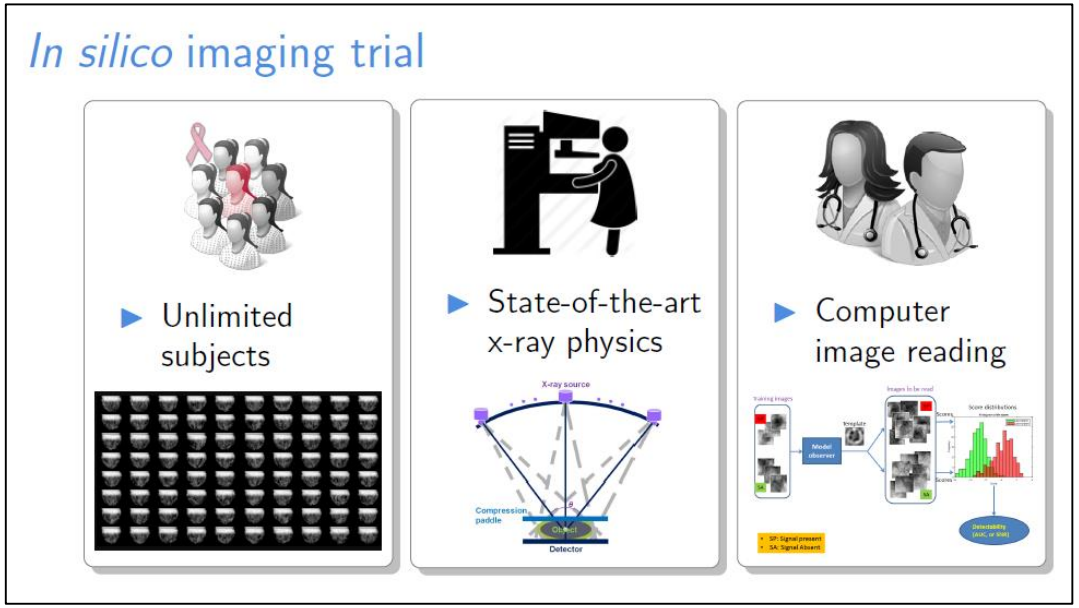
Division of Imaging, Diagnostics, and Software Reliability (OSEL/CDRH/FDA)

FDA U.S. FOOD & DRUG ADMINISTRATION

No humans were harmed in the design or production of this trial …



We believe in modeling.
Physics is evidence.



8749 scans total
7744 scans downloaded per week (avg.)
(LINK to VICTRE data on TCIA)

Badano, A.; Graff, C. G.; Badal, A. & et al (2018), 'Evaluation of digital breast tomosynthesis as replacement of full-field digital mammography using an in silico imaging trial', *JAMA Network Open* ***1(7), e185474-.***

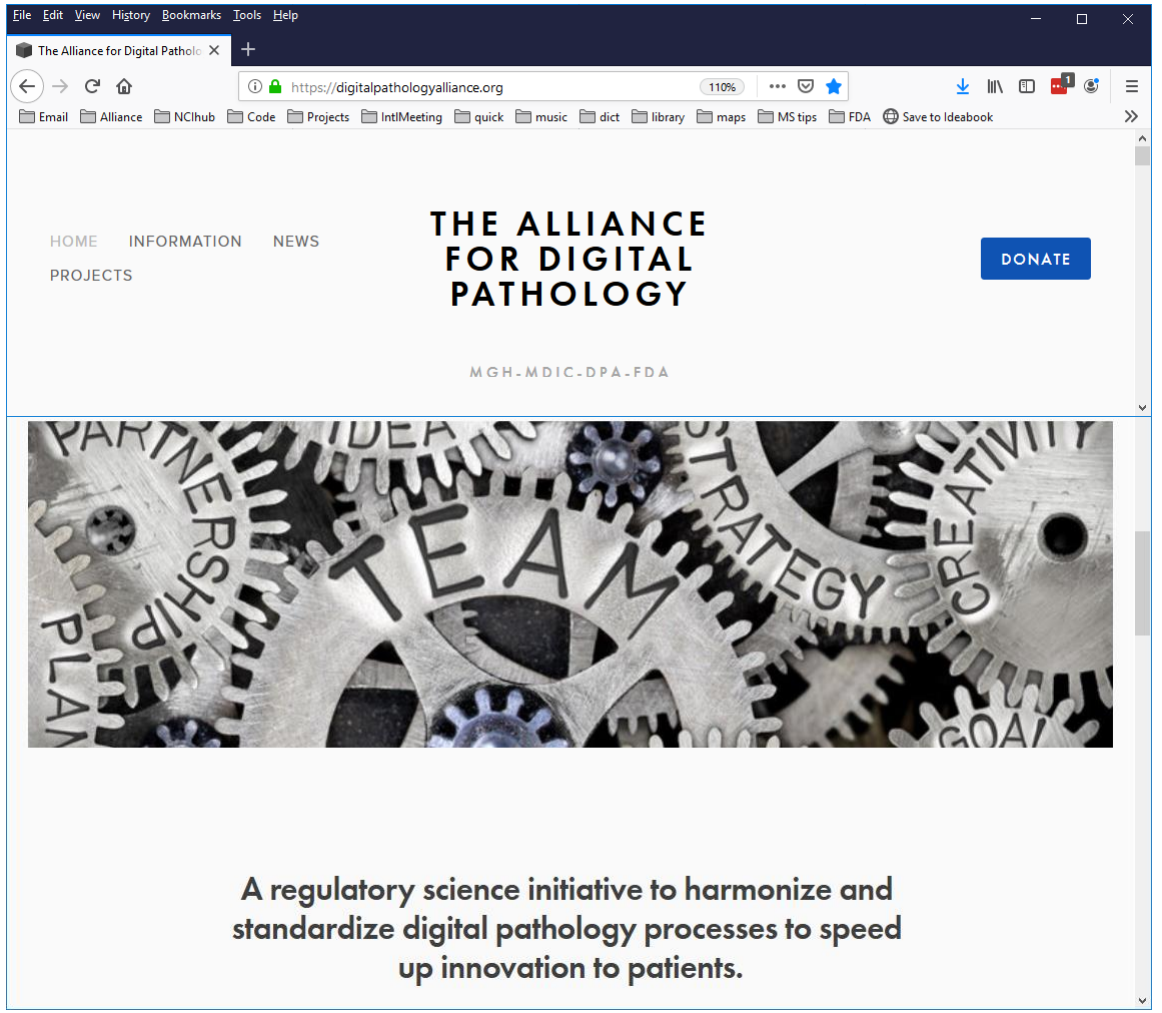
Forming a Collaboration Alliance

Collaboration Alliance

- Kick-off meeting July 18, 2019

- **Objectives:**
  - Clarify and improve regulatory pathways
  - Develop evaluation tools, methods, and standards
  - Tackle large-scale projects in pre-competitive space

https://digitalpathologyalliance.org/

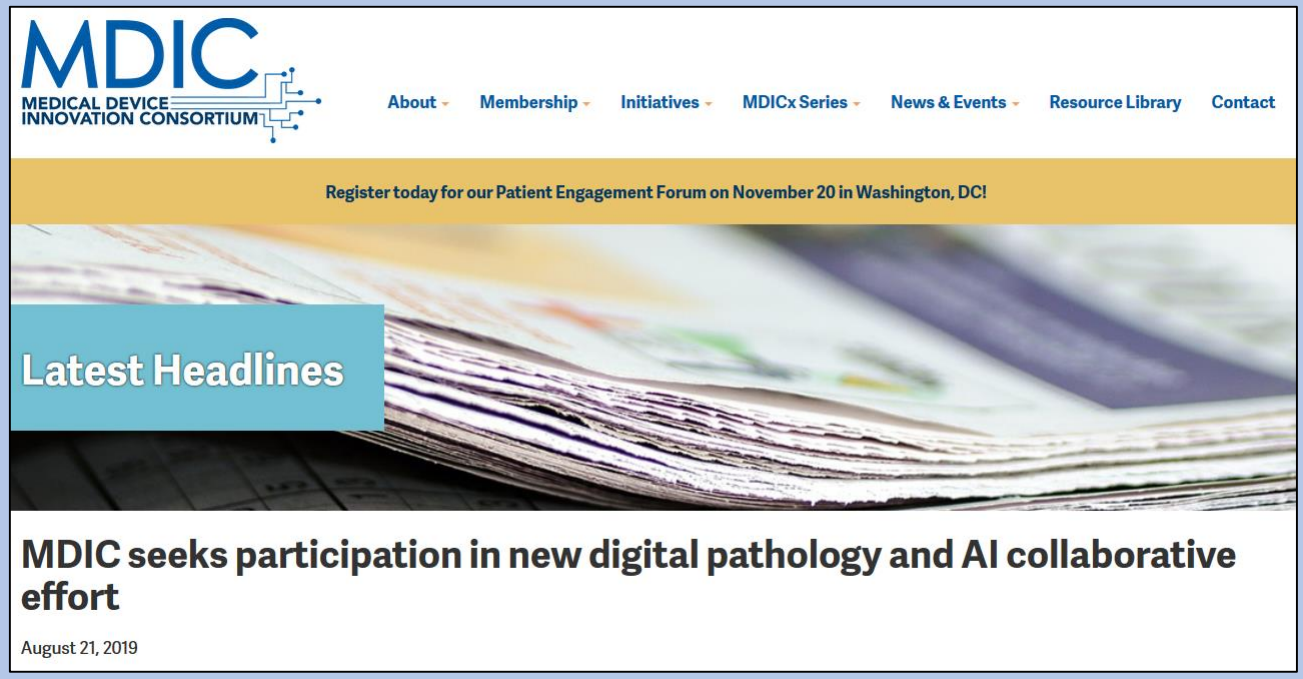


Collaboration Alliance

- Kick-off meeting July 18, 2019

- Stakeholder participants (>50):
  - FDA, NIH, MDIC

https://mdic.org/mdic-seeks-participation-in-new-digital-pathology-collaborative-effort/



MDIC: Public-private partnership with the sole objective of advancing medical device regulatory science for patient benefit.

Collaboration Alliance

- Kick-off meeting July 18, 2019

- Stakeholder participants (>50):
  - FDA, NIH, MDIC
  - Clinical societies: pathology (DPA, CAP) **and radiology (ACR)!**
  - Academic and clinical subject matter experts
  - Patient advocates

DPA: Digital Pathology Association

CAP: College of American Pathologists

https://mdic.org/mdic-seeks-participation-in-new-digital-pathology-collaborative-effort/

https://digitalpathologyassociation.org/dpa-mdic-fda-alliance-meeting

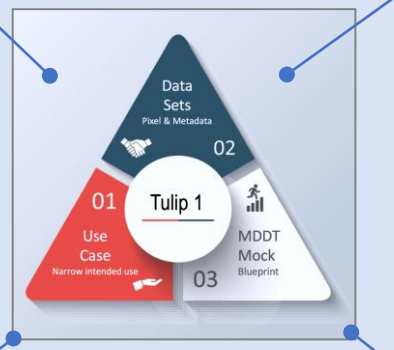
Logos used without permission

Collaboration Alliance: Purpose/Role

Alliance Members

Industry

FDA



SME's

Patients

IDEAS

BIG IDEAS

Alliance Steering Committee

Governance
Prioritization
Administration

ALLIANCE PROJECTS

Potential Project Homes

MDIC MEDICAL DEVICE INNOVATION CONSORTIUM

COLLEGE of AMERICAN PATHOLOGISTS

Association of Directors of Anatomic and Surgical Pathology

DATA SCIENCE INSTITUTE™ AMERICAN COLLEGE OF RADIOLOGY

DIGITAL PATHOLOGY DPA ASSOCIATION

api PATHOLOGY INFORMATICS

IHE INTERNATIONAL Integrating the Healthcare Enterprise

CMS CENTERS for MEDICARE & MEDICAID SERVICES

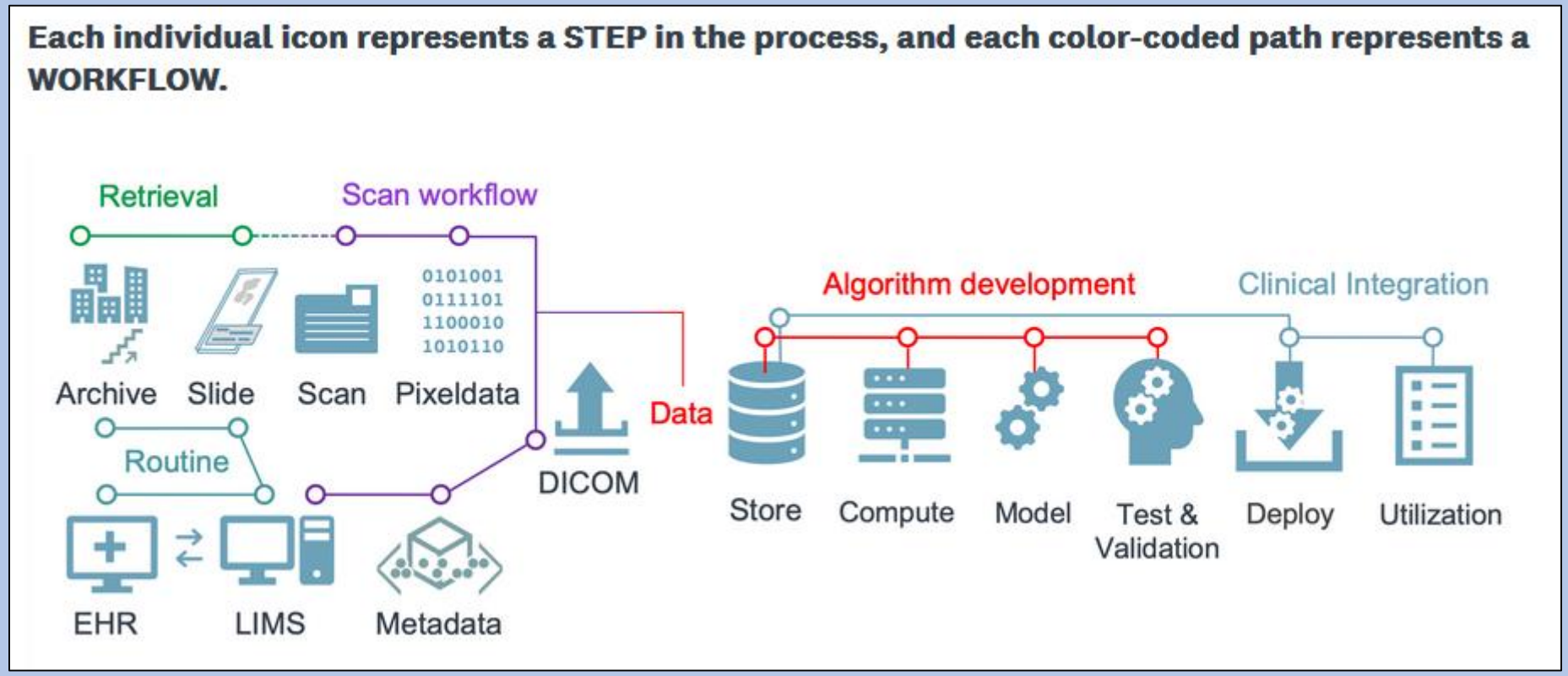
FDA

USCAP

INTERNATIONAL ACADEMY OF PATHOLOGY

Collaboration Alliance: Survey

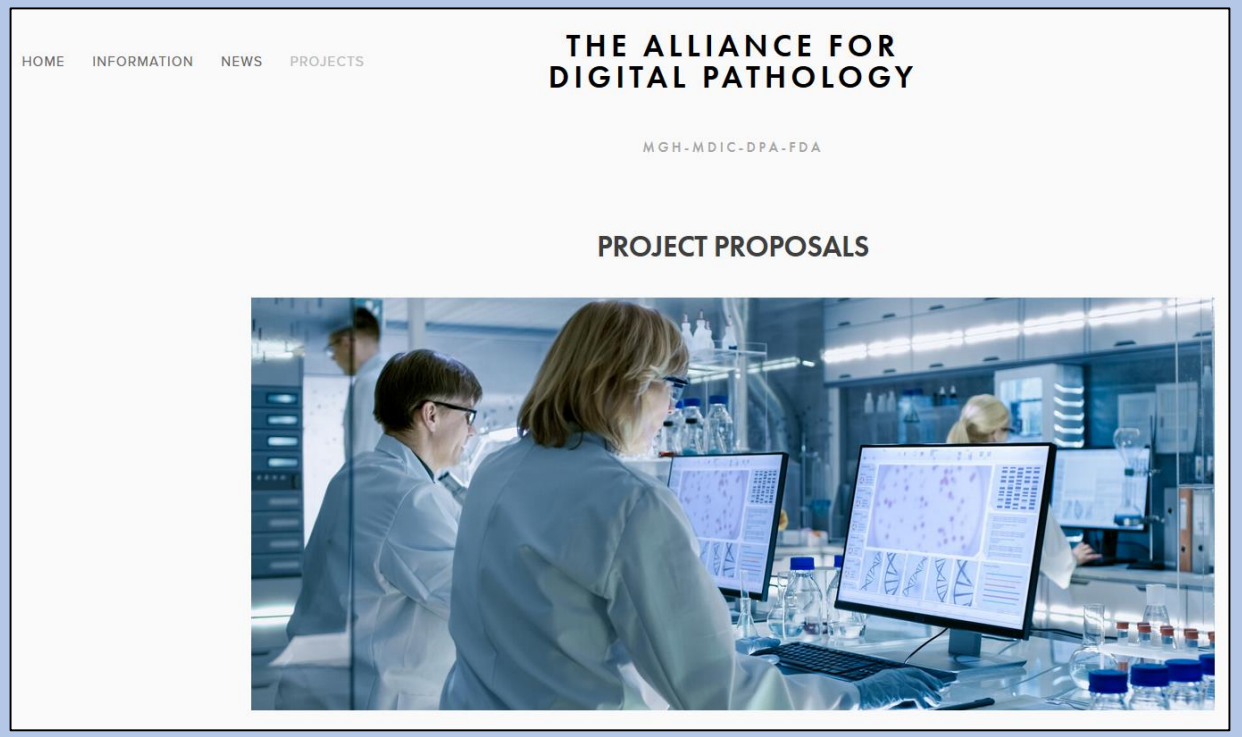
- Express desire to participate
- Express areas of interest



Collaboration Alliance: Project Proposals

- "Who are you looking for in terms of collaborators, supporters, stakeholders?"
- "What is the current challenge? What is your problem statement?"
- What deliverable(s) will your project produce?
- "How will the proposed project be valuable to:"
  - Clinical implementation
  - Regulatory business
  - Research and Development

https://digitalpathologyalliance.org/s/Alliance-Project-Proposal-Blueprint.docx



https://digitalpathologyalliance.org/projects

Collaboration Alliance

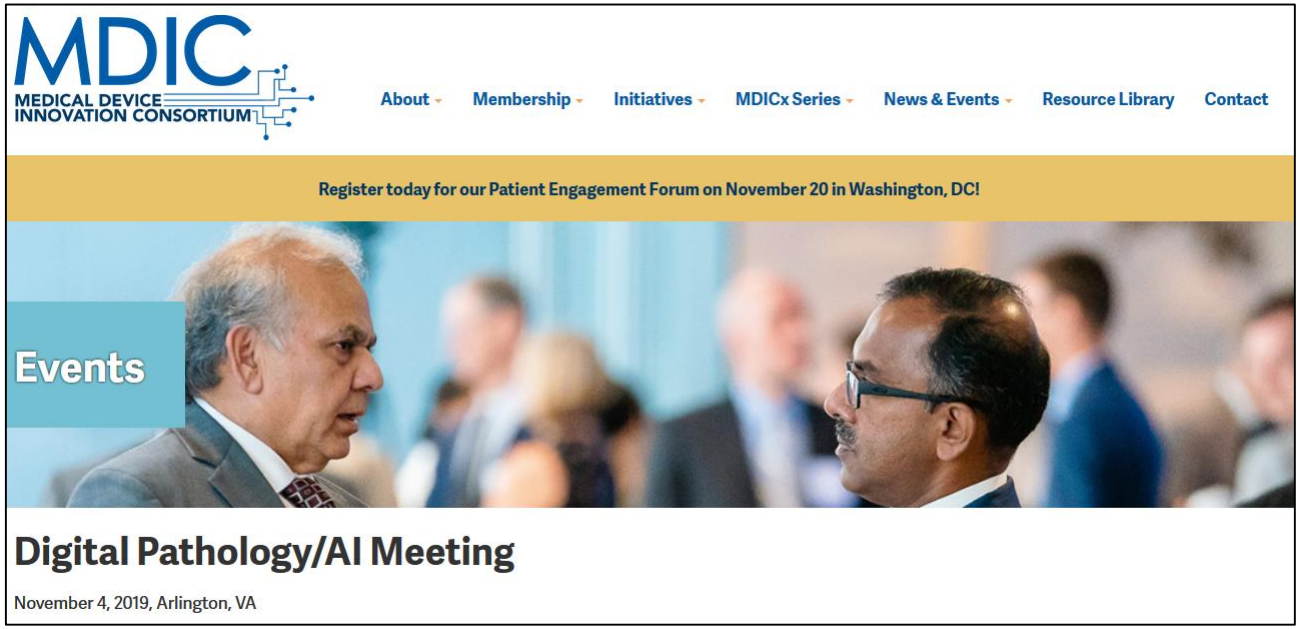
- Kick-off meeting July 18, 2019

Results: Commitments to future meetings

- Brainstorming, spread-the-word meeting
  - Hosted by DPA at their Pathology Visions meeting
  - TODAY in Orlando, FL!

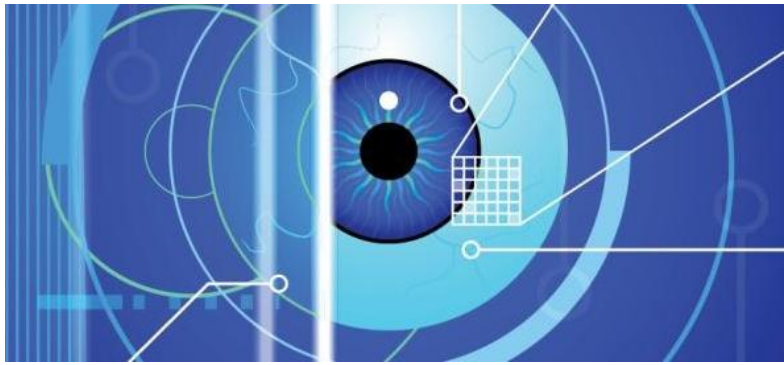
- MDIC Executives and Fellows meeting
  - Engage industry
  - November 4, Arlington, VA.

https://mdic.org/event/digital-pathology-ai-exe/ Meeting website



https://digitalpathologyassociation.org/dpa-mdic-fda-alliance-meeting

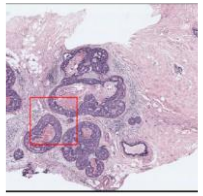
Cognition and Medical Image Perception at NIH THINK TANK, September 12-13, 2019



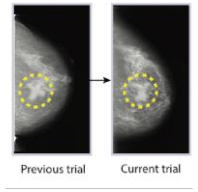
Goal: Reduce diagnostic errors by understanding the role of *human cognitive and perceptual limitations* in medical image interpretation

Sample projects

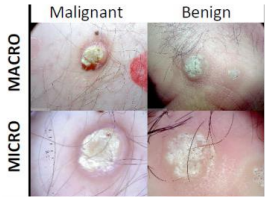
- Reader Accuracy in Pathology Interpretation and Diagnosis: Perception and Cognition (RAPID-PC)



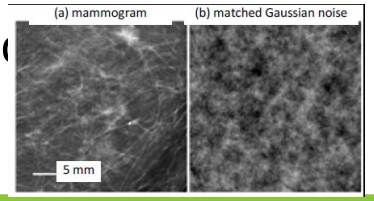
- Isolating and mitigating sequentially dependent perceptual errors in clinical visual search



- Perceptual and Adaptive Learning in Cancer Image Interpretation



- Perceptual sensitivity to anatomical backgroun statistics in mammography



How to encourage collaboration?

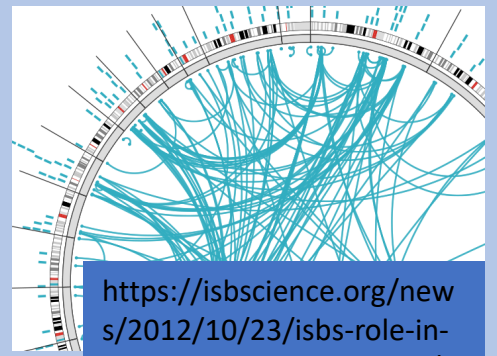
- Pop-up labs at professional conferences
- Embedding psychologists in radiology & pathology departments
- Interested? Ideas? Contact Todd Horowitz todd.horowitz@nih.gov

Hot Topic in Alliance

- Databases for training and testing algorithms
- Need to make use of open platforms, distributed/federated
  - NIH/NCI: The Cancer Genome Atlas (TCGA)
  - NIH/NCI: The Cancer Imaging Archive (TCIA)

https://www.cancer.gov/about-nci/organization/ccg/research/structural-genomics/tcga

https://www.cancerimagingarchive.net/

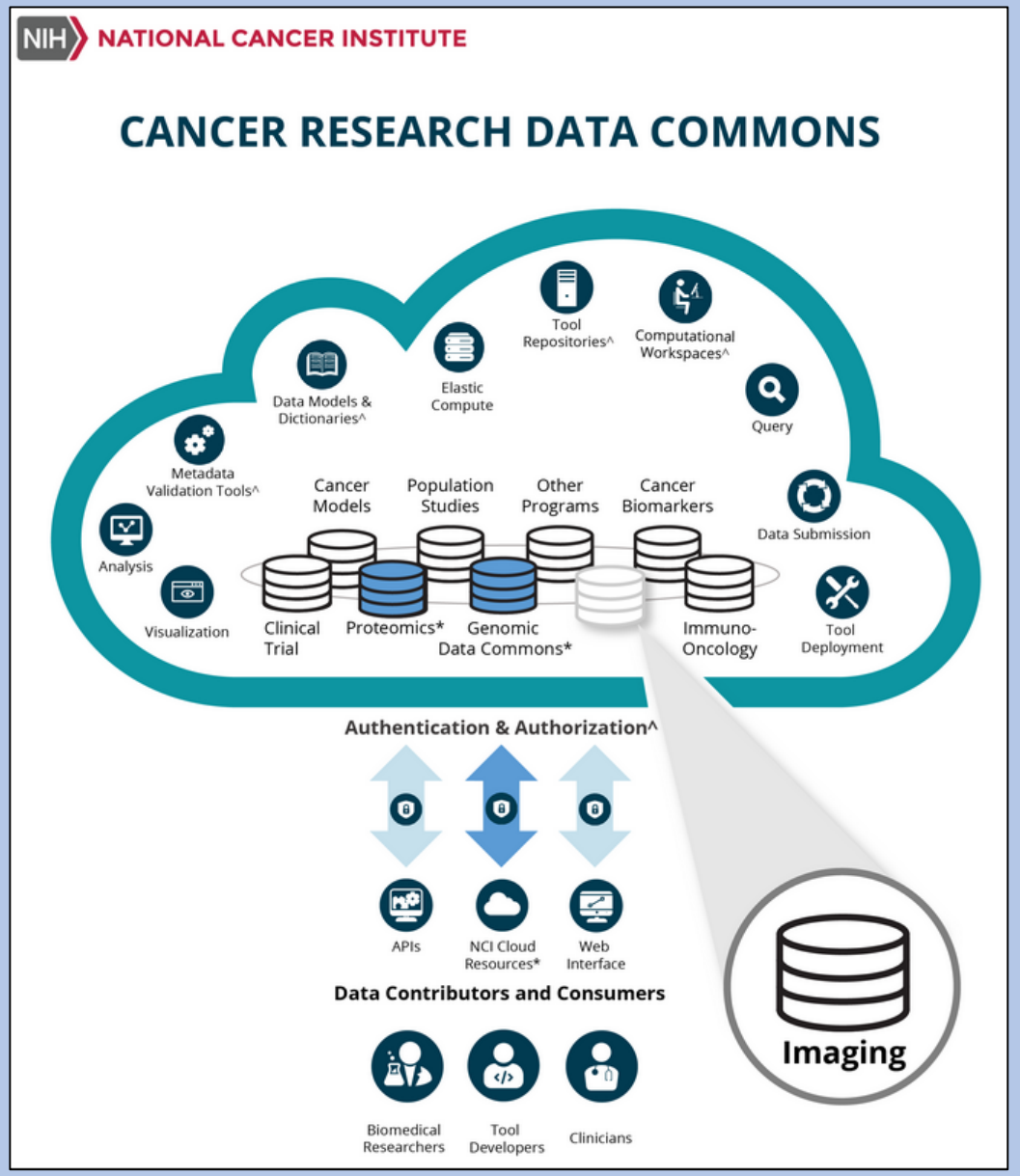


https://isbscience.org/news/2012/10/23/isbs-role-in-the-cancer-genome-atlas/

https://datascience.cancer.gov/data-commons

Hot Topic in Alliance

- Databases for training and testing algorithms
- Need to make use of open platforms, distributed/federated
  - NIH/NCI: The Cancer Genome Atlas (TCGA)
  - NIH/NCI: The Cancer Imaging Archive (TCIA)
  - NCI: Imaging Data Commons

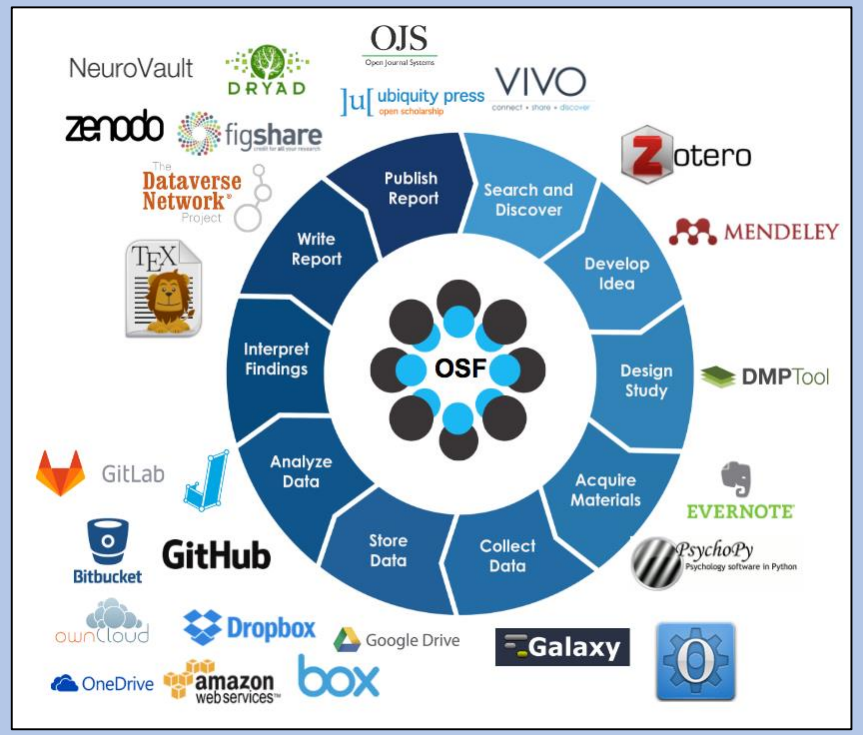


https://datascience.cancer.gov/news-events/blog/award-imaging-data-commons-bringing-multi-modal-imaging-data-cancer-research

Hot Topic in Alliance

- Databases for training and testing algorithms
- Need to make use of open platforms, distributed/federated
  - NIH/NCI: The Cancer Genome Atlas (TCGA)
  - NIH/NCI: The Cancer Imaging Archive (TCIA)
  - NCI: Imaging Data Commons
  - Center for Open Science: Open Science Framework (OSF)
  - Project Data Sphere
  - NEST and more

https://guides.nyu.edu/data_management/osf

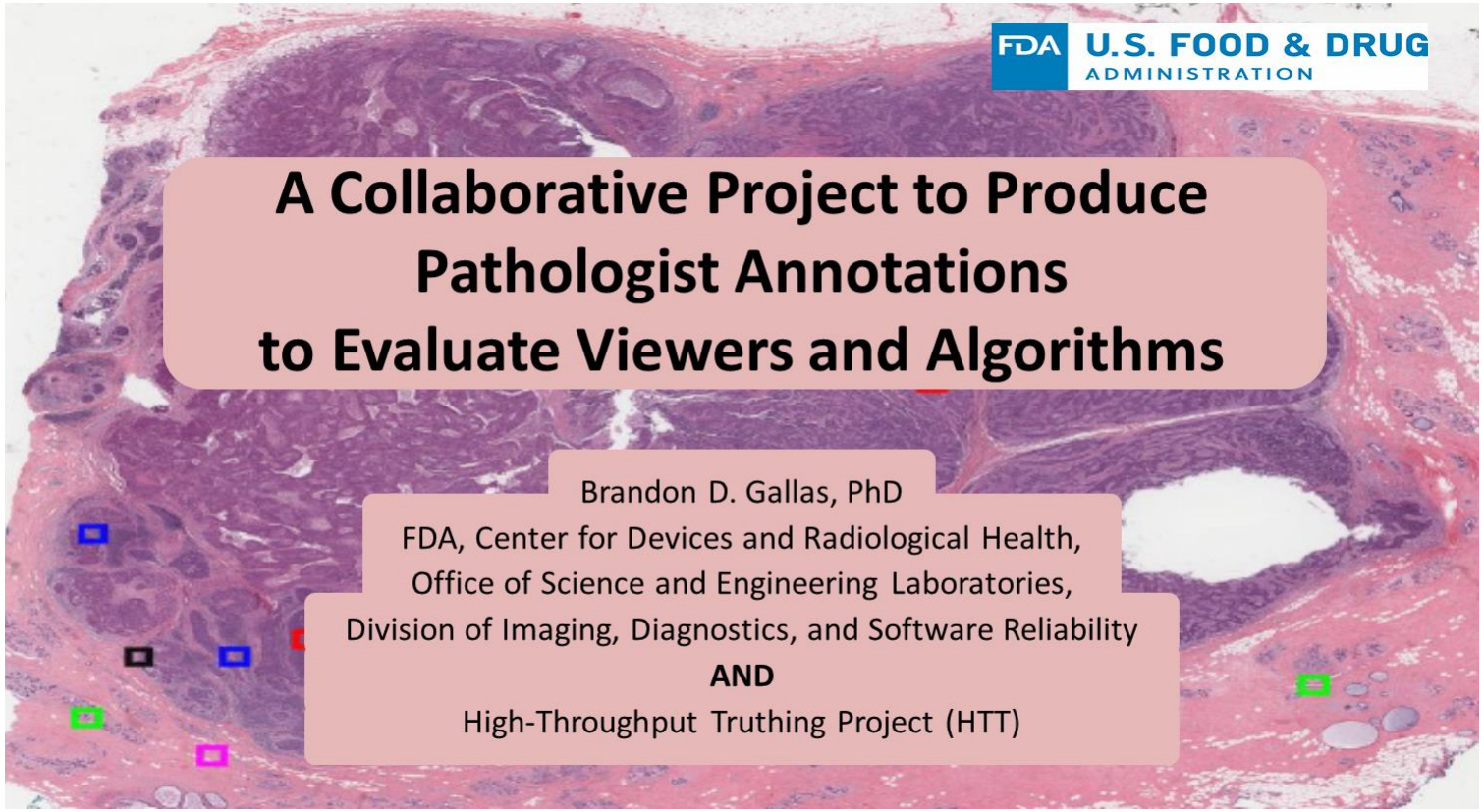


https://www.projectdatasphere.org/projectdatasphere/html/home

https://nestcc.org/

Data MDDT: High-throughput truthing project (HTT)

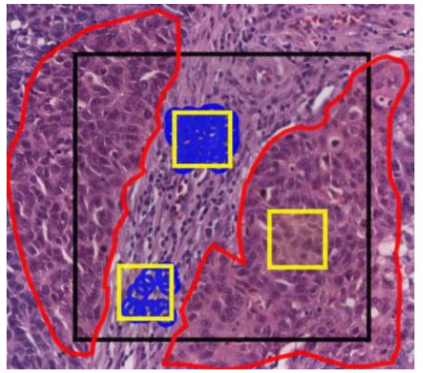
GOAL: Pursue an **MDDT** (*Medical Device Development Tool*) qualification for slides, images, and annotations

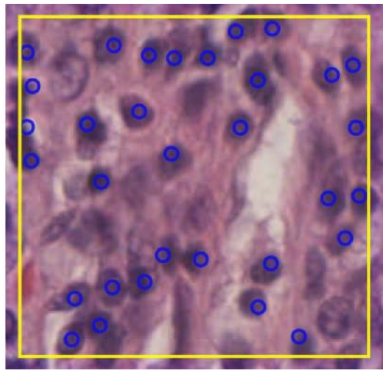


https://ncihub.org/groups/eedapstudies/wiki/HighthroughputTruthingYear2

Data as Tool:

To be available to any algorithm developer to be used to validate their algorithm in a submission to the FDA





Data MDDT: High-throughput truthing project (HTT)

Data MDDT:

- ***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 medical device submissions by multiple sponsors

Building a pathway

Build consensus. Build tools. Disseminate.

Data MDDT: High-throughput truthing project (HTT)

Data MDDT:

- ***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 medical device submissions by multiple sponsors

Building a pathway

Build consensus. Build tools. Disseminate.

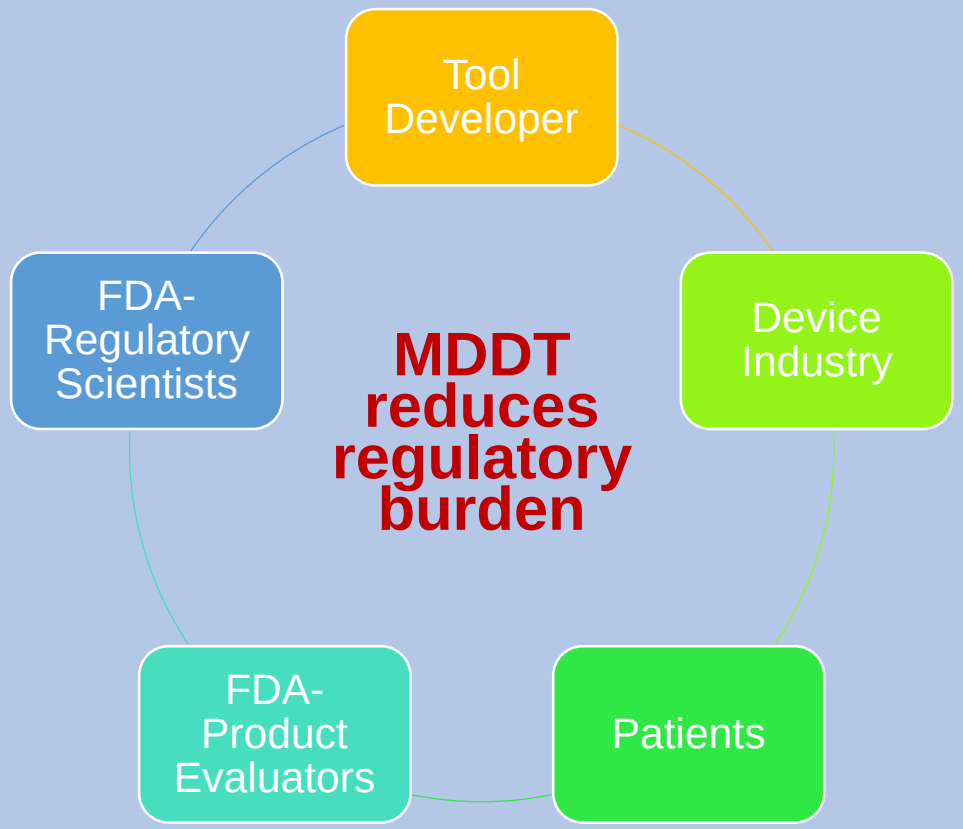
- High-throughput data-collection tools and protocols
- Standardize annotation formats for humans and algorithms
- Statistical methods and software for algorithm performance evaluation

Improve submissions.

Support and enable interoperability.

Medical Device Development Tool Program

Promotes Efficient Medical Device Development



Benefit of Qualifying Tools

- Fosters innovation
- Encourages collaboration
- Reduces resource expenditure
- Qualified MDDT applied in multiple device submissions
- Promotes efficiency in CDRH regulatory review resources
- Minimizes uncertainty in regulatory review process

What Is A Qualified MDDT?

- **Medical Device Development Tool (MDDT)** is a method, material, or measurement used to assess effectiveness, safety, or performance of a medical device
  - MDDT Categories: Clinical Outcome Assessment (COA), Biomarker Test (BT), Nonclinical Assessment Model (NAM)
  - A MDDT is scientifically validated and qualified for a specific ***Context Of Use*** (COU) on the way the MDDT should be used
  - Qualification is a FDA conclusion that within the COU a MDDT has a specific interpretation and application in medical device development and regulatory review

Website:
http://www.fda.gov/MedicalDevices/ScienceandResearch/MedicalDeviceDevelopmentToolsMDDT/default.htm

Questions? email: MDDT@fda.hhs.gov

Mock Submissions:

Representation of a premarket application

- PMA, 510(k), or IDE
- Hypothetical device with hypothetical characteristics and companion information

- ***Reduce uncertainty for sponsors***
  - ***Clarify pathway to market***

- ***FDA may join submission team (consultant) and creates regulatory review team***
  - ***Firewall between two groups***

Building a pathway

Build consensus. Build tools. Disseminate.

- High-throughput data-collection tools and protocols
- Standardize annotation formats for humans and algorithms
- Statistical methods and software for algorithm performance evaluation

Improve submissions.

Support and enable interoperability.

Mock submission history

- **Protein-based multiplex assays**
  - *2008-2010*
  - **IOTF MDx**: Interagency oncology task force, molecular diagnostics subcommittee
  - Origin: IOTF MDx workshop 2008
  - NCI was the sponsor/submitter
- **Virtual patient**
  - *2015-2017*
  - **MDIC**: Medical Device Innovation Consortium
  - Origin: MDIC computational modeling and simulation group
  - MDIC was the sponsor/submitter

- Essential to have FDA review division on board
  - Sees value in devoting resources to mock review
- Essential to have many stakeholders involved
  - Extensive interactions
- Sections submitted:
  - Intended Use
  - Device description
  - Analytical studies
  - Clinical trial protocol
  - Statistical evaluation plans

Summary

- Regulatory science and decisions are built on
  - Sound arguments that demonstrate safety and effectiveness
  - Documentation
  - Generalizability
  - Reproducibility
- DIDSR has been in radiology's business for a long time
  - Work is relevant today
  - Consensus building
- Physics, modeling, and simulation are forms of evidence
- Recent research emphasizes sharing data and digital tools

  … reducing regulatory burden
- We survive on collaborations and look forward to big projects
- We want to inform the community

Thank You!

- Wiki page: links to guidance documents, special controls, and examples
  - https://ncihub.org/groups/eedapstudies/wiki/DeviceAdvice
  - Post our slides here … at the bottom
- Collaboration Alliance "Executives and Fellows" meeting at MDIC
  - Nov. 4, Arlington, VA
  - Relevant and aligned with ACR/DSI
- FDA Public Workshop: Applications of AI-Assisted Radiology
  - Opportunity to work with stakeholders: e.g., industry, clinical practice, academia, government agencies, and patients

Search "NCIhub device advice"

This talk is based on FDA's Current Thinking

Our current thinking changes over time just like science!