Decision Summary

Lessons from the first FDA Approval of an AI Tool in Diagnostic Pathology

Featuring

Paige

MAYO CLINIC

FDA

Jointly presented by

DIGITAL PATHOLOGY DPA ASSOCIATION

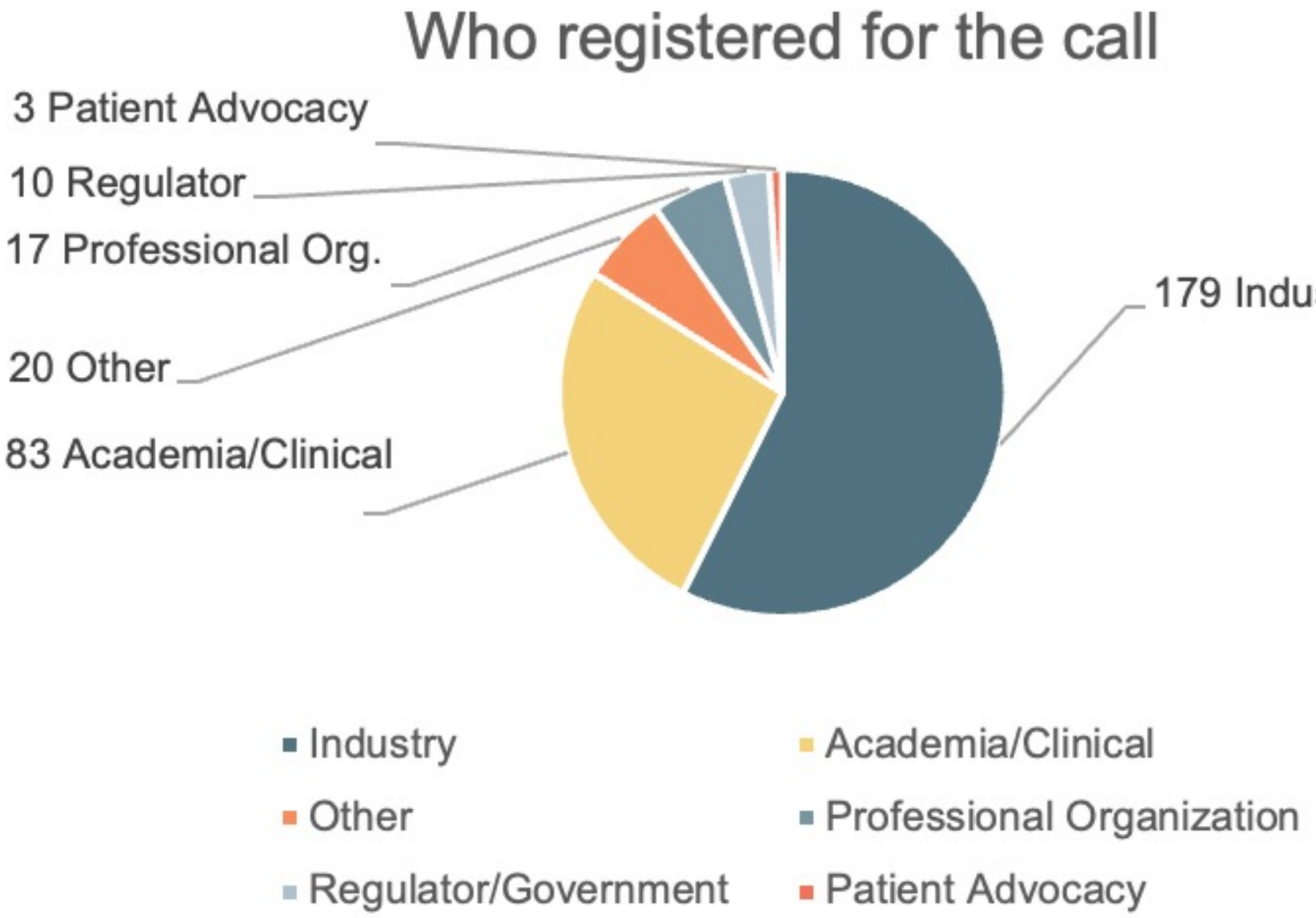
Pathology Innovation Collaborative Community PIcc

- Regulatory science initiative
- Facilitate Innovation
- Advance safety and effectiveness evaluations
- Harmonize approaches to speed delivery to patients
- Collaboration in the **pre-competitive space**
- Open to all stakeholders

https://pathologyinnovationcc.org/

N=324

Thank you

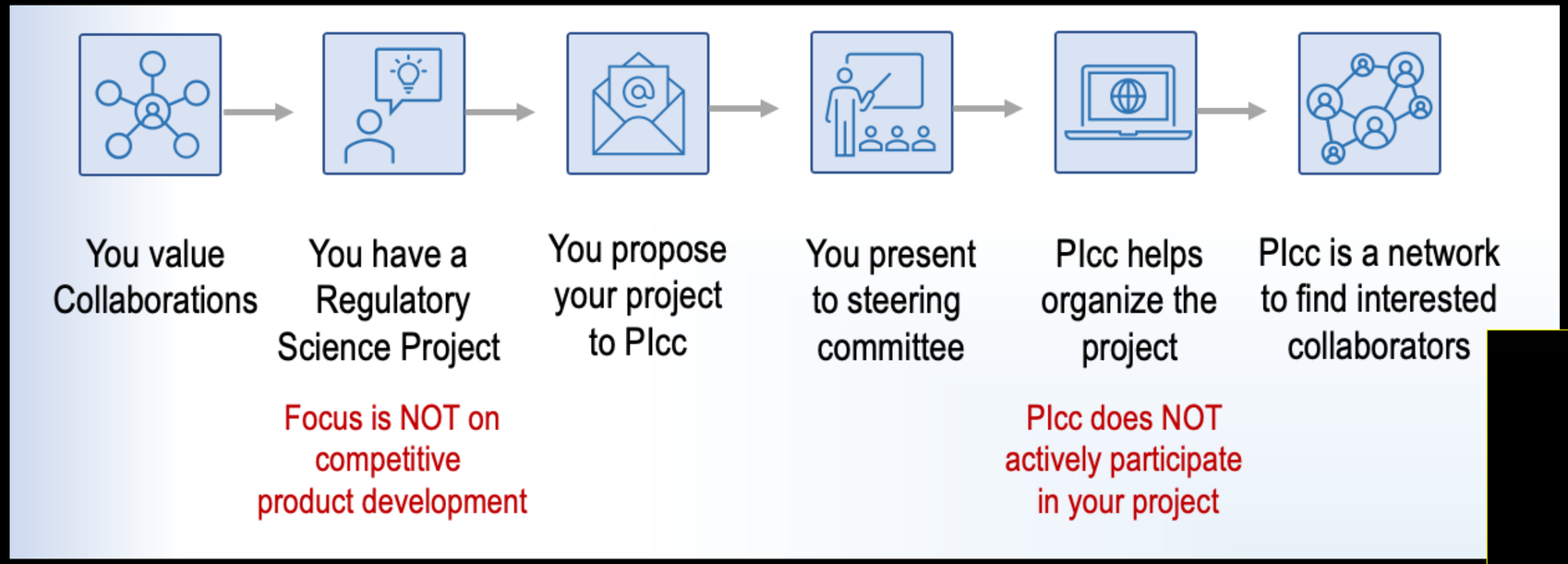


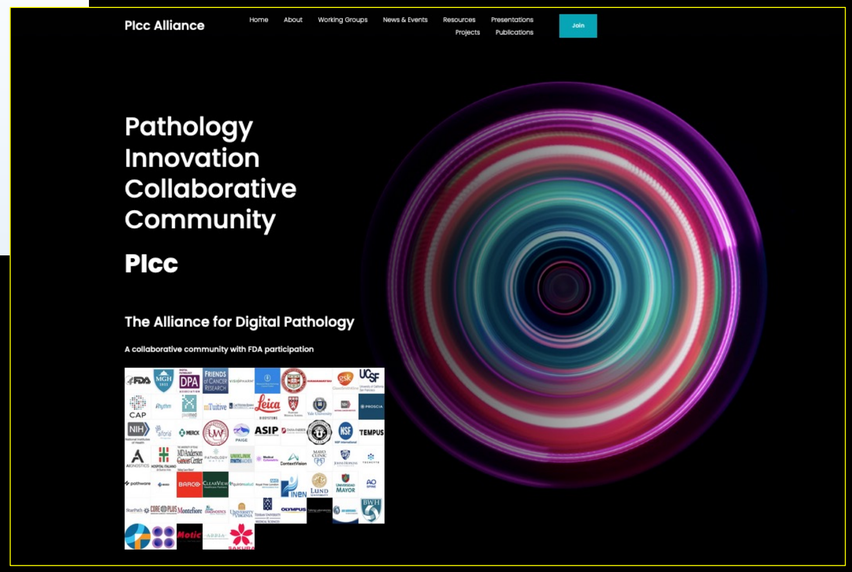
Antitrust monitoring is needed

These meetings need to stay within protected subject matters and need to be monitored so that they do not stray off into inappropriate areas, such as:

- Pricing and price terms
- Sales and service territories for particular products
- Customers and customer territories
- Each company's individual decisions regarding selection of suppliers or customers
- Marketing plans and especially future marketing plans or new product offerings
- Other proprietary or competitively sensitive information

PIcc provides the infrastructure to connect all stakeholders





https://pathologyinnovationcc.org/

DIGITAL PATHOLOGY

DPA

ASSOCIATION

Mission: facilitate awareness, education and adoption of digital pathology and AI applications in healthcare and life sciences.

DPA fosters an exchange of ideas helping members understand, navigate, and influence the future of pathology.

JOIN OUR COMMUNITY & CONNECT WITH 2,600+ DIGITAL PATHOLOGY PROFESSIONALS!

DPA COLLABORATE

The online community for members to connect, innovate, and learn.

Share ideas, pose questions, and network with your peers

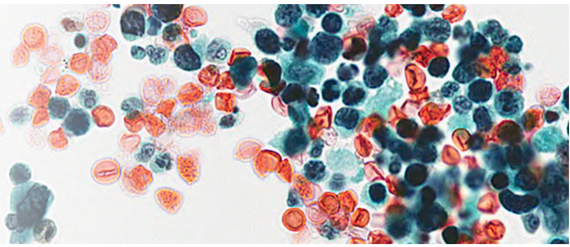
LOG IN

DIGITAL ANATOMIC PATHOLOGY ACADEMY (DAPA)

WSI EDUCATIONAL PLATFORM PROVIDED BY THE DPA FOR ITS MEMBERS

Cloud-based platform provides annotated digital slides with diagnosis and relevant information of morphology and ancillary testing

MEMBER LOGIN



DIGITAL PATHOLOGY DPA ASSOCIATION

Beyond the Scope

a podcast focusing on the hot topics in digital pathology

Listen on Apple Podcasts

Listen on Spotify

THE LENS OF INNOVATION

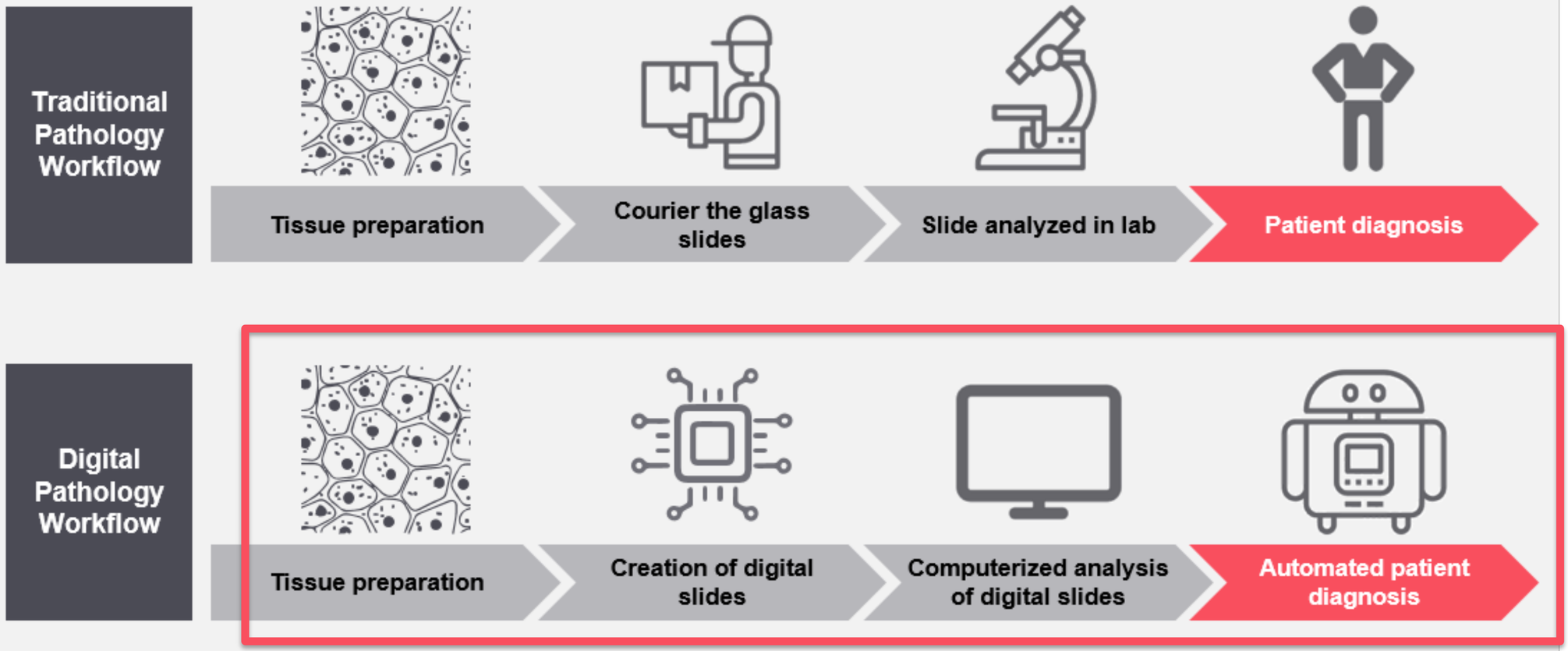
OCTOBER 16-18 | MGM GRAND | LAS VEGAS, NV

#PathVisions22

PATHOLOGY VISIONS

Keep it simple

EXHIBIT 1: Comparison between Manual Approach and Digitalization followed in Pathology



Source: Pathkids.com

Decision Summary

Number of FDA approvals with AI-based algorithms is increasing – currently >80 algorithms

npj | Digital Medicine

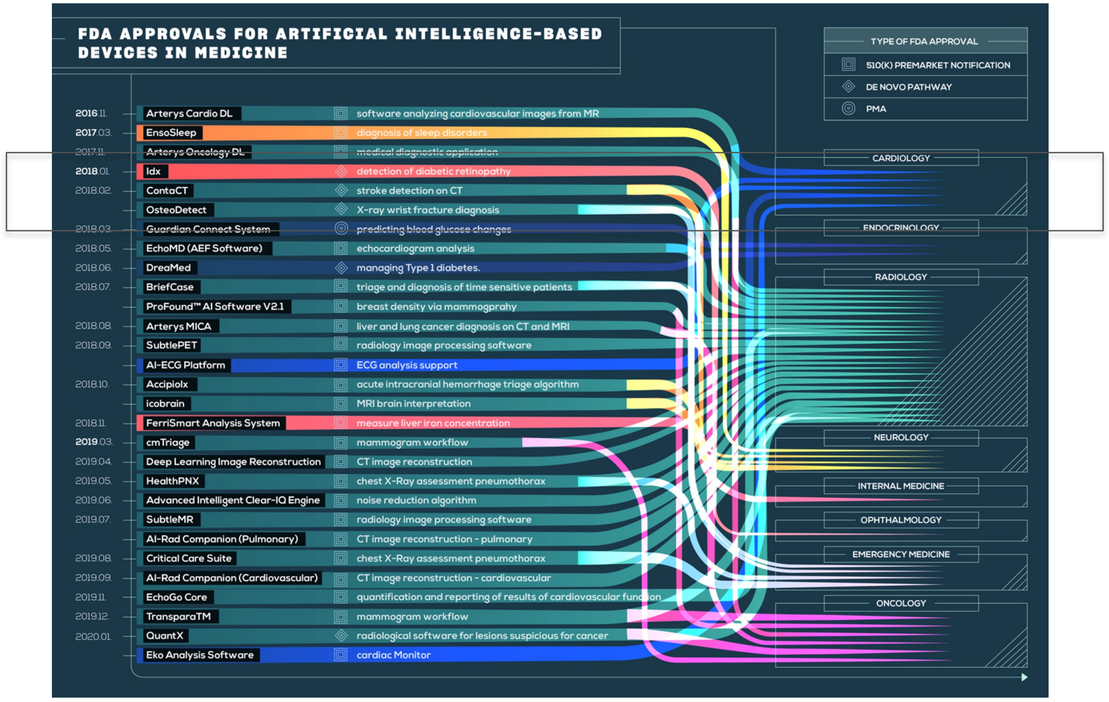
www.nature.com/npjdigitalmed

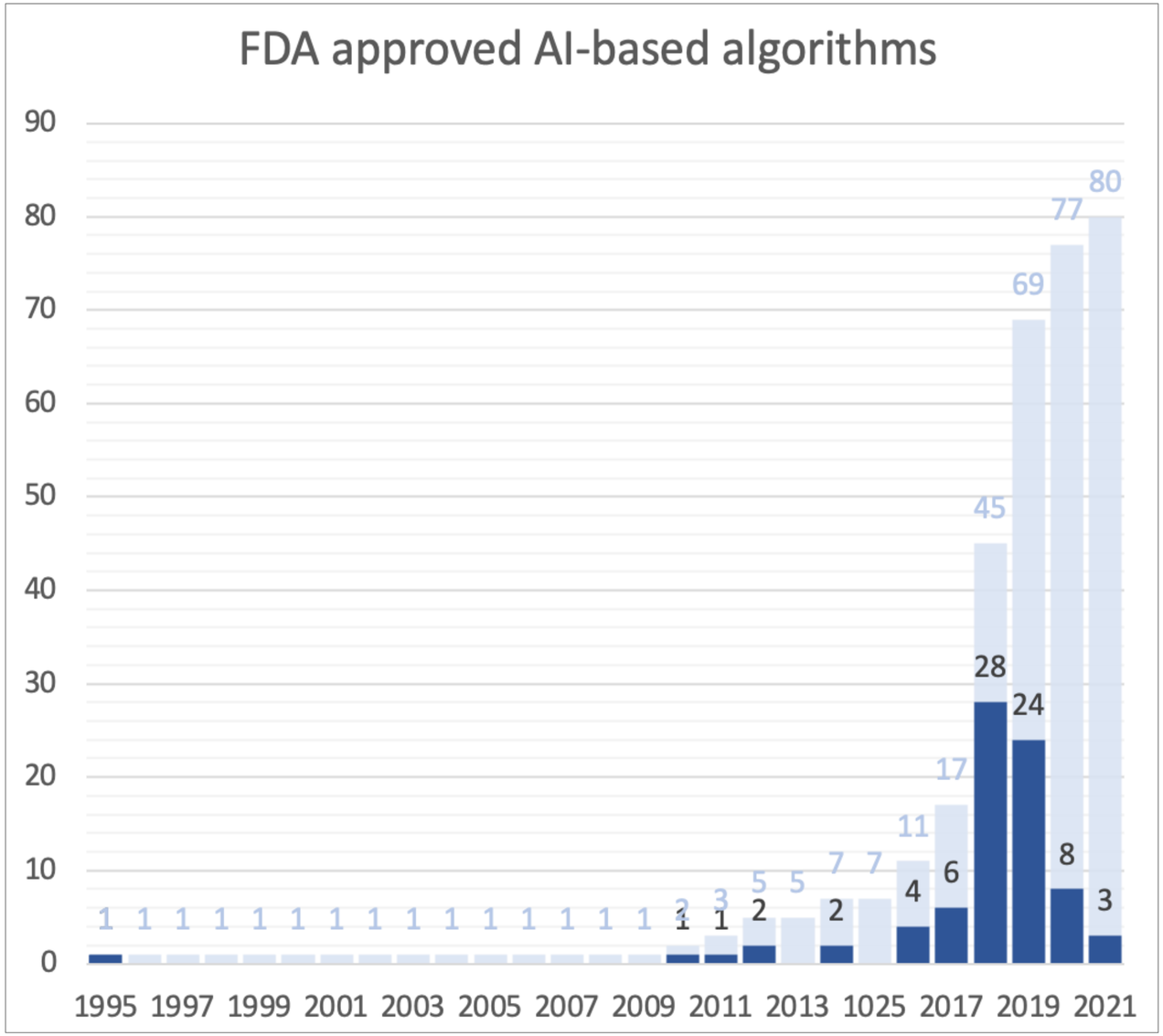
ARTICLE OPEN

Check for updates

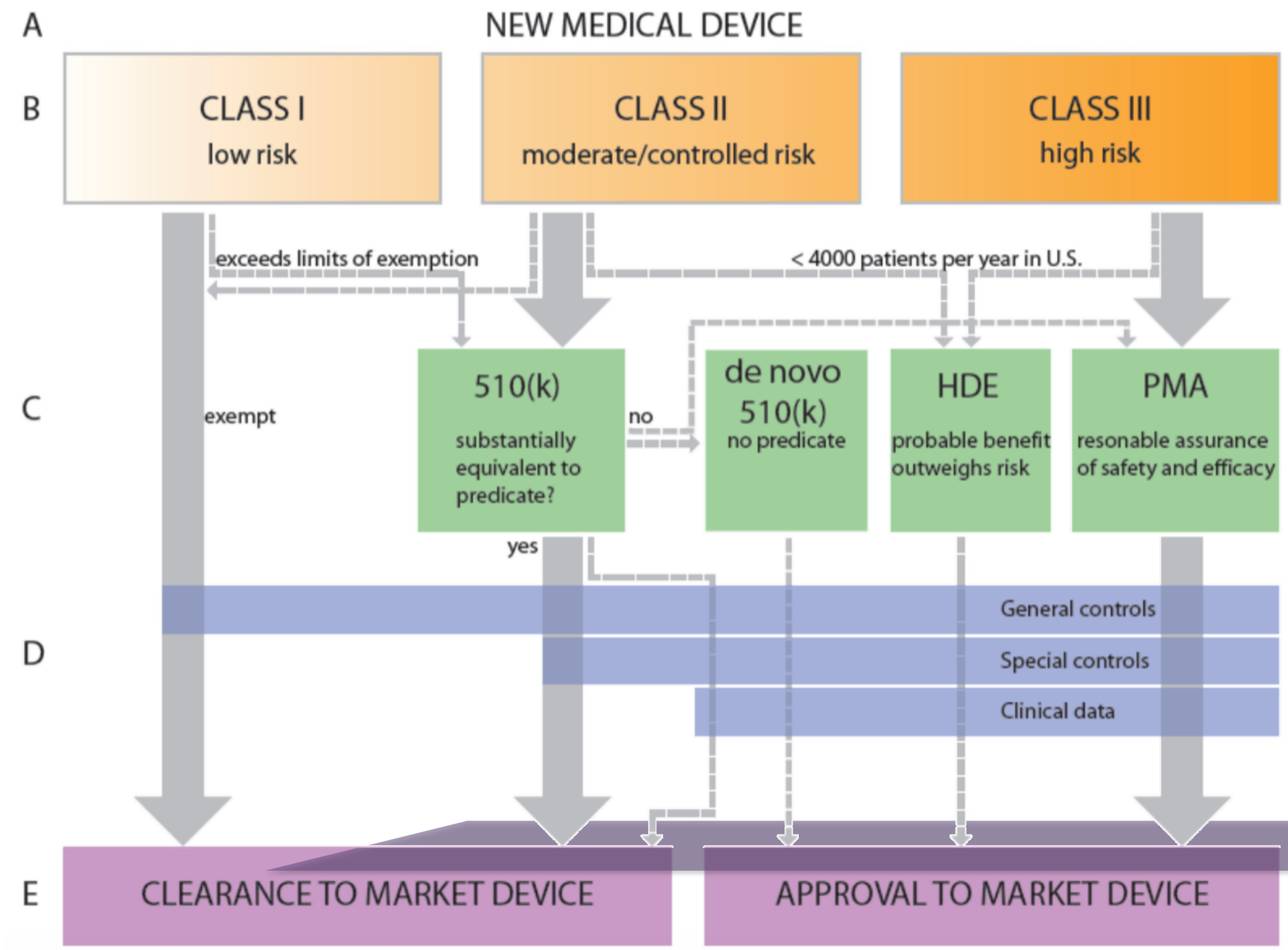
The state of artificial intelligence-based FDA-approved medical devices and algorithms: an online database

Stan Benjamens[1,2], Pranavsingh Dhunnoo[3] and Bertalan Meskó[3,4]✉





Decision Summary is one of the Authorization Documents



Authorization Documents

- Approval Letter
- **Decision Summary**
  - Summary of Safety and Effectiveness
- Patient Labeling
  - PMA Database Entry

...

U.S. FOOD & DRUG ADMINISTRATION

Decision Summaries and Paige Prostate

Peter J. Yang, PhD, RAC
De Novo Program Lead

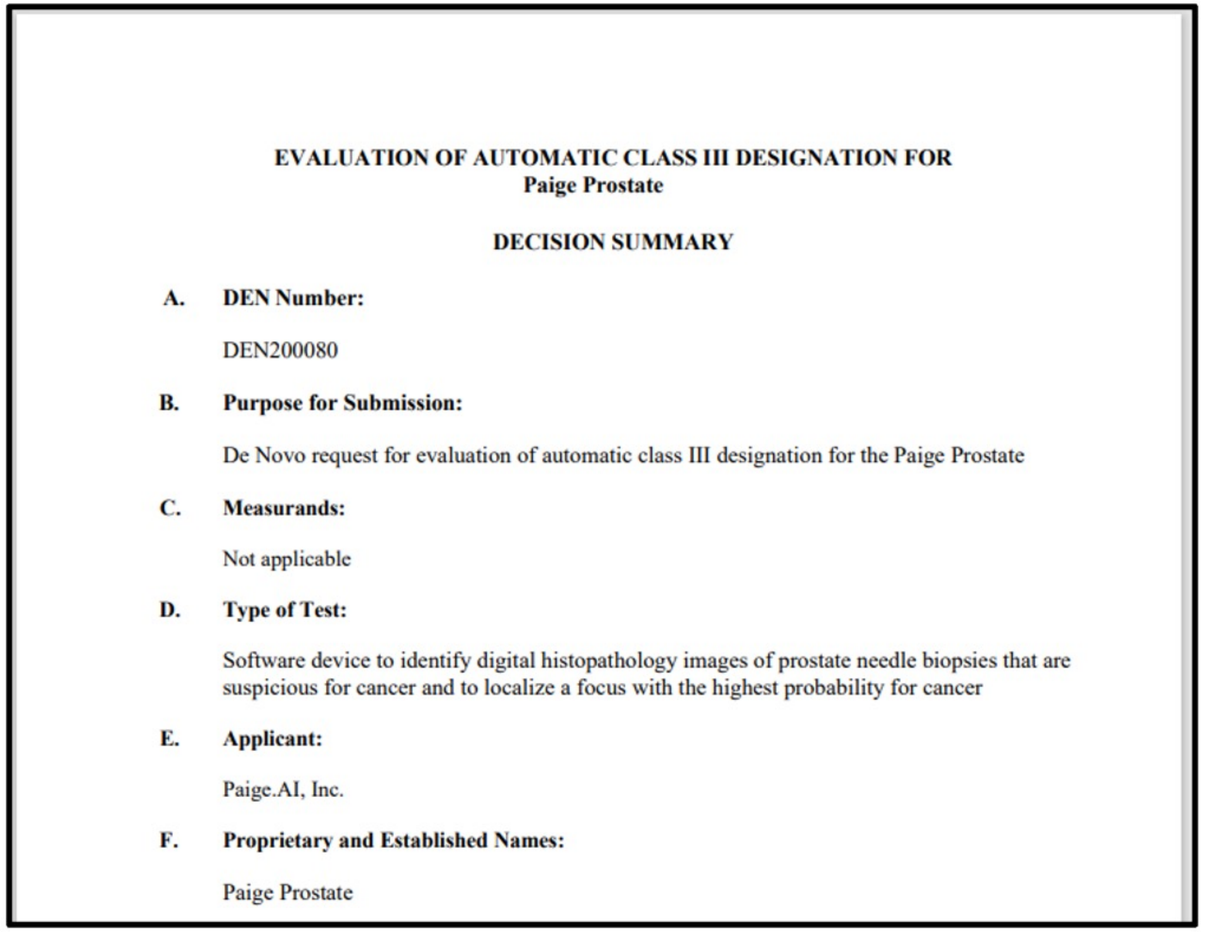
Office of Regulatory Programs (ORP)
Office of Product Evaluation and Quality (OPEQ)
Center for Devices and Radiological Health
Food and Drug Administration

What Are We Talking About Here?

- FDA assures the safety and effectiveness of medical devices, including many in vitro diagnostics (IVDs) for use in pathology.
- FDA adopts a risk-based classification process to support predictable and least burdensome requirements for the data needed for FDA to permit marketing of medical devices.
- FDA is committed to transparency in decision making to foster innovation.
- Transparency for FDA's decision making for the Paige Prostate device is found in a **Decision Summary** document.
- I provide programmatic and regulatory oversight for devices reviewed through the "**De Novo request**" regulatory pathway, a common pathway to market for many novel types of devices.

Paige Prostate Decision Summary

- Search for “FDA” and “Medical Device Databases”
- Search the De Novo database for Paige Prostate and click on the Decision Summary link
- https://www.accessdata.fda.gov/cdrh_docs/reviews/DEN200080.pdf



EVALUATION OF AUTOMATIC CLASS III DESIGNATION FOR
Paige Prostate

DECISION SUMMARY

A. DEN Number:

DEN200080

B. Purpose for Submission:

De Novo request for evaluation of automatic class III designation for the Paige Prostate

C. Measurands:

Not applicable

D. Type of Test:

Software device to identify digital histopathology images of prostate needle biopsies that are suspicious for cancer and to localize a focus with the highest probability for cancer

E. Applicant:

Paige.AI, Inc.

F. Proprietary and Established Names:

Paige Prostate

Decision Summaries

- Decision summary formats depend on submission type
  - 510(k) premarket notifications: 510(k) Summary
  - De Novo requests: De Novo Decision Summary
  - Premarket approvals: Summary of Safety and Effectiveness Data (SSED)
- Purpose of decision summary:
  - Provide transparency into FDA's decision making
  - Serve as comparison and reference for future submissions

Broad Regulation Designed to Enable Pathology Innovation

- 21 CFR 864.3750 **Software algorithm device to assist users in digital pathology.** A software algorithm device to assist users in digital pathology is an in vitro diagnostic device intended to evaluate acquired scanned pathology whole slide images. The device uses software algorithms to provide information to the user about presence, location, and characteristics of areas of the image with clinical implications. Information from this device is intended to assist the user in determining a pathology diagnosis.

Special Controls (Class II)

- Special controls are legal requirements for all devices in the regulation and are written into the new classification regulation
- Special controls include:
  - Non-clinical (analytical) validation requirements
  - Clinical validation requirements
  - Labeling requirements
- **The De Novo device must meet its own special controls**

De Novo Decision Summary

- The Decision Summary, combined with the De Novo granting letter, tells FDA's risk-based classification "story"
  - New regulation (number, name, and identification)
  - Risk/mitigation table
  - Special controls (if class II)
  - Device description
  - Non-clinical and clinical data summaries
  - Benefit-risk discussion
- Demonstrates how special controls were met
- Serves as reference to support future 510(k) submissions

Decision Summaries for AI

- Ideally, a Decision Summary should discuss:
  - The general overview of AI model development
  - The dataset that was used to train the model
  - The validation process and dataset (separate and distinct) that was used to validate the model for real-world use
  - Any warnings, precautions, or limitations for using the AI software
  - Any information needed for ensuring correct use, including inputs and processing

How to Use A Decision Summary

- Read the Decision Summary and seek to:
  - Understand what the sponsor needed to do to get their device granted/cleared/approved
  - Understand how the sponsor met the special controls for the device type (if any) and what risks FDA is trying to address by requiring certain information
  - Align your own testing strategy to meet FDA's requirements
  - Assemble a testing strategy document for FDA to review in the context of a Pre-Submission (recommended)

Peter J. Yang, PhD, RAC
De Novo Program Lead
Office of Regulatory Programs (ORP)
Office of Product Evaluation and Quality
Center for Devices and Radiological Health
Peter.Yang@fda.hhs.gov


FDA
U.S. FOOD & DRUG
ADMINISTRATION

The First AI Algorithm in Pathology

Emre Gültürk

Vice President of Regulatory Affairs and Quality Assurance at Paige

INDUSTRY PERSPECTIVE
DECISION SUMMARY SCOPE AND GOAL

Decision Summary is a high-level output from a much larger submission effort:

- ~2000 pages of documentation submission throughout its review

Its main goal is to:

- Provide key details to the predicate device manufacturers to establish fundamentals, but also reserve specific details to encourage the industry to engage with the FDA to fine-tune device functionality and study designs etc. appropriate to the device indication/intended use

INDUSTRY PERSPECTIVE

DECISION SUMMARY KEY COMPONENTS

1. Intended Use/Indications for Use

- Locks the intended use, and requires device manufacturer to work with FDA to review for certain modifications
- In pathology, this is specifically true for the tissue type, device output, compatibility (scanner & image viewer), use setting (diagnostic aid).

2. Development Dataset Distribution and Its Diversity

Indications for use:

Paige Prostate is a software only device intended to assist pathologists in the detection of foci that are suspicious for cancer during the review of scanned whole slide images (WSI) from prostate needle biopsies prepared from hematoxylin & eosin (H&E) stained formalin-fixed paraffin embedded (FFPE) tissue. After initial diagnostic review of the WSI by the pathologist, if Paige Prostate detects tissue morphology suspicious for cancer, it provides coordinates (X,Y) on a single location on the image with the highest likelihood of having cancer for further review by the pathologist.

Paige Prostate is intended to be used with slide images digitized with Philips Ultra Fast Scanner and visualized with Paige FullFocus WSI viewing software.

Paige Prostate is an adjunctive computer-assisted methodology and its output should not be used as the primary diagnosis. Pathologists should only use Paige Prostate in conjunction with their complete standard of care evaluation of the slide image.

tissue

output

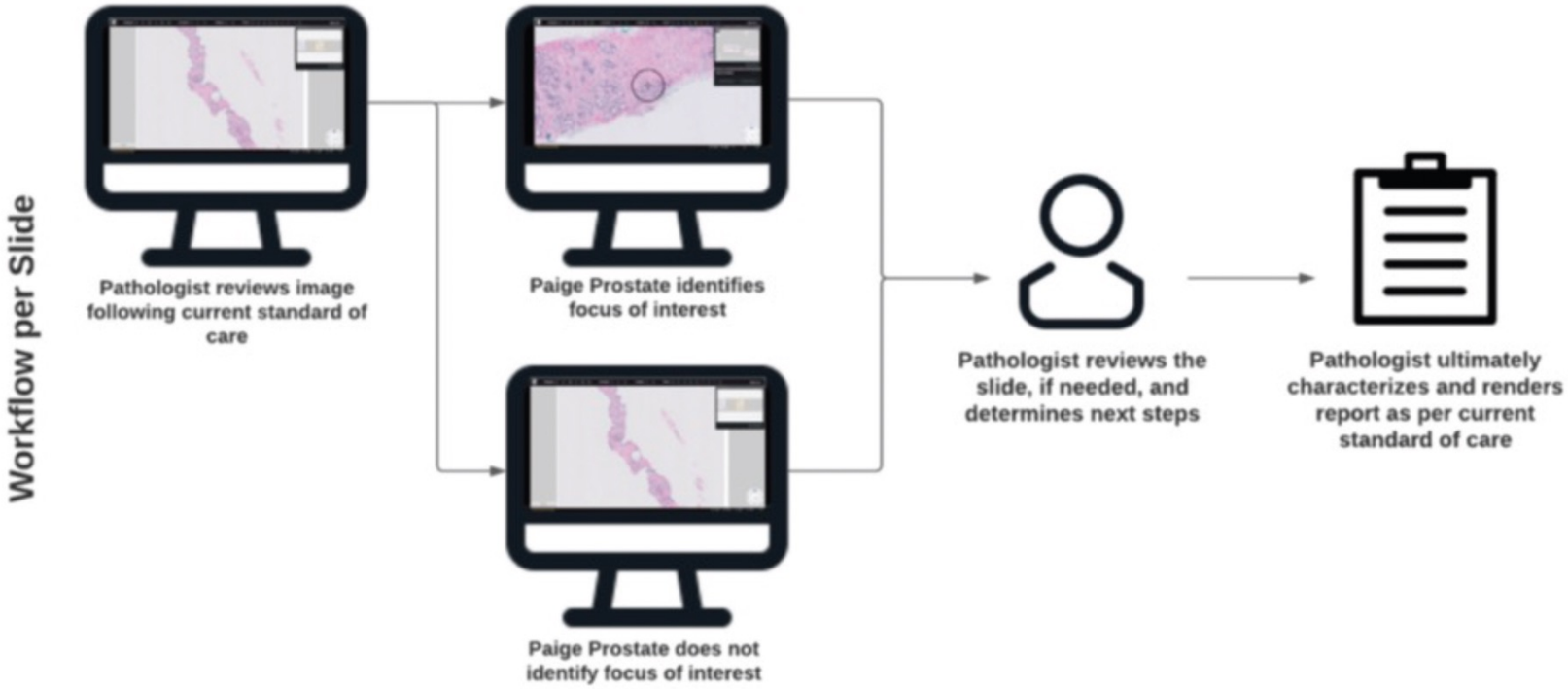
compatibility

use setting

INDUSTRY PERSPECTIVE

DECISION SUMMARY KEY COMPONENTS (CONTD.)

3. Mechanism of Action and Principles



Approved Clinical Workflow

INDUSTRY PERSPECTIVE

DECISION SUMMARY KEY COMPONENTS (CONTD.)

4. Performance Characteristics

- Study Data Characteristics and Its Diversity
- Analytical Performance Study Design Overview and Results (Accuracy, Localization, and Precision for Intra/Inter Scanner/Operator Variability)
- Clinical Performance Study Design Overview and Results (Reader Diagnostic Accuracy)

		Classification for unassisted read			
		Cancer	**Deferred**	**No Cancer**	**Total**
Classification for assisted read	**Cancer**	4.31 (1.2%)	1.12 (0.3%)	0.81 (0.2%)	6.25 (1.76%)
	Deferred	3.19 (0.9%)	22.75 (6.4%)	5.19 (1.5%)	31.12 (8.74%)
	No cancer	0.69 (0.2%)	9.06 (2.5%)	308.87 (86.8%)	318.62 (89.5%)
	Total	8.19 (2.3%)	32.94 (9.2%)	314.87 (88.45%)	356 (100%)

Numbers in grey colors are numbers of slide images with the same classification in assisted and unassisted reads. Numbers in green color, 0.69 (0.2%) and 9.06 (2.5%), present a reduction in the number of false positive results for the benign slide images because of use of the Paige Prostate device. Numbers in orange color, 0.81 (0.2%) and 5.19 (1.5%), present an increase in the number of false positive results because these benign slide images had "Cancer" assisted reads (0.2%) or "Deferred" assisted reads (1.5%) but had "No Cancer" for unassisted reads. Overall difference in the number of false positive slide images was 3.75 slides [=(0.69+9.06)-(0.81+5.19)] what is 1.05% (=3.75/356). Difference in the number of false positives slides of 1.1% with 95% CI: (-0.7%; 3.4%) was not statistically significant.

INDUSTRY PERSPECTIVE

DECISION SUMMARY KEY COMPONENTS (CONTD.)

5. Benefit/Risk Determination

- Discussion for favorable benefits when compared to its risks

Paige Prostate appears to provide a reasonable assurance of safety and effectiveness for diagnostic use by its intended users after taking into consideration the special controls. The clinical and analytical studies have shown that the risk of accuracy loss resulting in a false positive or false negative diagnosis, is minimal relative to the patient safety benefits, including new findings that would contribute to the correct diagnosis. This is contingent on the device being used according to the approved labeling, particularly that the end user must be fully aware of how to interpret and apply the device output.

The potential for false negative and false positive results is mitigated by special controls. Labeling requirements, which include certain device description information as well as certain limitations, ensure that users will employ all appropriate procedures and safeguards as specified, including use of the device as an adjunct rather than as the sole basis of making the diagnosis. In addition, design verification and validation includes data on software performance as supported by the underlying software design, as well as software algorithm training and validation within the limits of the specified intended use. This also includes analytical validation (including precision studies) and clinical validation (including user validation studies and performance studies) studies.

The probable clinical benefits outweigh the potential risks when the standard of care is followed by qualified users, and appropriate mitigation of the risks is provided for through implementation of and adherence to the special controls. The combination of the general controls and established special controls support the assertion that the probable benefits outweigh the probable risks.

INDUSTRY PERSPECTIVE

SPECIAL CONTROLS AND THEIR IMPORTANCE

Regulation is currently not published under Code of Federal Regulations (CFR) but is already available in Paige Prostate Reclassification Order document. Document provides key guidelines on:

- Intended Use
- Device Label
- Details about performance testing
- Device limitations
- Device verification and validation (analytical and clinical testing)

COLLEGE of AMERICAN PATHOLOGISTS

A proposed framework for deploying AI/ML in the clinical laboratory

Jansen N Seheult, *MB BCh BAO, MSc, MS, MD, FCAP*
Member, Machine Learning Working Group
Member, Artificial Intelligence Committee

Senior Associate Consultant & Assistant Professor,
Divisions of Hematopathology and Computational Pathology & AI,
Department of Laboratory Medicine and Pathology,
Mayo Clinic - Rochester

June 22, 2022

Terminology

- Verification: The process by which a laboratory determines that an unmodified FDA-cleared/approved test performs according to the specifications set forth by the manufacturer when used as directed
- Validation: The process used to confirm with objective evidence that a laboratory-developed or modified FDA-cleared/approved test method or instrument system delivers accurate and reliable results for the intended application

Understanding the Decision Summary

- Intended use – claimed model purpose
- Indications for use – A general description of the disease or condition the device will diagnose, treat, prevent, cure, or mitigate, including a description of the patient population for which the device is intended. Any differences related to gender, race/ethnicity, etc. should be included in the labeling.
- Sample size and distribution of data used in the model training and validation
- Inclusion and exclusion criteria
- How “ground truth’ was determined
- Performance claims
- Data compatibility, including how missing data are handled

Case Study: Prostate biopsy WSI analysis

- **FDA approved - For in vitro diagnostic (IVD) use only**
- **Indications for use:**
  - **Software only device intended to assist pathologists in the detection of foci that are suspicious for cancer during the review of scanned whole slide images (WSI) from prostate needle biopsies prepared from hematoxylin & eosin (H&E) stained formalin-fixed paraffin embedded (FFPE) tissue.**
  - **The software is intended to be used with slide images digitized with Scanner X and visualized with Vendor X's WSI viewing software.**
  - **The software is an adjunctive computer-assisted methodology and its output should not be used as the primary diagnosis. Pathologists should only use the software in conjunction with their complete standard of care evaluation of the slide image.**

Case Study: Prostate biopsy WSI analysis

- Training dataset: ~35,000 de-identified slides from single US laboratory between 2013-2017 and imaged with scanner Y
- Tuning dataset: ~6,000 slides prepared and stained at a single site and imaged with Scanner Y
- **Test datasets:**
  - **Tuning dataset (~6,000 slides) imaged with Scanner X**
  - **~11,000 slides prepared at >200 external sites but diagnosed at internal site and imaged with Scanner Y**
- **Approximately 80% of slides in training, tuning and testing datasets were collected from Caucasian patients, with approximately 8-9% from Black/ African American patients and 3% from Asians.**

Case Study: Accuracy characteristics

- **Accuracy study: Cancer (n = 311) and Benign (n = 417)**
  - **Sensitivity = 94.5% (95% CI: 91.4 – 96.6%)**
  - **Specificity = 94.0% (95% CI: 91.3 – 95.9%)**
  - **Accuracy = 94.2% (95% CI: 92.3 – 95.7%)**
- **Clinical study: Cancer (n = 171) and Benign (n = 356) read by 16 pathologists**
  - **Assisted macro-averaged sensitivity = 96.8%**
  - **Assisted macro-averaged specificity = 89.5%**
- **Case breakdown:**
  - **Cancer: ~50% had tumor size ≤ 0.5mm & 50% > 0.5mm, ~2% with PIN, ~3-4% ASAP**
  - **Benign: ~88% without atrophy, PIN or treatment effects**

Case Study: Precision characteristics

- Cancer (n = 35) and Benign (n = 36)
- Within-scanner: Slides scanned three times (3 reps) using one scanner/ operator
  - Cancer: 99.0% (95% CI: 94.8 – 99.8%) of all scans and 97.1% (34/35) of all slides produced correct results
  - Benign: 94.4% (95%CI: 88.4 – 97.4%) of all scans and 88.9% (32/36) of all slides produced correct results
- Reproducibility: Slides scanned once with three different scanners at different locations and by three different operators (one operator per scanner)
  - Cancer: 100.0% (95% CI: 96.5 – 100.0%) of all scans and 100.0% (35/35) of all slides produced correct results
  - Benign: 93.5% (95%CI: 87.2 – 96.8%) of all scans and 88.9% (30/36) of all slides produced correct results

Verifying manufacturer's accuracy claim

- H_0: Accuracy = P_0 versus H_1: Accuracy < P_0 (or Accuracy = P_1)
- With (1 – α)% confidence level and (1 – β)% power for detecting an effect of $P_1 - P_0$, the required sample size for cases is obtained from:

$$n = \frac{\left[Z_\alpha * \sqrt{P_0(1-P_0)} + Z_\beta * \sqrt{P_1(1-P_1)}\right]^2}{(P_1 - P_0)^2}$$

- For example, if the laboratory wishes to compare locally determined accuracy of a software or algorithm to the manufacturer's claim of 94.2%, the sample size required to have 95% confidence and 80% power to detect a difference of 5% from the claimed accuracy of 94.2% would be:

$$n = \frac{\left[1.645 * \sqrt{0.942(1-0.942)} + 0.84 * \sqrt{0.892(1-0.892)}\right]^2}{(0.892 - 0.942)^2} = 166$$

- To detect a difference of 10% from the claimed accuracy with 95% confidence and 80% power:

$$n = \frac{\left[1.645 * \sqrt{0.942(1-0.942)} + 0.84 * \sqrt{0.842(1-0.842)}\right]^2}{(0.842 - 0.942)^2} = 48$$

Verifying manufacturer's accuracy claim

- Dataset balance of cancer versus benign may influence choice of evaluation metric and required sample size
- How similar is your verification dataset to the manufacturer's accuracy and clinical study sets?
  - Your verification dataset should reflect your local patient population
  - Failure to verify manufacturer's stated claim may be driven by systematic differences in study sample characteristics

Verifying manufacturer's accuracy claim

- **Accuracy verification study performed with 166 samples (cancer = 83, benign = 83)**

$X^2 = 5.37$ P = 0.15	Vendor n = 728	Local lab n = 166
White	598	140
Black/ AA	58	18
Asian	22	3
Other	50	5

$X^2 = 7.69$ P = 0.05	Vendor n = 728	Local lab n = 166
Tumor ≤ 0.5mm	147	35
Tumor > 0.5mm	153	48
Benign (no atrophy/ PIN/ tx)	366	67
Other benign	51	16

- **Observed accuracy: 91.6% (152/166)**
- **Using one-sample test of proportion versus manufacturer's claim of 94.2%, p = 0.084 for one-sided alternate hypothesis**

Verifying manufacturer's precision claim

- **Simple precision**
  - **Repeatability: 10 slides (1:1 ratio of cancer: benign) scanned three times (3 reps) using one scanner/ operator**
  - **Reproducibility: 10 slides (1:1 ratio of cancer: benign) scanned once with different scanners, at different locations, by different operators (as appropriate)**
  - **Two-sample test of proportions:**
    - **Repeatability for cancer: 96.7% (95% CI: 83.3 – 99.4%) of 30 local scans compared with 99.0% (95% CI: 94.8 – 99.8%) of manufacturer's 105 scans (test if observed proportion significantly lower than manufacturer's claim: $p = 0.178$ at $\alpha = 0.05$)**
    - **Repeatability for benign: 86.7% (95%CI: 70.3 – 94.7%) of 30 local scans compared with 94.4% (95%CI: 88.4 – 97.4%) of manufacturer's 108 scans (test if observed proportion significantly lower than manufacturer's claim: $p = 0.075$ at $\alpha = 0.05$)**
- **Complex precision (ISO 16140)**

Case Study: What's next?

- **Verification of accuracy and precision claims are not the end of your responsibilities as a Laboratory Director**
  - **Think PARR for method verification: Precision, Accuracy, Reportable range, Reference interval**
- **Try to break the model to understand its limitations**
- **Equipment qualification**

COLLEGE of AMERICAN PATHOLOGISTS

The Clinical Outlook for AI in Pathology

David S. Klimstra, MD
Founder & Chief Medical Officer at Paige

CLINICAL OUTLOOK

APPROVAL OF AI FOR PATHOLOGY DECISION SUPPORT

Product Classification

Software algorithm device to assist users in digital pathology (21 CFR Part 864.3750)

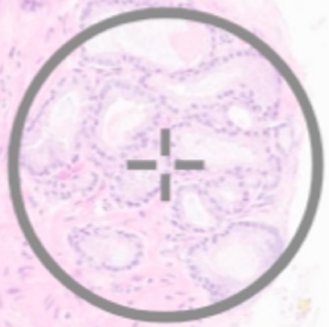
- A software algorithm device to assist users in digital pathology is an in vitro diagnostic device intended to evaluate acquired scanned pathology whole slide images. ***The device uses software algorithms to provide information to the user about presence, location, and characteristics of areas of the image with clinical implications***. Information from this device is intended to assist the user in determining a pathology diagnosis.

- Broad device classification
- Can apply to an array of AI tools for digital pathology
- Establishes the predicate for other approvals (regular 510(k) rather than de novo

CLINICAL OUTLOOK

APPROVAL OF AI – MEANING FOR OUR INDUSTRY

- Enables confident use of AI-based devices that help pathologists arrive at the correct diagnosis
- Ensures that AI generalizes across practice settings and laboratory variables
- Creates a path for clearances of future products in the category
- Sets a quality bar for future products
- Interoperability exists but is limited to FDA-cleared devices (scanner, viewer, monitor)
- Even within FDA cleared devices, each compatibility expansion effort would require 510(k) premarket review
- FDA intends to enforce regulations rigorously



CLINICAL OUTLOOK

APPROVAL OF AI – MEANING FOR PATHOLOGISTS

- Clinical-grade AI has been defined as an aid to pathologists
- Added scrutiny on slides provide pathologists greater confidence in their diagnoses
- Pathologists can focus their attention on the most critical aspects of establishing the diagnosis
- Compatibility responsibility is with the manufacturer
- Quality control rests with the manufacturer
- Unmodified use will scale easily among sites
- Verification much simpler, compared with LDT
- Cleared devices are “locked down” and require FDA review upon modification
- Modifications to the end-to-end test will require an LDT for the specific modification

CLINICAL OUTLOOK
WHERE ARE WE HEADED?

- Development of additional pathology decision support AI for defined use cases
- Process established to validate algorithms and ensure generalizability
- FDA clearance will establish “clinical grade” quality and ease implementation for users
- Increasing availability of decision support AI across pathology will help motivate transition to digital pathology
- Increased use of digital pathology opens the door for more advanced computational tools, such as digital biomarkers and multimodal data analytics

DISCUSSION

Key discussion points.