USE CASE
The FIND validation platform: Manufacturer independent evaluations for CXR-CAD

Matthew Arentz

MD MPH; Lead – Digital Health Validation Program

Senior Technical Consultant

FIND

13-14 September
INTELLIGENT HEALTH 2023
Basel, Switzerland

PRODUCED BY INSPIRED MINDS!

Accredited cpd Event

@IntHealthAI
#IntelligentHealthAI
#SaveLivesWithAI

HEADLINE PARTNER
NOVARTIS | The Novartis Foundation

PLATINUM SPONSOR
INNOVATION & TECHNOLOGY
Deloitte.

CITY PARTNER
Basel
basel.ch

FIND

MANUFACTURER INDEPENDENT EVALUATION FOR CXR CAD

14 Sept 2023

FIND

1

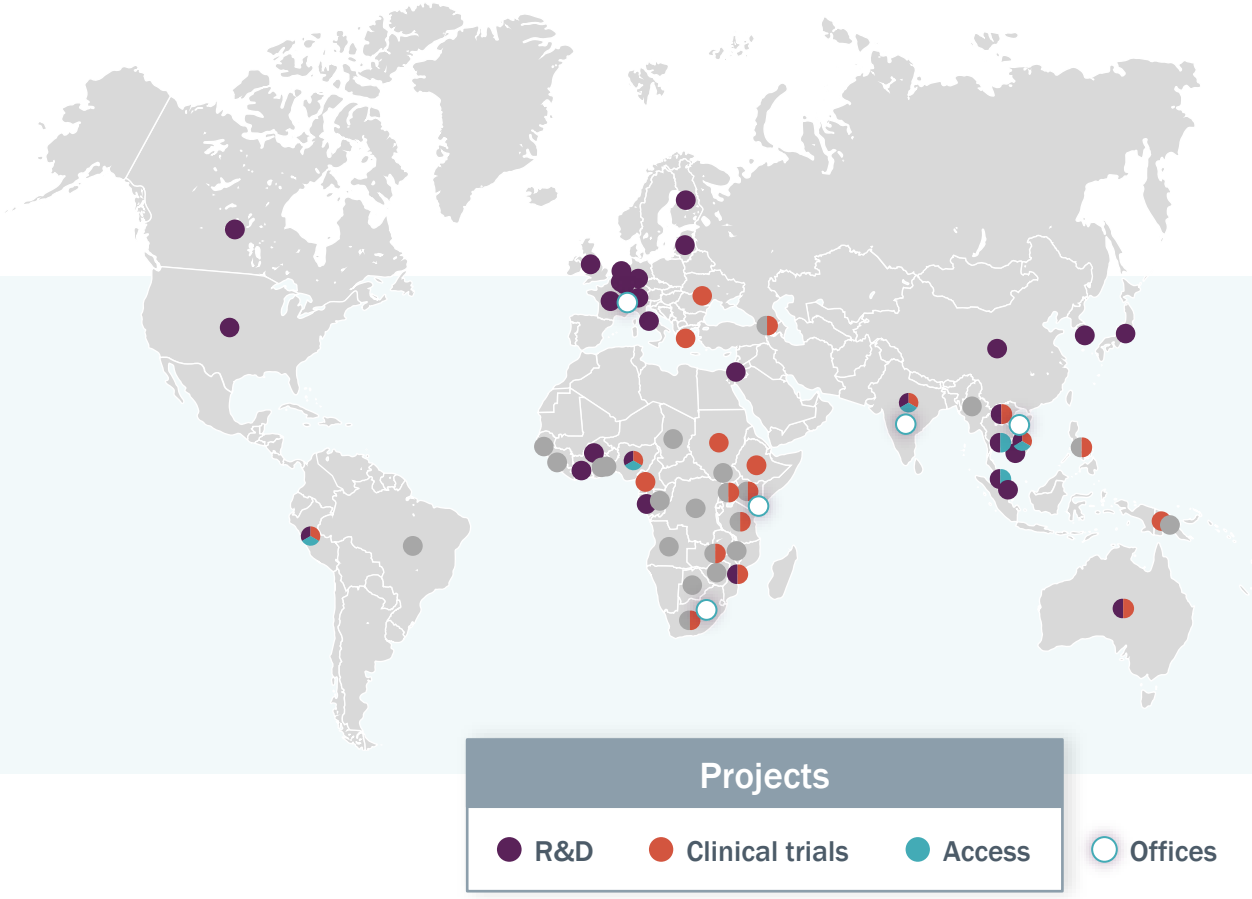
INTRODUCTION

FIND, THE GLOBAL ALLIANCE FOR DIAGNOSTICS

We connect countries and communities, funders, decisionmakers, healthcare providers and developers to spur diagnostic innovation and make testing an integral part of

sustainable, resilient health systems

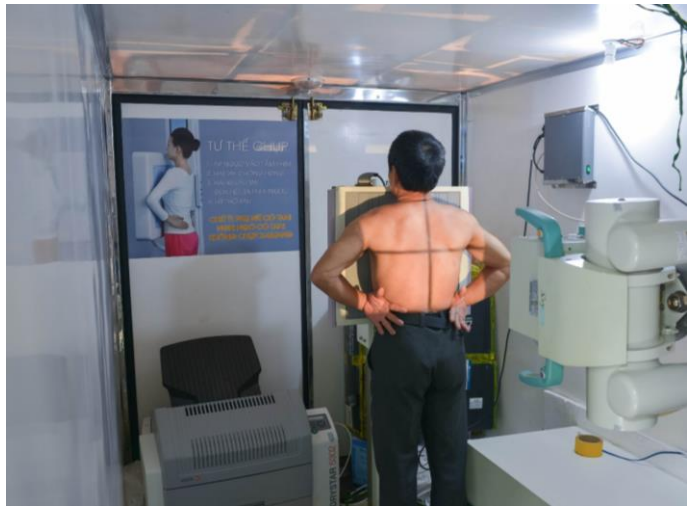
- Established in 2003 as a not-for-profit product development & delivery partnership
- Co-convener of the Access to COVID-19 Tools (ACT) Accelerator Diagnostic Pillar
- WHO Collaborating Centre for Laboratory Strengthening & Diagnostic Technology Evaluation



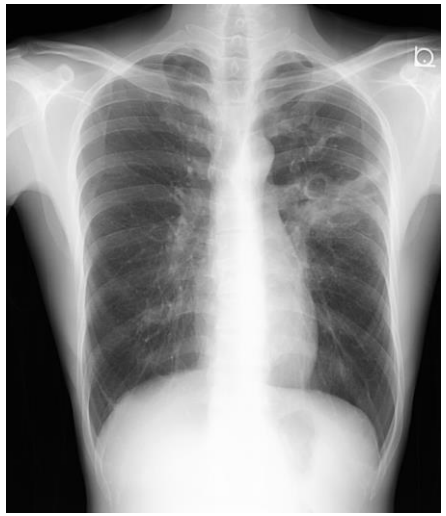
@FINDdx @FINDdiagnostics finddx FINDdiagnostics @FINDdx #DiagnosisForAll

CXR CAD TECHNOLOGY PERFORMANCE REVIEW FOR WHO TB SCREENING GUIDELINE

FIRST EVALUATION OF AN AI-BASED TB DIAGNOSTIC

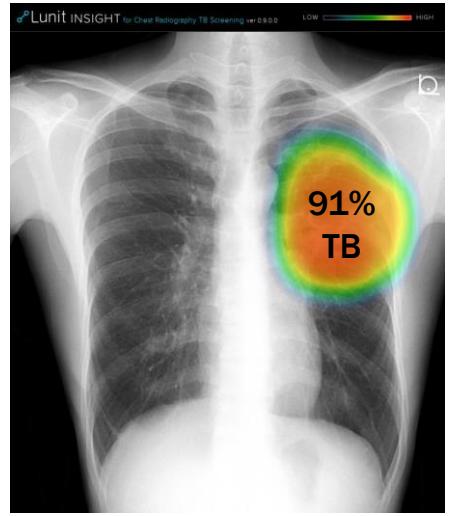


Input:
Digital AP CXR

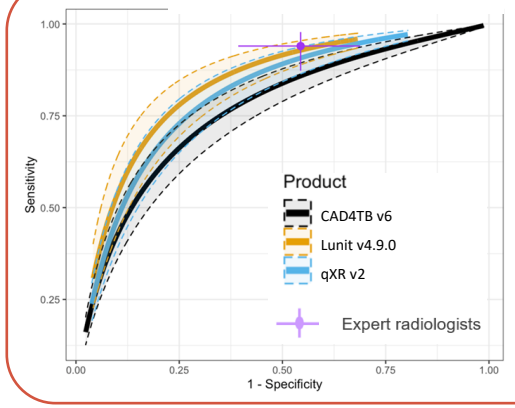


AI algorithm

Output:
Abnormality score & overlay

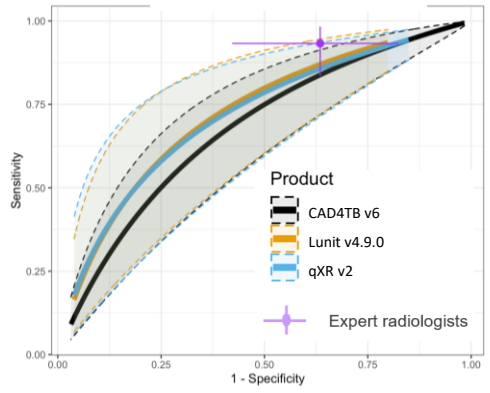


SCREENING USE-CASE



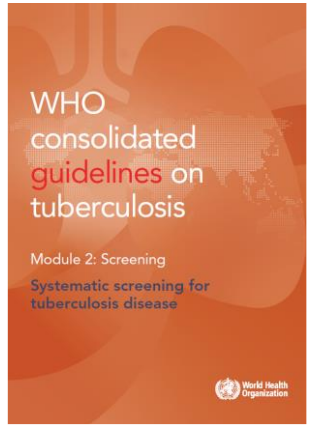
FIND, Kik et al (under review)

TRIAGE USE-CASE



FIND, Kik et al (in preparation)

Recommendation 2021:
"CAD software programmes may be used in place of human readers for interpreting digital chest X-rays for screening and triage for TB disease, among individuals aged 15 years and older."



CXR CAD FOR TB

A RAPIDLY GROWING FIELD

2017

1 TB CAD product

2018

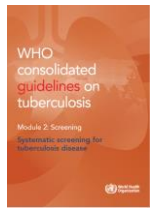
2019

2020

TB CAD site

WHO review evidence on 3 CE marked CAD products

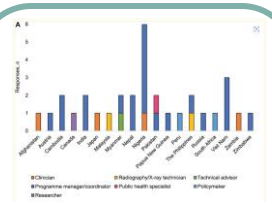
2021



WHO policy published

Tech landscape of CAD and CXR products

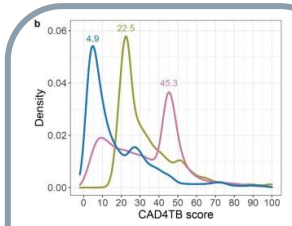
2022



Implementation and user experience

Landscape identifies > 22 TB CAD products

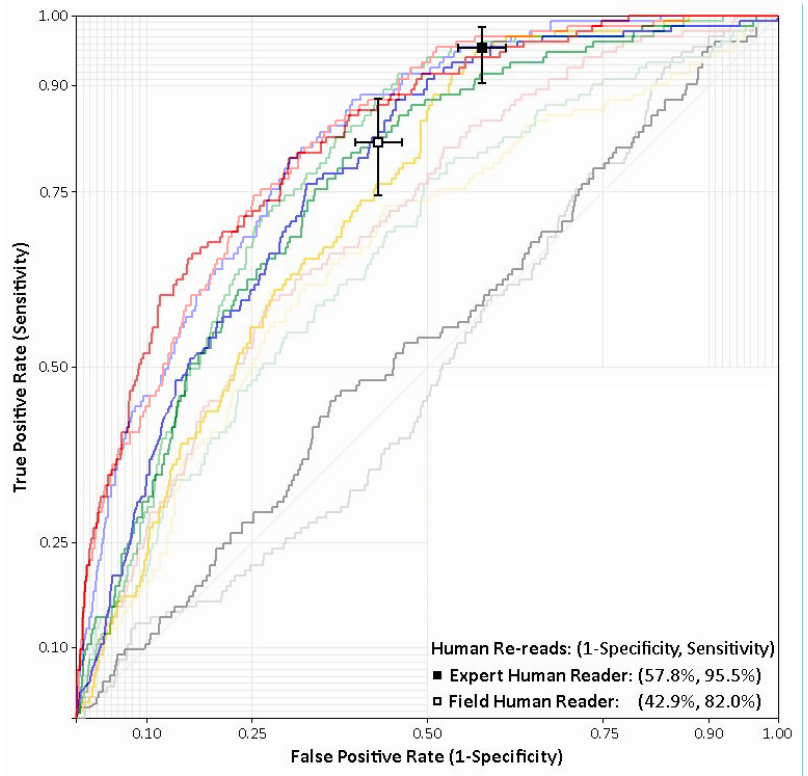
2023

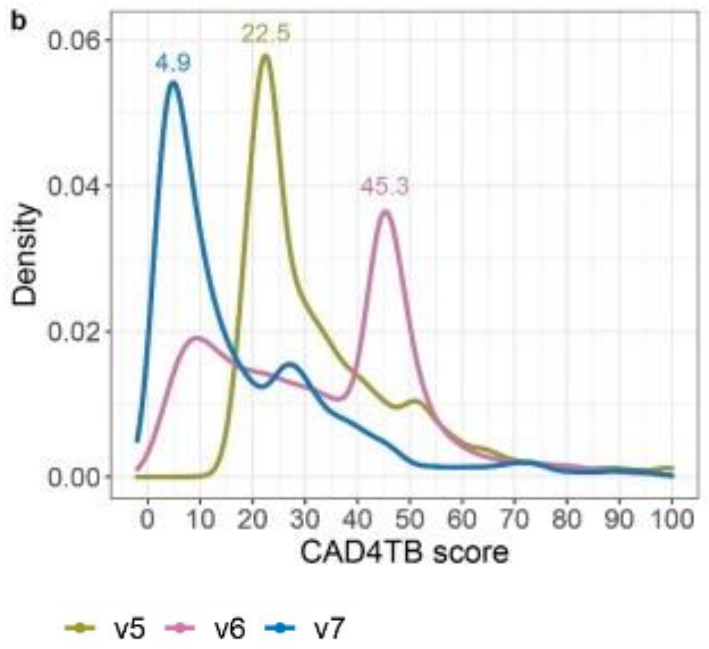


New versions, with change in performance, & thresholds

Ref: WHO, Barrett/StopTB, AI4hlth.org, , Fehr, IJTLD 2023,

VARIABLE PERFORMANCE AND DESIGN CHANGES





Need for independent performance evaluation to confirm accuracy of products covered under WHO class recommendation

- To allow for quick evaluations when new products/versions are available
- Using independent, representative datasets (not shared with manufacturers, global representation and applicable for use-case)

Ref: Codling AJ, Nat. Sci Rep 2021, Fehr, IJTLD 2023

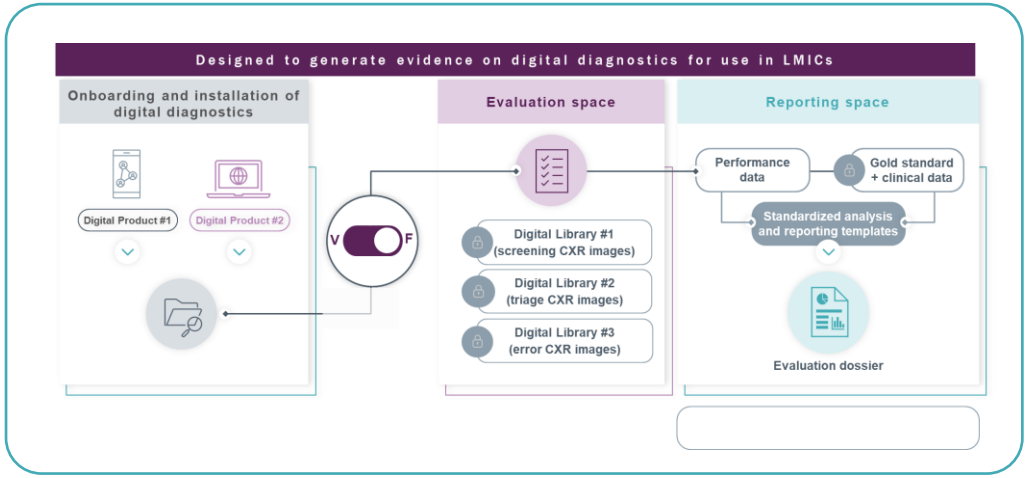
FIND VALIDATION PLATFORM FOR DIGITAL DIAGNOSTICS

The FIND Validation Platform

- A digital infrastructure for fast and standardized in-silico assessment of digital diagnostics (CAD software, RDT reader software, cough apps, AI based sequencing interpretation software…)

1st iteration is TB CAD for WHO PQ

- Independent performance evaluations of new TB-CAD software and versions
- Describe their performance in line with the original technology class evaluation by the WHO



FIND

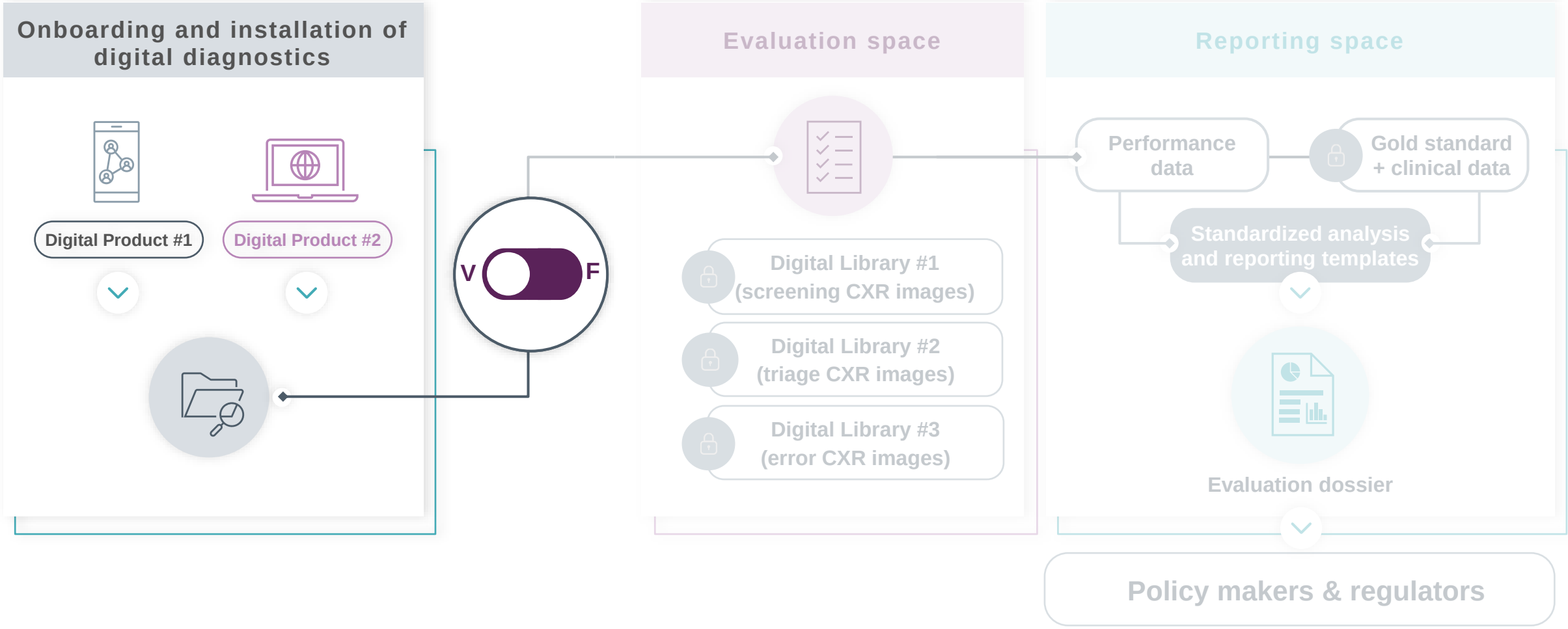
2

FIND VALIDATION PLATFORM: STRUCTURE AND SOFTWARE ONBOARDING

VALIDATION PLATFORM

FOR ONBOARDING AND EVALUATION OF AI BASED DIAGNOSTICS

Designed to generate evidence on digital diagnostics for use in LMICs

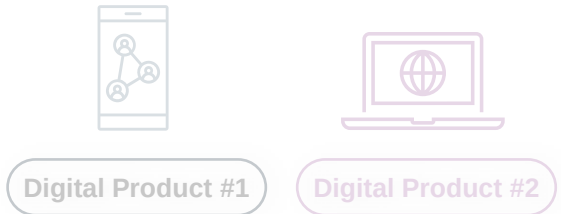


VALIDATION PLATFORM

FOR ONBOARDING AND EVALUATION OF AI BASED DIAGNOSTICS

Designed to generate evidence on digital diagnostics for use in LMICs

Onboarding and installation of digital diagnostics



V F

Evaluation space

- Digital Library #1 (screening CXR images)
- Digital Library #2 (triage CXR images)
- Digital Library #3 (error CXR images)

Reporting space

Performance data

Gold standard + clinical data

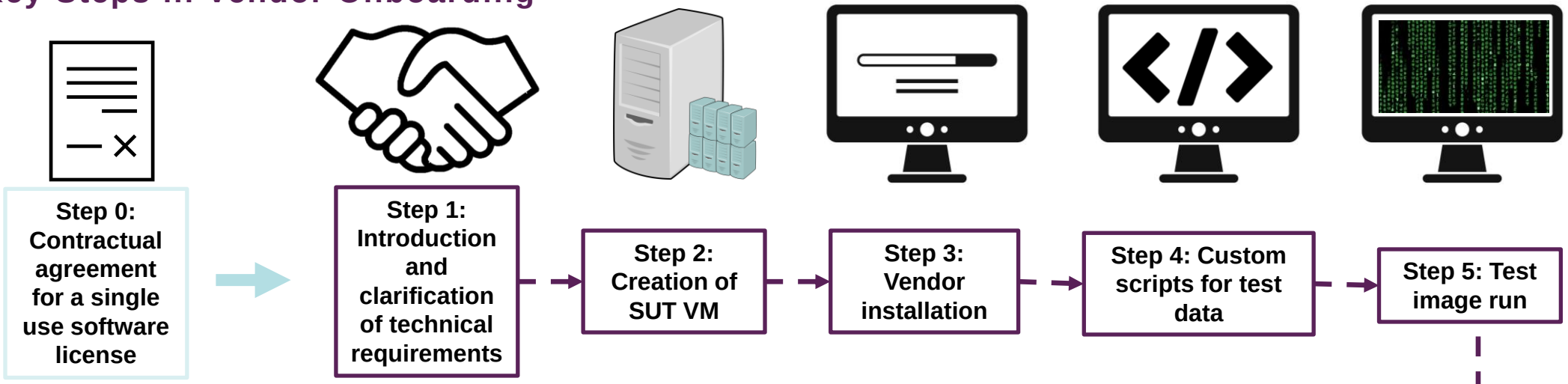
Standardized analysis and reporting templates

Evaluation dossier

Policy makers & regulators

VALIDATION PLATFORM

Key Steps in Vendor Onboarding



Step 0: Contractual agreement for a single use software license

Step 1: Introduction and clarification of technical requirements

Step 2: Creation of SUT VM

Step 3: Vendor installation

Step 4: Custom scripts for test data

Step 5: Test image run

Step 6: FIND Confirm Test Images are available and usable

Step 7: FIND gets software training

Step 8: Backup of virtual computer (*optional)

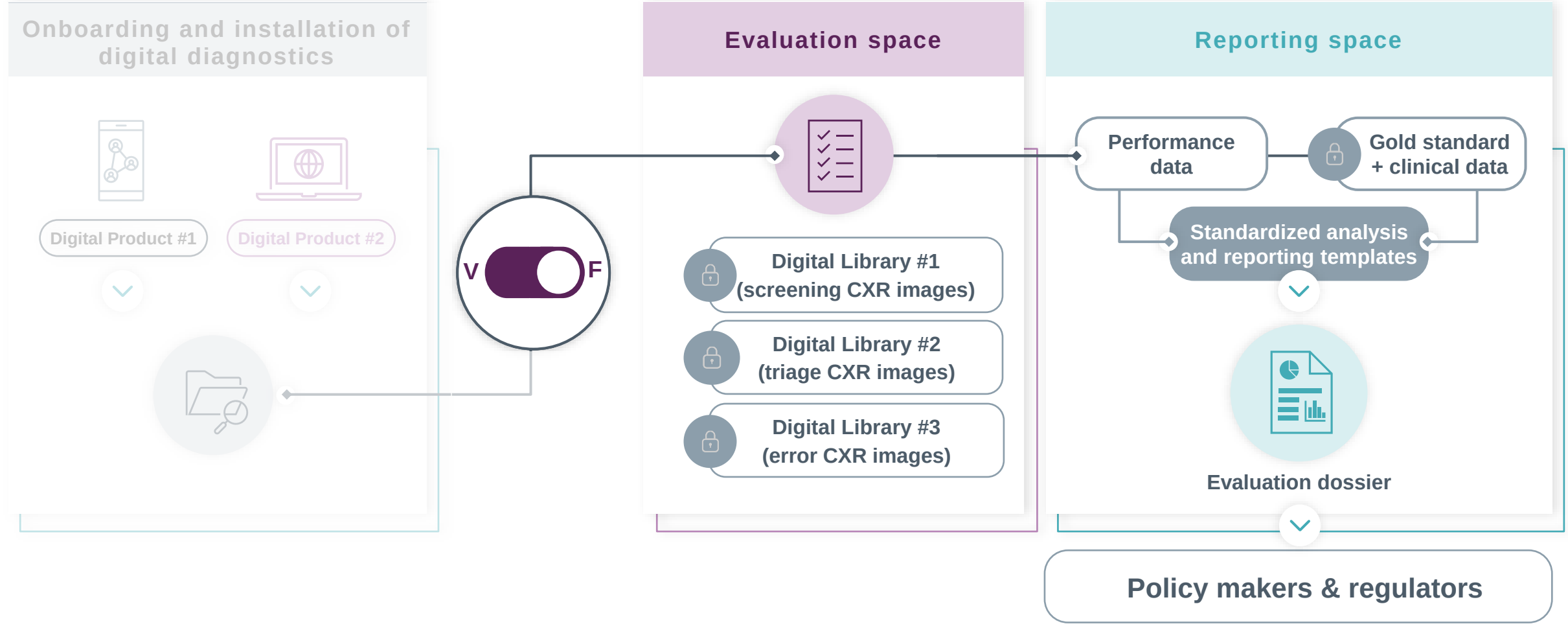
Step 9: Removal of vendor access

Step 10: Mutual agreement that software is ready for evaluation

VALIDATION PLATFORM

FOR ONBOARDING AND EVALUATION OF AI BASED DIAGNOSTICS

Designed to generate evidence on digital diagnostics for use in LMICs



FIND

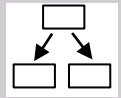
3

FIND VALIDATION PLATFORM: A TOOL FOR INDEPENDENT EVALUATION OF CXR CAD PRODUCTS

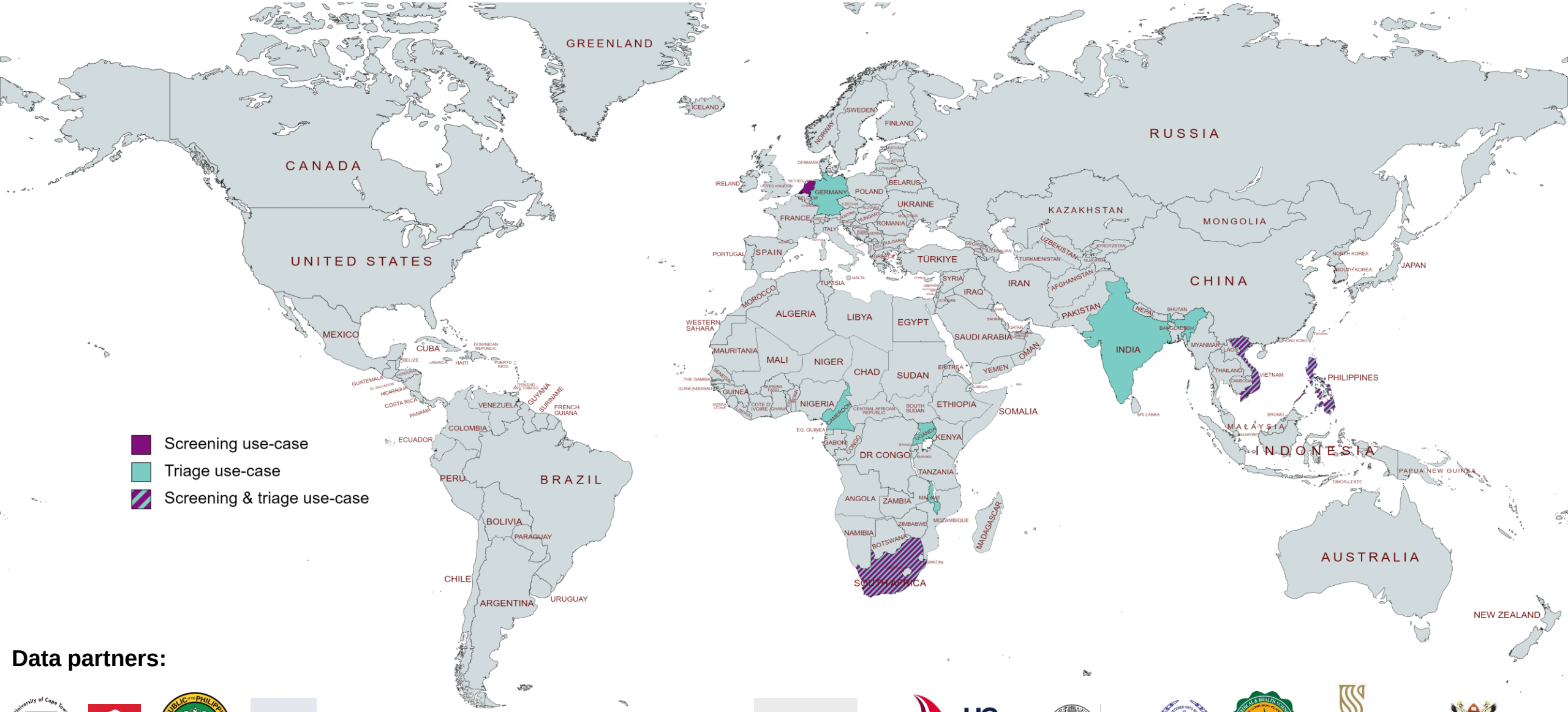
VALIDATION PLATFORM

Principles of independent performance evaluation conducted by FIND

- FIND curated images and clinical data from different countries and use-cases
- Relevant ethical permissions obtained
- None of the images included in the validation datasets have been shared with CAD developers for training
- Performance evaluations will be carried out for use-cases that have been approved by WHO

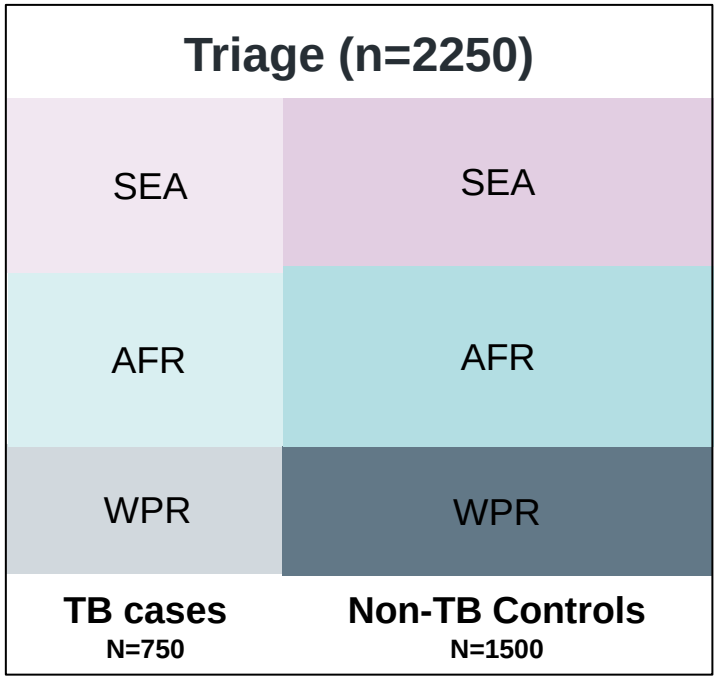
Validation datasets:	
	Screening and triage use cases (adults)
	Images from ≥ 2 WHO regions
	Sampled in case-cohort design
1:2	Fixed ratio of cases and non cases from each study
	All images linked to a microbiologic reference standard result
	Independently interpreted by an expert radiologist

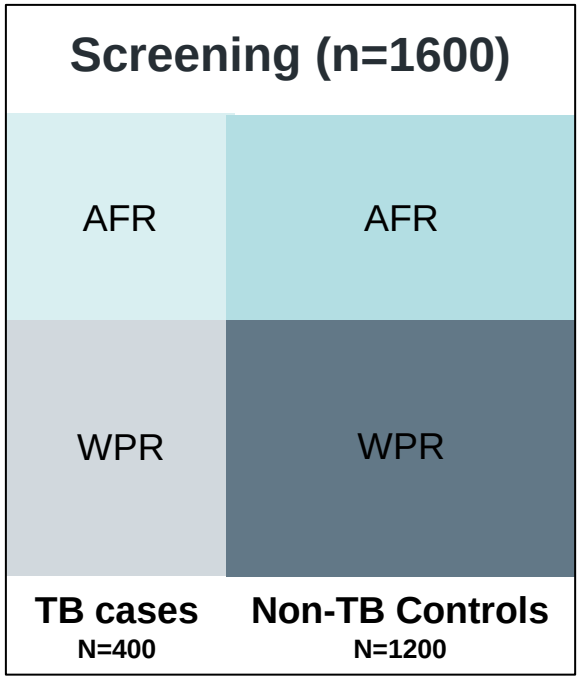
FIND'S GLOBAL TB VALIDATION ARCHIVE:
GROWING NUMBER OF COUNTRIES SHARING SCREENING AND TRIAGE DATA



Data partners:

FIND'S GLOBAL TB VALIDATION DATASET V1.0: MULTIPLE WHO REGIONS AND USE CASES

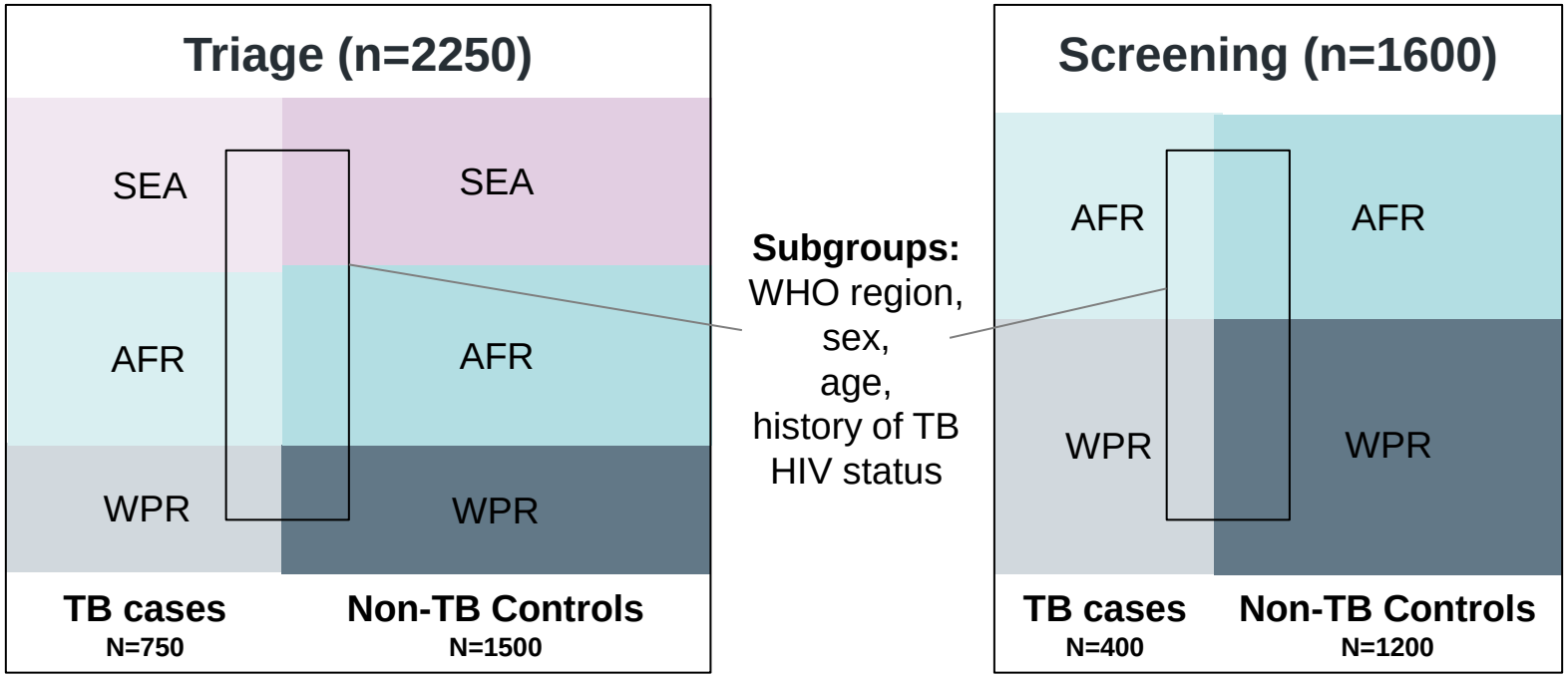




Triage use-case test:
A test that can be rapidly conducted among **people presenting to a health facility** to differentiate those who should have further diagnostic evaluation for TB from those who should undergo further investigation for non-TB diagnoses.

Screening use-case test:
A test that can be rapidly conducted to systematically test **people who are at risk for TB disease**, in a predetermined target group. For those who screen positive, the diagnosis needs to be established by one or several diagnostic tests and additional clinical assessments.

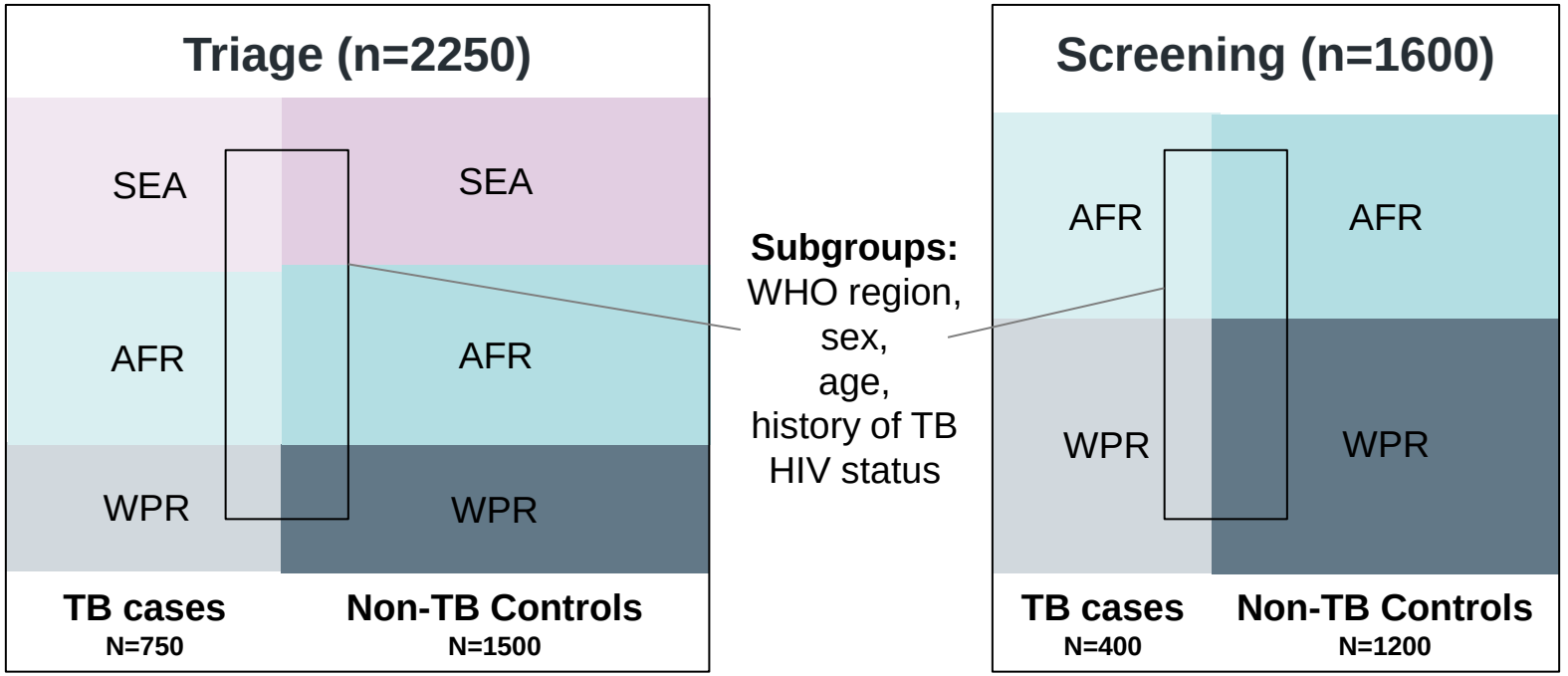
FIND'S GLOBAL TB VALIDATION DATASET V1.0: MULTIPLE WHO REGIONS AND USE CASES



Triage use-case test:
A test that can be rapidly conducted among **people presenting to a health facility** to differentiate those who should have further diagnostic evaluation for TB from those who should undergo further investigation for non-TB diagnoses.

Screening use-case test:
A test that can be rapidly conducted to systematically test **people who are at risk for TB disease**, in a predetermined target group. For those who screen positive, the diagnosis needs to be established by one or several diagnostic tests and additional clinical assessments.

FIND'S GLOBAL TB VALIDATION DATASET V1.0: MULTIPLE WHO REGIONS AND USE CASES



Repeatability (n=120)
Global

Error handling (n=60)
Global

Triage use-case test:
A test that can be rapidly conducted among **people presenting to a health facility** to differentiate those who should have further diagnostic evaluation for TB from those who should undergo further investigation for non-TB diagnoses.

Screening use-case test:
A test that can be rapidly conducted to systematically test **people who are at risk for TB disease**, in a predetermined target group. For those who screen positive, the diagnosis needs to be established by one or several diagnostic tests and additional clinical assessments.

FIND

4

STANDARDIZED OBJECTIVES & REPORTING

VALIDATION PLATFORM

FIND'S APROACH TO KEY PERFORMANCE EVALUATION:

Primary Objective #1	Primary Outcomes (performance)
Secondary Objectives #2-3	Secondary Outcomes (performance)
Secondary Objectives #4	Secondary Outcomes (repeatability)
Secondary Objectives #5	Secondary Outcomes (error handling)

VALIDATION PLATFORM

FIND'S APROACH TO KEY PERFORMANCE EVALUATION:

Primary Objective #1

To assess the clinical performance of TB CAD software in **screening** and **triage use** cases for adults, using a microbiological reference standard.

Primary Outcomes (performance)

Demonstrate that performance of products undergoing evaluation are non-inferior to products that were reviewed by the WHO GDG in 2020.

Conclusion will inform WHO and policy makers

Secondary Objectives #2-3

Secondary Outcomes (performance)

Secondary Objectives #4

Secondary Outcomes (repeatability)

Secondary Objectives #5

Secondary Outcomes (error handling)

VALIDATION PLATFORM

FIND'S APROACH TO KEY PERFORMANCE EVALUATION:

Primary Objective #1

Primary Outcomes (performance) 

Secondary Objectives #2-3

Secondary Outcomes (performance) 

To describe additional clinical performance measures of CXR CAD software

To conduct **subgroups analysis by:** gender, age group, HIV status, TB history & WHO region

Provide additional descriptive information about products which may be of relevance for end-users to understand the differences between different products

Secondary Objectives #4

Secondary Outcomes (repeatability)

Secondary Objectives #5

Secondary Outcomes (error handling)

VALIDATION PLATFORM

FIND'S APROACH TO KEY PERFORMANCE EVALUATION:

Primary Objective #1

Primary Outcomes (performance)

Secondary Objectives #2-3

To describe additional clinical performance measures of CXR CAD software

To conduct **subgroups analysis by:** gender, age group, HIV status, TB history & WHO region

Secondary Outcomes (performance)

Provide additional descriptive information about products which may be of relevance for end-users to understand the differences between different products

Secondary Objectives #4

Secondary Outcomes (repeatability)

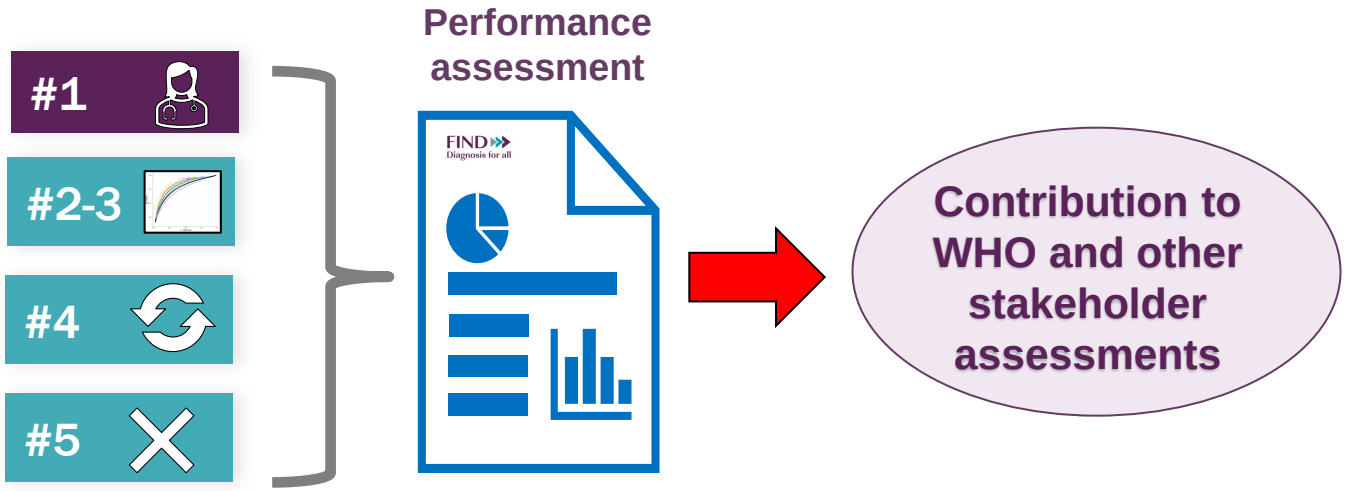
To assess the repeatability of CXR CAD software.

Secondary Objectives #5

Secondary Outcomes (error handling)

To assess error handling for a given CXR CAD software when exposed to inadequate or challenging CXR images.

STANDARD METRICS ACROSS SOFTWARE



- FIND will produce a standard report on the evaluation for each CAD software assessed on the validation platform
- Reports will clearly indicate details including: product name and version, testing date, the content of the library/ies and outputs of evaluation
- Reports will be shared with WHO PQ, stakeholders, and developers and ultimately be published

FIND

5

EVIDENCE GENERATION FOR FUTURE USE CASES

EVIDENCE GENERATION

EXPANDING DATA AND USE CASES

FIND is proactively expanding its archives to assess CAD performance for new use cases:

- **Pediatric TB**

  Evaluate TB CAD performance in pediatric use cases

- **PLHIV**

  Evaluate TB CAD performance in PLHIVs (in- and out-patients)

- **Non-TB CXR findings and non-TB indications**

  Evaluate CXR CAD performance in evaluation of pneumonia, COVID-19, and other radiographic findings

FIND'S GLOBAL TB VALIDATION ARCHIVE:
EXPANDING THE ARCHIVE TO ALLOW FOR EVALUATION OF ADDITIONAL USE-CASES FOR TB

Data partners:

Status: Q2 data at FIND

FIND'S GLOBAL TB VALIDATION ARCHIVE:
EXPANDING THE ARCHIVE TO ALLOW FOR EVALUATION OF OTHER RESPIRATORY FINDINGS BEYOND TB

Data partners:

EPFL

UNIVERSITÄT HEIDELBERG ZUKUNFT SEIT 1386

CLINTON HEALTH ACCESS INITIATIVE

Swiss TPH
Swiss Tropical and Public Health Institute

Status: Q2 data at FIND

ACKNOWLEDGEMENTS

Data contributors

- University of Cape Town Lung Institute, Cape Town, South Africa
- Aurum Institute, South Africa
- National Lung Hospital and Hanoi Lung Hospital, Vietnam
- Woolcock Institute of Medical Research, Glebe, Australia
- University Medical Center Groningen, Groningen, Netherlands & GGD Groningen, Groningen, Netherlands
- Research Center Borstel, Sülfeld, Germany
- Bamenda Regional Hospital, Bamenda, Cameroon & Tuberculosis Reference Laboratory Bamenda, Bamenda, Cameroon
- National Tuberculosis Programme, Philippines
- Clinton Health Access Initiative (CHAI)
- National Capital Territory of Delhi, India
- Liverpool School of Tropical Medicine, UK
- University College London, UK
- University of California San Francisco, USA
- University Hospital of Heidelberg, Germany
- Christian Medical College, Vellore, India
- De La Salle Medical and Health Sciences Institute, Philippines
- Stellenbosch University, South Africa
- Makerere University College of Health Sciences, Uganda

- **Digital health infrastructure and onboarding:**
  - Nikhil Jagtiani
  - Nick Banks,
  - Sophie Crettaz
  - Jean-Philippe Lutz
  - Dennis Walusimbi
- **Data formatting and analysis:**
  - Aurelie Kamoun
  - Sam Linsen
  - Nathalie Frey
- **Technical partner infrastructure:**
  - eSHIFT
- **Radiologist reading:**
  - IOM
- **Scientific oversight and coordination**
  - Matthew Arentz
  - Sandra Kik
  - Victor Anyebe
  - Mikashmi Kohli
  - Morten Ruhwald
  - Rigveda Kadam
  - Stefano Ongarello

FIND

THANK YOU