1
2

TECHNICAL SPECIFICATION SERIES FOR SUBMISSION TO WHO PREQUALIFICATION - MEDICAL DEVICE ASSESSMENT SCREENING AND TRIAGE OF TB USING COMPUTER AIDED DETECTION SOFTWARE FOR INTERPRETATION OF CHEST X-RAY IMAGES

DRAFT 2024/02/09

REV 11

Abbreviations

The following abbreviations have been used in this document.

Acronym	Name
AI	Artificial Intelligence
CAD TB	Computer Aided Detection Software for Tuberculosis
CAPA	Corrective and Preventive Actions
CE Mark	Conformité Européene or European Conformity
CXR	Chest x-ray
DICOM	Digital Imaging and Communications in Medicine
DHF	Design History File
EIR	Establishment Inspection Report
EMC	Electromagnetic Compatibility
EP	Essential Principles
EU	European Union
FDA	Food and Drug Administration
FSCA	Field safety corrective action
GDG	Guideline Development Group
GMLP	Good Machine Learning Practice
IEC	International Electrotechnical Commission
IFU	Instructions for Use
IMDRF	International Medical Device Regulators Forum
ISO	International Organization for Standardization
IVD	Invitro Diagnostic
LMIC	Low and middle income countries
MDR	Medical Device Regulation
MDSAP	Medical Device Single Audit Program
ML	Machine learning
MRMC	Multiple reader multiple case
N	Number
PQ	Prequalification
QMS	Quality Management System
RFID	Radio-frequency identification
ROC	Receiver Operating Characteristic
SOP	Standard Operating Procedures
SOUP	Software of unknown provenance
SRS	Software Requirements Specification
SW	Software
TB	Tuberculosis
TGA	Therapeutic Goods Administration
TSS	Technical Specifications Series
UDI	Unique Device Identification
UN	United Nations
WHO	World Health Organization

Definitions

The following terms have been used in this document. The associated definitions have been taken from WHO Consolidated Guidelines on Tuberculosis Module 2: Screening 9789240022676-eng.pdf (who.int)

Term	Definition
Screening test, examination, or other procedure for TB disease	A test, examination, or other procedure used to distinguish people with a high likelihood of having TB disease from people who are highly unlikely to have TB. A screening test is not intended to be diagnostic. People with positive results on a screening test should undergo further evaluation, depending on the screening algorithm used.
Triage	The process of deciding the diagnostic and care pathways for people based on their symptoms, signs, risk markers and test results. Triaging involves assessing the likelihood of various differential diagnoses as a basis for making clinical decisions. It can follow more- or less-standardized protocols and algorithms, and it may be done in multiple steps.

Summary of Sections for Comments

SECTION 1 –ADMINISTRATIVE	
1.01	Quality Management System, Full Quality System or Other Regulatory Certificates
1.02	Free Sale Certificate/ Certificate of Marketing authorization
SECTION 2 – SUBMISSION CONTEXT	
2.01	Device Description
2.01.01	Comprehensive Device Description and Principle of Operation
2.01.02	Description of Device Packaging
2.01.03	History of Development
2.01.04	Reference and Comparison to Similar and/or Previous Generations of the Device
2.02	Indications for Use and/or Intended Use and Contraindications
2.02.01	Intended Use; Intended Purpose; Intended User; Indications for Use
2.02.02	Intended Environment/ Setting for use
2.02.03	Contraindications For Use
2.03	Global Market History
2.03.01	Global Market History
2.03.02	Global Incident Reports and Recalls
2.03.03	Sales, Incident and Recall Rates
2.03.04	Evaluation/Inspection Reports
SECTION 3 – NON-CLINICAL EVIDENCE	
3.01	Risk Management
3.02	Essential Principles (EP) Checklist
3.03	Electrical Systems: Safety, Mechanical and Environmental Protection, and Electromagnetic Compatibility
3.04	Software/Firmware
3.04.01	Software/Firmware Description
3.04.02	Hazard Analysis
3.04.03	Software Requirement Specification
3.04.04	Architecture Design Chart
3.04.05	Software Design Specification
3.04.06	Traceability Analysis
3.04.07	Software Development Environment Description
3.04.08	Software Verification and Validation
3.04.09	Revision Level History
3.04.10	Unresolved Anomalies (Bugs or Defects)
3.04.11	Cybersecurity
3.04.12	Interoperability
3.05	Other non-clinical Evidence
SECTION 4 – CLINICAL EVIDENCE	
4.01	Clinical Evidence
4.02	Usability/Human Factors
SECTION 5 – LABELLING AND PROMOTIONAL MATERIAL	
5.01	Product/Package Labels
5.02	Package Insert/Instructions for Use
5.03	e-labelling
5.04	Technical/Operators Manual
5.05	Product Brochures
5.06	Other Labelling and Promotional Material
SECTION 6A – QUALITY MANAGEMENT SYSTEM PROCEDURES Administrative information needed to evaluate the premarket submission related to the QMS (not device specific)	
6A.01	Product Descriptive Information
6A.02	General Manufacturing Information
6A.03	Quality management system procedures
6A.04	Management responsibilities procedures
6A.05	Resource management procedures
6A.06	Product realization procedures
6A.07	QMS measurement, analysis and improvement procedures
6A.08	Other Quality System Procedures Information

SECTION 6B – QUALITY MANAGEMENT SYSTEM DEVICE SPECIFIC INFORMATION	
6B.01	Quality management system information
6B.02	Management responsibilities information
6B.03	Resource management information
6B.04	Product realization information
6B.05	QMS measurement, analysis and improvement procedures
6B.06	Other Device Specific Quality Management System Information
ANNEXES	
Annex A	Considerations for Risk Management to ensure destination countries specificities, including LMICs, are properly addressed
Annex B	WHO TB CAD: Justification of sample size
Annex C	Annex C: Referenced documents

1
2

Introduction to TSS

The Technical Specifications Series documents (TSS) set out the performance evaluation criteria for meeting prequalification requirements. Each TSS document provides information on the minimum performance requirements for WHO prequalification that should be met by a manufacturer to ensure that the medical device that is being submitted for prequalification is safe and performs optimally. Compliance with prequalification technical specifications is verified during surveillance audits. Failure to comply with the relevant technical specifications will result in the delisting of the product concerned from the WHO List of Prequalified Medical Devices. TSS are not designed to be used for procurement.

Purpose of this TSS

The purpose of this document (Screening and Triage of TB using Computer Aided Detection Software for Interpretation of Chest X-ray Images) is to provide technical guidance to manufacturers who create software that is a medical device, that intend to seek WHO prequalification of Computer Aided Detection Software (CAD) that are interpreting Chest X-ray images (CXR) for Tuberculosis (CAD TB) and that can replace an interpreting radiologist. Minimum performance requirements for prequalification are summarised herein.

This document applies to CAD TB products when the intended use is for screening or triage of tuberculosis as recommended by WHO in 2021 (ref guidelines). This document only addresses those CAD TB medical devices that are intended to replace a human reader. Devices that are intended to "assist", rather than "replace", trained/qualified human readers are out of scope of this TSS.

With regard to the devices being used to 'replace a human reader', it is important to note that this refers to the decision step involving a trained radiologist. It is the intention that these products will be used in a clinical setting, by trained radiographers who are responsible for and familiar with, operation of digital chest radiograph equipment, capture and interpretation of adequate quality CXR images, identification of foreign bodies, implanted medical devices, or other external factors that may impact effective use of CAD TB software.

Some CAD TB may be accompanied by hardware that are either accessories or essential components of the system. The requirements for establishing the safety and performance of the hardware are out of the scope of this TSS. However, certain documents that support their safety and performance when used together with the CAD software as a system may be requested by WHO.

Where possible, WHO performance conditions are aligned with published guidance, standards and/or regulatory documents. Although references to source documents are provided, in some cases WHO has additional requirements. These requirements are associated with ensuring the safe and effective use of these devices in jurisdictions with more challenging conditions. A list of considerations relating to such conditions can be found in Annex A. For prequalification purposes, manufacturers must provide evidence in support of the clinical performance of CAD TB which can demonstrate that reasonable steps have been taken to ensure that a properly manufactured device, being correctly operated in the hands of the intended user, will detect the condition and fulfil its indications for use consistently. The 2021 WHO guidelines that recommended the use of CAD TB was based on evidence showing that the performance of this technology for screening and triage of TB was comparable to a human reader.

Prequalification requirements summarised in this document do not extend to the demonstration of clinical utility, i.e. the effectiveness and/or benefits of CAD TB, relative to and/or in combination with other measures, as a tool to inform clinical intervention in a given population or healthcare setting. To demonstrate clinical utility, a separate set of studies is required. Clinical utility studies usually inform programmatic strategy and are thus the responsibility of programme managers, Ministries of Health and other related bodies in individual WHO Member States. Such studies do not fall under the scope of prequalification but are considered by WHO when issuing or updating its technical guidelines.

		General Requirements	WHO Comments/proposals
SECTION 1 –ADMINISTRATIVE			
1.01	Quality Management System, Full Quality System or Other Regulatory Certificates including product related Certificates	Provide Objective evidence of Quality System implementation and regulatory authorisation. It can take the form of conformity assessment procedure and type of application (new, change or recertification)), current certification details, manufacturer details, critical supplier details and device details including classification as well as other official means of demonstrating regulatory authorisation. For example: Objective evidence of certification or regulatory approval such as but not limited to: 1. ISO 13485 certificate, 2. CE Marking certificates, 3. MDSAP certificate, 4. TGA certificates. 5. FDA Premarket Approval Letter (PMA) letter 6. Japan PMDA issued certificates 7. Japan Registered Certification Body (RCB) issued certificates 8. FDA 510(k) Summary 9. FDA De Novo Decision Summary 10. Japan PMDA Ministers Approval/ "Shonin Approval" of High Risk (generally class C+D) devices. 11. Health Canada Medical Device Licence (MDL)	Acceptable standards for Quality Management System 21 CFR Part 820 ISO 13485:2016 (EN ISO 13485:2016+AC:2018) Article 10 EU MDR
1.02	Free Sale Certificate/ Certificate of Marketing authorization	Document/certificate/letter issued by a Regulatory Authority or an official representative (e.g.: Notified Body) for the same version that is being submitted where the medical device is marketable.	Ensure certificate and/or license are available for the jurisdiction(s) where the device is marketed.
SECTION 2 – SUBMISSION CONTEXT			
2.01	Device Description		
2.01.01	Comprehensive Device Description and Principle of Operation	a) A general description of the device, including: i. A statement of the device name (see Notes 1+2) ii. What the device does iii. Who uses it and for what (High level statement) iv. Risk class of the device (for each jurisdiction where device is marketed) v. Where to use it (places/environment where the device is intended to be used)	For example, but not limited to: ● Commercial name, ● Purpose and function, ● Classification in the different jurisdictions, ● Indication for use, ● User requirements, ● Targeted population requirements,

		General Requirements	WHO Comments/proposals
		vi. How it works. Including a description of the features/variants/operating modes that enable the device to be used for indications/intended use (principle of operation/mechanism of action) and if not readily apparent or typical for the device type, a brief description of the underlying science/technology, design concepts, and/or theoretical principles supporting the device's function. vii. If applicable, labelled pictorial representation (diagrams, photos, drawings). viii. If system, how the components relate ix. identify the role of any software/firmware. Note 1: For the AI algorithm and any other software that comes with it, the version should be specified, plus a date when the version was released. Note 2: The "device" may include the software and associated hardware. Note 3: Further details about the software are requested in *section 3.04.* b) Product specification, including: i. Physical characteristics or relevance to the end user (dimensions, performance specifications, weight) ii. Features and operating modes iii. Input specifications (e.g., electrical power requirements, settings and associated allowable ranges/limits) (see Note 4) iv. Output and performance characteristics (see Note 5)) v. Overall Device Performances Processing time c) List of accessories intended to be used in combination with the devices. d) Indication of any other medical devices or general product intended to be used in combination with the medical device. Note 4: This includes input requirements for images that can be processed by the software. Note 5: The output results that are generated by the software should be included here.	• Copy of labelling (IFU + labels) in English and availability of translations, • Functional diagram, • Description of interactions with associated Medical Device, • Description of options and accessories, • Description of Software structure, Software distribution and versioning method, • Description of physical and electrical characteristics, • Description of environmental conditions for storage, installation, and operational usage, • Description of data inputs requirements, • Description of data output standard, • Output performance characteristics, • Combinations and interactions. • Modality: There might be different modalities available for how the product is deployed. Provide a description of your device for each of the modalities (e.g. a hardware box, laptop or tablet with the product installed on for offline use, a cloud based installation of the product, a local server or cloud based installation of the product etc).
2.01.02	Description of Device Packaging	a) Information regarding the packaging of the devices, including, when applicable, primary packaging, secondary and any other packaging associated. b) Specific packaging of accessories marketed together with the medical devices must also be described;	If applicable, packaging must include information regarding the requirements for transport, storage, and operational use. All options and accessories must include the same information.
2.01.03	History of Development & validation	Provide an overview of the previous generation or generations of the device produced by the manufacturer, where such devices exist.	The multi generation description of the device must include the changes conducted as well as the benefit for users and patients. With each new generation the risk management file, the usability study and the residual risk report must be updated.

		General Requirements	WHO Comments/proposals
		For any device versions/prototypes and any software version referenced in the evidence presented in the submission, a table describing the version/name, with 4 columns (Device Name and/or Version; Software Name and Version; Description of changes from previous row; motivation for the change; list of verification/validation activities, including clinical studies conducted using this version). For any design verification or validation activities presented in this submission (including clinical studies) performed on any earlier versions of the subject device and software, include a justification for why the changes do not impact the validity of the data collected under those activities in supporting the safety and effectiveness of the final device design.	Change controls must be established since the beginning of verification and validation. Verification and validation must be conducted on devices representative of the production version. Verification and validation shall be conducted considering the options and the accessories of the device. The device should be tested to the limits for verification and validation.
2.01.04	Reference and Comparison to Similar and/or Previous Generations of the Device	a) A list of similar devices (available on local and international market) and/or previous generation of the devices (if existent) relevant to the submission. This should include any similar/previous generation devices that were previously reviewed and refused by the subject regulator. b) Description of why these similar products were selected. c) A key specification comparison, preferably in a table, between the references (similar and/or previous generation) considered and the device.	If applicable, comparisons can be used to support the safety and effectiveness of the subject device. Similar Medical Devices should be reviewed for equivalence and predicates must be properly documented and evaluated. Benefit/risk balance should be evaluated and documented.
2.02	**Indications for Use and/or Intended Use and Contraindications**		
2.02.01	Intended Use; Intended Purpose; Intended User; Indications for Use	This section should include, as appropriate: a) **Intended Use:** The statement of intended use must specify the clinical function provided by the device. b) **Intended Purpose:** What is expected with the use of this medical device? Which results are expected? c) **Intended user** and skills**/**knowledge/training that the user should have to operate or use the device. d) **Indications for Use:** i. Disease or medical condition that the device will screen for or triage, parameters to be monitored and other considerations related to indication for use. ii. If applicable, information about patient selection criteria. iii. If applicable, information about intended patient population (e.g. adults symptomatic, adult asymptomatic,) or a statement that no subpopulations exist for the disease or condition for which the device is intended. **NOTES:** i. The statements of intended use and purpose and the intended user and indications for use must be **<u>as presented in the labelling</u>**.	To meet WHO recommendations for CAD for TB, the manufacturer's intended use must be aligned to the following use-case: *Among individuals aged 15 years and older in populations in which testing for TB is recommended, computer-aided detection software programmes may be used to interpret abnormalities on chest radiographs that are suggestive of TB in place of human readers for screening and triage for TB disease. The results are expressed as abnormality scores as well as binary or categorical outcomes. Its use should be limited to the interpretation of plain CXRs for pulmonary TB in individuals aged 15 years or older.* The intended use statement must reflect the need to replace the review of CXRs by trained/qualified human readers for the detection of TB. Devices intended only to assist human readers are out of scope of WHO PQ. The CAD system should only process images generated on a Digital CXR imager.

		General Requirements	WHO Comments/proposals
		ii. If more than one device is included, the information should be provided for each device. iii. It is expected the use of the CAD software will be limited to use of digital CXR images as opposed to scanned CXR film.	
2.02.02	Intended Environment/ Setting for use	a) The setting where the device is intended to be used (e.g., domestic use, hospitals, medical/clinical laboratories, ambulances/mobile vans or trucks, medical offices, rural, urban, ambulatory, primary or secondary settings). Multiple options can be indicated. b) If applicable, environmental conditions that can affect the device's safety, handling, and/or performance (e.g., temperature, humidity, power, internet connectivity, pressure, movement).	The clinical setup in which the Medical Device is planned to be used must be defined and documented. The description should integrate any specificity for countries/regions in which the Medical Device is marketed. The environmental conditions in which the Medical Device is planned to be used must be defined and documented. Ensure the description integrates any specificity for countries/regions in which the Medical Device is marketed (e.g., temperature, humidity, altitude (pressure), power, telecommunication, mobility (rail, road as example)).
2.02.03	Contraindications and Limitations for Use	If applicable, specify the disease or medical conditions that would make use of the device inadvisable due to an unfavourable risk/benefit profile (contraindications). In addition, specify any recognised factor or condition that may restrict or affect the optimal performance of the CAD software (limitation).	For instance, CAD TB may not operate correctly in the presence of medical conditions that impede a clear imaging of the lung fields, such as extreme obesity or severely deformed thoracic cage. Ensure the description integrates any specificity for countries in which the Medical Device is marketed. **NOTE:** The statements concerning contraindications and limitations for the device must be as presented in the labelling.
2.03	Global Market History		
2.03.01	Global Market History	a) With respect to CAD for TB, provide an up to date indication of the markets (all countries or jurisdictions) where the device is approved for marketing, including any marketing under compassionate use regulations. b) Should include history of the marketing of the device by any other entity in as much detail as possible, acknowledging that detailed information may not be available in all cases. c) If the subject device is different in any way (e.g., design, labelling, specifications) from those approved or marketed in other jurisdictions, the differences should be described. d) The month and year of market approval in each country or jurisdiction where the device is marketed. If the device has been marketed for greater than 10 years, a statement of greater than 10 years can be made. e) For each of the markets listed in (a) above, a statement of the commercial names used in those markets OR a clear statement that the commercial names are the same in all jurisdictions.	Ensure the Medical Device distribution is controlled for the countries where the Medical Device is marketed.

		General Requirements	WHO Comments/proposals
		f) State the date of data capture for the market history data.	
2.03.02	Global Incident Reports and FSCAs	a) List incidents (safety, quality) associated with the device and a statement of the period associated with this data. b) If the number of incidents is voluminous, provide a summary that states the number of incidents events for each medical device problem type. c) List of the medical device field safety corrective actions and/or advisory notices, and a discussion of the handling and solution given by the manufacturer in each case. d) A description of any root cause analysis, any corrections made, and/or corrective actions undertaken in response to items listed above. e) If there have been no incidents, field safety corrective actions and/or advisory notices to date, provide an attestation from device owner on company letterhead, that there have been no incidents, field safety notices and/or advisory notices since commercial introduction of the device. **NOTE** It is acknowledged that the definition of recall may vary from one jurisdiction to another.	Post market activities must be established and effective for all countries where the Medical Device is to be marketed. Procedural requirements must be established and documented for Post Market activities in countries where the Medical Device is marketed. When no specific jurisdictional requirements have been established, ensure there is a system in place to process feedback activities. Ensure proper definition and documentation of requirements has been made when dealing with subcontractors and/or agents. Ensure monitoring of third parties is established and effective. Ensure mis-use cases are documented, evaluated and proper actions have been established for records issued from the different countries where the Medical Device is marketed.
2.03.03	Sales, Incident and FSCA Rates	a) A summary of the number of units sold (devices and software stratified by version) in each country/region and a statement of the period associated with this data. b) Provide the rates calculated for each country/region, for example: i. Incident rate = # incidents divided by # units sold, expressed as a percentage ii. FSCA rate = N FSCA divided by N units sold, expressed as a percentage Rates may be presented in other appropriate units such as per patient year of use or per use. In this case, methods for determining these rates should be presented and any assumptions supported. c) Critical analyses of the rates calculated (e.g. Why are they acceptable? How do they break down in terms of incidents? Is there some outlier data that has driven the rates up? Are there any trends associated with any sub-groups of the devices that are subject of the submission (e.g. size, version)?). **NOTES** i. Sales in this context should be reported as the number of units sold.	Ensure records issued from all countries where the Medical Device is marketed are monitored and properly reviewed by top management of the manufacturer.

		General Requirements	WHO Comments/proposals
		ii. The summary of sales should be broken down by components when appropriate.	
2.03.04	Evaluation/Inspection Reports	Copies of Evaluation certificates /Inspection Reports from other parties (e.g., Notified Body inspection reports/MDSAP reports/ CE certification reports).	Ensure reports from evaluation/inspection bodies are available for review. For example, review reports from: ● EIR for FDA, ● MDSAP report for Audit Organization, ● CE Marking audit report for EU, ● ISO 13485 for Notified bodies.
SECTION 3 – NON-CLINICAL EVIDENCE			
3.01	Risk Management	a) A summary of the risks identified during the risk analysis process and how these risks have been controlled to an acceptable level. b) The results of the risk analysis should provide a conclusion with evidence that remaining risks are acceptable when compared to the benefits. Where a standard is followed, identify the standard	The Risk Management file shall cover the specific Hazards & Hazardous Situations experienced in destination countries, refer to Annex A. Refer to ISO 14971:2019 For Risk Management. Verify it is implemented by the Manufacturer. It is recommended to refer to the consensus report, AAMI CR34971:2022 "Guidance on the Application of ISO 14971 to Artificial Intelligence and Machine Learning". Verified implementation of Usability (ISO 62366-1:2015). Verify how Usability connects to Risk Management (separate or integrated). Residual risks must be properly identified, evaluated and deemed acceptable (benefit/risk balance). Ensure the risk of poor-quality images is considered in both the design and the warnings provided in the instructions for use. The medical device should be capable of identifying and notifying the user of poor quality and inappropriate images that are inadequate for interpretation. Data protection including non-authorised physical access to the system in the setting of use should be considered. It is recommended that the IMDRF document "Principles and Practices for Medical Device Cybersecurity "dated 18-Mar-2020 is considered in the risk assessment. In addition, the requirements of The Health Insurance Portability and Accountability Act of 1996 (HIPAA) (a US based federal law that are in place to protect sensitive patient health information from being disclosed without the patient's consent or knowledge) and similarly, the

		General Requirements	WHO Comments/proposals
			General Data Protection Regulation (EU GDPR) to protect patient data access for both Manufacturers and Health establishments are considered in the risk assessment. The principles applied in these two regulations are supported by WHO PQ.
3.02	Essential Principles (EP) Checklist	i. An EP checklist (such as that established by IMDRF) established for the medical devices, information about method(s) used to demonstrate conformity with each EP that applies, references for the method adopted and identification of the controlled document with evidence of conformity with each method used. ii. For the controlled documents indicated which are required for inclusion in the submission: a cross-reference of the location of such evidence within the submission. iii. If any EP indicated in the checklist does not apply to the device: a documented rationale of the non-application of each EP that does not apply. **NOTE:** i. Methods used to demonstrate conformity may include one or more of the following: ii. conformity with standards; iii. conformity with a commonly accepted industry test method(s); iv. conformity with an in-house test method(s); v. the evaluation of pre-clinical and clinical evidence; comparison to a similar device already available on the market. vi. Where standards and guidance are referenced, please ensure the following information is provided: a) the standard organisation, standard number, standard title, year/version, and if full or partial compliance. b) list of relevant guidance documents published by regulators and referenced in the design and/or manufacture of the device with the jurisdiction of publication, publication date and title identified c) If applicable, a list of relevant clinical guidelines referenced in the design and/or manufacture of the device, the publisher, publication date and title identified.	Refer to the ISO 60601 series and specially for requirements relating to the image acquisition device. Ensure clarity of combination of device and appropriate handling of patient workflow. Ensure a list of appropriate standards are defined and documented. Ensure relevant guidelines and standards for ML have been considered and documented as part of the Design and in Risk Management. Where hardware is an integral part of the CAD TB system, the checklist should include reference to evidence that the combination of the hardware with the software does not result in impaired performance, and that the system is safe. Any restrictions on use applying to such combinations shall be indicated on the label and/or in the instructions for use.
3.03	Electrical Systems: Safety, Mechanical and Environmental Protection, and Electromagnetic Compatibility	Evidence supporting electrical safety, mechanical and environmental protection, and electromagnetic compatibility are to be included in this section. This should include: a) A summary of the non-clinical evidence that falls within this category b) A discussion of the non-clinical testing considered for the device and support for their selection or omission from the verification	Refer to ISO 60601 series for applicability for the medical device. As an example, ensure EMC, Electrical, Mechanical, Chemical, Biological risks have been considered and properly retired. Ensure Cyber Security Hazards have been sufficiently covered in the Design. Refer IMDRF "Principles and Practices for Medical Device Cybersecurity '' dated 18-Mar-2020

		General Requirements	WHO Comments/proposals
		and validation studies conducted in this category (i.e. what tests were considered and why they were or were not performed) c) Discussion to support why the evidence presented is sufficient to support the application. **OR** d) A statement of why this category of study is not applicable to this case.	
3.04	Software/Firmware		
3.04.01	Software/Firmware Description	a) Specify the name of the software. b) Specify the version of the software - The version tested must be clearly identified and should match the release version of the software, otherwise justification must be provided. c) Provide a description of the software including the identification of the device features that are controlled by the software, the programming language, hardware platform, operating system (if applicable), use of Off-the-shelf software (if applicable), a description of the realisation process. Provide a statement about software version naming rules; specify all fields and their meanings. The initial distribution mode of the software shall be defined and documented as well as the mechanisms for release updates. A summary of the device's maintenance plan should be submitted that describes the post-market processes by which the manufacturer intends to ensure the continued safety and performance of the device throughout its life cycle.	Refer to IEC 62304:2006 (confirmed in 2021) Medical device software — Software life cycle processes. Ensure Software identification and version is clearly displayed. Ensure software compatibility and interoperability has been addressed with the Medical Device combinations. Ensure DICOM version compatibility is defined and documented. Ensure Software distribution (e.g. CD, DVD, Cloud) is defined and documented. (Refer note (i) below in this section). Ensure software tested for entry errors (Example: verification of patient age with intended use, values outside limits etc.) Ensure the Machine Learning algorithm has been identified and documented including use of open source elements. Ensure Software of unknown provenance (SOUP) is identified and processed as such. For example, documenting the software component version, maintaining version control and maintaining the list of residual anomalies. Ensure the Data Sets for training, verification and validation are well defined, selected, evaluated for relevance, maintained.(Refer note (ii) below in this section).Ensure the Data Set is consistent with the population of patients in all countries where the Medical Device is marketed. The training data set must not be the same as that used for verification and validation. Ensure the Medical Device performance is aligned with expected outcome in all countries where the Medical Device is marketed. Note (i): JPEG and PNG image format may be considered as output format for report but not as an input format for the CAD Software. Only DICOM, which is the open medical imaging standard, shall be considered for image inputs. DICOM is an open standard dedicated for medical images including specific medical metadata (which is not the case for JPEG and PNG).

		General Requirements	WHO Comments/proposals
			Also, it is important to use image compression which does not result in loss of critical information. This is not the case for JPEG format, the use of which is not recommended by WHO because of the potential loss of critical information. Note (ii): Refer to ""Good Machine Learning Practice for Medical Device Development: Guiding Principles October 2021" from FDA, HC & MHRA Guiding Principles are: 1. Multi-Disciplinary Expertise Is Leveraged Throughout the Total Product Life Cycle 2. Good Software Engineering and Security Practices Are Implemented 3. Clinical Study Participants and Data Sets Are Representative of the Intended Patient Population 4. Training Data Sets Are Independent of Test Sets 5. Selected Reference Datasets Are Based Upon Best Available Methods 6. Model Design Is Tailored to the Available Data and Reflects the Intended Use of the Device 7. Focus Is Placed on the Performance of the Human-AI Team 8. Testing Demonstrates Device Performance During Clinically Relevant Conditions 9. Users Are Provided Clear, Essential Information 10. Deployed Models Are Monitored for Performance and Re-training Risks Are Managed The initial distribution mode of the software shall be defined and documented as well as the mechanisms for release updates. A summary of the device's maintenance plan describing the post-market processes by which the manufacturer intends to ensure the continued safety and performance of the device throughout its life cycle.
3.04.02	Hazard Analysis	The Hazard Analysis should take into account all device hazards associated with the device's intended use, including both hardware and software hazards. **NOTE:** i. This document can be in the form of an extract of the software-related items from comprehensive risk management documentation, described in ISO 14971:2019; "Medical Devices: Application of risk management for medical devices" and in AAMI/CR34971:2022; "Guidance on the Application of ISO 14971 to Artificial Intelligence and Machine Learning."	Ensure Hazardous situations cover all specifics in the countries where the Medical Device is marketed. Ensure sources of potential bias have been evaluated and taken into consideration as necessary.

		General Requirements	WHO Comments/proposals
		ii. Hazard analysis should address all foreseeable hazards, including those resulting from intentional or inadvertent misuse of the device.	
3.04.03	Software Requirement Specification	The Software Requirements Specification (SRS) documents the requirements for the software. This typically includes functional, performance, interface, design, developmental, and other requirements for the software. In effect, this document describes what the Software Device is supposed to do. For example, hardware requirements, programming language requirement, interface requirements, performance and functional requirements, A complete SRS document should be provided. The documentation should include a description of the software requirement identification and tracking methodology used to support the traceability of the requirements.	Ensure there are specific requirements for the Medical Device including Software. Ensure Software modules have been defined and documented. Ensure Software has been classified. Ensure Software Of Unknown Provenance (SOUP) is properly identified, documented and managed. For example ensure the SRS documents the requirements for the software which typically specifies inputs and outputs, functions that the software will perform, hardware, programming language, compiler version, performance, interfaces, user interaction, error definition and handling, response times, intended operating environment, safety related requirements derived from a risk assessment and all ranges, limits, defaults, and specific values that the software will accept. Ensure each requirement (e.g.,hardware, software, user, operator interface, and safety) identified in the software requirements specification is evaluated for accuracy, completeness, consistency, testability, correctness, and clarity. The software must be capable of determining whether the input CXR image is in a supportive format e.g. DICOM.
3.04.04	Architecture Design Chart	Detailed diagrams of the modules, layers, and interfaces that comprise the device, their relationships, the data inputs/outputs and flow of data, and how users or external products (including IT infrastructure and peripherals) interact with the system and software. The purpose of the system and software architecture diagram is to present a roadmap of the device design to facilitate a clear understanding of 1) the modules and layers that make up the system and software, 2) the relationships among the modules and layers, 3) the data inputs/outputs and flow of data among the modules and layers, and 4) how users or external products, including IT infrastructure and peripherals (e.g., wirelessly connected medical devices) interact with the system and software. Define and document the distribution methodology (CD, USB Key, local/server, cloud, …) A summary of the device's maintenance plan describing the post-market processes by which the manufacturer intends to ensure the	Ensure a high-level block diagram that describes the different functions and modules of the Medical Device is properly documented. The design may be totally independent of the distribution technology. However the decisions made and associated warnings must be defined and documented. To the extent appropriate, the system and software architecture diagram can be communicated in one or more diagrams and in one or more formats and may convey different dimensions of the system and software (e.g., cybersecurity architecture, state diagram). If more than one diagram is used, the applicant should provide a high-level diagram that communicates the overview and points to the other diagrams that provide additional detail. The relationship between diagrams should also be clearly communicated. In general, the applicant should take note of the following visual, language, and reference considerations when developing an effective system and software architecture

		General Requirements	WHO Comments/proposals
		continued safety and performance of the device throughout its life cycle shall be provided.	diagram.
3.04.05	Software Design Specification	The Software Design Specification (SDS) describes the implementation of the requirements for the Software Device. The SDS describes how the requirements in the SRS are implemented.	Ensure the SDS properly covers for the ML specificity and in particular the algorithm selected and how the Data Set is maintained.
3.04.06	Traceability Analysis	A Traceability Analysis links together your product design requirements, design specifications, and testing requirements. It also provides a means of tying together identified hazards with the implementation and testing of the mitigations.	Ensure traceability is maintained, especially when using ML approaches.
3.04.07	Software Development Environment Description	A summary describing the software development life cycle and the processes that are in place to manage the various life cycle activities.	Ensure specific tools used for the design and development of the Software and especially around ML are identified and documented. Special attention should be made on the Data Set selection and control. For instance, special attention should be paid to the training data set to ensure populations (including demographics), use cases, reference standards, annotation definitions, and CXR hardware/devices used. These should be documented.
3.04.08	Software Verification and Validation	This heading should include: a) An overview of all verification, validation and testing performed prior to final release. b) For each test presented, identify the testing environment (e.g., in-house, in a simulated or actual user environment). c) Discussion to support why the evidence presented is sufficient to support the application. **OR** d) A statement of why this is not applicable to this case. **NOTE** i. Discussion should address all the different hardware configurations and, where applicable, operating systems identified in the labelling. Applicant to provide a description of the testing activities at the unit, integration, and system levels. System level test protocol including expected results, observed results, pass/fail determination, and system level test report. This should include a summary description of the testing activities at the unit, integration, and system levels.	Software verification involves evaluating the consistency, completeness, and correctness of the software and its supporting documentation, as it is being developed, and provides support for a subsequent conclusion that software is validated. Software testing is one of several verification activities intended to confirm that the software development output meets its input requirements. Other verification activities include walk-throughs, various static and dynamic analyses, code and document inspections, module level testing and integration testing. Software validation is a part of design validation of the finished device. It involves checking for proper operation of the software in its actual or simulated use environment, including integration into the final device where appropriate. Software validation is highly dependent upon comprehensive software testing and other verification tasks previously completed at each stage of the software development life cycle. Special attention should be made to ensure the software has been tested at the limits, such as demonstrating the ability of the software to identify images of insufficient quality in all or part of the image.

		General Requirements	WHO Comments/proposals
			Verification and validation activities should ensure that the whole combination of CAD, the DICOM images and any other part of the system, shall be safe and shall not impair the specified performance of the devices. Any restrictions on use applying to such combinations shall be indicated on the label and/or the instructions for use.
3.04.09	Revision Level History	Revision history log, including release version number and date. The documentation should include the history software revisions generated during product development. This typically takes the form of a line-item tabulation of the major changes to the software during the development cycle, including date, version number, a brief description of the changes in the version relative to the previous version, and an indication of the version testing was performed on, including bench testing, animal testing, and clinical testing, if applicable. The last entry in the list should be the final version to be incorporated in the released device. This entry should also include any differences between the tested version of software and the released version, along with an assessment of the potential effect of the differences on the safety and effectiveness of the device. If a manufacturer's development practices use an iterative methodology, information on any changed software requirements and how they continue to meet the system requirements or design inputs should be provided.	In the Machine Learning context, a special attention should be made regarding how change control is handled regarding Data Sets which are used to train, verify, and validate the software as a Medical Device. It is recommended that the draft FDA guidance "Marketing Submission Recommendations for a Predetermined Change Control Plan for Artificial Intelligence/Machine Learning (AI/ML)-Enabled Device Software Functions" dated 03_Apr-2023 is taken into consideration.
3.04.10	Unresolved Anomalies (Bugs or Defects)	All unresolved anomalies in the release version of the software should be summarised, along with a justification for acceptability (i.e. the problem, impact on safety and effectiveness, and any plans for correction of the problems). An anomaly is any condition that deviates from the expected behaviour based on requirements specifications, design documents, standards, or from someone's perceptions or experiences. For each unresolved anomaly, indicate the problem; the impact on device performance; and any plans or timeframes for correcting the problem (where appropriate). Each item should be annotated with an explanation of the impact of the anomaly on device safety or effectiveness, including operator usage and human factors issues. If the resolution of any unresolved anomalies will be deferred, a risk-based rationale for why each unresolved anomaly would not impact device safety or effectiveness should be provided. A list of unresolved anomalies should be communicated to end user(s) as appropriate to assist in the proper operation of the device. In all instances where it is practical to do so, any mitigations or possible	Ensure unresolved anomalies are properly documented, evaluated, and resolved if possible. Ensure Residual anomalies are part of the Instructions For Use (IFU).

		General Requirements	WHO Comments/proposals
		workarounds for unresolved anomalies should be included in the submission.	
3.04.11	Cybersecurity	Evidence to support the cybersecurity solutions should be provided here. For example, but not limited to: a) Cybersecurity vulnerabilities and risks analysis b) Cybersecurity controls measures Traceability matrix linking cybersecurity controls to the cybersecurity vulnerabilities and risks Applicants should provide the following information related to the cybersecurity of their medical device: Hazard analysis, mitigations, and design considerations pertaining to intentional and unintentional cybersecurity risks associated with the device, including: A specific list of all cybersecurity risks that were considered in the design of the device; A specific list and justification for all cybersecurity controls that were established for the device. A traceability matrix that links the actual cybersecurity controls to the cybersecurity risks that were considered; A summary describing the plan for providing validated software updates and patches as needed throughout the lifecycle of the medical device to continue to assure its safety and effectiveness. A summary describing controls that are in place to assure that the medical device software will maintain its integrity (e.g. remain free of malware) from the point of origin to the point at which that device leaves the control of the manufacturer; and Device instructions for use and product specifications related to recommended cybersecurity controls appropriate for the intended use environment (e.g. anti-virus software, use of firewall).	Ensure that the list of cybersecurity risks is consistent with the specificities of the different settings of use, in the different countries where it is marketed. See IMDRF "Principles and Practices for Medical Device Cybersecurity " dated 18-Mar-2020. Ensure cybersecurity hazards have been sufficiently covered in the design
3.04.12	Interoperability	If the device can communicate with other devices, equipment or systems (e.g. CAD software, PACS systems, digital X-ray systems, LIMS systems), evidence to support the interoperability should be provided.	Ensure these considerations have been identified, designed and tested, with particular focus around input/output format and compatibility with DICOM for example. Ensure there are documented contraindications, warnings, and precautions on the use of the exchanged information with different types of combined Medical Device for image acquisition. For example, with the intended user in mind they may include specific instructions to verify the correct configuration and operation. Ensure the need to develop different instructions per use dependent on the end user.

		General Requirements	WHO Comments/proposals
			Considering the intended interoperability scenarios verify there are specific warnings, precautions, or contraindications in device labelling, and they are included to reduce risks to acceptable levels. Ref FDA guidance "Design Considerations and Premarket Submission Recommendations for Interoperable Medical Devices" 2016
3.05	Other non-clinical Evidence	Other information, if any, that may be important to the submission but that does not fit in any of the other headings. This section is specifically intended for tests performed to ensure the safety and/or effectiveness of the device **that are not delineated in the rest of this section.** This should include a) A description of the purpose of the test, the risk/safety issue the test is addressing; the test methods and results of the test	Ensure any other non-clinical have been considered and if they have been handled properly. As an example, but not limited to how portable / mobile units have been specified, designed and tested.
SECTION 4 – CLINICAL EVIDENCE			
4.01	Clinical Evidence	Results from performance testing that demonstrate performance and safety of the device for screening and/or triage of TB in the intended use population when used in accordance with the instructions for use. The performance assessment must be based on appropriate diagnostic accuracy measures. These are considered to be receiver operator characteristic (ROC) plots, sensitivity and specificity and where appropriate predictive values and diagnostic likelihood ratios. Agreement studies, which do not use a reference standard, are not sufficient. The test dataset must contain a sufficient number of cases from important cohorts (e.g., subsets defined by clinically relevant confounders, effect modifiers, concomitant diseases, and subsets defined by image acquisition characteristics) such that the performance estimates and confidence intervals of the device for these individual subsets can be characterised for the intended use population and imaging equipment. (REF USA CFR Title 21 : § 892.2070 H PART 892 Subpart B Subpart B - Diagnostic Devices § 892.2070 Medical image analyzer.)	Clinical evidence must be generated on a data set independent of that used for training of the software. The performance of devices must be compared with the interpretation of a human reader. The results of this comparison should be reported alongside the performance of the device against the microbiological reference. The proficiency of the human readers used should also be summarized (eg, years of experience, level of training, degrees.) Sensitivity and specificity should be provided including the performance at the manufacturers recommended threshold. There must be evidence of external validation (i.e. the process of evaluating the performance of a predictive or statistical model using data that was not used in the development (or training) of the model) of CAD TB devices for each intended use (screening as well as triage). Testing data set should be representative of the target population. This should include data for different geographical settings (minimum of 2 WHO regions) and implementation sites. There should also be consideration of the population, including: ● race ● different age groups (greater than 14 years old and ensuring coverage of the spectrum of ages), ● different sexes, ● obesity,

		General Requirements	WHO Comments/proposals
			● smoking status, ● previous TB diagnosis ● low body mass index ● people living with HIV ● other comorbidities. Justification should be provided to demonstrate how the data reflects what is expected for the target population with respect to disease occurrence, characteristics, practice of medicine, and clinician competency. Clinical evidence shall demonstrate performance under conditions of clinical relevance to an LMIC setting. Although a prospective clinical trial is not essential, sufficient clinical evidence should consider all steps of data acquisition, CAD and clinical interpretation of CAD results with the local staff who would be involved. Performance evaluation should include all different modalities and use, such as online and offline functionality - including time to results.
	General requirements for accuracy evaluations	● Diagnostic accuracy (sensitivity and specificity) of CAD software shall be determined from PA or AP chest X-ray images (DICOMs) which are obtained from fully digital x-ray systems (i.e. not scans of film images) ● representative CXR images for the use-case ● Reference standard 1: Microbiologically confirmed TB i.e. liquid automated and/or solid culture with speciation culture and/ or molecular WHO-recommended rapid diagnostics for screening[1] ● Reference standard 2:suggestive clinical presentation of pulmonary TB	*Reference standards* The accuracy of CAD systems and Human Readers must be established against the microbiological reference standard as well as the clinical reference standard. Justification should be given if alternative reference standards are used, such as: ● the output from another CAD device; ● a follow-up clinical imaging examination; ● a follow-up medical examination other than imaging; or ● an interpretation by reviewing clinician(s) (i.e., clinical truther(s)), or ● a combination of the above. A clear process for defining clinical truth should be provided. *Comparator* Performance of CAD systems must be evaluated by comparison with trained radiologists. A panel should contain at least 3 radiologists.

[1] 3.1.4 Molecular WHO-recommended rapid diagnostics for screening | TB Knowledge Sharing (tbksp.org)

		General Requirements	WHO Comments/proposals
			Clinical Evidence generation & reporting There should be consideration to the principles and requirements. of ISO 14155: 2020 “Clinical investigation of medical devices for human subjects — Good clinical practice” Please consider minimal sets of testing cohorts that will create evidence that is generalisable to cover intended use in LMIC settings. Paired analysis should be conducted, where both tests are compared on the same persons, as well as a reference standard to identify true TB status of persons. The use of prospective cross-sectional study designs is preferred. However, well-designed, retrospective case-control studies that minimise bias, may be used to supplement data generated from prospective studies, where justified. Appropriately powered (≥90% power) study sample sizes should be justified by manufacturers depending upon the study design, non-inferiority margin, performance of human readers in a given context, the claimed performance of the device and the acceptable difference of CAD to human readers. Sample size considerations are presented in Annex B. For case control designs of screening and triage tests it is recommended that a minimum number of 760 cases (confirmed TB person-scans) and >1500 controls (confirmed without TB person-scans) are included. For cross sectional Triage studies (assuming a 10% prevalence) it is recommended that >3100 participants are included. Whilst, for cross sectional Screening studies (assuming a 2% prevalence) it is recommended that >15350 participants are included. Details of the recommended threshold should be provided by the manufacturer and include whether different thresholds are recommended for different groups/populations/countries etc. Clarify which clinical studies were led by the manufacturer and which clinical studies were done independent from the manufacturer. There should be consideration of ISO 14155: 2020 principles and requirements.

		General Requirements	WHO Comments/proposals
			At least one clinical study in the field and with the intended users and therefore also in LMICs should be considered. This should also report on usability aspects along the whole chain from participant recruitment, CXR acquisition, CXR processing by CAD, CXR interpretation by user, follow-on testing and final treatment decision/initiation. This does not need to be conducted in a prospective patient cohort. All relevant clinical studies that have been conducted should be reported, both by the manufacturer and those conducted independently. Study designs should minimise sources and risk of bias, taking into account relevant guidance documents, for example STARD-2015, TRIPOD, QUADAS-2 and CLAIM (Reporting guidelines | EQUATOR Network (equator-network.org)). It is also recommended that best practices of reporting guidelines specifically for artificial intelligence-centred diagnostic test accuracy studies are considered, although these are still at early stages of development and consensus is yet to be reached. (e.g."Developing a reporting guideline for artificial intelligence-centred diagnostic test accuracy studies: the STARD-AI protocol"). The assessment of the human readers for validation should ensure that the data set is clinically representative and there is consideration of performance across varied image quality (for example contrast, resolution & artefacts.). For each study please define the study cohorts including the intended user of the CAD. *Reporting of clinical data* Reporting should follow best practice, as described by CLAIM[2], STARD-2015[3] and TRIPOD[4]. When reporting on the study the qualifications and experience of any intended human reader should be documented. This should include their relevant education and qualifications, and the number of years of experience with TB.

[2] Checklist for Artificial Intelligence in Medical Imaging (CLAIM): A Guide for Authors and Reviewers | EQUATOR Network (equator-network.org)
[3] STARD 2015: An Updated List of Essential Items for Reporting Diagnostic Accuracy Studies | EQUATOR Network (equator-network.org)
[4] Transparent reporting of a multivariable prediction model for individual prognosis or diagnosis (TRIPOD): The TRIPOD statement | EQUATOR Network (equator-network.org)

		General Requirements	WHO Comments/proposals
			The sources of images must be defined & documented. Details should include: the input of images: PA or AP (DICOMs) or fully digital systems. such as but not limited to : - AP or PA (Antero Posterior or Postero Anterior) - type of source (CXR machine, DICOM level, proprietary format) Report on all steps of data acquisition, CAD and clinical interpretation of CAD results and identify all actors involved in this pathway. Please report on the following: - flow of participants, using a diagram. - clinical setting; country, location where study is done i.e. at a peripheral level or hospital level - Baseline demographic and clinical characteristics of participants (refer Section 4.01 “Clinical Evidence”) - Type and severity of pulmonary involvement in those with the target condition eg bilateral or miliary or cavitary disease - Distribution of alternative diagnoses in those without the target condition - Time interval and any clinical interventions between CXR and reference standard Test - Cross tabulation of the CAD results (or their distribution) by the results of the reference standard - Estimates of test performance such as sensitivity and specificity and the associated 95% confidence intervals.
4.02	Usability/Human Factors	Studies specifically assessing the instructions and/or device design in terms of impact of human behaviour, abilities, limitations, and other characteristics on the ability of the device to perform as intended should be included here. This should include: a) A summary of the non-clinical evidence that falls within this category b) A statement of the test environment and relation to the intended use environment c) A discussion of the non-clinical testing considered for the device and support for their selection or omission from the verification	Define usability issues that are important for the safe and effective use of CAD TB medical devices in LMICS. For example: ● The CAD TB should be able to replace the role of a radiologist. ● clinical setting infrastructure ● Localization/Internationalization: Language Support ● Accessibility: WCAG or other compliance if any ● Error Handling: How errors and handled and users provided feedback ● Training Material for Installation and Usage

		General Requirements	WHO Comments/proposals
		and validation studies conducted in this category (i.e. what tests were considered and why they were or were not performed) i. If a clinical study has been conducted that includes human factors/usability endpoints, reference to the studies and endpoints should be made, but full results do not need to be repeated. These results should be included in Chapter 4– Clinical Evidence. d) Discussion to support why the evidence presented is sufficient to support the application. **OR** e) A statement of why this category is not applicable to this case.	Manufacturers should demonstrate the software/device interfaces usability, such as ease and clarity in communicating results to all intended users. This should include consideration of user training requirements, errors in use/interpretation among users, time to result (observed), time to reporting (observed), and online vs. offline differences (observed). Studies should address the importance of CAD threshold calibration and describe the methods used to select the threshold Studies should demonstrate that user training requirements are sufficient to ensure that the demands associated with the use of the software do not exceed the expected user's capabilities Controls: Usability studies should include the identification of radiological findings that produce an alert or "raise a flag" following the recording of an additional abnormal finding, image artefacts or image of insufficient quality. Identify other aspects, such as CAD operation or patient eligibility (e.g. age) that will result in a warning flag. Study design should consider the fact that the user will not be a trained radiologist.
SECTION 5 – LABELLING AND PROMOTIONAL MATERIAL			
5.01	Product/Package Labels	Samples of the primary and secondary packaging labels. Labelling must include the following, if applicable: ● The information required on the label should be provided in a label on the device itself. If this is not practicable or appropriate (for example, for small-size devices, contact lenses, bone cement, software, etc.), some or all the information may appear on the packaging for each unit, and/or on the packaging of multiple devices. If UDI is required, it should follow the requirements of the appropriate UDI-issuing agency/entity. The UDI should be on the label and on all device packages.	Ensure packaging and labelling are designed and manufactured to ensure product preservation regardless the variable conditions in all countries the device is marketed

		General Requirements	WHO Comments/proposals
		● The label on the outside packaging should include any special handling measures or permissible environmental conditions. ● The label should contain the brand or trade name of the medical device. ● The label should be provided in a human-readable format but may be supplemented by machine-readable forms, such as radio-frequency identification (RFID) or bar codes. ● There should be only one machine-readable format on the label. If there are multiple, there should be a clear indication to anyone relying on capture/use of this format throughout distribution and use, including the provider of care, which machine-readable format to scan when and for what purpose. ● If a catalogue number is used to identify the device, the label should include this catalogue number. ● The label should contain the name and address of the manufacturer in a format that is recognizable and allows the location of the manufacturer to be established. ● If the label includes symbols and safety-related identification colours, the marking should be explained, where necessary ● The label should include an identifier, such as version, revision level or date of build release/issue. ● The label should include any warnings or precautions to be taken that need to be brought to the immediate attention of the user, and to any other person where appropriate. This information may be kept to a minimum, such as using symbols, in which case more detailed information should appear in the instructions for use.	
5.02	Package Insert/Instructions for Use	Instructions for use (IFU) should be written in terms readily understood by the intended user and, where appropriate, supplemented with drawings and diagrams near the corresponding text. Instructions for use may be provided to the user in paper or electronic format or both. They may be supplied by various means either with the medical device or separate from it. Examples of other means are information displayed on a screen incorporated into the medical device, information downloaded from the manufacturer's website, and machine-readable sources. The means chosen should be appropriate for the use environment and accessible to the anticipated user population. Any updates to the instructions for use need to be consistent across paper and electronic formats whether they are retrospective or batch specific. The instructions for use may also be made available on that website. In this situation, the medical device packaging should include a means	The instructions for use must reflect the use cases supported by the WHO GDG guidelines: *Among individuals aged 15 years and older in populations in which TB screening is recommended, computer-aided detection software programmes may be used to interpret abnormalities on chest radiographs that are suggestive of TB in place of human readers for screening and triage for TB disease. The results are expressed as abnormality scores. Its use should be limited to the interpretation of plain CXRs for pulmonary TB in individuals aged 15 years or older.* Ensure the IFU media, format, language is aligned with the different user profiles and settings where the device is planned to be used when appropriate for all countries where the device is marketed.

		General Requirements	WHO Comments/proposals
		for the user to easily access the appropriate electronic instructions for use via inclusion of a web address or other information. Where instructions for use are provided on a medium other than paper, the manufacturer should ensure the user has information on how to: view the instructions for use; access the correct version of the instructions for use; and obtain a paper version of the instructions for use. The instructions for use should contain the name or trade name of the medical device. The instructions for use should include a description of the medical device and how it is intended to be used. The instructions for use should contain the name and address of the manufacturer in a format that is recognizable and allows the location of the manufacturer to be established, together with contact information (e.g., telephone number, fax number, website, or email address) to obtain technical assistance. The instructions for use should state the medical device's intended use/purpose, including the indications for use, intended user (e.g., professional, or lay person), and intended use environment, as appropriate. This should include the following: (i) A detailed description of the patient population for which the device is indicated for use. (ii) A detailed description of the intended reading protocol. (iii) A detailed description of the intended user and user training that addresses appropriate reading protocols for the device. The instructions for use should state the performance of the medical device claimed by the manufacturer. This should include a detailed summary of the performance testing, including test methods, dataset characteristics, results, and a summary of sub-analyses on case distributions stratified by relevant confounders, such as lesion and organ characteristics, disease stages, and imaging equipment. The instructions for use should include a detailed description of the device inputs and outputs and a detailed description of compatible imaging hardware and imaging protocols. The instructions for use should include any specifications the user requires to use and maintain the device appropriately. For example, if the medical device performs any measurements, the instructions for use should include the claimed limits of accuracy. The instructions for use should include information that allows the user and/or patient to be sufficiently informed of any warnings, precautions, measures to be taken and limitations of use regarding the medical device. For instance the instructions for use should define which checks should be conducted by the user before using the software, to confirm that the appropriate 'input' files are submitted	The instructions for use should clearly identify the intended setting and user profile. The appropriateness of the instructions must be verified. The instruction manual should reflect each software distribution modality. Formats not supported by the software must be identified. The instructions should include any information relevant to the continued safe use and performance following version change. The instructions for use should include information that allows the user to be sufficiently informed of any further measures to be taken following the generation of a result.

		General Requirements	WHO Comments/proposals
		for processing in understandable language appropriate for the user. Other cautionary information should cover, where appropriate: warnings, precautions and/or measures to be taken in the event of malfunction of the medical device or changes in its functionality that may affect safety or performance; warnings, precautions and/or measures to be taken in regard to the exposure to reasonably foreseeable external influences or environmental conditions, such as magnetic fields, external electrical and electromagnetic effects, electrostatic discharge, radiation associated with diagnostic or therapeutic procedures, pressure, humidity, or temperature; warnings, precautions and/or measures to be taken in regard to the risks of interference posed by the reasonably foreseeable presence of the medical device during specific diagnostic investigations, evaluations, therapeutic treatment or use (e.g. electromagnetic interference emitted by the device affecting other equipment); and precautions related to materials incorporated into the device that are potentially carcinogenic, mutagenic or toxic, or could result in sensitization or allergic reaction for the patient or user. Discussion of warnings, precautions, and limitations must include situations in which the device may fail or may not operate at its expected performance level (e.g., poor image quality or for certain subpopulations), as applicable. The instructions for use should include any recommended quality control procedures to be taken to verify that the medical device performs as intended, including the following if applicable: The procedures for using any available controls; instructions recommending the frequency of use; the limitations of the quality control procedure; how the user should interpret the quality control procedure results, including a description of whether test results can or cannot be accepted; and the actions to be taken if there is a failure of any of the controls. The instructions for use should identify information for safety including any relevant residual risks, contraindications, and any expected and foreseeable adverse events, including information to be conveyed to the patient in this regard. The instructions for use should include the details of any preparatory treatment or handling of the medical device before it is ready for use (e.g., identification of other necessary equipment not provided with the medical device, final assembly, calibration). The instructions for use should include any requirements for special facilities, or special training, or qualifications of the user and/or third parties. The instructions for use should contain any information needed to verify that the medical device is properly installed and ready to perform safely and as intended by the manufacturer, including when	

		General Requirements	WHO Comments/proposals
		applicable: details and frequency of preventive and regular maintenance; necessary calibration information; and methods for mitigating risks encountered during cleaning, installation, calibration, or servicing. For medical devices intended for use together with other medical devices, and/or general-purpose equipment, the instructions for use should include sufficient information to identify such devices or equipment, in order to obtain a safe combination, and/or information on any known restrictions to combinations of medical devices and equipment. The instructions for use should state the date of issue or latest revision of the instructions for use and, where appropriate, an identification number. (REF IMDRF Principles of Labelling for Medical Devices and IVD Medical Devices, IMDRF/GRRP WG/N52 FINAL:2019 USA CFR Title 21 : § 892.2070 H PART 892 Subpart B Subpart B - Diagnostic Devices § 892.2070 Medical image analyzer.)	
5.03	e-labelling	The following should be provided: a) For eligible medical devices and stand-alone software, the applicant needs to identify which form of e-labelling is being used in case of e-labelling (e.g., electronic storage system or built-in system, website). b) Details of risk management in relation to e-labelling. If this is part of the overall risk management, refer to it here. c) A description of the procedure and operations on providing IFU's when requested. d) Written information for user Information on webpage where IFU and further information can be found in relevant languages. e) A description on how the requirements detailed for the website have been met. f) If a video/App is available to demonstrate how the test is to be performed and interpreted, provide a link as well as details about how it is maintained and updated throughout the life cycle of the device.	Ensure a workaround is possible if digital access through a network is not possible at all times.
5.04	Technical/Operators Manual	Labelling directed the technical users and operators of medical devices focusing on the proper use and maintenance of the device	Ensure the labelling is aligned with user profiles and user situations.
5.05	Product Brochures	All product brochures used to promote the CAD TB medical device should be submitted. No unverified claims should be made on these brochures.	Ensure all claims are verifiable and quantifiable.
5.06	Other Labelling and Promotional Material	Other information that may be important to the submission but that does not fit in any of the other headings of this section	Ensure limitations of use are clearly defined (operators, patients, situations).
SECTION 6A – QUALITY MANAGEMENT SYSTEM PROCEDURES			

		General Requirements	WHO Comments/proposals
Administrative information needed to evaluate the premarket submission related to the QMS (not device specific)			
6A.01	Product Descriptive Information	Abbreviated description of the device, operating principles, and overall manufacturing methods	
6A.02	General Manufacturing Information	a) Address and contact information for all sites where the device or its components are manufactured. b) Where applicable, addresses for all critical subcontractors, such as outsourced production, critical component or raw material production (e.g. animal tissue, drugs), and sterilisation, will need to be provided.	These are general requirements. Ensure distributors are clearly defined in all countries where the medical device is marketed.
6A.03	Quality management system procedures	High level quality management system procedures for establishing and maintaining the quality management system such as the quality manual, quality policy, quality objectives, and control of documents and records ***ISO 13485 Elements– SOPs to satisfy clause 4***	These are general requirements. Describe the policies regarding the company's application of standards and guidance used in product development and post-market activities for those countries without stringent medical device regulations (e.g, reference is made to IMDRF principles).
6A.04	Management responsibilities procedures ● Quality Policy ● Planning ● Quality objectives ● Quality objectives ● Responsibility, authority and communicatio n ● management review	Procedures that document the management commitment to the establishment and maintenance of the QMS by addressing quality policy, planning, responsibilities/authority/communication and management review. ***ISO 13485 Elements – SOPs implementing clause 5***	Ensure the roles of representative, importers and distributors are properly defined and documented for all countries where the medical device is marketed. Ensure repackaging and relabelling activities are properly established for all countries where the medical device is marketed. Establish as defined by 21 CFR 820.3
6A.05	Resource management procedures ● Provision of resources ● Human resources ● Infrastructure	Procedures that document the adequate provision of resources to implement and maintain the QMS including human resources, infrastructure and work environment. ***ISO 13485 Elements – SOPs implementing clause 6***	These are general requirements.

		General Requirements	WHO Comments/proposals
6A.06	Product realization procedures ● Planning of product realisation ● Customer-related processes ● Design and development procedures ● Purchasing procedures ● Production and service controls procedures ● Control of monitoring and measuring devices procedures	High level product realization procedures such as those addressing planning and customer related processes ***ISO 13485 Elements – SOPs implementing clause 7, including, but not limited to:*** *Design and development procedures* Procedures that document the systematic and controlled development of the device design from initiation of the project to transfer to production. a) Design Control Procedure(s) b) Design & Development Planning Procedure(s) c) Design Input Procedure(s) d) Design Output - Procedure(s) e) Design Review Procedure(s) f) Design Verification Procedure(s) g) Design Validation Procedure(s) h) Risk Analysis Procedure(s) i) Design Transfer Procedure(s) j) Design Changes Procedure(s) k) Design History File Procedure(s) ***ISO 13485 Elements – SOPs for implementing sub clause 7.3*** *Purchasing procedures* Procedures that document that purchased products/services conform to established quality and/or product specifications. ***ISO 13485 Elements – SOPs to implement sub clause 7.4*** *Production and service controls procedures* Procedures that document the production and service activities are carried out under controlled conditions. These SOPS address issues such as cleanliness of product and contamination control; installation and servicing activities; process validation; identification and traceability; etc. ***ISO 13485 Elements – SOPs implementing sub clause 7.5*** *Production and service controls procedures* Procedure that documents that monitoring and measuring equipment used in the QMS is controlled and continuously performing per the established requirements. ***ISO 13485 Element- SOPs for implementing sub clause 7.6***	These are general requirements.

		General Requirements	WHO Comments/proposals
6A.07	QMS measurement, analysis and improvement procedures ● monitoring and measurement ● control of non-conforming product ● analysis of data ● improvement	Procedures that document how monitoring, measurement, analysis and improvement to ensure the conformity of the product and QMS, and to maintain the effectiveness of the QMS. a) Explain how complaint handling ties to MDR procedures b) Explain how risk management is tied to the CAPA activities c) CAPA Subsystem Procedures d) Nonconforming Product Procedure(s) e) Complaint Handling Procedures f) Quality Audit Procedures ***ISO 13485 Element – SOPS for implementing clause 8***	These are general requirements.
6A.08	Other Quality System Procedures Information	Other information that may be important to the submission but that does not fit in any of the other headings of this section.	These are general requirements. Ensure there is a list of jurisdictional requirements for all countries where the medical device is marketed.
SECTION 6B – QUALITY MANAGEMENT SYSTEM DEVICE SPECIFIC INFORMATION			
6B.01	Quality management system information	Documentation and records specific to the subject device that results from the high-level quality management system procedures for establishing and maintaining the quality management system such as the quality manual, quality policy, quality objectives, and control of documents, noted in the previous section. ***ISO 13485 Elements – documentation specific to the subject device for the implementation of clause 4***	These are general requirements.
6B.02	Management responsibilities information	Documentation and records specific to the subject device that result from the implementation of the management responsibilities procedures noted in the previous section. ***ISO 13485 Elements – documentation specific to the subject device for the implementation of clause 5***	These are general requirements.
6B.03	Resource management information ● Provision of resources ● Human resources	Documentation and records specific to the subject device that result from the implementation of the resource management procedures noted in the previous section. ***ISO 13485 Elements – documentation specific to the subject device for the implementation of clause 6***	These are general requirements.

		General Requirements	WHO Comments/proposals
	● Infrastructure		
6B.04	Product realization information ● Planning of product realisation ● Customer-related processes ● Design and development procedures ● Purchasing procedures ● Production and service controls procedures ● Control of monitoring and measuring devices procedures	Product realisation information Documentation and records specific to the subject device that results from the implementation of the high-level product realization procedures noted in the previous section. ***ISO 13485 Elements – documentation specific to the subject device for the implementation of clause 7 including, but not limited to:*** *Design and development information* Documentation and records specific to the subject device that results from the implementation of the design and development procedures noted in the previous section. **NOTE:** 1) Design Control Information includes. a) Design Outputs - List of Essential Design Outputs b) Design Validation- Justification for use of non-production units in validation testing, if applicable 2) The source of this information is the Design and Development Records (e.g., DHF - Design History File). ***ISO 13485 Elements – documentation specific to the subject device for the implementation of sub clause 7.3*** *Purchasing information* Documentation and records specific to the subject device that results from the implementation of purchasing procedures noted in the previous section. NOTE: Include a list of suppliers of goods or services that affect product conformity with requirements (critical suppliers) and a description of how purchasing requirements are fulfilled for these suppliers. ***ISO 13485 Elements – documentation specific to the subject device for the implementation of sub clause 7.4*** *Production and service controls information* a) Detailed Manufacturing Flow Diagram b) Summary of in-process acceptance activities for subject device c) Process Validation Master Plan d) List of processes that have not been validated. e) For each process validation considered critical to the safety and effectiveness of the device: i. Protocols/Procedures for the validated process ii. Process validation report iii. The procedures for monitoring and controlling the process parameters of a validated process should be fully described. iv. State the frequency of re-validation.	Ensure all specifics concerning countries where the device is marketed are properly documented in risk management, design inputs, usability, and cybersecurity. Ensure these inputs are properly processed, verified, and validated. Ensure these inputs are properly translated into specific requirements for packaging and labelling of the device. Ensure these inputs are properly translated into specific requirements for installing and servicing the device. Ensure Medical Device description includes combination with other Medical Devices in order to fulfil the overall intended function. For example, to define the requirement for the data and image input. Ensure the different categories of users (operators, radiographer, radiologist, technician, …) and associated education and training requirements. Ensure the targeted population is clearly defined and documented.

		General Requirements	WHO Comments/proposals
		ISO 13485 Elements – documentation specific to the subject device for the implementation of sub clause 7.5 *QMS measurement, analysis and improvement information:* Documentation and records specific to the subject device that results from the implementation of the control of monitoring and measuring device procedures noted in the previous section. ***ISO 13485 Elements – documentation specific to the subject device for the implementation of sub clause 7.6***	
6B.05	QMS measurement, analysis and improvement procedures ● monitoring and measurement ● control of non-conforming product ● analysis of data ● improvement	Documentation and records specific to the subject device that results from the implementation of the QMS measurement, analysis and improvement procedures noted in the previous section. ***ISO 13485 Elements – documentation specific to the subject device for the implementation of clause 8***	Ensure specific entries concerning countries where the medical device is marketed are properly documented, evaluated and proper actions are taken to address them. Example: Review samples of complaint records from selected countries and assess actions taken to address them.
6B.06	Other Device Specific Quality Management System Information	Other information that may be important to the submission but that does not fit in any of the other headings of this section.	Identify and evaluate conformity to claim additional regulation and standard (radiation emitting device, radioprotection, DICOM, Software as a medical device, AI, …).

Annex A: Considerations for Risk Management to ensure destination countries specificities, including LMICs, are properly addressed

1. General considerations for risk management and testing inputs for CAD TB

1.1. Environmental conditions

- Temperature & humidity
- Altitude
- Dust
- Mechanical constraints
- Electrical: Power Supply, EMC
- Chemicals
- Biological agents (cleaning)
- Cyber Security

1.2. Operator - related

- Radiation Emitting Devices education (use and protection)
- Radiological education (what is a good Xray image)
- Using a dedicated Medical Device
- Clinical environment

1.3. Infrastructure - related

- Fixed vs Mobile setup
- Power supply (including availability, stability, backup, …)
- Networking
- Maintenance (preventive and curative)
- Software traceability and updates
- Cyber Security

1.4. Clinical

- Population distribution (age, morphology, …)
- Other disease prevalence

2. Artificial Intelligence (AI)/Machine Learning (ML) considerations

These considerations are based on the 10 guiding principles that can inform the development of Good Machine Learning Practice (GMLP).

1. Multi-disciplinary expertise is leveraged throughout the total product life cycle.
2. Good software engineering and security practices are implemented.
3. Clinical study participants and data sets are representative of the intended patient population.
4. Training data sets are independent of test sets.
5. Selected reference datasets are based upon best available methods.
6. Model design is tailored to the available data and reflects the intended use of the device.
7. Focus is placed on the performance of the human-AI team.
8. Testing demonstrates device performance during clinically relevant conditions.
9. Users are provided clear, essential information.
10. Deployed models are monitored for performance and re-training risks are managed.

Annex B: WHO TB CAD: Justification of sample size

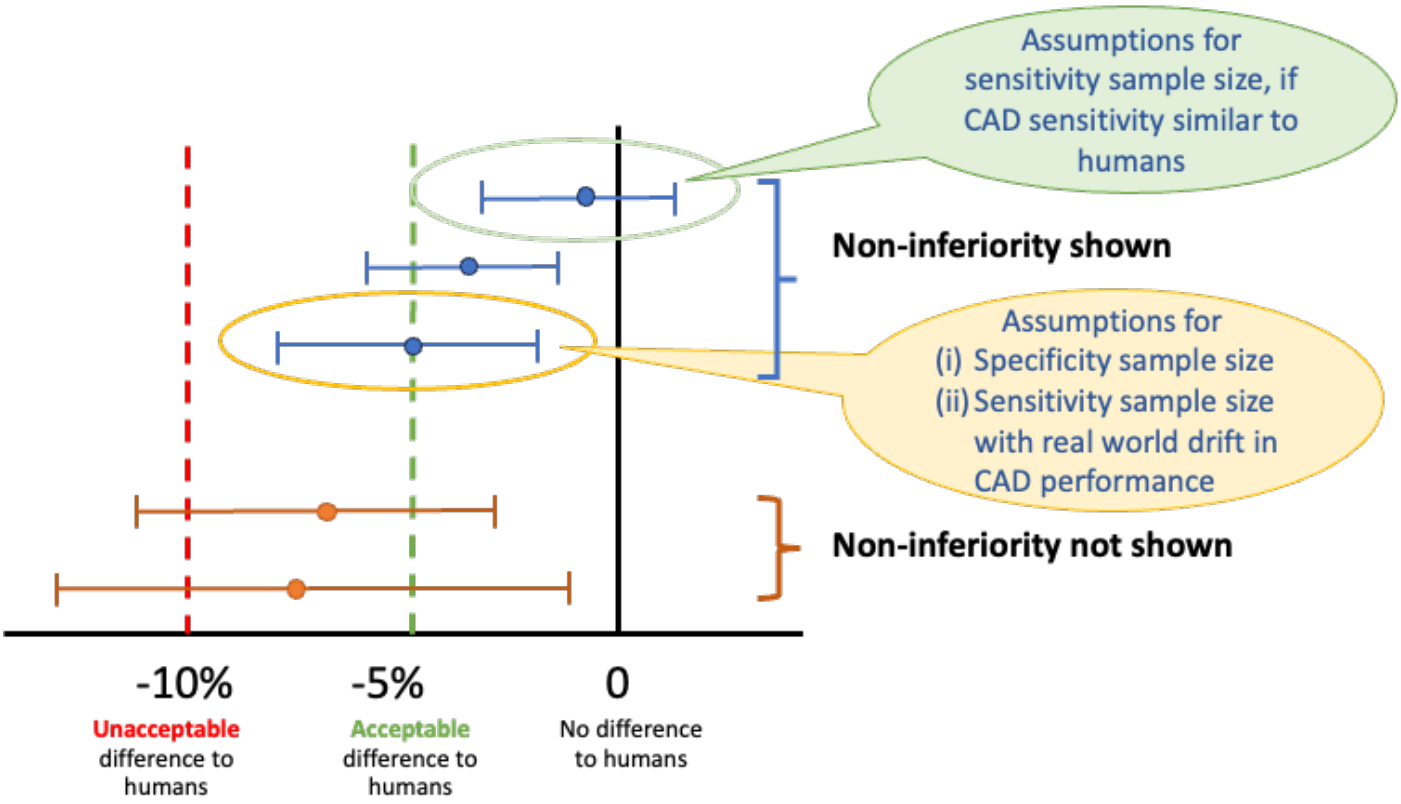
1.RESEARCH QUESTION

1.1 Main aim

What is the minimum sample size needed to compare CAD device performance to human reader

- GDG suggests that the metric should be "not inferior to the products reviewed by the GDG in 2020" by 10% to 20%
- As CAD threshold is set so the sensitivity of CAD is typically slightly superior to humans (90% sensitivity) this typically results in a lower specificity to humans. This similar sensitivity to humans is unlikely to be maintained in a real world setting where the threshold will drift due to mis-calibration.
- We have included two sample size calculations, showing sample sizes if the sensitivity threshold of CAD were maintained to be similar to humans, and in a more realistic setting where the CAD threshold drifts such that it could be 5% inferior to human interpretation.
- We note that for screening use it might be better for the CAD threshold to be set to be similar to specificity for humans, as the additional burden of further diagnostic work up in LMIC health systems resulting from false positives is often prohibitive to CAD introduction for screening.

1.2 Diagram showing non-inferiority and different assumptions for sensitivity and specificity



1.3 Additional aspects

The Technical Specification document for WHO prequalification states that is important for the test dataset to include sufficient numbers of cases for important cohorts as stated below.

This sample size calculation document sets out the basic sample size calculation corresponding to a single population subgroup required, see below. Careful considerations will be needed to understand any cross sectional population subgroups (e.g. people living with HIV) that require meeting the basic sample size.

Important population subgroups are specified in section “4.01 Clinical Evidence” of the Technical Specification document, which states:

“The test dataset must contain a sufficient number of cases from important cohorts (e.g., subsets defined by clinically relevant confounders, effect modifiers, concomitant diseases, and subsets defined by image acquisition characteristics) such that the performance estimates and confidence intervals of the device for these individual subsets can be characterised for the intended use population and imaging equipment.

There should also be consideration of the population, including:

- race
- different age groups (greater than 14 years old and ensuring coverage of the spectrum of ages),
- different sexes,
- obesity,
- smoking status,
- previous TB diagnosis
- low body mass index
- people living with HIV
- other comorbidities. “

2. STUDY DESIGN

All these sample sizes apply to studies where the same persons have both tests being compared i.e. scans are interpreted independently by both humans and CAD, as well as a reference standard to identify true TB status of persons.

Table 1: Summary of study design for sample size calculations

Aspect of design	Notes	Evidence
Study design	Paired analysis, where both tests are compared on same persons, as well as a reference standard to identify true TB status of persons.	
CXR and demographics	A previous assessment of using CAD for automated interpretation of digital CXRs for TB by WHO determined that in order to adequately assess diagnostic accuracy, it was necessary to evaluate CAD software using a standard panel of CXR files with associated demographic and clinical data, including TB diagnosis, drawn from a representative population for the corresponding use case for the technology.	P27 GDG
Readers	Typical readers of those who would conduct both human and CAD interpretation in practice should be used. There is no need for an MRMC study, as these are usually recommended only where it is difficult to obtain patient images. MRMC studies have the disadvantage that they do not typically enable a wide range of person characteristics to be included in the study.	
Independent blinded test interpretations	The two tests (human and CAD) should be interpreted blinded to results from both the other test and the reference standard	

3. STUDY DESIGN WHERE TWO TESTS ARE COMPARED ON SAME PATIENT

3.1 Screening sample size assumptions

Table 2: Assumptions for screening sample size

Parameter	Estimates used in sample size	Evidence for assumptions
Prevalence TB	<5%	P27 GDG doc
Sensitivity human readers	0.82 to 0.93	P27 GDG doc
Specificity human readers	0.14 to 0.63	P27 GDG doc
Sensitivity CAD	0.90 to 0.92	P27 GDG doc. Note the threshold for CAD was set on 90% sensitivity
Specificity CAD	0.23 to 0.66	P27 GDG doc
Unacceptable difference of CAD to human readers	10% non-inferiority 20% non-inferiority	P27 GDG doc
Acceptable difference of CAD to human readers	1% for sensitivity 5% for sensitivity (more real world) 5% for specificity	No reference in documents to minimally acceptable difference. Sensitivity threshold for CAD set at similar sensitivity to humans, so 1% inferior acceptable. There is an expectation that specificity will be inferior to humans, so a 5% acceptable reduced specificity with a non-acceptable 10% difference is plausible.
Prevalence of important population subgroups	10% 2%	For each important population subgroups (e.g. subgroups listed on p27 GDG report, plausible relevant local prevalences would need to be estimated for calculation of sample size. We exemplify sample sizes for two prevalences (2% and 10%) to provide worked examples.
non-inferiority margin with FIND/WHO	10% 20%	Discussion WHO/IVD-ACT

3.1.1 Case-control retrospective design where two tests compared in the SAME patients

- CAD performance in original studies were set at a sensitivity of 90%, as the threshold for CAD positive was set retrospectively to achieve 90% sensitivity.
- It is unclear what the sensitivity of CAD would be if the threshold was pre-specified, and so we have allowed this to range from 60% to 90%

Table 3: Sample size for screening study using case control design

Case control design	Acceptable (i.e. expected) difference	10% non-inferiority	10% non-inferiority	10% non-inferiority	10% non-inferiority	20% non-inferiority	20% non-inferiority	20% non-inferiority	20% non-inferiority
Sensitivity Human performance		60%	70%	80%	90%	60%	70%	80%	90%
Sensitivity CAD performance	1% lower sensitivity	50%	60%	70%	80%	40%	50%	60%	70%
Number of confirmed TB person-scans required*		576	470	364	257	162	138	114	90
Specificity Human performance		60%	70%	80%	90%	60%	70%	80%	90%
Specificity CAD performance	1% lower specificity	50%	60%	70%	80%	40%	50%	60%	70%
Number of confirmed without TB person-scans required*		576	470	364	257	162	138	114	90
Sensitivity Human performance		60%	70%	80%	90%	60%	70%	80%	90%
Sensitivity CAD performance	5% lower sensitivity	50%	60%	70%	80%	40%	50%	60%	70%
Number of confirmed TB person-scans required*		1788	1445	1102	758	252	213	175	136
Specificity Human performance		60%	70%	80%	90%	60%	70%	80%	90%
Specificity CAD performance	5% lower specificity	50%	60%	70%	80%	40%	50%	60%	70%
Number of confirmed person-scans without TB required		1788	1445	1102	758	252	213	175	136

*for 90% power (assuming 100% prevalence of persons with TB for sensitivity, or 100% prevalence of persons without TB for specificity)

3.1.2 Cross-sectional prospective design where two tests compared in the SAME patients

- Sample size calculations for cross sectional prospective designs take into account the uncertainty in prevalence
- CAD performance in original studies were set at a sensitivity of 90%, as the threshold for CAD positive was set retrospectively to achieve 90% sensitivity.
- It is unclear what the sensitivity of CAD would be if the threshold was pre-specified, and so we have allowed this to range from 60% to 90%

Table 4: Sample size for screening study using cross-sectional prospective design

Cross-sectional prospective design	Acceptable (i.e. expected) difference	10% non-inferiority	10% non-inferiority	10% non-inferiority	10% non-inferiority	20% non-inferiority	20% non-inferiority	20% non-inferiority	20% non-inferiority
Human performance		60%	70%	80%	90%	60%	70%	80%	90%
CAD performance	1% lower sensitivity	50%	60%	70%	80%	40%	50%	60%	70%
Total person (5% prevalence)		12,920	10,600	8,480	6,140	4,040	3,560	3,080	2,400
Total person (2% prevalence)		32,300	26,500	20,700	15,350	10,100	8,900	7,700	6,000
Total person (1% prevalence)		65,600	54,000	41,400	30,700	20,200	17,800	14,400	12,000
Human performance		60%	70%	80%	90%	60%	70%	80%	90%
CAD performance	5% lower sensitivity	50%	60%	70%	80%	40%	50%	60%	70%
Total person (5% prevalence)		37,960	30,900	24,040	16,560	6,040	5,060	4,300	3,320
Total person (2% prevalence)		94,900	77,750	59,600	41,900	15,100	13,150	10,750	8,300
Total person (1% prevalence)		190,800	155,500	119,200	84,800	30,200	25,300	21,500	16,600

3.2 Triage based on signs and symptoms

WHO GDG defines triaging as the process of deciding the diagnostic and care pathways for people based on their symptoms, signs, risk markers and test results

Table 5: Assumptions for triage sample size

Parameter	Estimates used in sample size	Evidence for assumptions
Prevalence TB	10-20%	P27 GDG
Sensitivity human readers	0.89 to 0.96	P27 GDG
Specificity human readers	0.36 to 0.63	P27 GDG
Sensitivity CAD	0.90 to 0.91	P27 GDG
Specificity CAD	0.25 to 0.79	P27 GDG
Acceptable difference of CAD to human readers	1% for sensitivity 5% for sensitivity (more real world) 5% for specificity	No reference in documents to minimally acceptable difference. Sensitivity threshold for CAD set at similar sensitivity to humans, so 1% inferior acceptable. There is an expectation that specificity will be inferior to humans, so a 5% acceptable reduced specificity with a non-acceptable 10% difference is plausible.
Prevalence of important population subgroups	10% 2%	For each important population subgroups (e.g. subgroups listed on p27 GDG report, plausible relevant local prevalences would need to be estimated for calculation of sample size. We exemplify sample sizes for two prevalences (2% and 10%) to provide worked examples.

3.2.1 Case-control retrospective design where two tests compared in the SAME patients

- CAD performance in original studies were set at a sensitivity of 90%, as the threshold for CAD positive was set retrospectively to achieve 90% sensitivity.
- It is unclear what the sensitivity of CAD would be if the threshold was pre-specified, and so we have allowed this to range from 60% to 90%

Table 6: Sample size for triage study using case control design

Case control design	Acceptable (i.e. expected) difference	10% non-inferiority	10% non-inferiority	10% non-inferiority	10% non-inferiority	20% non-inferiority	20% non-inferiority	20% non-inferiority	20% non-inferiority
Sensitivity Human performance		60%	70%	80%	90%	60%	70%	80%	90%
Sensitivity CAD performance	1% lower sensitivity	50%	60%	70%	80%	40%	50%	60%	70%
Number of confirmed TB person-scans required*		576	470	364	257	162	138	114	90
Specificity Human performance		60%	70%	80%	90%	60%	70%	80%	90%
Specificity CAD performance	1% lower specificity	50%	60%	70%	80%	40%	50%	60%	70%
Number of confirmed person-scans without TB required*		576	470	364	257	162	138	114	90
Sensitivity Human performance		60%	70%	80%	90%	60%	70%	80%	90%
Sensitivity CAD performance	5% lower sensitivity	50%	60%	70%	80%	40%	50%	60%	70%
Number of confirmed TB person-scans required*		1788	1445	1102	758	252	213	175	136
Specificity Human performance		60%	70%	80%	90%	60%	70%	80%	90%
Specificity CAD performance	5% lower specificity	50%	60%	70%	80%	40%	50%	60%	70%
Number of confirmed person-scans without TB required*		1788	1445	1102	758	252	213	175	136

*for 90% power (assuming 100% prevalence of persons with TB for sensitivity, or 100% prevalence of persons without TB for specificity)

3.2.2 Cross-sectional prospective design where two tests compared in the SAME patients

- Sample size calculations for cross sectional prospective designs take into account the uncertainty in prevalence
- CAD performance in original studies were set at a sensitivity of 90%, as the threshold for CAD positive was set retrospectively to achieve 90% sensitivity.
- It is unclear what the sensitivity of CAD would be if the threshold was pre-specified, and so we have allowed this to range from 60% to 90%

Table 7: Sample size for triage study using cross-sectional prospective design

Cross-sectional prospective design	Acceptable (i.e. expected) difference	10% non-inferiority	10% non-inferiority	10% non-inferiority	10% non-inferiority	20% non-inferiority	20% non-inferiority	20% non-inferiority	20% non-inferiority
Human performance		60%	70%	80%	90%	60%	70%	80%	90%
CAD performance	1% lower sensitivity	50%	60%	70%	80%	40%	50%	60%	70%
Total person (10% prevalence)		6,460	5,300	4,140	3,070	2,020	1,780	1,440	1,200
Total person (2% prevalence)		32,300	26,500	20,700	15,350	10,100	8,900	7,700	6,000
Human performance		60%	70%	80%	90%	60%	70%	80%	90%
CAD performance	5% lower sensitivity	50%	60%	70%	80%	40%	50%	60%	70%
Total person (10% prevalence)		19,080	15,450	11,920	8,280	2,920	2,530	2,150	1,760
Total person (2% prevalence)		94,900	77,750	59,600	41,900	15,100	13,150	10,750	8,300

4. METHODS

Sample size method (Paired test comparison based on McNemar’s test)

Lu Y, Bean JA. On the sample size for one-sided equivalence of sensitivities based upon McNemar's test. Stat Med. 1995 Aug 30;14(16):1831-9. doi: 10.1002/sim.4780141611. PMID: 7481214.

.

Annex C: Referenced documents

Software as a Medical Device

Software as a Medical Device (SaMD): Key Definitions
18 December 2013
Technical document - IMDRF/SaMD WG/N10

Software as a Medical Device: Possible Framework for Risk Categorization and Corresponding Considerations
18 September 2014
Technical document - IMDRF/SaMD WG/N12

Software as a Medical Device (SaMD): Application of Quality Management System
2 October 2015
Technical document - IMDRF/SaMD WG/N23

Software as a Medical Device (SaMD): Clinical Evaluation
21 September 2017
Technical document - IMDRF/SaMD WG/N41

Good Machine Learning Practice for Medical Device Development:
Guiding Principles
https://www.fda.gov/media/153486/download
https://www.fda.gov/medical-devices/software-medical-device-samd/artificial-intelligence-and-machine-learning-aiml-enabled-medical-devices

Machine Learning-enabled Medical Devices: Key Terms and Definitions
https://www.imdrf.org/sites/default/files/2022-05/IMDRF%20AIMD%20WG%20Final%20Document%20N67.pdf