

**UNIVERSIDADE FEDERAL DA GRANDE DOURADOS  
FACULDADE DE CIÊNCIAS DA SAÚDE  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS DA SAÚDE**

**Uso de método de interpretação automatizada de raio-x para triagem em  
massa de tuberculose em prisões**

**THIEGO RAMON SOARES**

**Dourados - MS  
2021**

THIEGO RAMON SOARES

**Uso de método de interpretação automatizada de raio-x para triagem em massa de tuberculose em prisões**

Dissertação apresentada ao Programa de Pós-Graduação em Ciências da Saúde da Faculdade de Ciências da Saúde da Universidade Federal da Grande Dourados (UFGD), para obtenção do título de Mestre em Ciências da Saúde

Área de concentração: Doenças Crônicas e Infecto-Parasitárias ou Farmacologia

Orientador: Prof. Dr. Julio Henrique Rosa Croda

Dados Internacionais de Catalogação na Publicação (CIP).

S676u Soares, Thiego Ramon

Uso de método de interpretação automatizada de raio-x para triagem em massa de tuberculose em prisões [recurso eletrônico] / Thiego Ramon Soares. -- 2021.  
Arquivo em formato pdf.

Orientador: Julio Henrique Rosa Croda.

Dissertação (Mestrado em Ciências da Saúde)-Universidade Federal da Grande Dourados, 2021.

Disponível no Repositório Institucional da UFGD em:  
<https://portal.ufgd.edu.br/setor/biblioteca/repositorio>

1. Triagem em massa. 2. Tuberculose. 3. Prisão. 4. Raio-X. 5. Interpretação automatizada. I. Croda, Julio Henrique Rosa. II. Título.

Ficha catalográfica elaborada automaticamente de acordo com os dados fornecidos pelo(a) autor(a).

©Direitos reservados. Permitido a reprodução parcial desde que citada a fonte.



**UFPGD**

**MINISTÉRIO DA EDUCAÇÃO**  
**UNIVERSIDADE FEDERAL DA GRANDE DOURADOS**  
**PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIAS**  
**DA SAÚDE**

ATA DA DEFESA DE DISSERTAÇÃO DE MESTRADO APRESENTADA POR, THIEGO RAMON SOARES DO PROGRAMA DE PÓS-GRADUAÇÃO *STRICTO SENSU* EM CIÊNCIAS DA SAÚDE, ÁREA DE CONCENTRAÇÃO “DOENÇAS CRÔNICAS E INFECTO-PARASITÁRIAS”.

Aos vinte e três dias do mês de junho de dois mil e vinte e um, às 09 horas, em sessão pública, realizou-se por videoconferência a defesa de dissertação de Mestrado intitulada “**Uso de método de interpretação automatizada de raio-X para triagem em massa de tuberculose em prisões**” apresentada pelo aluno **Thiego Ramon Soares**, do Programa de Pós-Graduação em Ciências da Saúde, à Banca Examinadora constituída pelos membros: Prof. Dr. Júlio Henrique Rosa Croda - PPGCS/UFPGD (presidente), Dr.<sup>a</sup> Andrea da Silva Santos/UFPGD e Dr. Afranio Lineu Kritski/UFRJ. Iniciados os trabalhos, a presidência deu a conhecer ao candidato e aos integrantes da Banca as normas a serem observadas na apresentação da dissertação. Após o candidato ter apresentado a sua explanação, os componentes da Banca Examinadora fizeram suas arguições. Terminada a Defesa, a Banca Examinadora, em sessão secreta, passou aos trabalhos de julgamento, tendo sido o candidato considerado APROVADO, fazendo *jus* ao título de **MESTRE EM CIÊNCIAS DA SAÚDE**. **O presidente da banca abaixo-assinado atesta que os doutores Andrea da Silva Santos e Afranio Lineu Kritski participaram de forma remota desta defesa de dissertação, conforme o § 3º do Art. 1º da Portaria RTR/UFPGD n. 200, de 16/03/2020 e a Instrução Normativa PROPP/UFPGD Nº 1, de 17/03/2020, considerando o candidato APROVADO, conforme declarações anexas.** Nada mais havendo a tratar, lavrou-se a presente ata, que vai assinada pelo presidente da Comissão Examinadora.

Dourados, 23 de junho de 2021.

Dr. Júlio Henrique Rosa Croda - PPGCS/UFPGD \_\_\_\_\_  
Dr.<sup>a</sup> Andrea da Silva Santos/UFPGD (participação remota)  
Dr. Afranio Lineu Kritski/UFRJ (participação remota)

ATA HOMOLOGADA EM: \_\_\_\_ / \_\_\_\_ / \_\_\_\_, PELA PRÓ-REITORIA DE ENSINO DE PÓS-GRADUAÇÃO E PESQUISA / UFPGD.

Pró-Reitoria de Ensino de Pós-Graduação e Pesquisa  
Assinatura e Carimbo

## **AGRADECIMENTOS**

Traçar objetivos, planos e metas sempre foi um percurso que segui seja na vida pessoal como na vida profissional. A carreira acadêmica faz parte dos meus planos de vida, onde busco realização profissional e pessoal. Buscarei sempre caminhar para executar todos meus planos e alcançar meus objetivos.

E desta forma encerro um desses objetivos, e todo fim de jornada é inevitável olhar para trás refletir e agradecer à todos que contribuíram, vibraram, ajudaram para a conclusão desta etapa. Primeiramente sou grato a Deus por toda benção e luz no meu caminho. Imensa gratidão a minha esposa Luma, pela parceria incentivo, companhia, torcida em todas as etapas sempre foi minha maior incentivadora para minhas conquistas. Toda honra a minha mãe (Nubia), meu pai (Francisco) minha irmã (Nivia), que sempre acreditaram e vibraram por mim.

Gratidão ao Professor Julio pela oportunidade de ser seu orientando, muito obrigado pelos ensinamentos e direcionamento, ao amigo Roberto e Andres que muito me ensinaram e ajudaram na conclusão desta jornada, todas as etapas eles foram primorosos na parceria.

## LISTA DE ABREVIATURAS E SÍMBOLOS

TB	Tuberculose
PPL	População Privada de Liberdade
Raio-x	RXT
<i>MTB</i>	<i>Mycobacterium tuberculosis</i>
ILTBI	Infecção Latente da tuberculose
OMS	Organização Mundial de Saúde
HIV	Vírus da Imunodeficiência Humana
AIDS	Síndrome da Imunodeficiência Adquirida
PNCT	Programa Nacional de Controle da Tuberculose
BAAR	Bacilo Álcool-Ácido-Resistente
PCR	Reação em Cadeia de Polimerase
IA	Inteligência Artificial
CAD4TB	Computer Aided Detection for Tuberculosis (Detecção de tuberculose auxiliada de computador)
LUNIT	Larning unit (Unidade de aprendizagem)
REDCap	Research Eletronic Data Capture (Captura Eletrônica de Dados de Pesquisa)
AUC	Area Under the Curve (Área sob a Curva)

## RESUMO

**Introdução:** A tuberculose ainda é um grave problema de saúde pública global, além de em 2019 ter sido a principal causa de morte, por uma única doença, dentre as doenças infecciosas. A incidência de tuberculose na população privada de liberdade é 10 vezes maior que na população em geral e, o diagnóstico precoce pode ser uma estratégia de controle nestes ambientes, considerados de alta carga da doença. O uso do raio-x do tórax (RXT), associando à inteligência artificial, pode acelerar a detecção de casos de tuberculose pulmonar durante a triagem em massa nas prisões. Nesta perspectiva, neste estudo o objetivo foi comparar quatro métodos de interpretação automatizada de raio-x para triagem em massa de tuberculose em prisões. **Métodos:** Foi realizado um estudo transversal, em três penitenciárias do estado do Mato Grosso do Sul, entre os anos de 2017 a 2019. Um questionário foi aplicado para obter dados demográficos, histórico de prisão, estilo de vida, histórico de saúde e sintomas sugestivos de TB. Em seguida foi realizado o RXT, e as imagens foram transferidas eletronicamente para análise automatizada pelo CAD4TB versões 5 e 6 e para o sistema LUNIT INSIGHT CXR2 versões Tb e MCA. É atribuído um escore de 0 a 100 para as versões do CAD4-TB, e foi considerado alterado o score  $\geq 60$ . Para o LUNIT o escore varia de 0 a 1, sendo alterado, o valor  $\geq 0.88$ . Após o RXT foi coletado uma amostra de escarro para realização de teste molecular Xpert MTB/RIF. O banco de dados foi gerenciado por meio do REDcap®. Para as análises calculou-se a sensibilidade, a especificidade e a área sob a curva-AUC, e os resultados comparados por meio do intervalo de confiança. **Resultados:** Foram analisadas 7081 imagens de RXTx e observou-se uma AUC de 0,92 para o LUNIT-TB, 0,88 para os CAD4TB v5 e v6. Identificou-se que os indivíduos com radiografias alteradas no LUNIT-TB, a sensibilidade foi de 33,4% e a especificidade de 94,8%. Além disso, o LUNIT- TB teve um desempenho mais potente ao diferenciar os indivíduos com pontuação baixa e alta quando comparado com as versões do CAD4-TB. **Conclusão:** Os resultados sugerem que o uso de RXT auxiliado com interpretação automatizada, pode ser uma estratégia eficiente de rastreamento em massa em ambientes de alto risco para tuberculose.

**Palavra-chave:** Triagem em massa. Tuberculose. Prisão. Raio-X. Interpretação automatizada. Diagnóstico

## ABSTRACT

**Introduction:** Tuberculosis remains a serious global public health problem and it is the leading isolated cause of death among infectious diseases. The incidence of tuberculosis in the population deprived of liberty is 10 times higher than in the general population, early diagnosis is a control strategy in these high-incidence environments. The use of x-ray combined with artificial intelligence can accelerate the detection of active tuberculosis cases, during mass screening in prisons. In this perspective, this study aimed to compare four automated x-ray interpretation methods for mass tuberculosis screening in prisons. **Methods:** A cross-sectional study was carried out in three prisons in the state of Mato Grosso do Sul, from which the data were collected from 2017 to 2019. Demographic data, prison history, lifestyle, health history and symptoms suggestive of tuberculosis were collected through a questionnaire. Then, the x-ray was performed, and the images were electronically transferred for automated analysis by CAD4-TB, versions 5 and 6, and to the LUNIT INSIGHT CXR2 system, versions Tb and MCA. For both CAD4-TB versions, with score range from 0 to 100, scores  $\geq 60$  was considered altered. While for LUNIT system, the score ranges from 0 to 1, being altered, the value  $\geq 0.88$ . After the x-ray, a sputum sample was collected to be processed in the Xpert MTB/RIF. All information was recorded in REDcap® and sensitivity, specificity and area under the AUC-curve were calculated, and the results was compared using the confidence interval. **Results:** 7081 x-ray images were analyzed and observed an AUC of 0.92 was observed for LUNIT-TB, 0.88 for CAD4TB v5 and v6. It was identified that 33.4% of individuals with altered radiographs in LUNIT-TB were diagnosed with tuberculosis and specificity of 94.8%. In addition, in LUNIT-TB it performed more powerfully to differentiate individuals with low and high scores when compared to the CAD4TB versions. **Conclusion:** The results suggest that the use of chest radiography supported by automated interpretation can be an efficient strategy for tuberculosis mass screening in high-risk environments.

**Keywords:** Mass screening. Tuberculosis. Prison. X-ray. Automated interpretation. Diagnosis.



## SUMÁRIO

1 INTRODUÇÃO .....	9
2 REVISÃO DE LITERATURA.....	10
2.2 Epidemiologia da tuberculose .....	11
2.3 Sistema prisional e tuberculose .....	12
2. 4 Diagnóstico da tuberculose .....	14
3 OBJETIVO .....	17
3.1 Específicos .....	17
4 REFERÊNCIAS BIBLIOGRÁFICAS .....	18
5 APÊNDICE .....	21
5.1 Artigo: Evaluation of chest X-Ray with automated interpretation algorithms for mass tuberculosis screening in prisons .....	22
6 CONCLUSÕES .....	57

## 1 INTRODUÇÃO

A tuberculose (TB), apesar de ser uma doença milenar, ainda representa um importante problema de saúde pública mundial, acometendo principalmente países de baixa e média renda (SULIS *et al.*, 2014). É uma doença ligada diretamente as condições sociais e econômicas dos indivíduos, portanto, se apresenta de forma heterogênea na população, resultando maior incidência em alguns subgrupos populacionais, como, por exemplo, na população privada de liberdade (PPL) (BASU; STUCKLER; MCKEE, 2011).

A prisão é um ambiente ideal para a transmissão do bacilo da TB, por apresentar condições ambientais favoráveis (ambientes lotados, mal ventilados com pouca incidência de luz solar), soma-se a isso condições individuais (usuários de tabaco, álcool, de drogas ilícitas, nutrição precária) e ainda, acesso limitado aos cuidados de saúde adequados (CORDS *et al.*, 2021).

No Brasil, o conselho nacional de política criminal e penitenciária, recomendou a busca ativa para TB, no momento da entrada do indivíduo na prisão, e posteriormente, uma vez ao ano, usando a baciloscopia e a radiografia de tórax (RXT). Porém, devido aos custos e logística envolvidos nessa triagem, a grande maioria das unidades prisionais não seguem integralmente essa recomendação (BRASIL, 2010).

O RXT é uma ferramenta para diagnóstico de imagem, rápida, pouco invasivo, com uma alta sensibilidade e uma baixa especificidade para lesões por TB (JANSEN; LOPES; DOMENICO CAPONE, 2019). Embora tenha resultados satisfatórios na detecção da doença, exige tempo e trabalho, pois necessita de um profissional médico especialista para interpretar e emitir laudos dos resultados destes exames, ou seja, um custo financeiro elevado, principalmente para países que vivenciam a limitação destes profissionais. Nos últimos anos discutiu-se a utilização de programas de interpretação automatizados de RXT, com a finalidade de auxiliar no diagnóstico da TB, já que esse mecanismo poderá reduzir o tempo de diagnóstico principalmente em locais com acesso restrito aos serviços de saúde, como as prisões (QIN *et al.*, 2019).

Portanto, o uso do RXT com interpretação automatizado para triagem em massa de TB, em ambiente de alta incidência, pode ser uma oportunidade de agilizar o diagnóstico e controle da doença. Este estudo teve por objetivo comparar 4 métodos de interpretação automatizada de RXT para triagem em massa de TB em prisões.

## 2 REVISÃO DE LITERATURA

### 2.1 *Mycobacterium tuberculosis* e a tuberculose

A TB, no homem, geralmente é causada pelo *Mycobacterium tuberculosis* (*Mtb*), descoberto por Roberto Koch em 1882. É uma bactéria aeróbica, imóvel, em forma de bastonete, da família *Mycobacteriaceae*. Pertence ao um grupo complexo, formado por várias espécies de micobactérias, agentes causais da TB em diferentes hospedeiros: *M. tuberculosis*, *M. africanum*, *M. canetti*, *M. microti*, *M. bovis*; *M. pinnipedi*, *M. caprae* e a cepa da vacina BCG (Bacilo de Calmette e Guérin). Importante salientar que as diferentes cepas bacterianas do complexo *Mtb* estão intimamente relacionadas, pois possuem similaridade de 99,9% no material genético, mas apresentam forte especificidade de hospedeiro (SMITH, 2003).

As principais características do *Mtb* são: o alto teor de guanina-citosina em seu DNA; o longo tempo de geração (divisão celular entre 12 à 24 horas sob condições favoráveis) e a parede celular especial, compartilhada por todas as espécies do complexo *Mtb*. O que torna essa parede especial é sua composição: peptidoglicanos altamente reticulados e ácido micólico, tornando a bactéria altamente hidrofóbica e ácido-resistente, característica essa utilizada no diagnóstico da TB. A parede celular rica em lipídios forma a base para a resistência intrínseca do *Mtb* para a maioria dos antibióticos (PALOMINO; LEAO; RITACCO, 2007).

Quando o *Mtb* entra no organismo, ele é fagocitado pelos macrófagos. No entanto, pode ocorrer a inibição da fusão fagossoma-lisossoma, que leva ao impedimento da digestão da bactéria. Os macrófagos infectados nos pulmões, atraem monócitos, linfócitos e neutrófilos inativados, que resultam na formação de lesões focais granulomatosas (SMITH, 2003).

A princípio cinco categorias são possíveis dentro da patogênese da TB: na primeira o sistema imunológico do indivíduo supera as bactérias, que são mortas e eliminadas completamente; na segunda ocorre uma infecção latente (ILTBI), na qual o *Mtb* permanece viável no organismo, sem causar manifestação clínica, havendo um equilíbrio entre o sistema imunológico e as bactérias; na terceira, tem-se a TB incipiente, quando ocorre uma infecção pelo *Mtb* e progride para doença ativa, mas não produz sintomas clínicos, anormalidade radiológicas ou evidências microbiológicas; na quarta, tem-se a TB subclínica, da qual não apresenta sintomas, porém se tem anormalidades radiológicas e evidências microbiológicas; na quinta tem se TB ativa, nesta conjuntura ocorre sintomas clínicos e evidências microbiológica e radiológicas. Em 90-95% dos casos as bactérias permanecem em estado latente e nunca se

reativam. Nos 5-10% remanescentes a reativação das bactérias ocorre depois de alguns anos, levando a doença clínica, chamada TB secundária (DRAIN *et al.*, 2018; PAI, 2015).

A transmissão da TB acontece por meio de indivíduos infectados pelo *Mtb*, apresentando TB pulmonar ou laríngea, que expõem aerossóis durante a fala e, principalmente, no momento da tosse. As gotículas maiores caem no chão e em um rápido processo de evaporação, permanecem suspensas no ar até serem inaladas ou carregadas para fora do ambiente, portanto, um ambiente mal ventilado favorece que as partículas fiquem suspensas no ar por um período significativo, favorecendo a transmissão entre os contatos. O sintoma mais comum da TB é a tosse produtiva e persistente por várias semanas, servindo, assim, como um indicador para triagem da doença em países endêmicos. É frequentemente acompanhada por outros sintomas inespecíficos como fadiga, febre, suor noturno e, posteriormente, perda de peso, dor no peito e dificuldade respiratória (WORLD HEALTH ORGANIZATION, 2018; YATES *et al.*, 2016).

O tratamento da TB é realizado com uma associação de fármacos, que tem como objetivo quebrar a cadeia de transmissão e prevenir a seleção de cepas naturalmente resistente. No Brasil é realizado de acordo com as recomendações do Ministério da Saúde e compreende duas fases; uma intensiva de 2 meses com o esquema RHZE (rifampicina, isoniazida, pirazinamida, etambutol) sob a forma de dose fixa (DFC) combinada, cada comprimido contém 150 mg de rifampicina, 75 mg de isoniazida, 400 mg de pirazinamida e 275 mg de etambutol; seguido por uma fase de manutenção de 4 meses com esquema RH (rifampicina, isoniazida) em DFC, sendo utilizado para todas as formas da doença em pessoas acima de 10 anos. A exceção é nos casos de pacientes com TB meníngea e osteoarticular, os quais na fase de manutenção, são tratados por 10 meses e com a associação de um corticosteroide oral (prednisona, na dose 1-2 mg/kg/dia por 4 semanas) ou um corticosteroide endovenoso (dexametasona, na dose de 0,3-0,4 mg/kg/dia, por 4-8 semanas) (BRASIL, 2019; RABAHI *et al.*, 2017).

## **2.2 Epidemiologia da tuberculose**

A TB é uma doença que acomete a humanidade há muito tempo. Achados demonstram existência de micobactérias em humanos há 9.000 anos, identificados em dois corpos em Haifa, uma cidade no norte de Israel. Em meados dos séculos XVIII e XIX a TB surge de forma epidêmica na Europa e na América, na Guiana Britânica a incidência aumentou de 240/100 mil habitantes em 1873 para 391/100mil habitantes em 1893 e, inicialmente, este aumento foi

explicado pela falta de imunidade hereditária dos povos colonizados (HERSHKOVITZ *et al.*, 2008; MEDCALF *et al.*, 2013).

Estima-se que 10 milhões de pessoas adoeceram de TB no mundo em 2019. No entanto os casos de TB estão diminuindo globalmente na maioria dos países, mas não o suficiente para alcançar as metas da estratégia “Stop TB” da Organização Mundial de Saúde (OMS), que era obter uma queda na incidência de 20% de 2015 até 2020. A redução geral neste período foi de 9% (de 140 para 130 novos casos por 100 mil habitantes), incluindo uma redução de 2,3% de 2018 para 2019. A Região Europeia teve bons avanços e quase atingiu a meta, obtendo uma redução de 19% na taxa de incidência de TB entre 2015 e 2019, e na África a redução da incidência foi de 16% (WORLD HEALTH ORGANIZATION, 2020).

A OMS estimou 289.000 casos novos de TB para a região das Américas em 2018 (3% dos casos de TB do mundo), representando uma incidência de 28,7 por 100.000 habitantes. Entre 2017 e 2018 o número estimado de casos de TB nas Américas, aumentou 2,5% (aproximadamente 8000 novos casos de TB), isso foi atribuído ao aumento da estimativa para o Brasil (4.000), Peru (2.000), México (1.000) e Venezuela (1.000) (PAN AMERICAN HEALTH ORGANIZATION, 2020).

No Brasil, em 2020 foram diagnosticados 66.819 casos novos de TB, que correspondem a uma incidência de 31,6 por 100.000 habitantes, com um coeficiente de mortalidade de 2,2 óbitos por 100.000 habitantes. No Mato Grosso do Sul foram notificados 900 casos novos de TB, com uma incidência de 32 por 100.000 habitantes e mortalidade de 1,8 por 100.000 habitantes (BRASIL, 2021).

A TB está estritamente ligada a condição social e estilo de vida da população, e um difícil obstáculo para o controle da TB no Brasil e em outros países de média e baixa renda, pode estar relacionado as populações vulneráveis como os indígenas, moradores de rua, pessoas vivendo com o vírus da imunodeficiência humana/ síndrome da imunodeficiência adquirida (HIV/AIDS) e a PPL, segmentos sociais que funcionam como reservatórios ou amplificadores para a transmissão da doença (BASU; STUCKLER; MCKEE, 2011).

### **2.3 Sistema prisional e tuberculose**

No Brasil existem 748.009 mil pessoas presas, o que torna o país o terceiro em termos de população prisional no mundo, mantendo uma taxa de aprisionamento de 359,40 pessoas presas por 100 mil habitantes, representando um aumento de 150% em todo o país. O Mato

Grosso do Sul tem terceira maior taxa de aprisionamento do país, estando com 641,74 por 100 mil habitantes (BRASIL, 2019b).

Nos últimos anos, a prevalência de TB na PPL foi 26,4 vezes maior do que a população geral. Estudo realizado por (URREGO *et al.*, 2015) em prisões brasileiras, mostrou que o sistema prisional apresenta um ambiente ideal para a disseminação da doença, devido às condições de superlotação, pouca ventilação e baixa incidência de luz solar no interior das celas. Tendo em vista essas características, o Programa Nacional de Controle da Tuberculose (PNCT) categorizou a PPL como população mais vulnerável à infecção por TB, assim como os indivíduos em situação de rua, indígenas e pessoas vivendo com HIV (BRASIL, 2011; MABUD *et al.*, 2019).

Uma meta-análise, recentemente publicada, avaliou a incidência global da infecção por *MTB*, além da prevalência e incidência da TB nas prisões. Os autores identificaram que a incidência de TB foi consideravelmente mais alta nas prisões do que na população geral, com uma taxa de incidência de 26,9% na América do Sul. Além disso, foi identificada uma incidência combinada (infecção de *MTB* e TB ativa) de 40 casos por 100.000 pessoas-ano, nos países com incidência de TB de 0-9 casos por 100.000 pessoas-ano, já nos países com incidência de TB de 250 casos por 100.000 pessoas-ano, a incidência combinada entre prisioneiros foi de 2.090 casos por 100.000 pessoas-ano (CORDS *et al.*, 2021).

Estudo realizado no Brasil, especificamente no estado do Mato Grosso do Sul, observou que a taxa de incidência da TB aumentou de 111 por 100.000 pessoas, por ano no momento da prisão, para 1.303 por 100.000 pessoas por ano, em 5,2 anos, evidenciando que o indivíduo entra na prisão com baixo risco de TB e que este aumenta rapidamente com o passar dos anos. Na saída da prisão, a incidência é reduzida de 229 casos por 100.000 pessoas ao ano para 42 por 100.000 em 7 anos de avaliação. Esta alta incidência de TB no ambiente prisional representa uma ameaça para o controle da TB na população em geral (MABUD *et al.*, 2019).

Em um estudo caso-controle associado a métodos moleculares, (SACCHI *et al.*, 2015), examinaram o papel da prisão na carga de TB na população de Dourados. Entre 240 casos de TB notificados ao longo de um período de 46 meses, 180 (75%) ocorreram na população geral e 60 (25%) na população prisional, mostrando um risco (incidência) 40 vezes maior entre os presos de adquirir TB. Entre os casos de TB da população geral, 23% ocorreram entre pessoas que estiveram previamente encarceradas. Encarceramento anterior foi o principal fator de risco para a TB na comunidade (OR ajustada, 28,5). Entre 97 isolados para o qual realizou-se a genotipagem por polimorfismo de fragmentos de restrição (RFLP IS6110), 79 poderiam ser classificados em 17 grupos. Dos quais 10 destes 17 conjuntos, compreenderam 82% dos

isolados em cluster, dois prisioneiros e membros da comunidade, incluindo os ex detentos. Esses resultados demonstram um elevado grau de ligação entre as linhagens nas prisões e na comunidade; e indicam que a abordagem da transmissão da TB nos presídios será fundamental para o controle da doença em meio a população geral.

## 2. 4 Diagnóstico da tuberculose

O diagnóstico rápido e preciso, associado a oferta de tratamento de acordo com as recomendações internacionais, impede morte e reduz os danos de saúde ocasionados pela TB na população. A OMS estipulou diagnosticar e tratar 40 milhões de pessoas, entre 2018-2022. E estabeleceu como meta, uma repartição anual de 7 milhões de casos diagnosticados em 2018, e 8 milhões de novas notificações para os anos subseqüente, para que desta forma alcance o objetivo proposto. Todo esforço se faz necessário, para buscar novas estratégias de diagnóstico, pois uma das lacunas para o controle da TB é a subnotificação dos casos (WORLD HEALTH ORGANIZATION, 2020).

Além do mais, é necessário o aumento dos casos diagnosticados bacteriologicamente, pois, dessa forma, mais pessoas seriam corretamente diagnosticadas e iniciariam um tratamento eficaz mais precoce possível. Em 2019, 7,1 milhões de casos novos e recidivas foram notificados, e 57% destes foram diagnosticados bacteriologicamente; observa-se um discreto aumento quando comparado a 2018, que teve 55% dos casos confirmado por exames bacteriológicos (WORLD HEALTH ORGANIZATION, 2020).

A TB pulmonar é a manifestação mais comum do *Mtb* em todo o mundo e a prioridade de diagnóstico da doença. Para a detecção desses casos, os métodos a serem utilizados necessitarão de especificidade, valor preditivo e sensibilidade elevada. Usualmente, os exames utilizados para o diagnóstico são: bacteriológicos, moleculares, bioquímicos, imunológicos, histopatológicos e radiológicos (PROCÓPIO, 2014).

Um dos métodos utilizado no serviço de saúde para diagnóstico é o bacteriológico, que consiste na identificação do Bacilo Álcool-Ácido-Resistente (BAAR), uma técnica desenvolvida em meados de 1880, e ainda bastante usada, por sua simplicidade, rapidez e baixo custo, com um preço de 1 a 6 dólares por amostra. No entanto apresenta uma sensibilidade de 20 a 80%, sendo necessário em média 10.000 bacilos/ml, e ainda não diferencia o *Mtb*, ou seja, qualquer bacilo do complexo *Mtb* é reagente na microscopia de escarro. Vale ressaltar que no

Brasil, 71,9% dos casos novos (58106) diagnosticados em 2020, foi por meio de método farmacológico (BRASIL, 2021; HANRAHAN; SHAH, 2014; PROCÓPIO, 2014).

A cultura do escarro se mantém como padrão ouro para diagnóstico de TB, a qual necessita de uma baixa carga bacilar e ainda é sensível para a identificação de uma cepa multidroga resistente, porém demanda em média de 2 a 6 semanas para seu resultado, um tempo considerável para se ter um diagnóstico final, e conseqüentemente, ocorre o atraso no início do tratamento. Apresenta um custo variando de 12 a 35 dólares por teste, e ainda é necessária uma rede laboratorial de classe nível 3 (NB-3, classe risco III), que é adequado para trabalho com microrganismo de alta infectividade (HANRAHAN; SHAH, 2014; LE PALUD *et al.*, 2014; PAI; SCHITO, 2015).

Testes moleculares realizados por reação em cadeia da polimerase (PCR) já estão disponíveis comercialmente, eles aliam um diagnóstico rápido à detecção de resistência a medicamentos por meio de amostra clínica, um desses é o Xpert<sup>®</sup> MTB/RIF. Trata-se de um teste molecular, restritamente automatizado, que fornece um diagnóstico em até 2h, que além de fornecer o diagnóstico para TB, verifica também a resistência à rifampicina. Quando comparado com a baciloscopia apresenta uma sensibilidade de 88% e especificidade de 98%. O uso do Xpert está indicado pela OMS desde 2010, em decorrência de seu bom desempenho e rapidez, possibilitando um início de tratamento precoce e melhor expectativa na quebra da cadeia de transmissão, com o tratamento adequado (HANRAHAN; SHAH, 2014).

O Xpert tem grande sensibilidade para paciente com alta carga bacilífera, sendo pouco sensível para indivíduos paucibacilares, portanto, nestes casos, a cultura torna-se necessário, uma outra limitação do Xpert é a resistência testada apenas para rifampicina. Estudos mostram que seu uso é eficaz para diagnóstico de TB extrapulmonar, incluindo TB meníngea (SHEN *et al.*, 2019).

A radiografia de tórax é utilizada em vários contextos para diagnóstico médico, trata-se de uma ferramenta de imagem rápida e precisa para identificação de inúmeras anormalidades pulmonares, como, por exemplo, TB pulmonar. É um exame relativamente barato, pouco invasivo, marcando uma sensibilidade e especificidade de 87% e 89% para lesões por TB, embora tenha resultados satisfatórios na detecção, exige tempo e trabalho, pois necessita de um profissional especialista para interpretação e laudo do exame, um custo financeiro elevado, principalmente para países onde existem limitações destes profissionais médicos (AYAZ; SHAUKAT; RAJA, 2021; HWANG *et al.*, 2019; VAN'T HOOG *et al.*, 2012).

Nos últimos anos, vem se discutindo o uso de sistemas de interpretação automatizada (inteligência artificial) de raio-x para diagnóstico de TB, que leva em consideração sua



capacidade de ofertar rapidez no diagnóstico, um custo relativamente baixo, e suprir a escassez de radiologista, principalmente em países de baixa e média renda (KHAN et al., 2020; RAHMAN *et al.*, 2017)

MELENDEZ e colaboradores., (2016), realizaram um estudo comparando, por meio de curva ROC, o uso de programa de interpretação automatizado de RXT verso laudo emitido por um profissional médico especializado, evidenciando resultados significativos, principalmente na localização da lesão, mostrando que um algoritmo automatizado pode superar médicos na detecção de lesão de TB.

O uso de programa automatizado tem um grande potencial de auxiliar no diagnóstico para TB, principalmente em locais onde a disponibilidade de especialistas é limitada. Recentemente, foi realizado um estudo realizado onde foram comparando 3 programas automatizados de leitura de RXT e observou-se que todos tiveram melhor resultado para detecção de lesão de TB, quando comparados aos laudos de médicos especialistas. (QIN *et al.*, 2019),

Em estudo realizado com pacientes no Paquistão, avaliou-se o desempenho do software CAD4-TB como instrumento de triagem na detecção de TB pulmonar, para isso foi utilizado teste rápido molecular (Xpert) como teste de referência, no qual mostrou capacidade de identificar corretamente uma alta proporção de pessoas que foram diagnosticadas com TB no Xpert, portanto teria potencial de reduzir o número de testes moleculares realizados (ZAIDI *et al.*, 2018),

Nesta perspectiva observa-se o quanto é importante um modelo confiável para triagem da TB. Santos e colaboradores (2020) avaliaram algoritmos para triagem em massa na prisão, e seus respectivos custos, para isso, utilizaram RXT, sintomas preconizados pelo OMS e o Xpert, dos quais obtiveram um custo variando entre US\$249 a US\$395. O RXT com uso de interpretação automatizada pelo CAD4-TB teve um custo calculado em US\$ 370 e um rendimento de 64%, a estratégia com menor custo e maior rendimento foi o uso do Xpert para todos os participantes capaz de produzir escarro, que obteve um rendimento de 74% ao custo de US\$ 249 (SANTOS et al., 2020).

Cabe ressaltar a agilidade e praticidade da utilização do RXT como estratégia de triagem na TB, com o avanço da leitura por meio de inteligência artificial torna essa ferramentacada vez mais promissora, permitindo escolhas claras e consistentes, diminuindo a proporção de pessoas encaminhadas para testes bacteriológicos, neste sentido um estudo recente evidenciou uma variação de 7,1% a 24% na redução de indivíduos encaminhados para realizar testes bacteriológico (FRASCELLA et al., 2020)

### **3 OBJETIVO**

Avaliar a acurácia de quatro programas baseados em inteligência artificial, para diagnóstico de imagem por RXT, na triagem em massa de TB em prisões.

#### **3.1 Específicos**

Comparar a acurácia de quatro métodos de interpretação automatizada de RXT para triagem em massa de TB na prisão.

Avaliar a aplicabilidade de inteligência artificial, no diagnóstico por RXT da TB.

#### 4 REFERÊNCIAS BIBLIOGRÁFICAS

AYAZ, M.; SHAUKAT, F.; RAJA, G. Ensemble learning based automatic detection of tuberculosis in chest X-ray images using hybrid feature descriptors. **Physical and Engineering Sciences in Medicine**, v. 44, n. 1, p. 183–194, 1 mar. 2021.

BASU, S.; STUCKLER, D.; MCKEE, M. Addressing Institutional Amplifiers in the Dynamics and Control of Tuberculosis Epidemics. **The American Journal of Tropical Medicine and Hygiene**, v. 84, n. 1, p. 30–37, 5 jan. 2011.

BRASIL. **Legislação em saúde no sistema penitenciário**. Brasília, DF: Ministério da Saúde, 2010.

BRASIL. **Manual de recomendações para o controle da tuberculose no Brasil**. 1a edição ed. Brasília, DF: Ministério da Saúde, Secretaria de Vigilância em Saúde, Departamento de Vigilância Epidemiológica, 2011.

BRASIL. **Manual de Recomendações para o Controle da Tuberculose no Brasil**. 2. ed. Brasília, DF: Ministério da saúde. Secretaria de Vigilância em Saúde Departamento de Vigilância das Doenças Transmissíveis, 2019a.

BRASIL. **Levantamento Nacional de Informações Penitenciária- INFOPEN**. Disponível em: <<https://app.powerbi.com/view?r=eyJrIjoiMmU4ODAwNTAtY2IyMS00OWJiLWE3ZTgtZGNjY2ZhNTYzZDliIiwidCI6ImViMDkwNDIwLTQ0NGMtNDNmNy05MWYyLTRiOGRhNmJmZThlMSJ9>>. Acesso em: 19 abr. 2021b.

BRASIL. **Boletim Epidemiológico Tuberculose/2021**. Brasília, DF: Ministério da Saúde, mar. 2021. Disponível em: <[https://www.gov.br/saude/pt-br/media/pdf/2021/marco/26/boletim\\_tuberculose\\_2021\\_24\\_03.pdf](https://www.gov.br/saude/pt-br/media/pdf/2021/marco/26/boletim_tuberculose_2021_24_03.pdf)>. Acesso em: 2 maio. 2021.

CORDS, O. et al. Incidence and prevalence of tuberculosis in incarcerated populations: a systematic review and meta-analysis. **The Lancet Public Health**, 22 mar. 2021.

DRAIN, P. K. et al. Incipient and Subclinical Tuberculosis: a Clinical Review of Early Stages and Progression of Infection. **Clinical Microbiology Reviews**, v. 31, n. 4, 18 jul. 2018.

FRASCELLA, B. et al. Subclinical tuberculosis disease - a review and analysis of prevalence surveys to inform definitions, burden, associations and screening methodology. **Clinical infectious diseases : an official publication of the Infectious Diseases Society of America**, 16 set. 2020.

HANRAHAN, C. F.; SHAH, M. Economic challenges associated with tuberculosis diagnostic development. **Expert review of pharmacoeconomics & outcomes research**, v. 14, n. 4, p. 499–510, ago. 2014.

HERSHKOVITZ, I. et al. Detection and Molecular Characterization of 9000-Year-Old Mycobacterium tuberculosis from a Neolithic Settlement in the Eastern Mediterranean. **PLOS ONE**, v. 3, n. 10, p. e3426, 15 out. 2008.

HWANG, E. J. et al. Development and Validation of a Deep Learning–based Automatic Detection Algorithm for Active Pulmonary Tuberculosis on Chest Radiographs. **Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America**, v. 69, n. 5, p. 739–747, 1 set. 2019.

JANSEN, J. M.; LOPES, A. J.; DOMENICO CAPONE. Diagnóstico por imagem da tuberculose pulmonar. p. 9, 2019.

KHAN, F. A. et al. Chest x-ray analysis with deep learning-based software as a triage test for pulmonary tuberculosis: a prospective study of diagnostic accuracy for culture-confirmed disease. **The Lancet. Digital Health**, v. 2, n. 11, p. e573–e581, nov. 2020.

LE PALUD, P. et al. Retrospective observational study of diagnostic accuracy of the Xpert® MTB/RIF assay on fiberoptic bronchoscopy sampling for early diagnosis of smear-negative or sputum-scarce patients with suspected tuberculosis. **BMC pulmonary medicine**, v. 14, p. 137, 12 ago. 2014.

MABUD, T. S. et al. Evaluating strategies for control of tuberculosis in prisons and prevention of spillover into communities: An observational and modeling study from Brazil. **PLOS Medicine**, v. 16, n. 1, p. e1002737, 24 jan. 2019.

MEDCALF, A. et al. **Tuberculosis: a short history = Tuberculose : uma breve história**. York]; Hyderabad: University of York, Centre for Chronic Diseases and Disorders ; Orient Blackswan, 2013.

MELLENDEZ, J. et al. An automated tuberculosis screening strategy combining X-ray-based computer-aided detection and clinical information. **Scientific Reports**, v. 6, n. 1, p. 25265, 29 abr. 2016.

PAI, M. Innovations in Tuberculosis Diagnostics: Progress and Translational Challenges. **EBioMedicine**, v. 2, n. 3, p. 182–183, 31 jan. 2015.

PAI, M.; SCHITO, M. Tuberculosis diagnostics in 2015: landscape, priorities, needs, and prospects. **The Journal of Infectious Diseases**, v. 211 Suppl 2, p. S21-28, 1 abr. 2015.

PALOMINO, J. C.; LEAO, S. C.; RITACCO, V. Tuberculosis 2007; from basic science to patient care. 2007.

PAN AMERICAN HEALTH ORGANIZATION. **Tuberculosis in the Americas. 2019 Regional Report**. [s.l.] Pan American Health Organization, 2020.

PROCÓPIO, M. J. **Controle da tuberculose: uma proposta de integração ensino-serviço**. 7. ed. [s.l.] Editora FIOCRUZ, 2014.

QIN, Z. Z. et al. Using artificial intelligence to read chest radiographs for tuberculosis detection: A multi-site evaluation of the diagnostic accuracy of three deep learning systems. **Scientific Reports**, v. 9, n. 1, p. 15000, 18 out. 2019.

RABAHI, M. F. et al. Tratamento da tuberculose. **Jornal Brasileiro de Pneumologia**, v. 43, n. 6, p. 472–486, dez. 2017.

- RAHMAN, M. T. et al. An evaluation of automated chest radiography reading software for tuberculosis screening among public- and private-sector patients. **The European Respiratory Journal**, v. 49, n. 5, maio 2017.
- SACCHI, F. P. C. et al. Prisons as Reservoir for Community Transmission of Tuberculosis, Brazil. **Emerging Infectious Diseases**, v. 21, n. 3, p. 452–455, mar. 2015.
- SANTOS, A. DA S. et al. Yield, Efficiency and Costs of Mass Screening Algorithms for Tuberculosis in Brazilian Prisons. **Clinical Infectious Diseases: An Official Publication of the Infectious Diseases Society of America**, 17 fev. 2020.
- SHEN, Y. et al. Diagnostic accuracy of the Xpert MTB/RIF assay for bone and joint tuberculosis: A meta-analysis. **PloS One**, v. 14, n. 8, p. e0221427, 2019.
- SMITH, I. Mycobacterium tuberculosis Pathogenesis and Molecular Determinants of Virulence. **Clinical Microbiology Reviews**, v. 16, n. 3, p. 463–496, 1 jul. 2003.
- SULIS, G. et al. Tuberculosis: Epidemiology and Control. **Mediterranean Journal of Hematology and Infectious Diseases**, v. 6, n. 1, 1 nov. 2014.
- URREGO, J. et al. The Impact of Ventilation and Early Diagnosis on Tuberculosis Transmission in Brazilian Prisons. **The American Journal of Tropical Medicine and Hygiene**, v. 93, n. 4, p. 739–746, 7 out. 2015.
- VAN'T HOOG, A. H. et al. Screening strategies for tuberculosis prevalence surveys: the value of chest radiography and symptoms. **PloS One**, v. 7, n. 7, p. e38691, 2012.
- WORLD HEALTH ORGANIZATION. **Global tuberculosis report 2018**. Geneva: World Health Organization, 2018.
- WORLD HEALTH ORGANIZATION. **Global tuberculosis report 2020**. Geneva: World Health Organization, 2020. Disponível em: <<https://www.who.int/publications-detail-redirect/9789240013131>>. Acesso em: 16 nov. 2020.
- YATES, T. A. et al. The transmission of Mycobacterium tuberculosis in high burden settings. **The Lancet Infectious Diseases**, v. 16, n. 2, p. 227–238, 1 fev. 2016.
- ZAIDI, S. M. A. et al. Evaluation of the diagnostic accuracy of Computer-Aided Detection of tuberculosis on Chest radiography among private sector patients in Pakistan. **Scientific Reports**, v. 8, n. 1, p. 12339, 17 ago. 2018.

## 5 APÊNDICE

## **5.1 Artigo: Evaluation of chest X-Ray with automated interpretation algorithms for mass tuberculosis screening in prisons**

### **Authors**

Thiego Ramon Soares<sup>1</sup>

Roberto Dias de Oliveira<sup>2</sup>

Yiran E. Liu<sup>3</sup>

Andrea da Silva Santos<sup>1</sup>

Paulo Cesar Pereira dos Santos<sup>1</sup>

Luma Ravena Soares Monte<sup>2</sup>

Lissandra Maia de Oliveira<sup>4</sup>

Chang Min Park<sup>5</sup>

Eui Jin Hwang<sup>5,6</sup>

Jason R. Andrews<sup>3\*</sup>

Julio Croda<sup>1,4,7,\*</sup>

\*Contributed equally

### **Affiliations**

<sup>1</sup>Faculty of Health Sciences of Federal University of Grande Dourados, Dourados, Mato Grosso do Sul, Brazil

<sup>2</sup>Nursing School, State University of Mato Grosso do Sul, Dourados, MS, Brazil

<sup>3</sup>Stanford University School of Medicine, Division of Infectious Diseases and Geographic Medicine, Stanford, CA, United States of America

<sup>4</sup>Oswaldo Cruz Foundation, Mato Grosso do Sul, Campo Grande, MS, Brazil

<sup>5</sup>Department of Radiology, Seoul National University College of Medicine, Seoul, Korea

<sup>6</sup>Department of Radiology, Seoul National University Hospital, Seoul, Korea

<sup>7</sup>*Department of Epidemiology of Microbial Diseases, Yale University School of Public Health, New*

Haven, United States of America

**Corresponding Author:** J. Croda, Oswaldo Cruz Foundation - Mato Grosso do Sul, Campo Grande-MS, Zip code 79074-460, Brazil. Phone: +55 67981229959 e-mail: [juliocroda@gmail.com](mailto:juliocroda@gmail.com)

**Author's contributions**

TRS, RDO, ASS, OCOS, and LRSM: acquisition, analysis, and interpretation of data for the work; drafting the work; final approval of the version to be published

YEL: drafting the work; revising it critically for important intellectual content; final approval of the version to be published

JRA: performed data analyses

CNP and EJH: final approval of the version to be published

JRA and JC: conception and design of the work; drafting the work; final approval of the version to be published

**Sources of support:** This study was supported by the US National Institutes of Health (grant numbers R01 AI130058 and R01 AI149620) and the State Secretary of Health of Mato Grosso do Sul.

**Running Head:** Chest x-ray automated interpretation for TB screening

**Descriptor number:** 11.1 Diagnosis of Tuberculosis or Latent Infection

**Word Count:** 3,086

**Tables:** 2

**Figures:** 4



## **At a Glance Commentary**

### *Scientific Knowledge on the Subject*

Automated interpretation of chest radiographs using deep learning algorithms have demonstrated promise for tuberculosis screening in resource-constrained environments. Studies among symptomatic individuals presenting to healthcare facilities have demonstrated variable accuracy as a screening or triage tool for pulmonary tuberculosis.

### *What This Study Adds to the Field*

We evaluated the accuracy of three deep learning algorithms tested on x-rays collected during a prospective, mass screening study in three Brazilian prisons, in which active case finding for tuberculosis was performed irrespective of symptoms. Two of the algorithms exceeded WHO accuracy benchmarks for a screening diagnostic, though all had markedly reduced specificity among individuals with previous tuberculosis.

**ABSTRACT (252 WORDS)**

**Rationale:** The World Health Organization (WHO) recommends systematic tuberculosis (TB) screening in prisons. Evidence is lacking for accurate and scalable screening approaches in this setting.

**Objectives:** To assess the diagnostic accuracy of three artificial intelligence-based chest x-ray interpretation algorithms for TB screening in prisons.

**Methods:** We conducted a prospective TB screening study in three male prisons in Mato Grosso do Sul, Brazil from October 2017 to December 2019. We administered a standardized questionnaire, performed chest radiography in a mobile unit, and collected sputum for confirmatory testing using Xpert MTB/RIF and culture. We evaluated x-ray images using three algorithms (CAD4TB version 6, LunitTB and qXR) and compared their diagnostic accuracy. We utilized multivariable logistic regression to assess the effect of demographic and clinical characteristics on algorithm accuracy. Finally, we investigated the relationship between abnormality scores and Xpert semi-quantitative results.

**Measurements and Main Results:** Among 2,075 incarcerated individuals, 259 (12.5%) had confirmed TB. All three algorithms performed similarly overall with AUCs of 0.87-0.91. At 90% sensitivity, only LunitTB and qXR met the WHO Target Product Profile requirements for a triage test, with specificity of 84% and 74%, respectively. All algorithms had variable performance by age, prior TB, smoking, and presence of TB symptoms. LunitTB was the most robust to this heterogeneity, but nonetheless failed to meet the TPP for individuals with previous TB. Abnormality scores of all three algorithms were significantly correlated with sputum bacillary load.

**Conclusions:** Automated x-ray interpretation algorithms can be an effective triage tool for TB screening in prisons. However, their specificity is insufficient in individuals with previous TB.

**Keywords:** active case finding, tuberculosis, prisons, X-ray, automated interpretation, triage, diagnostics.

## INTRODUCTION

Globally, tuberculosis (TB) incidence in prisons is more than ten times higher than the general population (1). This disparity is especially alarming in South America, where TB cases in prisons have more than doubled since 2000 amid rising incarceration rates (1, 2). Several factors contribute to the elevated risk of TB in prisons, including overcrowding, poor ventilation, high rates of smoking, drug use, and limited access to medical care, leading to delays in TB diagnosis (2).

Interventions to address this growing burden are urgently needed, including improvements in case detection. In 2021, the World Health Organization (WHO) released updated guidelines on screening for TB, upgrading to a strong recommendation that systematic screening be conducted in prisons and penitentiary institutions (3). However, the recommendation is based on “very low certainty of evidence”, and guidance on specific means for screening in this setting is lacking. Moreover, correctional health systems are often underfunded and poorly equipped, and few prison systems in low- and middle-income countries perform systematic screening for TB despite the widely acknowledged high burden. Therefore, effective, cost-efficient screening approaches are needed to bring case-finding to scale in these settings. An important part of such approaches is a point-of-care triage test that can substantially reduce the number of people who need further testing.

Chest radiography is among the oldest tools for pulmonary TB screening and historically played a major role in TB control programs in high burden settings (4, 5). However, by the 1970s, concerns were raised about the accuracy, logistics and personnel requirements for mass radiography, leading the WHO to conclude in its 9<sup>th</sup> expert committee report that “indiscriminate TB case finding by mobile mass radiography should be abandoned” (6).

Recently, there has been a resurgence in interest in the use of radiography as a screening tool for TB, leveraging recent advances in machine learning approaches to automate x-ray interpretation (7, 8). Clinic-based evaluations have demonstrated promising accuracy for several automated interpretation systems among individuals with TB symptoms. However, it is less clear how well these algorithms

will perform for the purpose of active case finding, irrespective of symptoms, in incarcerated populations with high prevalence of smoking, drug use, and history of TB. To address this knowledge gap, we evaluated the performance of three deep learning-based x-ray interpretation algorithms in the context of mass screening for TB in three high burden prisons in Brazil.

## **MATERIALS AND METHODS**

### ***Study Design and Participants***

We performed a cross-sectional study from October 2017 to December 2019 in three male prisons in Mato Grosso do Sul State, Brazil: Jair Ferreira de Carvalho Penitentiary (EPJFC), Campo Grande Penal Institute (IPCG), and Dourados State Penitentiary (PED). The prisons have a combined population of approximately 5,500 individuals. All incarcerated individuals in each prison were invited to participate in TB screening. Those who agreed to participate in the study provided written informed consent. The study was approved by the institutional review boards (IRBs) of the Federal University of Grande Dourados (UFGD) (#3.483.377) and Stanford University (#40285).

### ***Study Procedures***

We outfitted a Volkswagen Constellation 24-240 truck with a 9.8 x 2.5-meter container, lead covering, an access ramp, an x-ray machine (Altus ST 543 HF, Sawae<sup>®</sup>), an x-ray scanner and digitizer (Agfa 15-X CR, Mortsel Belgium) and a separate room for sputum processing with two 4-module GeneXpert machines (Cepheid, Sunnyvale, USA). The mobile screening team consisted of a nurse, a laboratory technician, and an x-ray technician, with a physician available for consultation.

Study nurses administered a structured questionnaire to obtain demographic data, incarceration history, lifestyle factors, health history, and TB symptoms as recommended by the World Health Organization (WHO) (6). A spot sputum sample was collected from all participants who were able to produce sputum, with a target volume of at least 2 ml. Sputum samples were tested by Xpert MTB/RIF G4 (Cepheid, Sunnyvale, USA) and, if volume was sufficient, culture on Ogawa-Kudoh media. *M. tuberculosis* growth in cultures was confirmed by an immunochromatographic assay (TB Ag MPT64 Rapid Test, Standard Diagnostics, Seoul, South Korea).

A posterior-anterior chest x-ray was performed for all participants and then scanned and digitized. The images were electronically transferred for automated analysis by Computer-Aided Detection for TB

version 6 (CAD4TBv6), developed by the Analysis Group at Radboud University Medical Center (Netherlands); Lunit INSIGHT CXR2 (hereinafter LunitTB) developed by the South Korean medical software company Lunit; and qXR, developed by Qure.ai in Mumbai, India. All information was recorded in Research Electronic Data Capture (REDCap<sup>®</sup>) (9, 10).

### ***Outcome Definitions and Analytic Approach***

We defined TB cases as individuals with a positive Xpert MTB/RIF or culture growing *M. tuberculosis*. We defined controls as individuals who had sputum testing for Xpert MTB/RIF and no positive result by Xpert or culture. As a secondary analysis, we included as controls all individuals who were screened for TB and did not have a positive test, regardless of whether they were able to provide sputum. Individuals already undergoing treatment for TB were excluded from all analyses.

For CAD4TBv6, which provides a score range of 0 to 100, we used a positivity threshold score of  $\geq 60$  through calibration with radiographic imaging data from a subset of participants with ( $n = 80$ ) and without ( $n = 200$ ) microbiologically confirmed TB. For LunitTB, which provides a score range of 0 to 1, we used a threshold of  $\geq 0.72$  as specified by the manufacturer and identified through prior calibration (10) (11). For qXR (score range 0 to 1), we used a threshold of  $\geq 0.5$  according to a previous study (11). We also evaluated the performance of each algorithm with the WHO's Target Product Profile (12) (TPP) for a triage test by identifying the threshold that achieved 90% sensitivity and examining the corresponding specificity.

For each algorithm we calculated the sensitivity, specificity, positive predictive value (PPV), negative predictive value (NPV), and area under the receiver operating characteristic (ROC) curve (AUC). We calculated exact binomial confidence intervals (CIs) for sensitivity and specificity. We compared algorithm AUCs using DeLong's test. For demographic and clinical characteristics, continuous variables were compared using the Mann-Whitney U test and categorical variables using the chi-square test. To assess the influence of demographic and clinical characteristics on algorithm performance, we conducted multivariable logistic regression controlling for age, race, drug use,

smoking, previous TB, and presence of any TB symptoms (cough, fever, night sweat, weight loss, loss of appetite, tiredness, and chest pain). We report predicted marginal estimates of specificity for each characteristic at the WHO TPP threshold of 90% sensitivity. Finally, we investigated the relationship between Xpert semi-quantitative result and x-ray algorithm score among confirmed TB cases using Spearman's rank correlation testing. Data were analyzed using SPSS version 25.0 and R version 4.0.3, including the pROC package (13).

## RESULTS

Between October 2017 and December 2019, we enrolled 7,081 participants across three male prisons in the Brazilian state of Mato Grosso do Sul. Sixty-six participants were excluded from further analyses as they were already under treatment for tuberculosis. Among the remainder, 2075 participants were able to produce valid sputum samples for Xpert and were included in the primary analysis (**Figure E1**). Participants in the primary analysis had a median age of 33 years (IQR 28-40) (**Table 1**). Compared with participants who did not produce a valid sputum, those who did had a higher prevalence of TB symptoms (73% vs 18%,  $p<0.001$ ), smoking (73% vs 55%,  $p<0.001$ ), illicit drug use (70% vs 54%  $p<0.001$ ), and previous tuberculosis (14% vs 5%,  $p<0.001$ ) (**Table E1**).

During the screening period, 259 (12.5%) participants were diagnosed with pulmonary TB, of which 113 were diagnosed by sputum Xpert alone, 17 were diagnosed through sputum culture alone, and 129 had positive Xpert and culture tests. The prevalence of any TB symptom did not differ between TB cases and controls (76% vs 73%,  $p=0.277$ ); however, cough was slightly more common among TB cases (66 vs 60%,  $p=0.037$ ). Smoking, drug use, history of incarceration, and history of TB were significantly more prevalent among TB cases compared to non-TB cases (**Table 1**).

Among TB cases, 209 (80.7%) were classified as positive by CAD4TBv6, 207 (79.9%) by LunitTB, and 193 (74.5%) by qXR when using pre-defined thresholds (**Table 2**). At 90% sensitivity, only LunitTB and qXR met the WHO's Target Product Profile (TPP) with specificity of 83.7% (95% CI 72.4-87.3) and 74.2% (95% CI 60.2-81.3), respectively. At a 4% prevalence of TB, LunitTB had the highest PPV (18.7%), followed by qXR (12.7%) and CAD4TBv6 (9.0%). Receiver operating characteristic (ROC) curves for each algorithm are shown in **Figure 1**. Compared with CAD4TBv6 (AUC 0.877), LunitTB (AUC 0.912,  $p=0.0003$ ) and qXR (AUC 0.901,  $p=0.0112$ ) had higher AUCs, though AUC did not differ between LunitTB and qXR ( $p=0.17$ ). In a secondary analysis of diagnostic accuracy in which we included the 4940 participants unable to provide sputum (total  $N=7081$ ), AUCs did not differ substantially from the primary analysis, with LunitTB at 0.926, qXR at 0.920, and CAD4TBv6 at 0.904.



We next performed multivariable logistic regression analysis to examine whether the performance of each algorithm varied by sociodemographic characteristics and risk factors, namely: age, race/ethnicity, current smoker, drug use, previous TB, and TB symptoms. Specificity of all three algorithms decreased with age and tended to be lower among current smokers and those without TB symptoms, compared to their respective counterparts (**Figure 2, Figure E2-E3**). LunitTB was the only algorithm that met WHO TPP criteria among individuals 45 years and older. Notably, specificity was under 50% across all three algorithms for individuals with a history of TB.

To further investigate diagnostic performance depending on history of previous TB, we analyzed the distribution of abnormality scores for TB cases versus non-TB cases as confirmed by sputum Xpert or culture, disaggregated by history of TB. We focused on LunitTB for this analysis given its superior overall performance and its relatively stable specificity by subgroup compared to the other two algorithms. Strikingly, the thresholds required to reach WHO TPP benchmarks of 90% sensitivity and 70% specificity varied dramatically by history of TB (**Figure 3**). In participants without previous TB, LunitTB score thresholds  $\geq 0.04$  had 70% specificity and those  $\leq 0.15$  had 90% sensitivity, providing a range of thresholds (0.04-0.15) meeting TPP benchmarks. Conversely, in participants with previous TB, a threshold of at least 0.73 was required for 70% specificity, and there was no score threshold to satisfy both TPP sensitivity and specificity. In participants with previous TB, neither CAD4TBv6 nor qXR had a score that satisfies both TPP sensitivity and specificity. (Figure S4-S5).

Finally, we assessed the relationship between sputum bacillary load and algorithm performance by examining x-ray abnormality scores by Xpert semi-quantitative result (negative, very low, low, medium, high). Among TB cases with a positive Xpert test for whom x-ray scores were available (188/242), all three algorithms yielded abnormality scores that were positively correlated with sputum Xpert semi-quantitative levels ( $p < 0.0001$ ) (**Figure 4**). Among the 67 participants with a medium or high Xpert result, CAD4TBv6 had 97% sensitivity (65/67) and LunitTB and qXR both had 96% sensitivity (64/67) at the 70% specificity threshold (**Table E2**).

## DISCUSSION

Active case finding for tuberculosis in high burden carceral settings is needed to address the substantial excess burden among incarcerated populations. However, despite WHO recommendations for routine TB screening in prisons, most facilities in low- and middle-income countries do not perform systematic active case finding, often citing resource and infrastructure constraints. Effective, cost-efficient screening strategies are needed to make active case finding more accessible in such environments. In this study, conducted via a nurse-led mobile diagnostic unit in three prisons in Central-Western Brazil, we found a very high prevalence of undiagnosed, microbiologically confirmed TB (3.7%). Algorithms for automated interpretation of x-rays achieved high sensitivity and specificity as a screening tool, with the LunitTB and qXR systems exceeding the WHO minimal TPP thresholds for a triage test. Sputum molecular testing is still needed to confirm TB, but a limiting factor in the speed and costs of screening has been the number of tests that can be run daily during mass screening of thousands of individuals. Our findings suggest that screening by mobile x-ray systems with automated interpretation could reduce the number of confirmatory tests required and enable screening to be more rapid and cost-effective in high burden TB settings, while still maintaining sufficient sensitivity.

Recent studies have evaluated x-ray interpretation algorithms among individuals presenting to clinics with TB symptoms, finding variable results. An individual participant meta-analysis found that none of the systems investigated met the WHO TPP criteria for triage, with specificities ranging from 54-61% at 90% sensitivity (14). By contrast, a study in Bangladesh found that the qXR system achieved 74% specificity at the same threshold, and that all algorithms outperformed interpretation by radiologists (15). Our study differed in that it was performed in the context of active case finding, irrespective of symptoms, which could affect estimates of diagnostic accuracy in several ways. For instance, the cases identified through systematic screening are often those in early stages of disease, with lower bacillary burden, as evidenced by the fact that 54% of confirmed cases in our cohort had low, very low, or negative Xpert results. Given the association we observed between sputum bacillary load and x-ray scores, this could have resulted in the algorithms having lower sensitivity in our cohort.

At the same time, we might expect higher specificity in the context of active case finding, regardless of symptoms, compared to use in clinics among those presenting with TB symptoms, as the latter setting may include more patients with other pulmonary diseases such as bacterial and viral pneumonias that can be challenging to distinguish from TB. Furthermore, we evaluated these algorithms in incarcerated populations, which tend to be younger, predominantly male, and with high prevalence of various risk factors for TB.

LunitTB was the best-performing algorithm in this cohort, with greater accuracy and generalizability among subgroups, with particularly superior robustness to age compared to the other two algorithms. Nonetheless, performance of all three algorithms varied by subgroup, with consistently lower specificity among older individuals and those with previous TB, corroborating previous findings (15, 16). We also found reduced specificity among current smokers and those without TB symptoms. Therefore, different thresholds may be considered for individuals with different demographic or clinical characteristics. Of note, our pre-defined thresholds for each algorithm led to overall sensitivity under 90%, suggesting that setting- or population-specific threshold calibration may be an important step in implementation.

Specificity of all three algorithms decreased considerably to less than 50% for those with previous TB, indicating failure to meet the WHO TPP for this subgroup. For the 3 analyzed software tools, the distribution of abnormality scores among non-TB cases was shifted higher for those with a history of TB, suggesting the algorithm may not distinguish active TB lesions from fibrous scarring of the lung parenchyma and other chest radiograph patterns indicative of previous TB (17). Thus, in populations with high prevalence of previous TB, Xpert may be more appropriate for screening (18).

We found that x-ray scores were higher—suggestive of more abnormalities—in individuals with high sputum bacillary loads. At the 70% specificity threshold, sensitivity for individuals with medium or high bacillary loads exceeded 96% for all three systems. Given that Xpert bacillary load correlates with smear status (19), and smear status predicts infectiousness (20, 21), it may be reasonable to infer that x-ray automated interpretation algorithms may be more sensitive in identifying the most infectious

individuals.

Even with the availability of automated interpretation algorithms, the cost-effectiveness of using x-rays for mass screening in prisons is still unclear. Previous work found that mass screening in prisons with sputum Xpert alone had high yield and was less costly than using x-ray and CAD4TB (version 5) for triage prior to confirmatory Xpert (19). However, the prior study used a single CAD4TB threshold for all individuals and evaluated an additional strategy where only individuals without symptoms were screened with x-ray and CAD4TB prior to confirmatory Xpert. Our present findings suggest that such strategies may be less effective due to the algorithms' variable performance by subgroup, particularly the reduced sensitivity specificity for individuals without TB symptoms. Furthermore, CAD4TB (version 6) was shown to have the lowest performance in this study; thus, screening with a more accurate algorithm like LunitTB could increase cost-effectiveness. Additionally, emerging technologies for portable, digital radiography could reduce consumable costs, making x-rays more accessible and affordable in resource-constrained environments.

This study has several limitations. First, in our primary analysis, we only included participants who were able to produce sputum for confirmatory testing, as sputum induction was not able to be undertaken in this environment. The excluded participants were less likely to be current smokers, to have TB symptoms, and to report previous TB; we expect that their inclusion may have affected overall estimates of algorithm performance in this population. In secondary analyses of the entire population, AUCs did not differ significantly. However, future research is needed to evaluate these x-ray interpretation algorithms on this group, given that a strength of x-ray screening is the lack of requirement for sputum. Second, we used solid media culture due to local availability costs; however, solid media culture is less sensitive than liquid media and could have led to missed cases.

Additionally, we note that while for LunitTB and qXR we used the manufacturers' recommended thresholds, for CAD4TBv6 we used a threshold determined from a subset of our population; therefore, the thresholds at 90% sensitivity may be more appropriate than our pre-defined thresholds for comparison of the three algorithms. Due to the low prevalence of HIV in our population, we did not consider HIV status in our study (22). Moreover, our study only included those in male prisons as

there are fewer than 10 cases annually among incarcerated women in this state; consequently, the performance of these algorithms for TB screening in female prisons remains unknown (16).

Overall, our results suggest that the use of chest x-rays and artificial intelligence-based interpretation algorithms can be part of an effective mass screening strategy in high-burden settings like prisons. Although Lunit-TB had the greatest accuracy and robustness in our cohort, all three algorithms exhibited similar performance, particularly as a rule-out-test, and could be used to reduce the need for universal molecular testing. However, our findings suggest the need for future optimization of these algorithms to improve generalizability across subgroups, especially for individuals with a history of TB. Nevertheless, given their high overall accuracy in this population, especially among cases with the greatest sputum bacillary load, automated interpretation algorithms could enable scaling of mass screening to help mitigate disparities in TB diagnosis among incarcerated populations.

## **ACKNOWLEDGMENTS**

The State Agency for the Administration of the Penitentiary System of the State of Mato Grosso do Sul (AGEPEN) for having authorized the study; the Coordination for the Improvement of Higher Education Personnel (CAPES) and the Federal University of Grande Dourados (UFGD) for having supported the study in their postgraduate programs.

## REFERENCES

1. Cords O, Martinez L, Warren JL, O'Marr JM, Walter KS, Cohen T, Zheng J, Ko AI, Croda J, Andrews JR. Incidence and prevalence of tuberculosis in incarcerated populations: a systematic review and meta-analysis. *Lancet Public Health* 2021;6:e300–e308.
2. Walter KS, Martinez L, Arakaki-Sanchez D, Sequera VG, Sanabria GE, Cohen T, Ko AI, García-Basteiro AL, Rueda ZV, López-Olarte RA, Espinal MA, Croda J, Andrews JR. The escalating tuberculosis crisis in central and South American prisons. *The Lancet* 2021;397:1591–1596.
3. World Health Organization. *WHO consolidated guidelines on tuberculosis. Module 2: screening – systematic screening for tuberculosis disease*. Geneva: World Health Organization; 2021.
4. Hermans SM, Andrews JR, Bekker L-G, Wood R. The mass miniature chest radiography programme in Cape Town, South Africa, 1948 - 1994: The impact of active tuberculosis case finding. *South Afr Med J Suid-Afr Tydskr Vir Geneeskde* 2016;106:1263–1269.
5. COMSTOCK GW, PHILIP RN. Decline of the tuberculosis epidemic in Alaska. *Public Health Rep Wash DC* 1896 1961;76:19–24.
6. Organization WH. WHO Expert Committee on Tuberculosis [meeting held in Geneva from 11 to 20 December 1973]: ninth report. 1974;
7. World Health Organization. *Systematic screening for active tuberculosis: an operational guide*. 2015. at [http://apps.who.int/iris/bitstream/10665/181164/1/9789241549172\\_eng.pdf?ua=1](http://apps.who.int/iris/bitstream/10665/181164/1/9789241549172_eng.pdf?ua=1).
8. Pinto LM, Pai M, Dheda K, Schwartzman K, Menzies D, Steingart KR. Scoring systems using chest radiographic features for the diagnosis of pulmonary tuberculosis in adults: a systematic review. *Eur Respir J* 2013;42:480–494.

9. Harris PA, Taylor R, Minor BL, Elliott V, Fernandez M, O'Neal L, McLeod L, Delacqua G, Delacqua F, Kirby J, Duda SN. The REDCap consortium: Building an international community of software platform partners. *J Biomed Inform* 2019;95:103208.
10. Harris PA, Taylor R, Thielke R, Payne J, Gonzalez N, Conde JG. Research electronic data capture (REDCap)—A metadata-driven methodology and workflow process for providing translational research informatics support. *J Biomed Inform* 2009;42:377–381.
11. Nash M, Kadavigere R, Andrade J, Sukumar CA, Chawla K, Shenoy VP, Pande T, Huddart S, Pai M, Saravu K. Deep learning, computer-aided radiography reading for tuberculosis: a diagnostic accuracy study from a tertiary hospital in India. *Sci Rep* 2020;10:210.
12. World Health Organization. *High-priority target product profiles for new tuberculosis diagnostics: report of a consensus meeting*. World Health Organization; 2014. at [http://apps.who.int/iris/bitstream/handle/10665/135617/WHO\\_HTM\\_TB\\_2014.18\\_eng.pdf;jsessionid=0CBFF2089B61E11AFF4C116E2CE6E273?sequence=1](http://apps.who.int/iris/bitstream/handle/10665/135617/WHO_HTM_TB_2014.18_eng.pdf;jsessionid=0CBFF2089B61E11AFF4C116E2CE6E273?sequence=1).
13. pROC: an open-source package for R and S+ to analyze and compare ROC curves. at <https://bmcbioinformatics.biomedcentral.com/articles/10.1186/1471-2105-12-77>.
14. Tavaziva G, Harris M, Abidi S, Geric C, Breuninger M, Dheda K, Esmail A, Muyoyeta M, Reither K, Majidulla A, Khan A, Campbell J, David P-M, Denkinger C, Miller C, Nathavitharana R, Pai M, Benedetti A, Ahmad Khan F. Chest X-ray Analysis With Deep Learning-Based Software as a Triage Test for Pulmonary Tuberculosis: An Individual Patient Data Meta-Analysis of Diagnostic Accuracy. *Clin Infect Dis* 2021;doi:10.1093/cid/ciab639.
15. Qin ZZ, Ahmed S, Sarker MS, Paul K, Adel ASS, Naheyan T, Barrett R, Banu S, Creswell J. Tuberculosis detection from chest x-rays for triaging in a high tuberculosis-



- burden setting: an evaluation of five artificial intelligence algorithms. *Lancet Digit Health* 2021;3:e543–e554.
16. Khan FA, Majidulla A, Tavaziva G, Nazish A, Abidi SK, Benedetti A, Menzies D, Johnston JC, Khan AJ, Saeed S. Chest x-ray analysis with deep learning-based software as a triage test for pulmonary tuberculosis: a prospective study of diagnostic accuracy for culture-confirmed disease. *Lancet Digit Health* 2020;2:e573–e581.
  17. Piccazzo R, Paparo F, Garlaschi G. Diagnostic accuracy of chest radiography for the diagnosis of tuberculosis (TB) and its role in the detection of latent TB infection: a systematic review. *J Rheumatol Suppl* 2014;91:32–40.
  18. Frascella B, Richards AS, Sossen B, Emery JC, Odone A, Law I, Onozaki I, Esmail H, Houben RMGJ. Subclinical tuberculosis disease - a review and analysis of prevalence surveys to inform definitions, burden, associations and screening methodology. *Clin Infect Dis* doi:10.1093/cid/ciaa1402.
  19. Beynon F, Theron G, Respeito D, Mambuque E, Saavedra B, Bulu H, Sanz S, Dheda K, Garcia-Basteiro AL. Correlation of Xpert MTB/RIF with measures to assess *Mycobacterium tuberculosis* bacillary burden in high HIV burden areas of Southern Africa. *Sci Rep* 2018;8:5201.
  20. Behr MA, Warren SA, Salamon H, Hopewell PC, Ponce de Leon A, Daley CL, Small PM. Transmission of *Mycobacterium tuberculosis* from patients smear-negative for acid-fast bacilli. *Lancet Lond Engl* 1999;353:444–449.
  21. Hernández-Garduño E, Cook V, Kunimoto D, Elwood RK, Black WA, FitzGerald JM. Transmission of tuberculosis from smear negative patients: a molecular epidemiology study. *Thorax* 2004;59:286–290.
  22. Sgarbi RVE, Carbone A da SS, Paião DSG, Lemos EF, Simionatto S, Puga MAM, Motta-Castro ARC, Pompilio MA, Urrego J, Ko AI, Andrews JR, Croda J. A Cross-Sectional

Survey of HIV Testing and Prevalence in Twelve Brazilian Correctional Facilities. *PloS One* 2015;10:e0139487.

## TABLES AND FIGURES

**Table 1:** Sociodemographic characteristics and risk factors for TB among study participants, stratified by TB status as determined by sputum Xpert or culture

**Table 2:** Sensitivity, Specificity, AUC, PPV and NPV of each algorithm at pre-defined thresholds or with thresholds adjusted to 90% sensitivity as required by the WHO Target Product Profile.

**Figure 1:** Receiver operating characteristic (ROC) curves for CAD4v6, LunitTB and qXR.

**Figure 2:** Specificity of LunitTB, by sociodemographic characteristics and risk factors. Shown are the predicted margins for specificity and 95% confidence intervals from a multivariable logistic regression, holding sensitivity at 90%.

**Figure 3:** Sensitivity and specificity according to LunitTB score threshold, with the WHO sensitivity (dashed line) and specificity (dotted line) benchmarks (top) among individuals without (A) and with (B) previous TB. Distribution of LunitTB scores for participants without (C) or with (D) previous TB (bottom).

**Figure 4.** Relationship between the bacillary load in the sputum and the performance of the algorithm through the stratification of the scores by the semi-quantitative Xpert result.

**Table 1** - Sociodemographic characteristics and risk factors for TB among study participants, stratified by TB status as determined by sputum Xpert or culture.

<b>Variables</b>	<b>Total N=2075 (%)</b>	<b>TB cases N=259 (%)</b>	<b>No TB N=1816 (%)</b>	<b>P Value</b>
Median age (IQR)	33 (28, 40)	33 (28, 39)	33 (28, 40)	0.497
Prison Unit				
PED	889 (42.8)	65 (25.1)	824 (45.0)	<0.001
EPJFC	840 (40.5)	144 (55.6)	696 (38.3)	
IPCG	346 (16.9)	50 (19.3)	296 (16.3)	
Race/ethnicity				
Mixed	1279 (61.6)	158 (61.0)	1121 (61.4)	0.192
White	508 (24.5)	56 (21.6)	452 (24.9)	
Black	253 (12.2)	38 (14.7)	215 (11.8)	
Indigenous	33 (1.6)	7 (2.7)	26 (1.4)	
Asian	2 (0.1)	-	2 (0.1)	
<8 years of schooling	1546 (74.5)	198 (76.4)	1348 (74.2)	0.230
Current smoker	1520 (73.3)	208 (80.3)	1312 (72.2)	0.006
Illicit drug use over the last year	1460 (70.4)	203 (78.4)	1257 (69.2)	0.003
Previously incarcerated	1557 (75.0)	214 (82.6)	1343 (74.0)	0.003
BCG vaccinated	1862 (89.7)	223 (86.1)	1639 (90.3)	0.039
Previous TB	293 (14.1)	55 (21.2)	238 (13.1)	0.000
Report any WHO TB symptoms	1512 (72.9)	196 (75.7)	1316 (72.5)	0.277
Report cough	1255 (60.5)	172 (66.4)	1083 (59.6)	0.037
TB contact	1565 (75.4)	211 (81.5)	1354 (74.6)	0.160

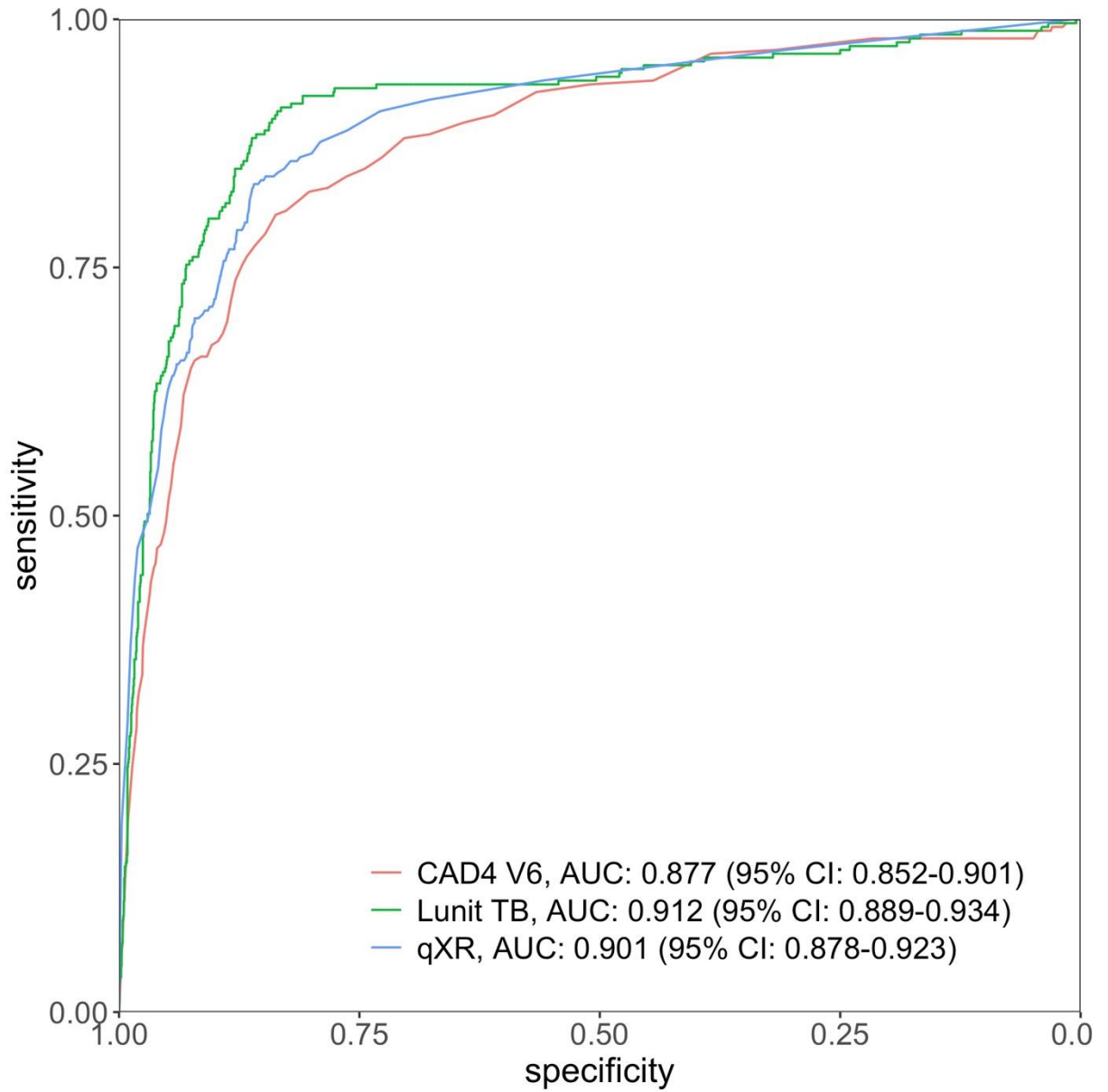
**Abbreviations:** BCG: Bacillus of Calmette and Guérin; IQR: interquartile range; TB: tuberculosis; WHO:

World Health Organization, y: years.

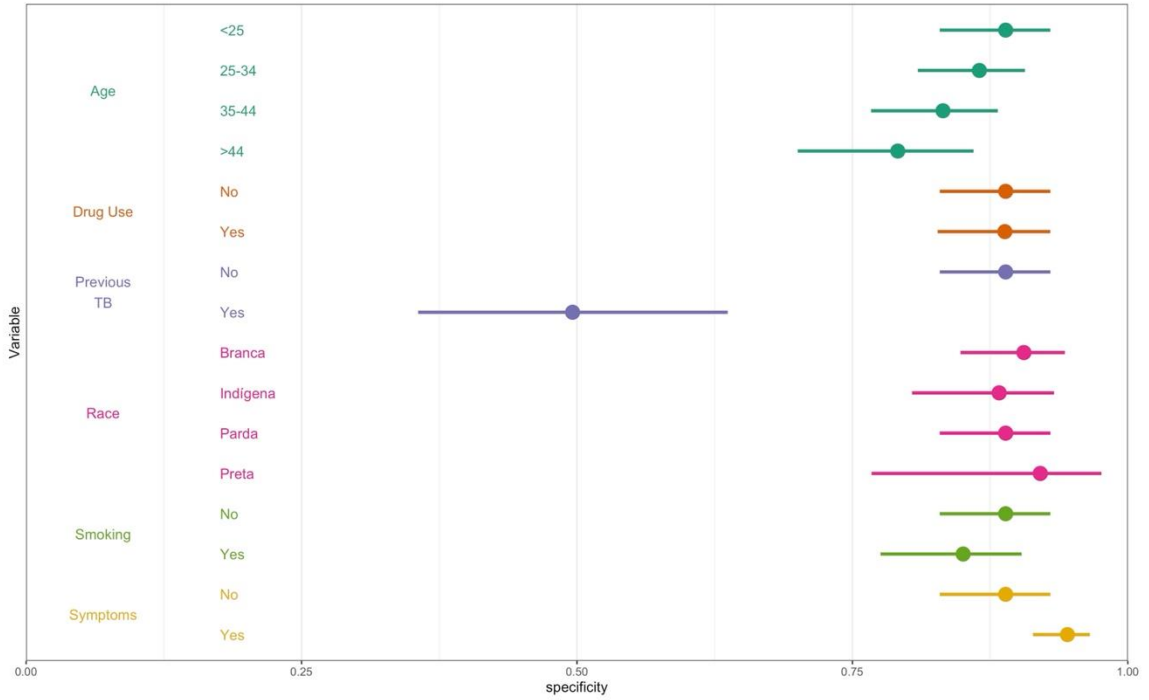
**Table 2** – Sensitivity, Specificity, Area Under the Curve (AUC), Positive Predictive Value (PPV) and Negative Predictive Value (NPV) of each algorithm at pre-defined thresholds or with thresholds adjusted to 90% sensitivity as specified by the WHO Target Product Profile minimum target.

System	AUC (95% CI)	At pre-defined thresholds		At 90% sensitivity, 4% prevalence		
		Sensitivity % (95% CI)	Specificity % (95% CI)	Specificity % (95% CI)	PPV %	NPV %
CAD4v6	0.87 (0.85-0.90)	80.7 (75.4-85.3)	82.7 (80.8-84.4)	62.3 (52.0-73.1)	9.0	99.3
LunitTB	0.91 (0.88-0.93)	79.9 (74.5-84.6)	89.8 (88.3-91.2)	83.7 (72.4-87.3)	18.7	99.5
qXR	0.90 (0.87-0.92)	74.5 (68.8-79.7)	89.4 (87.9-90.8)	74.2 (60.2-81.3)	12.7	99.4

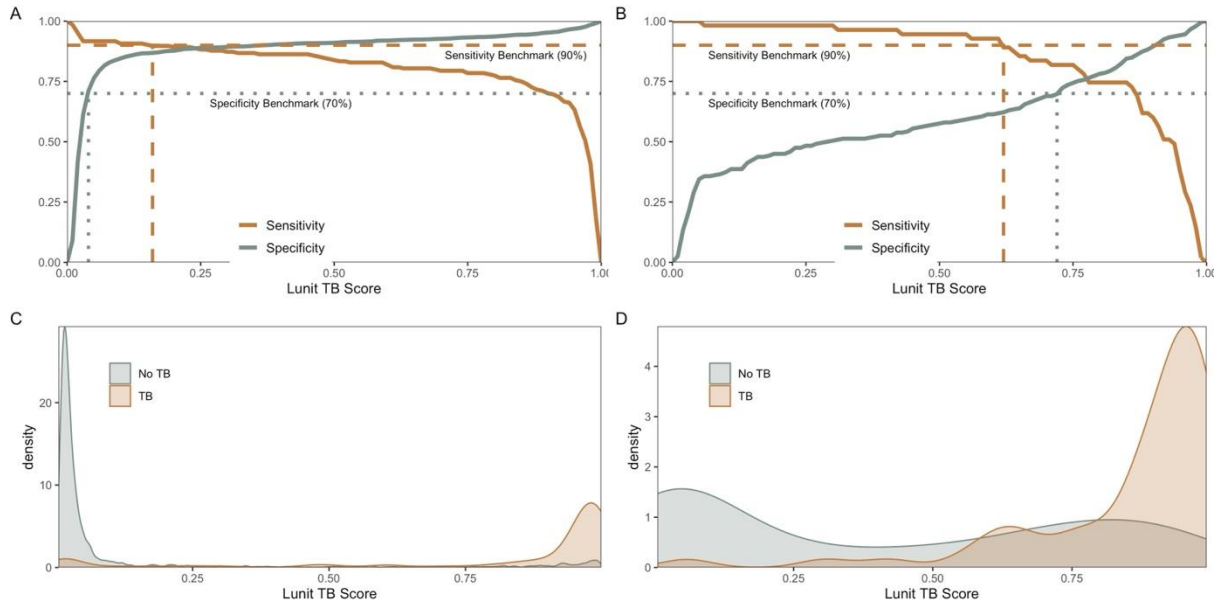
**Figure 1** – Receiver operating characteristic (ROC) curves for CAD4v6, LunitTB and qXR.



**Figure 2:** Specificity of LunitTB, by sociodemographic characteristics and risk factors. Shown are the predicted margins for specificity and 95% confidence intervals from a multivariable logistic regression, holding sensitivity at 90%.

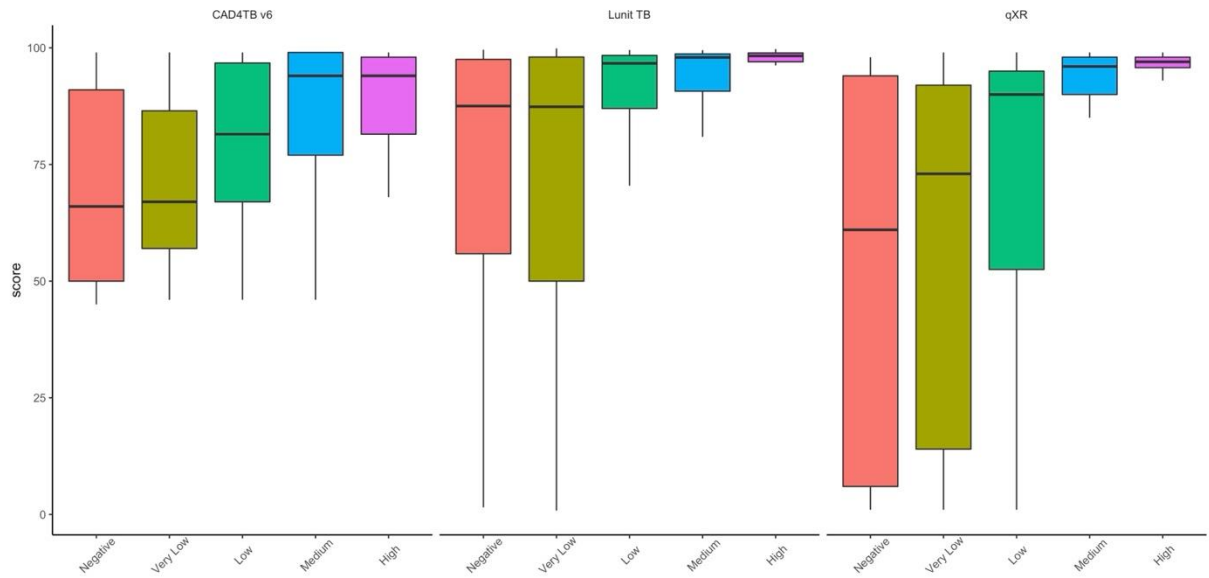


**Figure 3:** Sensitivity and specificity according to LunitTB score threshold, with the WHO sensitivity (dashed line) and specificity (dotted line) benchmarks (top) among individuals without (A) and with (B) previous TB. Distribution of LunitTB scores for participants without (C) or with (D) previous TB (bottom).





**Figure 4.** Relationship between the bacillary load in the sputum, measured by Xpert MTB/RIF semiquantitative result, and algorithm score.



## DATA SUPPLEMENT

**Table E1.** Sociodemographic characteristics and risk factors for TB by study inclusion/exclusion.

Excluded refers to those excluded due to lack of a confirmatory sputum Xpert or culture result.

**Table E2.** Sensitivity of x-ray interpretation algorithms according to Xpert MTB/RIF G4 semi-quantitative load, using a threshold for specificity of 70%.

**Figure E1.** Flow diagram of study participants in mass screening and inclusion of participants in x-ray evaluation.

**Figure E2:** Specificity of CAD4TBv6, by sociodemographic characteristics and risk factors. Shown are the predicted margins for specificity and 95% confidence intervals from a multivariable logistic regression, holding sensitivity at 90%.

**Figure E3:** Specificity of qXR, by sociodemographic characteristics and risk factors. Shown are the predicted margins for specificity and 95% confidence intervals from a multivariable logistic regression, holding sensitivity at 90%.

**Figure E4:** Sensitivity and specificity according to CAD4TBv6 score threshold, with the WHO sensitivity (dashed line) and specificity (dotted line) benchmarks (top) among individuals without (A) and with (B) previous TB. Distribution of CAD4TBv6 scores for participants without (C) or with (D) previous TB (bottom).

**Figure E5:** Sensitivity and specificity according to qXR score threshold, with the WHO sensitivity (dashed line) and specificity (dotted line) benchmarks (top) among individuals without (A) and with (B) previous TB. Distribution of qXR scores for participants without (C) or with (D) previous TB (bottom).

**Table E1.** Sociodemographic characteristics and risk factors for TB by study inclusion/exclusion.

Excluded refers to those excluded due to lack of a confirmatory sputum Xpert or culture result.

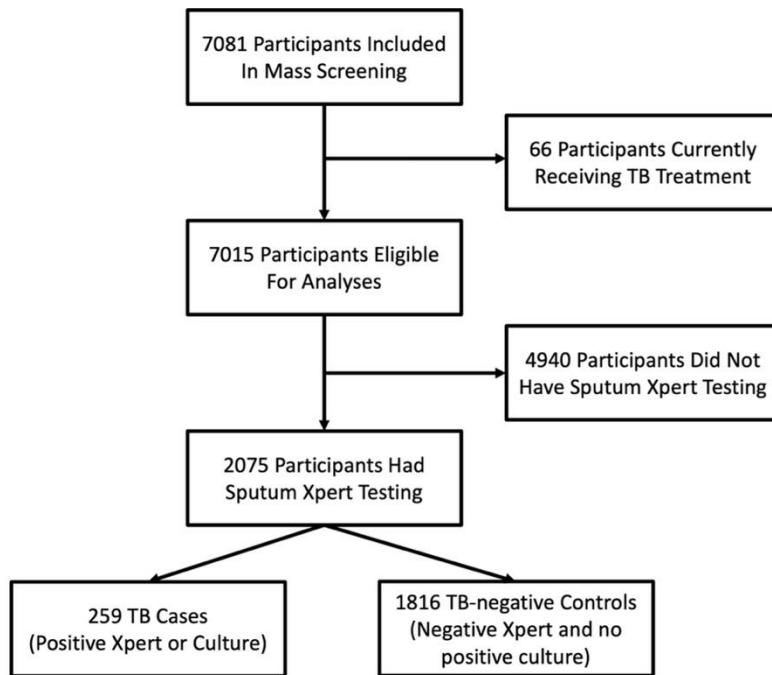
<b>Variables</b>	<b>Included N=2075 (%)</b>	<b>Excluded N=4940 (%)</b>	<b>P Value</b>
Median age, (IQR)	33 (28, 40)	34 (28, 40)	<0.001
Prison Unit			
PED	889 (42.8)	1444 (29.2)	<0.001
EPJFC	840 (40.5)	1940 (39.3)	
IPCG	346 (16.9)	1556 (31.4)	
Ethnic/skin color			
Mixed	1279 (61.6)	3068 (62.1)	0.428
White	508 (24.5)	1115 (22.6)	
Black	253 (12.2)	612 (12.4)	
Indigenous	33 (1.6)	138 (2.8)	
Asian	2 (0.1)	7 (0.1)	
<8y of schooling	1546 (74.5)	3511 (71.1)	0.003
Current smoker	1520 (73.3)	2737 (55.4)	<0.001
BCG vaccinated	1862 (89.7)	4378 (88.6)	0.197
Previous TB	293 (14.1)	268 (5.4)	<0.001
Report any WHO TB symptoms	1512 (72.9)	851 (17.2)	<0.001
Report cough	1255 (60.5)	536 (10.9)	<0.001
Illicit drug use over the last year	1460 (70.4)	2644 (53.5)	<0.001
Previously incarcerated	1557 (75.0)	3478 (70.4)	<0.001
TB contact	1565 (75.4)	3125 (63.3)	<0.001

**Abbreviations:** BCG: Bacillus of Calmette and Guérin; IQR: interquartile range; TB: tuberculosis; WHO: World Health Organization, y: years

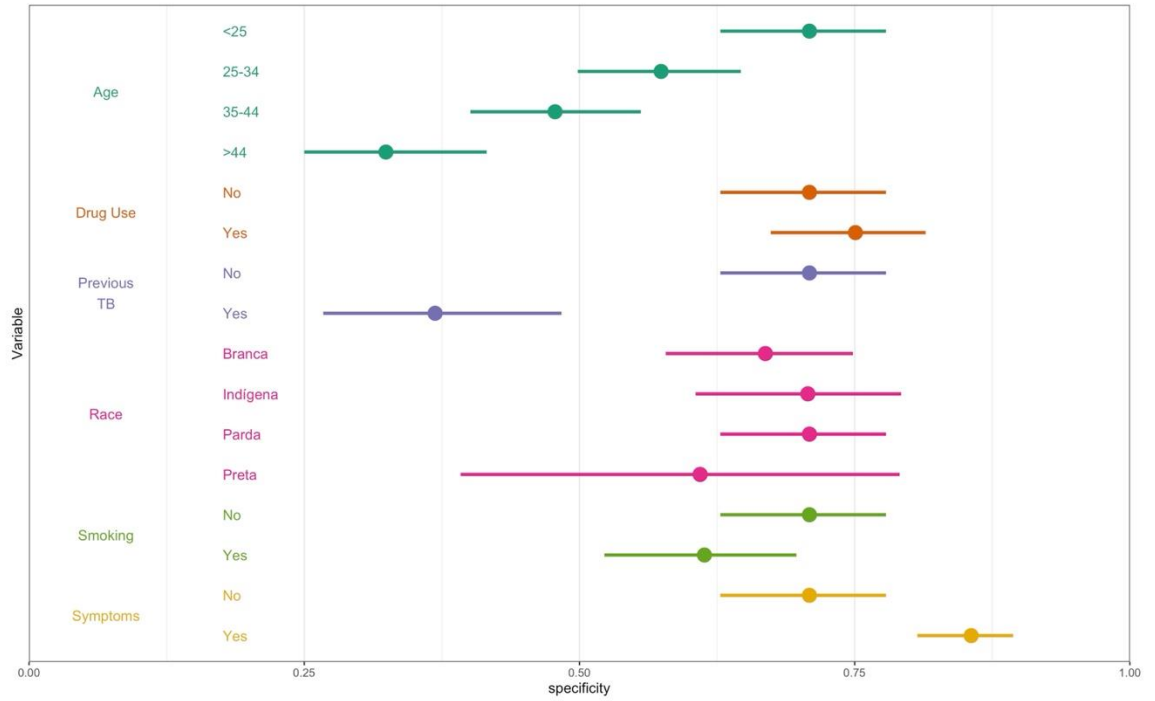
**Table E2.** Sensitivity of x-ray interpretation algorithms according to Xpert MTB/RIF G4 semi-quantitative load, using a threshold for specificity of 70%.

Sputum Xpert result	n	CAD4TBv6		LunitTB		qXR	
		Positive	Sens. (95% CI)	Positive	% Sens. (95% CI)	Positive	% Sens. (95% CI)
Negative/Very Low	60	47	78.3 (65.8-87.9)	53	88.3 (77.4-95.1)	49	81.7 (69.6-90.5)
Low	78	70	89.7 (80.8-95.5)	73	93.6 (85.7-97.9)	73	93.6 (85.7-97.9)
Medium/High	67	65	97.0 (89.6-99.6)	64	95.5 (87.5-99.1)	64	95.5 (87.5-99.1)

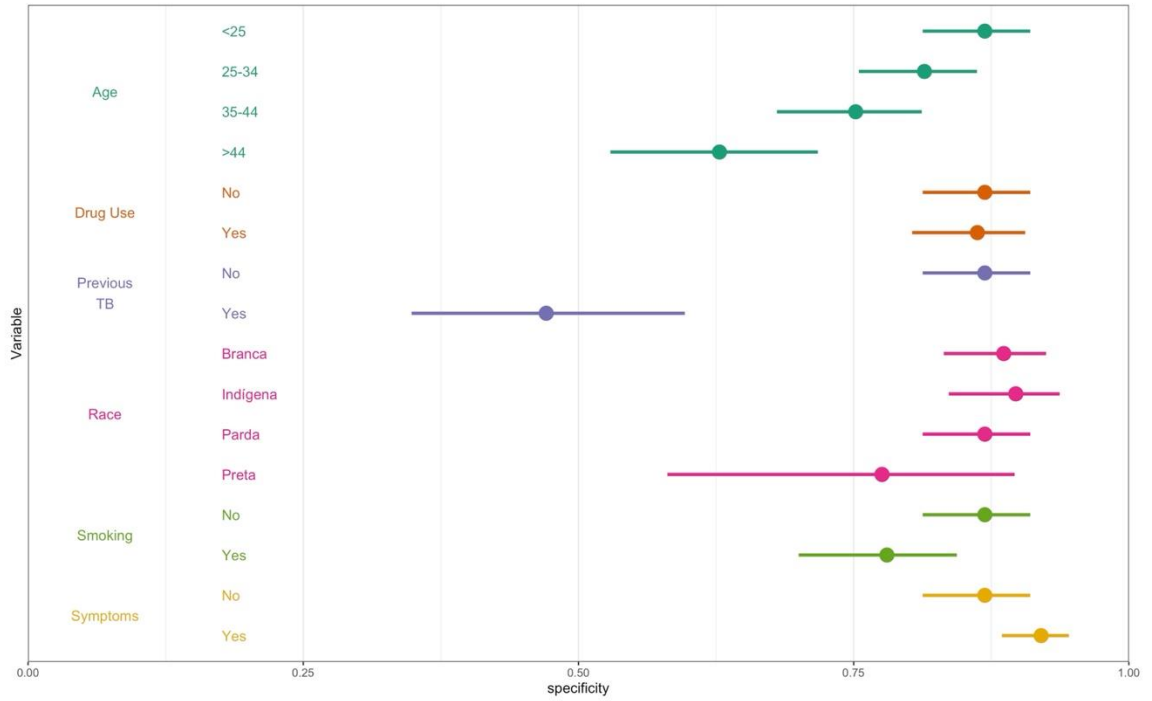
**Figure E1.** Flow diagram of study participants in mass screening and inclusion of participants in x-ray evaluation.



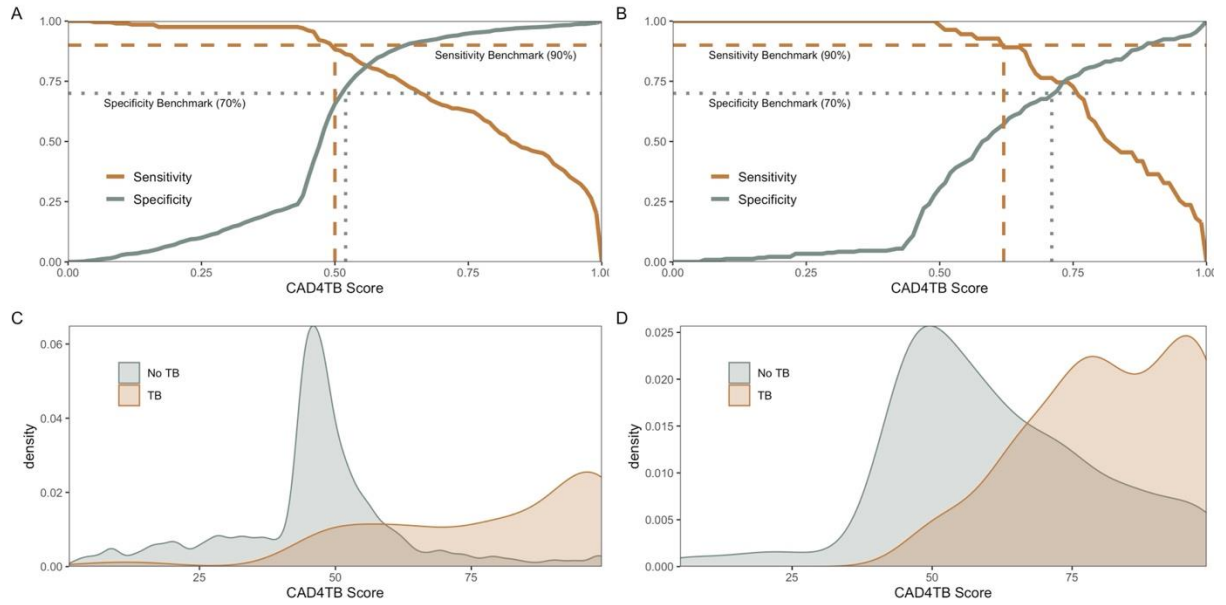
**Figure E2:** Specificity of CAD4TBv6, by sociodemographic characteristics and risk factors. Shown are the predicted margins for specificity and 95% confidence intervals from a multivariable logistic regression, holding sensitivity at 90%.



**Figure E3:** Specificity of qXR, by sociodemographic characteristics and risk factors. Shown are the predicted margins for specificity and 95% confidence intervals from a multivariable logistic regression, holding sensitivity at 90%.



**Figure E4:** Sensitivity and specificity according to CAD4TBv6 score threshold, with the WHO sensitivity (dashed line) and specificity (dotted line) benchmarks (top) among individuals without (A) and with (B) previous TB. Distribution of CAD4TBv6 scores for participants without (C) or with (D) previous TB (bottom).





**Figure E5:** Sensitivity and specificity according to qXR score threshold, with the WHO sensitivity (dashed line) and specificity (dotted line) benchmarks (top) among individuals without (A) and with (B) previous TB. Distribution of qXR scores for participants without (C) or with (D) previous TB (bottom).

## 6 CONCLUSÕES

Em suma, os resultados deste estudo sugerem que o uso do raio-x, aliado a métodos de inteligência artificial, pode ser uma estratégia de rastreamento em massa. Observa-se que o sistema LUNIT-TB apresentou a maior especificidade (94,8%) em relação aos demais sistemas, contudo os rendimentos foram semelhantes entre eles, desta forma, todos apresentam uma boa capacidade para serem utilizados como triagem, selecionando os indivíduos que deverão ser submetidos a outros métodos diagnósticos, como por exemplo, teste molecular. Vale ressaltar que foi observado uma potente capacidade do sistema LUNIT-TB, de discriminar os indivíduos com escores baixos, ou seja, aqueles que tiveram exame negativo para TB, dos escores mais altos, os quais foram diagnosticados com TB.

Destaca-se que o uso de raio-x interpretado por inteligência artificial, como método de triagem teve resultados promissores, uma vez que foram atendidos os critérios de especificidade para métodos de triagem recomendado pela OMS (especificidade > 70%), os quatro apresentaram especificidade acima de 70%, porém, é necessária cautela, visto que foi aplicado apenas em adultos, não sendo testado em outras faixas etárias.

A triagem em massa é um mecanismo importante para o controle da TB, principalmente, em ambientes de alta incidência, diversas pesquisas com mecanismos inovadores surgem com frequência. E neste estudo observou-se que o uso de inteligência artificial, seja com qualquer sistema, tem potencial promissor na triagem de indivíduos a serem submetidos a outros métodos diagnósticos, como o teste molecular.

Por fim, novas pesquisas para avaliar a sensibilidade são necessárias, para que melhoremos a generalização destas novas ferramentas tão promissoras, que ajudará na agilidade diagnóstica da TB, especificamente em ambientes de alta incidência e países com limitação de recursos em serviço de saúde.