2025
September

Impact Evaluation of use of MATCH-AI predictive modeling to target spots for TB active case finding in Pakistan

PROJECT TEAM

MERCY CORPS

Abdullah Latif, Nainan Nawaz, Frhan Ali

CGPH-PAKISTAN

Faran Emmanuel, Amna Mahfooz, Muhammad Asad Zaidi, Muhammad Wasim Ahmed, Tahira Ezra Reza, Fazal ur Rehman, Alina Shahid, Alizeh Faran, Faheem Baig

Contributions from team members involved in MATCH-AI implementation

Name	Designation	Organization
Dr. Adeel Tahir	Director TB Program	MC-Islamabad
Dr. Hasan Muhammad Tahir	Technical Advisor, ACF-TB Program	MC-Islamabad
Dr. Muhammad Umer Iqbal	Senior Program Manager - Case Finding	MC-Islamabad
Adnan Khaliq	Program Officer - ACF	MC-Islamabad
Muhammad Ishaq	Senior Program Officer - ACF	MC-Islamabad
Sanaullah Khan	Program Officer - ACF	MC-Islamabad
Shafqat Ali	Program Officer - ACF	MC-Islamabad
Dr. Shahid Naveed	Coordinator - ACF	MC-Islamabad
Dr Akmal Naveed	Director	ACD
Dr. Ali Ahmad	Manager TB PPM project	ACD
Pir Adnan	Project Coordinator	ACD
Nagina Said	Program Officer STM/ACF	ACD
Dr Muhammad Amir Khan	Chief Executive	ASD
Dr. Arshad Iqbal Syed	Program Manager NFM-III, PPM-TB Program	ASD
Farrukh Shaheen	Project Coordinator	ASD
Dr. Khalid Farough	General Manager, TB Project	GSM
Raja Farrukh	Assistant General Manager Project Operations	GSM
Muhammad Shoaib	Project Coordinator	GSM
Dr. Amber Elahi	Manager - Project Implementation Unit	MC-PIU
Temoos Ahmad	Project Coordinator	MC-PIU
Amjad Rafiq	Provincial Coordinator	MC-Sindh
Dr Ghulam Zainab	Regional Coordinator	MC-Sindh
Mehboob Ali Laghari	Project Coordinator	MC-Sindh
Dr. Sharaf Ali Shah	Executive Director	BCF
Dr Tajammul Baig	Program Manager - PPM	BCF
Altaf Ahmed	Project Coordinator	BCF
Dr. Saeedullah Khan	Team Leader (South)	MC-Quetta
Dr Amin Ullah	Project Manager	MC-Quetta
Syed Muhammad Kazim	Project Coordinator	MC-Quetta
Senator (r) Javed Jabbar	Chairperson	SPO
Mir Ahmed	Regional Manager	SPO
Muhammad Arsalan	Project Coordinator	SPO
Kashif Javed	Project Coordinator	SPO

Contents

LIST OF FIGURES

LIST OF TABLES

ABBREVATIONS

- **ACF:** Active Case Finding
- **AF-TB:** All-Forms Tuberculosis
- **AI:** Artificial Intelligence
- **ANOVA:** Analysis of Variance
- **B+:** Bacteriologically Confirmed (Tuberculosis)
- **CAD:** Computer-Aided Detection
- **CAD4TB:** Computer-Aided Detection for Tuberculosis
- **CGPH:** Center for Global Public Health
- **CMU:** Common Management Unit
- **COVID-19:** Coronavirus Disease 2019
- **DCNNs:** Deep Convolutional Neural Networks
- **DEPP-TB:** Digital Systems to Engage Private Providers for Tuberculosis
- **DFS:** District Field Supervisor
- **eCBS:** Electronic Case-Based Surveillance
- **EPCON:** (A commercial software firm, partner in developing MATCH-AI)
- **GB:** Gilgit-Baltistan
- **GF:** Global Fund
- **GIS:** Geographic Information System
- **GP:** General Practitioner
- **GPS:** Global Positioning System
- **ICC:** Intra-cluster Correlation
- **ICT:** Islamabad Capital Territory
- **IRR:** Incidence Rate Ratio
- **ITT:** Intention-to-Treat
- **JPRM:** Joint Program Review Mission
- **KIT:** KIT Royal Tropical Institute (partner in developing MATCH-AI)

- **KP:** Khyber Pakhtunkhwa
- **MATCH**: Mapping and Analysis for Tailored disease Control and Health system strengthening (approach developed by KIT)
- **MC:** Mercy Corps
- **MDR-TB:** Multidrug-Resistant Tuberculosis
- **NTP:** National Tuberculosis Control Program
- **PIU:** Provincial Implementation Unit
- **PR:** Principal Recipient
- **SDGs:** Sustainable Development Goals
- **SPOT-TB:** (Name of the trial evaluating MATCH-AI)
- **SRs:** Sub-Recipients
- **TB:** Tuberculosis
- **TWG:** Technical Working Group
- **UHC:** Universal Health Coverage
- **WHO:** World Health Organization
- **XDR-TB:** Extensively Drug-Resistant Tuberculosis
- **Xpert MTB/RIF:** A molecular diagnostic test for TB and rifampicin resistance, commonly referred to as "GeneXpert".
- **ACD:** Association for Community Development
- **ASD:** Association for Social Development
- **BCF:** Bridge Consultants Foundation
- **GSM:** Greenstar Social Marketing
- **MALC:** Marie Adelaide Leprosy Centre
- **MC Quetta:** Mercy Corps Quetta
- **MC Sindh:** Mercy Corps Hyderabad
- **MC-PIU:** Mercy Corps Project Implementation Unit
- **SPO:** Strengthening Participatory Organization

ACKNOWLEDGMENT

A heartfelt thanks to the National TB Control Program, Pakistan, for leading the initiative of evaluating MATCH-AI (Mapping and Analysis for Tailored disease Control using Artificial Intelligence) for its use in locating spots for TB screening activities. The MATCH-AI platform, co-developed by KIT (KIT Netherlands) with their partner EPCON, was a tremendous effort by both organizations and deserves special appreciation for introducing such an efficient platform in Pakistan.

Mercy Corps (MC) deserves special recognition for spearheading the implementation of the study and also contributing to the evaluation, along with the Center for Global Public Health (CGPH), Pakistan which played a crucial role for third party evaluation in the trial—from designing the stepped-wedge study, monitoring, data analysis, to the end line evaluation.

Special thanks to Dr. Faran Emmanuel, Principal Investigator, for his valuable time and expert guidance throughout the project. Heartfelt gratitude is also extended to all Sub-Recipients (SRs) involved (ACD-Association for Community Development, Association for Social Development-ASD, Green Star marketing-GSM, Bridge's Consultant's Foundation-BCF, Strengthening Participator Organization- SPO, Mercy Corps Project Implementation Unit-MC-PIU, Mercy Corps Quetta-MC-Quetta, Mercy Corps Hyderabad) for their coordination, support, and adherence to the stepped-wedge design throughout the implementation period.

The qualitative teams from CGPH and MC also deserve acknowledgment, as their work meaningfully complemented the study's quantitative components.

Finally, a special thanks to Mr. Abdullah, Senior Project Manager MC, for his leadership in guiding CGPH, the MC implementation team, SRs, and for effectively coordinating all aspects of the project with the Bill & Melinda Gates Foundation.

EXECUTIVE SUMMARY

The SPOT-TB trial was conducted to evaluate whether artificial intelligence-based (MATCH-AI) site selection could enhance tuberculosis case detection through active case finding chest camps in Pakistan, compared to conventional site selection methods based on local knowledge, staff experience, and historical data. The trial also aimed to strengthen the data management capacity of partner organizations by implementing Electronic Case-Based Surveillance (eCBS) and improving data entry, analysis, and routine programmatic use.

The study was implemented as a stepped wedge trial, with van teams (clusters) gradually transitioning from conventional to MATCH-AI-guided site selection. The trial was highly pragmatic, with the study being fully embedded within routine, programmatic implementation of ACF by Mercy Corps and its SRs. Implementation was carried out in three phases: (1) Preparatory design phase, (2) Pilot testing and Data collection phase and (3) Data analysis and dissemination phase which also included capacity building through workshops and manuscript development.

The use of AI-driven site selection demonstrated strong potential to significantly improve TB case detection compared to conventional methods, particularly when ACF camps were conducted within or near MATCH-AI-recommended sites. However, the Intention-to-Treat (ITT) analysis did not show statistically significant differences between the AI-guided and conventional arms, primarily due to implementation gaps, as many camps were held outside the 5 km buffer zone of AI-identified locations. MATCH-AI leverages TB notification records, population density, geospatial indicators, health system accessibility factors like immunization, local demographics, distribution of health services and countrywide estimates of disease risk to systematically identify potential high-efficiency active case finding sites, offering a data-driven alternative to conventional approaches that rely mainly on human judgment. Taken together, findings highlight that a hybrid approach integrating AI-generated insights with local knowledge and contextual expertise emerges as the most practical and impactful model for enhancing TB case detection.

Key Recommendations:

- Integrate MATCH-AI into national and provincial TB program planning, with clear operational guidelines and defined buffer zones around AI-recommended sites.
- Build trust and capacity among field teams through targeted training, engagement, and designation of AI “champions.”
- Assess cost-effectiveness of AI-guided approaches to support long-term sustainability.
- Ensure strong implementation support and adherence to protocols in the field.
- Conduct further research on long-term outcomes, including treatment success, reduced TB transmission, and efficiency gains.

› Regularly update MATCH-AI with high-quality, locally relevant data to maintain predictive accuracy.

1. INTRODUCTION & BACKGROUND

1.1. Tuberculosis in Pakistan

Pakistan is ranked as the 5th highest TB burden country and annually contributes to nearly 6.3% of new cases of TB worldwide. The incidence of TB per 100,000 population per year in the country is 277 and 48,600 deaths (WHO, 2025).

Of the 686,000 new cases estimated for 2023, a total of 480,136 (70%) were notified. Bridging this case-detection gap is crucial for the National TB Program and the National TB response. With 3619 laboratory confirmed cases of MDR-TB and another 982 patients identified with XDR-TB in 2023, Pakistan is also included in the list of 30 High Burden countries for drug resistant TB (WHO, 2025). Affected by the COVID19 pandemic, country's TB case notification dropped in 2020, however, in 2021, the health system is making efforts to revive its TB surveillance practices and the notification is steadily improving to reach 2019 levels.

1.2. The National response and Private sector engagement

The National response to TB in Pakistan is spearheaded by the National TB control program along with provincial TB control programs. Since 2001, the National TB Control Program Pakistan (NTP), has been implementing Global Fund supported grants for the public sector. The National TB Control Program and its implementing partners are working with a vision to reduce 50% prevalence of TB in the general population by the year 2025 through universal health coverage (UHC) and provision of quality TB care services on an equitable basis. The government of Pakistan is committed to end TB by 2030 as envisaged in Sustainable Development Goals (SDGs) and the World Health Organization (WHO) End TB Strategy.

Since 2007 Global Fund (GF) has also been supporting the private sector, with Mercy Corps as the first private Principal Recipient (PR) for GF TB grants. Mercy Corps has been implementing the Global Fund Grant in the four provinces (Punjab, Sindh, KP and Balochistan) and three regions (ICT, AJK and GB) of Pakistan. It is working through its sub-recipients and has intervened in the engagement of private health care providers and labs and active & enhanced case finding.

1.3. Active Case Finding for TB in Pakistan

Active case-finding (ACF) is defined by WHO as systematic screening for active TB, normally outside of health facilities but could also be undertaken at health facilities in a targeted population considered at higher risk of developing TB (WHO, 2021). WHO guidance states that systematic screening for active TB may be considered for geographically defined subpopulations with a high level of undetected TB (0.5% prevalence or higher), other

subpopulations that have poor access to health care, and vulnerable or marginalized groups (WHO, 2021).

The objectives of ACF are (i) targeted case-finding and (ii) prompt initiation of treatment to rapidly cure and render the disease non-infectious (WHO, 2021). Active case-finding (ACF) is an important component of the End TB Strategy, but it is considered resource-intensive ((Burke et al., 2021; Deya et al., 2022; Kranzer et al., 2013; WHO, 2021b). With more than three million people with incident TB not notified to public health authorities each year, closing the global case detection gap is essential to achieving the post-2015 End TB targets (WHO, 2024). Active case-finding (ACF)—that is, efforts to screen for TB in target populations (e.g., by geography and risk groups) outside of routine health services—may help to close the case-detection gap by linking patients to TB care early in their disease course. Recent evidence indicates that in addition to increasing case-detection in the short-term, ACF interventions, if sustained over time, can help reduce population-level TB incidence and prevalence (Burke et al., 2021; Corbett et al., 2010; Ortiz-Brizuela & Menzies, 2021; WHO, 2021b; Yuen et al., 2015). Further, there is an increasing recognition that a significant proportion of TB transmission takes place sub clinically, i.e., in the absence of any symptoms by the individual. This suggests that focusing on passive diagnosis alone may be insufficient for TB elimination.

Pakistan has invested significantly in improving its capacity for mass-screening and testing for TB including the use of mobile X-rays units with computer aided-detection (CAD) and molecular diagnostics, such as Xpert MTB/RIF. These interventions are currently led by Mercy Corps-Pakistan, a private-sector implementing partner of the NTP and its sub-recipients (SRs) and are supported by the Global Fund. While community-based ACF interventions (called “chest-camps” or “camps”) are now being conducted at scale in Pakistan, they are resource and labor intensive. Concerns remain regarding the yields and cost-effectiveness of systematic screening in programmatic settings, particularly given competing priorities within TB programs. A recent TB Joint Program Review Mission (JPRM) conducted by the Pakistan government, WHO and other technical agencies, identified low yields from ACF interventions and recommended the use of a geographically targeted approach for ACF in areas where TB prevalence is higher than the WHO threshold (JPRM, 2024).

1.4. Role of AI in improving planning and MATCH-AI strategy

During the last 10 years, artificial intelligence (AI) aided diagnostics systems have been developing, evolving at an unprecedented pace, which has led to its deployment and usage in clinical settings. Various medical image analyzing AI algorithms, based on deep learning and deep convolutional neural networks (DCNNs), have been utilized for radiographs reading at the same time(McClean et al., 2024; Rajpurkar et al., 2020). Such deep learning and DCNN algorithms are able to distinguish the features and characteristics of the TB-related abnormalities in chest X -rays. Considering the great improvement of AI-assisted TB diagnosis, computer aided TB screening software is a better substitute for physicians in digital chest

radiographs reading and analyzing, which was recommended by the updated guidelines of the World Health Organization (WHO) in March 2021. However, still there remains uncertainty about what kind of algorithm or AI should be developed and put into clinical practice.

In response to the JPRM recommendations (JPRM, 2024) and the emerging use of AI in screening and diagnostics Mercy Corps-Pakistan made a policy decision to utilize a geographically targeted approach towards ACF using MATCH-AI, an artificial intelligence (AI) software. This software utilizes a machine learning Bayesian modeling approach following the MATCH conceptual framework to identify potential high-efficiency active case finding sites to guide ACF site-selection (EPCON, 2022; KIT, 2025b; Zaidi et al., 2024) . The model was developed in 2019 and was being tested in Pakistan sporadically in different districts of Punjab with promising results. While the software may potentially support increase yields from ACF, it was not been rigorously evaluated in a prospective field-setting. The MATCH-AI trial (referred as SPOT-TB) therefore provided an opportunity to evaluate the impact of a geographically targeted approach towards ACF on TB detection through the use AI software in a real-world, programmatic setting (Zaidi et al., 2024).

2. OBJECTIVES OF THE TRIAL

2.1. Aims and objectives

The primary aim of the SPOT-TB trial was to measure the comparative yield (number of TB cases (B+) diagnosed) during ACF camps using a site selection approach based on predictions generated via an artificial intelligence software (MATCH-AI) versus the conventional approach of camp site selection using local knowledge, experience and analysis of historical data by field-staff. A secondary aim of this study was to strengthen capacity of the partnering organizations in routine programmatic data entry, management and analysis through implementation of existing electronic record systems (eCBS).

This study answered the following study questions:

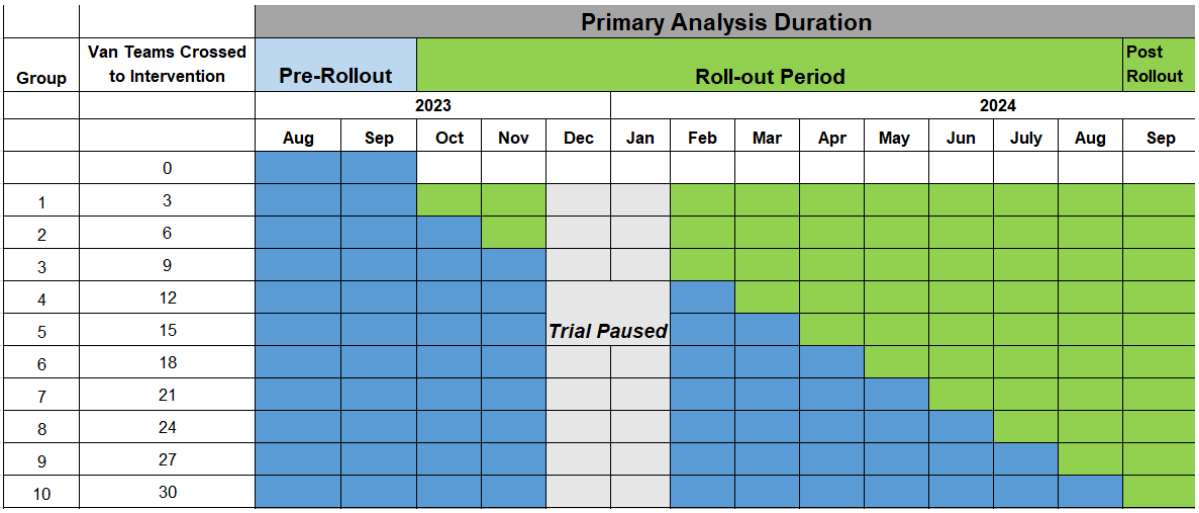
- Can computer algorithm-based targeting plans identify areas with high expected Bac+ TB rates among screened better than empiric targeting of communities for chest camps?
- Which targeting approach i.e., computer algorithm-based targeting OR case-finding through routine surveillance plans may be more effective in finding missing TB cases?

3. METHODOLOGY

3.1. Trial Design

The SPOT-TB trial was designed as a stepped wedge trial (Figure 1: Study roll out period). With this study design, clusters (van teams in this case) sequentially switch from control to intervention in a randomized manner. Initially all clusters are part of the control group, gradually shifting to the intervention group and by the conclusion of the trial all clusters are part of the intervention group. The SPOT-TB trial was designed to be highly pragmatic, with the study being fully embedded within routine, programmatic implementation of ACF by Mercy Corps and its SRs.

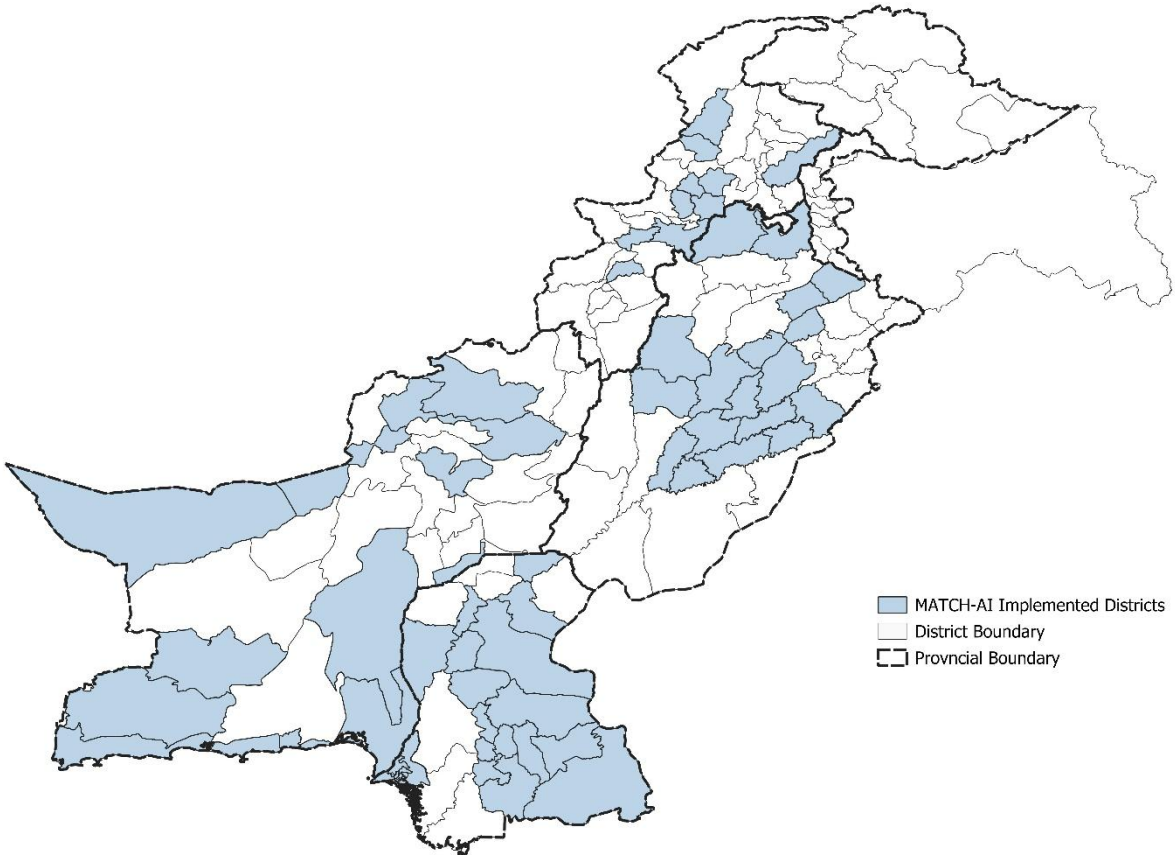
Figure 1: Study roll out period



3.2. Study Setting

The study was conducted in all four provinces of Pakistan within 68 districts (figure 2).

Figure 2: Study districts



3.3. Randomization

The unit of randomization in this trial was mobile X-ray unit and its associated team (van-team). Since each van-team works in more than one district, it was more appropriate to randomize the vans instead of individual districts (figure 3). The number of districts covered by each van-team varied from 1-3 depending on the size and population of the districts. A total of 30 van-teams were randomized to the intervention through a computer-generated list at the start of the trial. Randomization process and schedule were carried out independently by the research team.

The total proposed length of the study was 12 months (data collection period). At the start of the trial, no clusters were utilizing MATCH-AI. A two-month pre-rollout period was utilized to collect baseline control data and conduct pilot testing of the intervention. From month-3 onwards, 3 van teams were randomly selected to start using MATCH-AI for camp selection. By

the end of the trial, all 30 teams were using MATCH-AI for camp planning. The randomization was carried out through an online random sequence generator.

Figure 3: Randomization table

Van Registration Number										
Oct-23	Nov-23	Feb-24	Mar-24	Apr-24	May-24	Jun-24	Jul-24	Aug-24	Sep-24	Oct-24
JV-4815	JV-4815	JV-4815	JV-4815	JV-4815	JV-4815	JV-4815	JV-4815	JV-4815	JV-4815	JV-4815
JV-9956	JV-9956	JV-9956	JV-9956	JV-9956	JV-9956	JV-9956	JV-9956	JV-9956	JV-9956	JV-9956
RIS-1706	RIS-1706	RIS-1706	RIS-1706	RIS-1706	RIS-1706	RIS-1706	RIS-1706	RIS-1706	RIS-1706	RIS-1706
	JV2913	JV2913	JV2913	JV2913	JV2913	JV2913	JV2913	JV2913	JV2913	JV2913
	JV-9943	JV-9943	JV-9943	JV-9943	JV-9943	JV-9943	JV-9943	JV-9943	JV-9943	JV-9943
	RIS-1685	RIS-1685	RIS-1685	RIS-1685	RIS-1685	RIS-1685	RIS-1685	RIS-1685	RIS-1685	RIS-1685
		TV-913	TV-913	TV-913	TV-913	TV-913	TV-913	TV-913	TV-913	TV-913
		JV-3423	JV-3423	JV-3423	JV-3423	JV-3423	JV-3423	JV-3423	JV-3423	JV-3423
		JV-4635	JV-4635	JV-4635	JV-4635	JV-4635	JV-4635	JV-4635	JV-4635	JV-4635
			JV-9939	JV-9939	JV-9939	JV-9939	JV-9939	JV-9939	JV-9939	JV-9939
			JV-4812	JV-4812	JV-4812	JV-4812	JV-4812	JV-4812	JV-4812	JV-4812
			JV-9935	JV-9935	JV-9935	JV-9935	JV-9935	JV-9935	JV-9935	JV-9935
				JV-4637	JV-4637	JV-4637	JV-4637	JV-4637	JV-4637	JV-4637
				JV-9954	JV-9954	JV-9954	JV-9954	JV-9954	JV-9954	JV-9954
				JV-9936	JV-9936	JV-9936	JV-9936	JV-9936	JV-9936	JV-9936
					JV-4809	JV-4809	JV-4809	JV-4809	JV-4809	JV-4809
					JV-1742 (JV-9957)	JV-1742 (JV-9957)	JV-1742 (JV-9957)	JV-1742 (JV-9957)	JV-1742 (JV-9957)	JV-1742 (JV-9957)
					JV-8519	JV-8519	JV-8519	JV-8519	JV-8519	JV-8519
						JV-4803	JV-4803	JV-4803	JV-4803	JV-4803
						TV-883	TV-883	TV-883	TV-883	TV-883
						JV-4814	JV-4814	JV-4814	JV-4814	JV-4814
							JV-8518	JV-8518	JV-8518	JV-8518
							JV-3422	JV-3422	JV-3422	JV-3422
							JV-3896	JV-3896	JV-3896	JV-3896
								JV-9945	JV-9945	JV-9945
								JV-9942	JV-9942	JV-9942
								TV-847	TV-847	TV-847
									JV4808	JV4808
									JV-3424	JV-3424
									JV-9949	JV-9949
3	*6*	*9*	*12*	*15*	*18*	*21*	*24*	*27*	*30*	*30*

Districts			
District-1	District-2	District-3	District-4
Gujrat	Mandi Bahuddin		
Sukkar			
Faisalabad	Chiniot		
Hyderabad			
Sahiwal	Okara	Pakpattan	
Mardan	Nowshera		
Multan	Lodhran		
Bannu	Kohat	Hungu	
Sanghar	Matiari		
Rawalpindi	Attock		
Larkana			
Bhakkar	Layyah		
Khairpur	Kashmore	Shikarpur	
Sibi	Naseerabad	Jaffarabad	
Loralai+Dukhi	Killa Saifullah		
MirpurKhas	Umerkot	Tando Allahyar	
Khuzdar	Lasbela		
Quetta	Pishin	Chaghi	Noshki
Badin	Tharparkar	Tando Muhammed Khan	
Gujranwala	Hafizabad		
Mansehra	Battagram		
Naushero Feroze	Dadu	Banzirabad	
Karachi South	Karachi west	Karachi Kemari	
Lower Dir	Upper Dir		
Karachi Central	Karachi East		
Toba Tek Singh	Jhang		
Vehari	Khanewal		
Peshawar	Charsadda		
Karachi Korangi	Karachi Malir		
Kech	Panjgur	Gwadar	

3.4. Study Intervention

3.4.1. Overview of MATCH-AI

The primary intervention in this study was the roll-out of MATCH-AI, an artificial intelligence software that models Bac+ TB rates among those screened in created Thiessen polygons. This software is used to guide site selection of ACF camps. The software has been developed by the KIT Royal Tropical Institute in the Netherlands and EPCON, a commercial software firm (EPCON, 2022; KIT, 2025a). The MATCH-AI tool uses a Bayesian modelling approach to models Bac+ TB rates among those screened to a resolution of 10,000 population polygons. The Bayesian modelling and reasoning are based on the observed Bac+ TB rates over the total population screened. Using these rates, the platform learns how the contextual parameters influence the observed rate. Understanding these underlying drivers and their influence on the observed rate, allows the platform to estimate a likely risk for other areas with similar geospatial patterns and context. Within these newly identified 10 population clusters of interest (high risk), the EPCON platform then uses a predefined location as recommendation. The model integrates data from a range of sources including historical TB facility notification data, previous ACF data as well as contextual factors such as demographics, income, population density, health indicators such as vaccination coverage and climate related

variables to predict Bac+ TB cases out of screened. The software features a dashboard that allows for field-staff to visualize maps that models Bac+TB rates among those screened in created Thiessen polygons and help identify camp locations. The platform is self-learning and incrementally improves prediction as further data is entered into the system.

During the trial, several model upgrades were implemented. These refinements focused on improving chest-camp geolocation accuracy (automatic mapping of correct coordinates), integrating distance between screening sites and 10,000-population Thiessen clusters, and incorporating algorithmic improvements learned across EPCON's deployments in other countries. Continuous learning occurred as feedback from real screening outcomes informed later model iterations and data-flow behavior. Also, the Adjusted Number Needed to Screen (ANNS) formula prevented over-saturation by setting ANNS to null (excluding the Thiessen from recommendations) when 1) ≥90% of the population has been screened (positives found in last 24 months + negatives screened in last 12 months ≥ 0.9 × total population), or 2) Cases found exceed predictions (positives in last 24 months ≥ predicted total cases for the Thiessen). For individual tiles, within high-risk Thiessens, the system uses observed 5km chest camp rates when: 1) NNS change ≤60% between consecutive chest camps (stable yield), 2) ≥100 people screened in the 36-month rolling window, 3) At least one positive case has been identified. When these criteria were met, the model-predicted rate was overridden by the observed 5km rate before ANNS calculation and ranking. Exclusion mechanism: Tiles were excluded from recommendations (not just switched to model predictions) when: NNS increases by >60% (indicating declining efficiency/yield), OR Screened volume exceeds tile population within the 5km radius at least 25 people are estimated by population density in a 500m tile. These thresholds were calibrated to balance continued focus on high-yield areas while preventing repeated targeting once yields decline or saturation is reached.

In the intervention arm, camps were conducted primarily in locations guided by MATCH-AI. Van-teams were provided with locations in created Thiessen polygons in each district predicted by MATCH-AI and moved sequentially through this list to conduct ACF activities. Once van-teams had switched over to the intervention group, they continued to utilize MATCH-AI to guide their decision-making through to the end of the trial. A minimum (unspecified) number of camps (not more than 10%) were allowed independent of the MATCH-AI predictions due to any prior commitments of van-teams with local communities or due to logistical constraints in reaching AI directed sites.

3.4.2. Current Standard of Care

In the control sites, field-teams continued to utilize existing approaches towards camp site-selection. While not utilizing software to guide decision-making, field-staff included targeting in their approach towards site-selection. This is based on local knowledge, field-experience from previous camps, analysis of historical data in paper-based TB registries (TB 03 form of NTP) and consultations with local stakeholders such as district TB coordinators and health

officers. Camp sites were often focused in slum-dwellings and occasionally target key populations such as injection drug users. In addition, camps are often conducted in collaboration with local nonprofits, and private healthcare facilities such as General Practitioner (GP) clinics. However, the current approach of camp site-selection did not systematically use geographical targeting or quantitative methods to identify TB hotspots.

3.4.3. Selection of MATCH-AI camp sites

Camp Planning and Implementation

The SPOT TB trial followed a structured and systematic process for planning community-based TB screening camps and managing related data across all implementing sites (in line with the existing procedures at MC). This approach ensured standardization, transparency, and data consistency for effective monitoring and analysis.

Camp Planning

Each month, the DEPP team downloaded the updated list of high-priority (Top 15 locations) Thiessen polygons identified through the MATCH-AI model. These lists were then provided to the sub-recipients (SRs) responsible for conducting mobile screening activities, specifically for the vans that had been selected through the randomization process.

Upon receiving the lists, SRs incorporated the assigned spots into their monthly movement plans for the upcoming TB screening camps. This process ensured that camp locations were strategically selected based on data-driven insights. In cases where SRs chose to include additional sites not identified through the AI-generated list, they were required to provide documented justification for these manually selected locations to maintain transparency and programmatic accountability.

To generate locations for each Thiessen in the country, a 2 km tile model was used in combination with OpenStreetMap roads, intersections, and landmarks to identify potential sites for chest camps. This process was carried out independently of the modelling. Only high TB-risk Thiessens would then receive a specific location (from the pre-generated list) for potential screening. Since these locations were an approximation, local expertise; including both the MC and CGPH teams determined that the coordinates could be better optimized for screening yield by allowing a 5 km buffer around each coordinate. Therefore, implementing partners were permitted to conduct camps within a 5 km radius of the assigned coordinates.

Once the plans were finalized, SRs mobilized field teams to conduct screening activities at the identified spots. The workflow emphasized clear communication and coordination between DEPP team, and SRs, enabling timely planning, implementation, and reporting of camp activities across all participating provinces.

3.4.4. Eligibility Criteria

A total number of 68 districts of Pakistan having mobile X-ray and Gene Xpert testing facilities were enrolled into the study. Districts where mobile X-ray units and Xpert machines were unavailable were excluded to ensure that camp procedures and algorithms were uniform across the intervention and control arms.

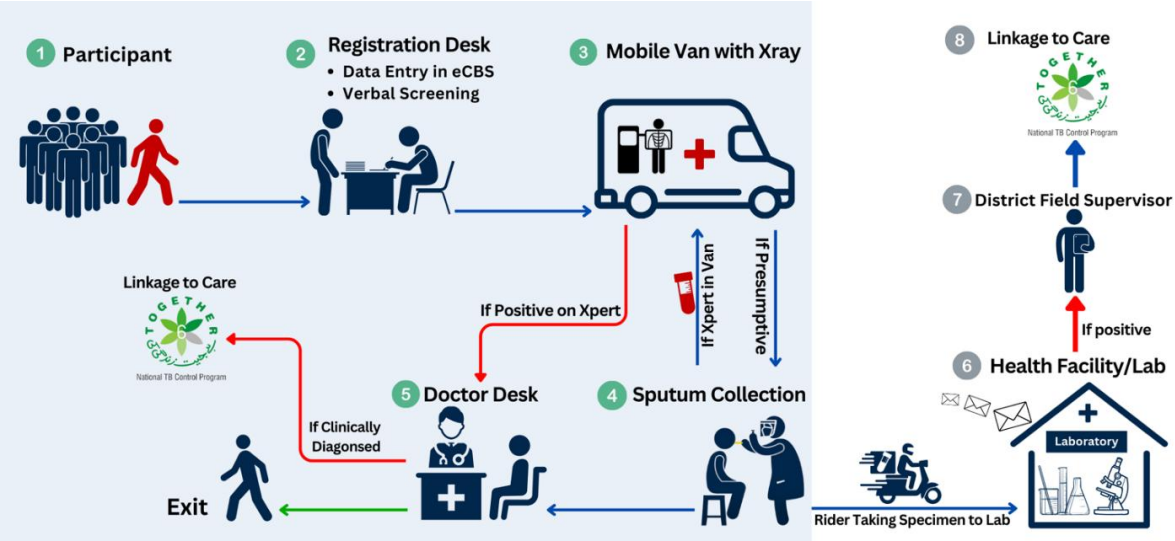
All individuals presenting to the camp sites were eligible to participate for screening, including those with previous history of TB disease as per routine ACF guidelines. Children <15 years of age and pregnant women were excluded from screening as per the NTP guidelines.

3.4.5. Participant Recruitment and Screening Algorithm

Participants were drawn from ACF camps conducted by Mercy Corps. Existing procedures for community mobilization, screening, diagnosis and treatment were utilized in both arms and remained unaffected by this trial. Once camp locations were finalized, meetings with "area notables," such as community elders and local politicians were conducted a few days prior to the activity. The goal of these meetings was to utilize their social networks and influence to increase community participation in the upcoming camp. Banners and posters were placed at prominent locations, pamphlets were distributed and announcements were made in mosques.

As shown in figure 4, on the day of the camp the mobile X-ray unit was brought to or near the intended camp location. Individuals were screened for TB using chest X-rays that were interpreted by computer-aided detection software (CAD4TB) by Delft Imaging. Individuals with the CAD4TB score of >60 were considered to be TB presumptive. Individuals with scores <60 were directed towards a medical officer for further evaluation. The medical officer may choose to request sputum investigation from these individuals on the basis of clinical examination. All TB presumptive (individuals with CAD4TB >60 or those identified by the medical officer), were encouraged to provide a sputum sample specimen onsite. Individuals who were not able to expectorate a sputum sample were provided containers with instructions on producing a morning specimen. No further workup is conducted for individuals that were not considered to be presumptive for TB. Collected sputum samples were tested using Xpert onsite in the mobile X-ray unit or transported to a nearby designated public or private-sector laboratory with Xpert testing facility. Individuals testing positive on Xpert or those diagnosed empirically by the medical officer were initiated on treatment.

Figure 4: Camp activity flowchart



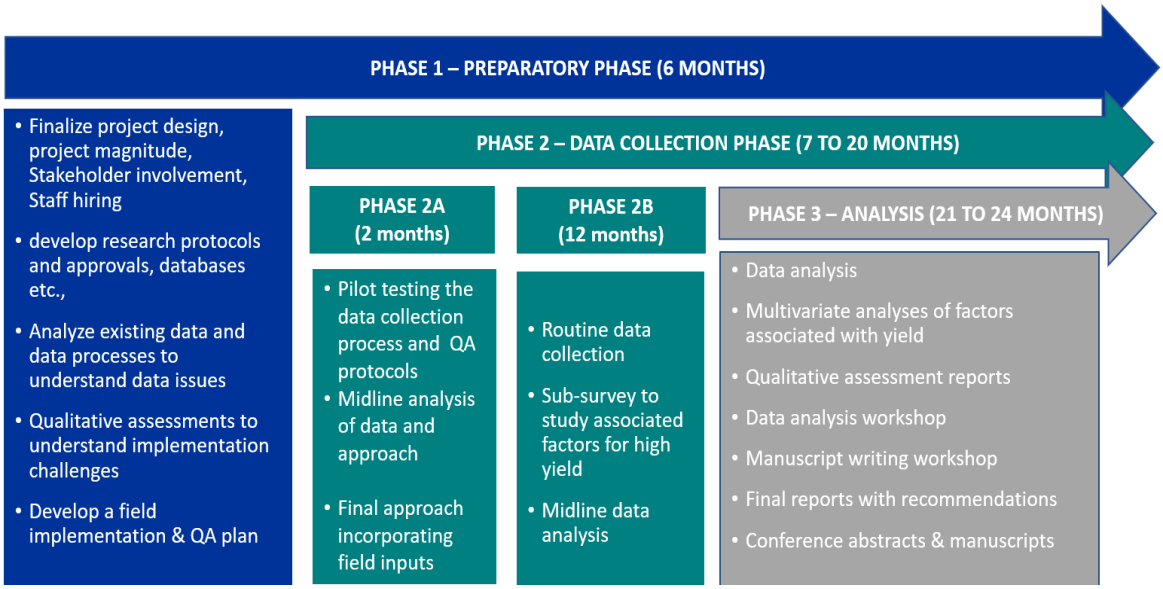
3.4.6. Informed Consent

Since the intervention involved a policy change in the approach to camp site selection, a consent form was developed in addition to the normal consent procedure adopted at the camp site for the participants, developed specifically for this change. The usual practice was followed: each participant received a registration slip at the registration desk with verbal consent taken at the registration desk plus the consent text was printed on the back. The consent was read out to participants, and their approval was obtained (consent form provided in the annexures).

4. IMPLEMENTATION PROCESS

Fig 5 shows the various phases of the study and the implementation schema of the project, highlighting various activities that each phase included.

Figure 5: Study phases and Implementation schema



5. PHASE I – Preparation phase (6 months)

The first phase established the logistical and conceptual foundations of the project. Several introductory and planning meetings were held with Mercy Corps, NTP, SRs, and other key stakeholders to share the project's purpose and scope, and to gather their input and support. This phase also included the development and finalization of the project work plan, along with defining activities, finalizing the project team, and clarifying staff roles. Following activities were conducted in Phase One:

5.1. Inception meeting

The inception meeting of the MATCH-AI Project was held to formally launch the initiative and bring all key stakeholders together. The session introduced the project's objectives, scope, and expected outcomes, with a focus on how MATCH-AI can support evidence-based site selection for Active Case Finding (ACF) in tuberculosis control. Roles, responsibilities, and timelines were discussed, along with mechanisms for coordination and communication. The meeting provided an opportunity to align expectations, address initial queries, and set the foundation for smooth implementation of the project.

5.2. Development of a technical working group

One of the key activities during this stage was the establishment of a Technical Working Group (TWG) to provide overarching support for the study. The TWG included technical staff from the National TB Control Program, Pakistan (Deputy National Coordinator CMU, Chief Research & Surveillance CMU and her team), the AI evaluation team from Mercy Corps—including the Digital Systems to Engage Private Providers (DEPP-TB) department led by Senior Project Manager, with his team from the ACF department—along with the CGPH team and an advisor from the University of Manitoba, Canada. The first TWG meeting was held on January 10, 2023, to introduce the project and present a tentative methodological approach for evaluation. This approach was refined with inputs from the TWG, donors, SRs, and PR, and was subsequently presented and formally approved in a TWG meeting on May 24, 2023.

5.3. Meeting with KIT and Gates Foundation technical team

Several meetings were held with the project donor, the Gates Foundation (GF), to discuss the project's modalities. These included a virtual call with Gates foundation, a meeting with KIT (who developed a first approach for the evaluation), and an in-country visit by Global Lead for TB Private Provider Engagement Gates Foundation in April, 2023. Based on discussions between GF, Mercy Corps (MC), and CGPH Pakistan, the evaluation methodology was revised

and strengthened. The updated protocol was shared with GF and, upon their approval, finalized for submission to the National TB Program's Ethical Review Board.

This phase also involved the development of research protocols for all project-related studies to be implemented in the subsequent phase. Ethical approval for these studies was formally submitted to the Ethical Review Board in Pakistan.

5.4. Camps visits

To develop a practical understanding of field activities and the implementation of ACF camps, a joint visit was conducted by CGPH and Mercy Corps to both a rural and an urban chest camp in Gujranwala, Punjab, in January 2023. These sites, randomly selected, were implemented by MC-PIU as the SR. During the visit, the District Field Supervisor (DFS) explained the complete process from registration through sputum collection. The team documented all protocols, met with camp personnel including the attending physician, and inquired about challenges faced by the DFS and community members attending the camps. Suggestions were also gathered to improve the overall performance and effectiveness of the camps.

Gujranwala camp visit by MC & CGPH

5.5. Meetings with SRs

In addition to visiting two ACF camps, multiple meetings were held with SRs across three of the four provinces and ICT to discuss the MATCH-AI study and obtain their input on the research methodology. During these meetings, SRs shared details of their organizational structures and explained the mechanisms of ACF implementation in their districts. All SRs confirmed their full participation in the research and endorsed the proposed approach.

SR visit (BCF Karachi) MC& CGPH

Two key decisions emerged from these discussions:

I. integration of additional research data collection into routine SR data collection processes, and
II. capacity-building sessions for staff on research methods, data analysis, and scientific writing. Meetings held with SRs during the inception phase included:

a) ACD Peshawar – May 5, 2023
b) MC-PIU – May 9, 2023
c) BCF Karachi – May 22, 2023
d) GSM Islamabad – May 26, 2023

5.6. Analysis of available data

Further activities during this phase included secondary analysis of available data to establish a baseline for the project, as well as qualitative assessments with the PR, SRs, and field staff. Mercy Corps shared aggregate data collected at chest camps from January to December 2022. The dataset did not include individual-level records of people screened or camps conducted. Analysis of the aggregate data revealed several quality concerns, including missing values, miscoded entries, and implausible data points. Nevertheless, the analysis provided a comprehensive overview of camps held across provinces and districts, along with recommendations for improvement.

In response to requests from SRs and suggestions from MC, capacity building was incorporated as a core component of the study. This initiative focused on strengthening the human resource capacity of each SR, equipping program staff with the MATCH-AI technique for location selection, and enabling them to self-evaluate performance and track program outcomes regularly.

The capacity-building initiative also aimed to familiarize PR and SR staff with the use of routine program data for learning, identifying areas for improvement, and setting goals. By leveraging existing management information systems, the initiative promoted continuous analysis and use of program data to inform decision-making and strengthen program performance.

6. PHASE II – (A and B)

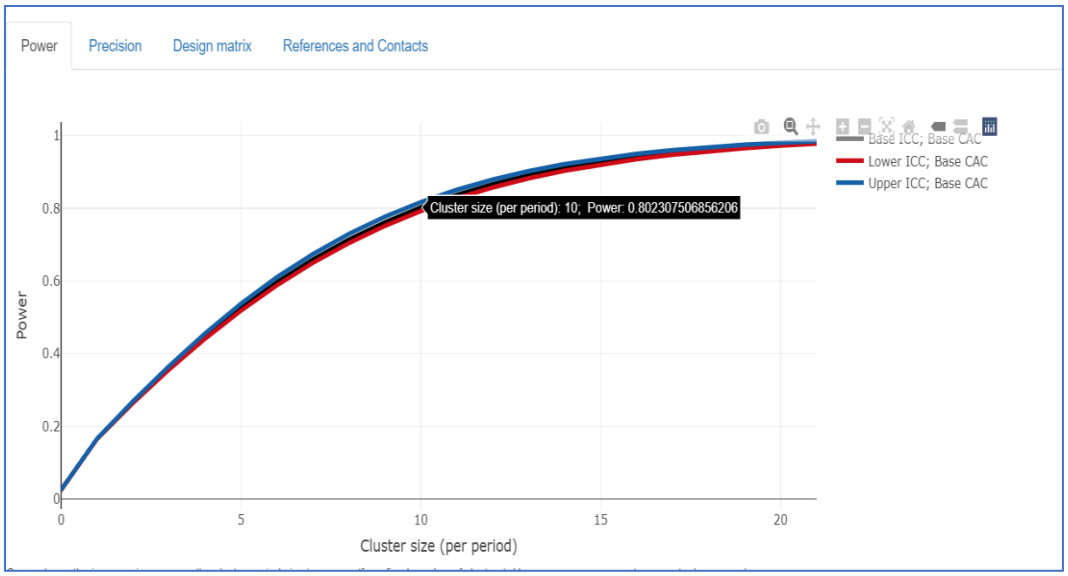
6.1. Sample Size

The sample size was calculated using prior data from the last five years (baseline data) from year 2019-2023, we calculated a mean of 0.88 B+ diagnosed per camp with an over dispersion of 1.34 (figure 6). We calculated intra-cluster correlation (ICC) of 0.12 using one-way analysis of variance (ANOVA) of B+ and van-clusters. We subsequently used 'The Shiny CRT X Calculator: Power and Sample size for Cluster Randomized Trials Test' to conduct the sample-size calculation for stepped-wedge trial using a cross-sectional sampling structure assuming an exchangeable correlation structure and significance level of 0.05.

We anticipated that each van would conduct approximately 10–13 camps per month (cluster-size), accounting for exclusion of camps in districts without Xpert testing and cancellations of camps due to external factors. With 10 steps and three vans switching per cluster, the trial was powered at 80% to detect a 20% increase in B+ in the intervention arm assuming an average of 10 camps per month or 3300 total camps over the study period. A power of 90% is achieved if 13 camps per month were undertaken during the study period or 4290 total camps over the study period.

By using ongoing implementation to conduct the trial, this study was able to achieve substantial precision to detect the effect of the intervention and be sufficiently powered to conduct subgroup analyses. Since this intervention was being rolled out programmatically, the sample size was not be capped and all camps included in the study period were included from vans meeting the eligibility criteria.

Figure 6: Sample size calculation and power estimation



6.2. Baseline data collection

Data collection for the trial began in August 2023 for all 30 vans selected for the study. Data collected in August and September were considered as baseline months. During this period, camp site selection was carried out in all districts by SRs using their existing, conventional methods. The data was recorded in the standard formats used by the SRs and reported to Mercy Corps. The baseline months were included in the trial statistical analysis period as control months.

6.3. Data collection from October 2023

From October 2023 onwards, a randomization plan was implemented for all 30 vans using a step-wedge design, with three vans randomized each month. Once randomized, all camps conducted by the selected vans were required to follow the recommendations provided by Mercy Corps using MATCH-AI. EPCON generated an updated list of recommended spots for each district, which were reviewed with MC and subsequently shared with the SRs. The list of recommendations remained unchanged for the duration of the trial and teams moved through the same list of locations once randomized to the intervention (approximately 10-15 initially and later up to 30 locations per district).

For vans entering the intervention phase, SRs were instructed to conduct majority approx. 90% of their camps using MATCH-AI recommendations. However, given concerns raised by SRs during initial meetings, around 10% flexibility margin was allowed for camps to be selected through their conventional methods but also instructed to follow as many recommendations as possible. In such cases, SRs were expected to record the reason for not following MATCH-AI recommendations.

There was a pause in camp activities during December 2023 and January 2024, as these were financial closing months for SRs and transition to the new grant. Screening activities resumed in February 2024, but the number of camps remained low due to the holy month of Ramadan followed by Eid.

6.4. Midline analysis

A comprehensive midline analysis of the MATCH-AI trial was conducted. The analysis compared AI-guided location selection with conventional site selection across all mobile X-ray screening camps in 68 districts. Results showed that adherence to AI recommendations—particularly within 5 km of suggested sites—improved TB detection yields. However, adherence to the intervention was lower than expected and teams were more proactively encouraged to follow the recommended locations.

6.5. Changes in study protocol and extension of project end date

The MATCH-AI project duration was extended from its original end date to September 2025. The extension was necessary to address operational challenges in data collection, including a pause in the trial for two-months during the new grant transition and elections in Pakistan.

7. PHASE III

7.1. Data management

Data management was coordinated across multiple institutional levels to ensure data quality and uniformity. Each month, the data unit of MC compiled data from all camps conducted by SRs nationwide. The MC team performed basic data cleaning, validation, and consistency checks before transmitting the consolidated datasets to DEPP.

DEPP subsequently shared the datasets with the CGPH. The CGPH team combined data received from all partners into a unified database, carried out detailed cleaning and harmonization of variables, and conducted analyses to generate monthly monitoring and performance reports. These reports provided key insights for program review, decision-making, and feedback to implementing partners, ensuring continuous improvement of camp planning and execution processes.

7.1.1. Initial screening and exclusion rules by CGPH

- Removed all records related to camps conducted using conventional method
- Removed non-study districts which were not included in the randomized van allocation
- Exclude records with zero participation (no attendees).
- Remove records where X-ray, Presumptive, and Xpert test fields were all zero.
- Exclude districts with only manual camps for analyses that require comparative MATCH-AI implementation.

7.1.2. Flagging and classification

- Camps conducted before the randomized allocation or outside the trial timeline were flagged as Manual Camps.
- Camps were classified as Control, ITT Manual, ITT Per Protocol (≤5 km), or Others (>5 km or missing distance) to support both ITT and adherence analyses.

7.1.3. Geospatial processing

- GPS coordinates of camp sites were validated (checking for broken links, obvious errors) and used to compute geodesic distances between MATCH-AI recommended coordinates and actual camp coordinates. The haversine formula or GIS library was used for accurate distance calculation. Camps within 5 km of a recommendation were marked validated per-protocol.

7.1.4. Master dataset formation

- After cleaning and deduplication, the final analytical dataset, with detailed indicators for participant counts, diagnostics (X-ray, Xpert Test), CAD4TB scores, and case outcomes. Subsets were created for ITT and per-protocol analyses and others were beyond 5-km. This included data from the extended project duration.

7.2. Qualitative study

Building on the SPOT-TB trial, a qualitative study was conducted with the following objectives:

1. The perspectives and experiences of healthcare workers while using MATCH-AI to identify the locations of ACF camps.
2. To compare the MATCH-AI campsite selection approach with the contemporary method of campsite selection (using routine TB notification-based targeting guided by local knowledge) and discuss the barriers and facilitators of both approaches with the implementers.
3. To investigate how the revised approach has improved staff efficiency in terms of ACF planning, workloads, processes, and data management.

Five sub-recipients (SRs) implementing ACF were purposively selected. Two districts were chosen under each SR to ensure representation of those with experience in both AI-assisted and routine ACF camps. Study participants included 30 in-depth interviews (IDIs) with Regional Coordinators (RCs), District Field Supervisors (DFSs), District Tuberculosis Coordinators (DTCs), and Project Managers (PMs). Two focus group discussions (FGDs) were held, one with 4 staff from the ACF unit at Mercy Corps and one with 8 project coordinators across all SRs.

Data was analyzed thematically using an inductive approach which involved familiarization with the transcripts, generating initial codes, grouping codes into categories, developing and refining themes, and reporting findings with supporting participant quotations. The Sociotechnical Systems (STS) framework was found to be a suitable lens for understanding the complex interplay of organizational, cultural, and human factors that influence the adoption and use of AI-assisted camp planning, providing a holistic view of MATCH-AI implementation in routine ACF delivery.

Six themes emerged from the analysis: MATCH-AI was perceived to reduce personal bias and political pressure in camp site selection. Staff appreciated introduction of new technology and saw potential in using AI for more transparent planning. Staff acknowledged that it made their work easier as they no longer needed to look at registers or ask doctors to decide a suitable location for camp. Many expressed interests in its continued use with improvements.

Understanding of MATCH-AI was uneven across staff levels. Many field staff, especially DFSs, were unclear about how the software functioned, leading to lower trust and limited ownership. It was also mentioned by senior participants that even if field staff is oriented, their limited technical backgrounds may still hinder comprehension.

While AI helped identify new locations and hence cases that would have otherwise gone missing, staff consistently relied on local insights to meet yield targets. Gender norms, safety, accessibility, terrain, and distance were seen as factors the AI did not fully account for.

As DFS were assigned clusters, some reported increased workload while some remained free if the AI location didn't fall in their designated district resulting in a disbalance of workload. DFSs faced increased travel requirements due to AI-recommended sites being farther and unfamiliar. Limited compensation, coupled with the dual burden of digital and paper-based reporting, added to their workload.

The transition to electronic case-based surveillance (eCBS) was hampered by poor connectivity in provinces like KPK and Balochistan. The application had a slow interface and not fully integrated with the system. Instead of improving efficiency, digital tools created frustration among users.

Pre-camp activities specially conducting notable meetings remained the single most critical determinant of camp success according to those interviewed. Engagement with local leaders (religious figures, teachers, tribal elders) was key for trust-building, turnout, and stigma reduction, regardless of whether the camp was AI-assisted or manual.

Overall, some staff liked MATCH-AI because it helped them decide where to set up camps more quickly, and reduced personal and political bias in site selection. They felt it saved time and made things more systematic. On the other hand, some staff members preferred manual methods because they trusted their local knowledge and felt the technology missed ground realities like distance, weather, security, or community needs. These staff members felt that MATCH-AI is useful for TB screening and selection of locations, but it works best when combined with local insight and involvement rather than used alone.

7.3. Workshops to enhance analytical and data utilization capacity

A series of workshops were organized to strengthen the research and data utilization capacity of PR and SR staff. This began with a one-day research seminar for all SRs, followed by three intensive workshops (3–5 days each) with selected PR/SR staff. These workshops, held at three-month intervals, provided hands-on training in data analysis, evidence generation, and scientific writing. Participants for the capacity-building workshops were identified in consultation with Mercy Corps and selected based on an eligibility criterion communicated to the SRs.

The three-day Data Analysis and scientific writing workshop was conducted from 29 to 31 July 2025 with the objective of strengthening participants' skills in data handling, statistical interpretation, and scientific writing for publication. The sessions were designed in a classroom teaching style to combine theoretical understanding with practical, hands-on exercises that ensured participant engagement.

7.4. Project closure and final report with draft manuscript

The MATCH-AI project was formally closed in September 2025. As part of the closure activities, a draft manuscript on the MATCH analysis was completed, and all project deliverables were ensured in line with agreed timelines. This milestone marks the successful conclusion of the trial and consolidation of outputs for dissemination and future programmatic use.

7.5. Dissemination

A comprehensive dissemination plan was developed to ensure that the findings and lessons from the MATCH-AI project are effectively shared with key stakeholders. The dissemination of the MATCH-AI evaluation will be carried out through an interactive multi-stakeholder workshop bringing together government representatives (national and provincial), implementing partners, developers of the model KIT and EPCON, NGOs, donors, academia, and technical AI partners. The workshop will present key quantitative and qualitative findings, highlight operational challenges, and engage participants in co-developing recommendations for the integration and scale-up of MATCH-AI in TB active case finding. Feedback will inform revisions to improve the system's functionality and guide updates to national guidelines. A consolidated workshop report and recommendations will be shared with all stakeholders. Additional dissemination activities include internal briefings with program teams, and multiple research publications and conference abstracts. Integration of MATCH-AI into national ACF guidelines and digital systems such as DHIS2 is also planned, ensuring sustainability and institutional ownership.

8. STATISTICAL ANALYSIS AND RESULTS

8.1. Analysis period

For the purposes of the final quantitative analysis, it was agreed by the Data Management Committee to utilize data from the 30 vans randomized to the intervention for a total study duration of 12-months, as originally specified in the published trial protocol. This included two months of baseline (control) data and one month where all 30 vans had crossed over to the intervention at the conclusion of the trial. The total trial duration for which the analysis has been presented is therefore from Aug 2023 – Sept 2024, with a pause for 2-months in Jan-Feb 2024, when no camps were conducted.

8.2. Analytical definitions

Since this was a pragmatic trial, the analysis was conducted using various approaches to better estimate the intervention effect. The following types of analyses were conducted:

- **Intention-to-Treat (ITT)**: Under the ITT analysis, all camps were included for analysis once van teams had crossed over to the intervention arm. This was irrespective of whether teams followed the recommendations. This analysis provides the "real-world" estimation of the benefit of the intervention since it considers programmatic realities and operational constraints that may limit adherence to the intervention. The ITT results were considered as the primary analysis as has been previously recommended for superiority trials (Ahn & Kang, 2023).

- **Per-Protocol**: The per-protocol analysis was restricted to only camps marked during reporting by the van-teams as being conducted using the MATCH-AI software for the intervention arm. During implementation, independent verification of camp GPS locations suggested that teams were frequently marking camps as being conducted using MATCH-AI during reporting, however, they were not being conducted at the recommended sites. The per-protocol camps were therefore not considered for a detailed analysis.

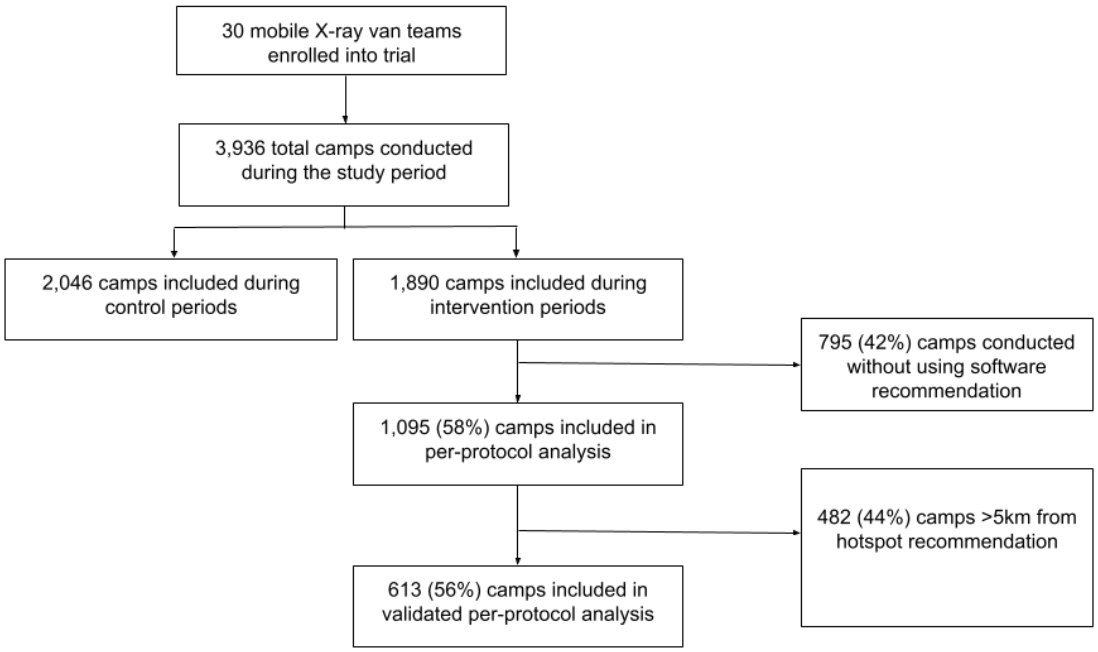
- **Validated Per-Protocol**: The "validated" per-protocol analysis was restricted to camps in the intervention group that were conducted within the 5-km radius of the recommended camp location, as verified independently using GPS coordinates by the study team. For the purposes of brevity, only the validated per protocol analyses are presented.

Camps in the control arm remained the same across all three analysis.

8.3. Descriptive analysis

Between August 2023 and September 2024, a total of 3,396 camps were conducted in the study districts of which 2,046 were conducted during the control periods and 1,890 were conducted during the intervention periods (Figure 7). Among camps conducted during the intervention periods, 1,095 (58%) were identified as being conducted with support of the AI software by field-teams and included in the per-protocol analysis. Among these, 613 camps (56%) were verified as being conducted within a 5-km radius of the recommended locations and were included in the validated per-protocol analysis.

Figure 7: Total number of camps conducted



We used frequency statistics to describe camps and the screening care cascade for individuals diagnosed in the study. We compared the means at each level of the care cascade to determine differences between the intervention and control groups using Student's t-tests. Due to challenges in the rollout of the electronic data-capture tool (eCBS) additional baseline factors such as healthcare access were not captured and could not be analyzed and compared between the study arms.

A total of 269,254 individuals participated in the camps, among whom 195,173 were screened for TB using CXR while the remaining were screened symptomatically (Table 1). The median number of camp participants and individuals screened with CXR was 68 (IQR 55 – 80) and 50 (IQR 39 – 62) in the control groups and 65 (IQR 54 - 78) and 47 (IQR 34 – 58) in the intervention group, respectively ($p < 0.01$). A greater proportion of individuals in the intervention group had an abnormal CXR (13.0%) relative to the control group (7.8%), whereas a greater proportion were symptom screen positive in the control group (18.0%), relative to the intervention group

(12.6%). The number of individuals with abnormal CXR per camp was higher in the intervention group (median = 3, IQR 1 – 5), relative to the control group (median = 2, IQR 0 – 5) (p <0.01), whereas the number of individuals that were symptom screen positive per camp was higher in the control group (median = 9, IQR 6 – 15) (p <0.01), relative to the intervention group (median = 6, IQR 3 – 10) (p <0.01). Both groups had a similar number of individuals that were tested on Xpert per camp (median = 9, IQR (6-13), p =0.94).

A total of 1,923 B+ TB cases were identified in the control group and 1,599 B+ TB cases were identified in the intervention group. A total of 5,053 AF-TB cases were identified in the control group and 4,249 AF-TB cases were identified in the intervention group. A total of 66 rifampicin-resistant cases were identified in total among both study groups (1.8% of B+TB cases).

Table 1: Screening and case-detection cascade

	Control Camps (N = 2,046)				**Intervention Camps (N= 1,890)**				**p-value**
	n	**%**	**Median (Per Camp)**	**(IQR)**	**n**	**%**	**Median (Per Camp)**	**(IQR)**	
Camp Participants	142,359		68	(55-80)	126,895		65	(54-78)	<0.01
Individuals Screened with CXR	105,965	74.4%	50	(39-62)	89,208	70.3%	47	(34-58)	<0.01
Abnormal CXR	8,223	7.8%	2	(0-5)	11,563	13.0%	3	(1-5)	<0.01
Symptom screen positive	25,558	18.0%*	9	(6-15)	15,969	12.6%*	6	(3-10)	<0.01
Total Presumptive	33,781	23.7%*	12	(9-18)	27,532	21.7%*	11	(8-16)	<0.01
Total Xpert Tests	21,009	62.2%	9	(6-13)	19,382	70.4%	9	(6-13)	0.94
Bacteriologically Confirmed TB	1,923	9.2%	1	(0-1)	1,599	8.2%	1	(0-1)	<0.01
All-Forms TB Cases	5,053	4.8%**	2	(1-3)	4,249	4.8%**	2	(1-3)	<0.01
Extra-pulmonary TB	188	3.7%	0	(0-0)	119	2.8%	0	(0-0)	0.03
Rifampicin Resistant TB	35	0.7%#	0	(0-0)	31	0.7%#	0	(0-0)	0.93

* Proportions among camp participants

** Proportion among individuals screened with CXR

Proportions among individuals with All-Forms TB

8.4. Primary analysis

The primary analysis utilized a mixed (multi-level) regression approach to allow for appropriate adjustments to the outcomes. A generalized linear mixed *negative binomial* regression was utilized to model the counts of bacteriologically positive TB diagnosed per camp whereas a generalized linear mixed *Poisson* regression was utilized to model counts of All-Forms TB. This was to support model goodness of fit in both instances. This was due to the fact there was over-dispersion of data for B+ TB (Figure 7), with a large proportion of camps where no cases detected whereas AF-TB camp outcomes were more normally distributed (Figure 9).

Figure 8: Distribution of bacteriologically confirmed TB diagnosed from mobile X-ray camps

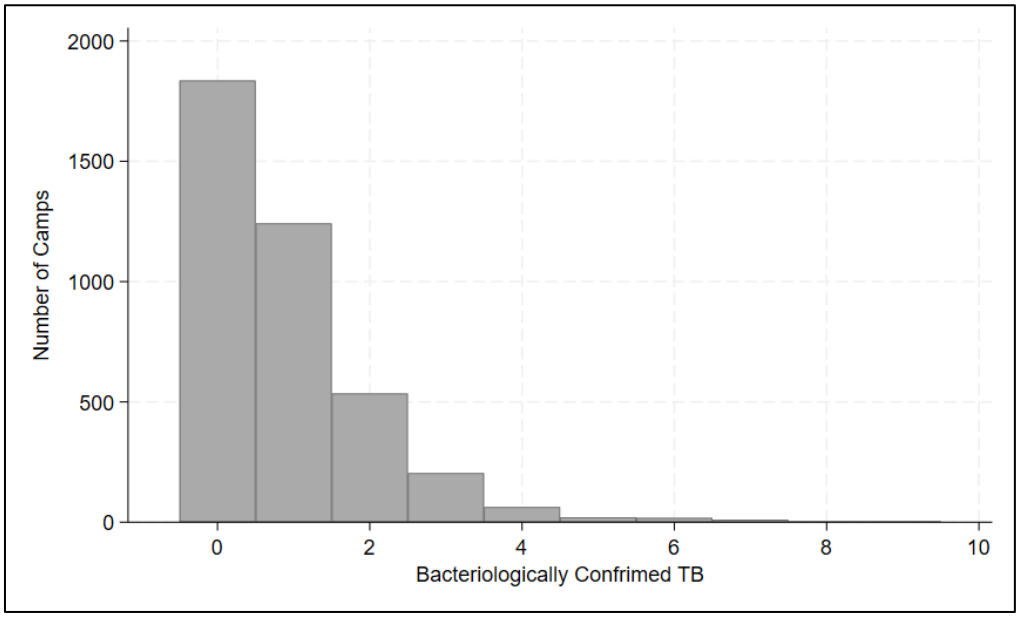
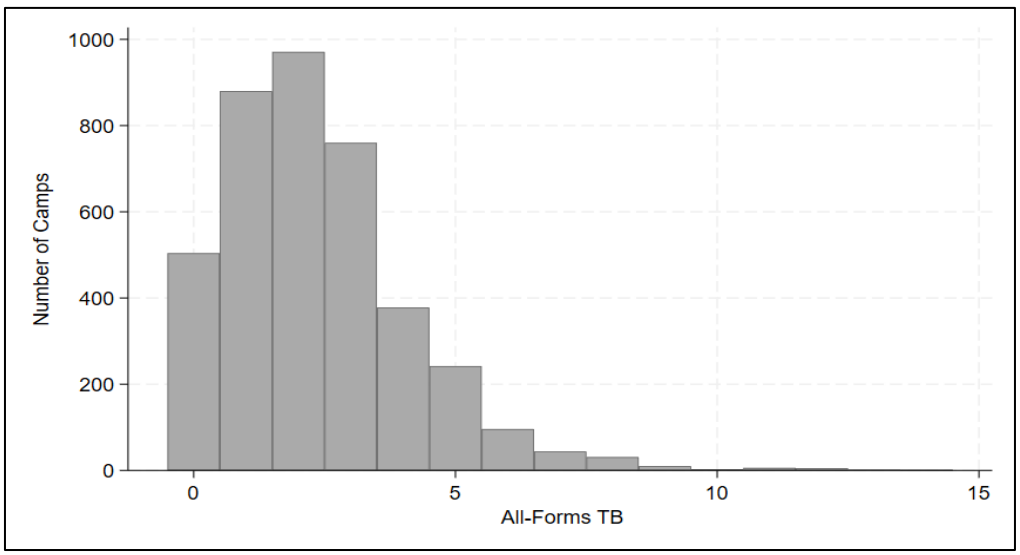
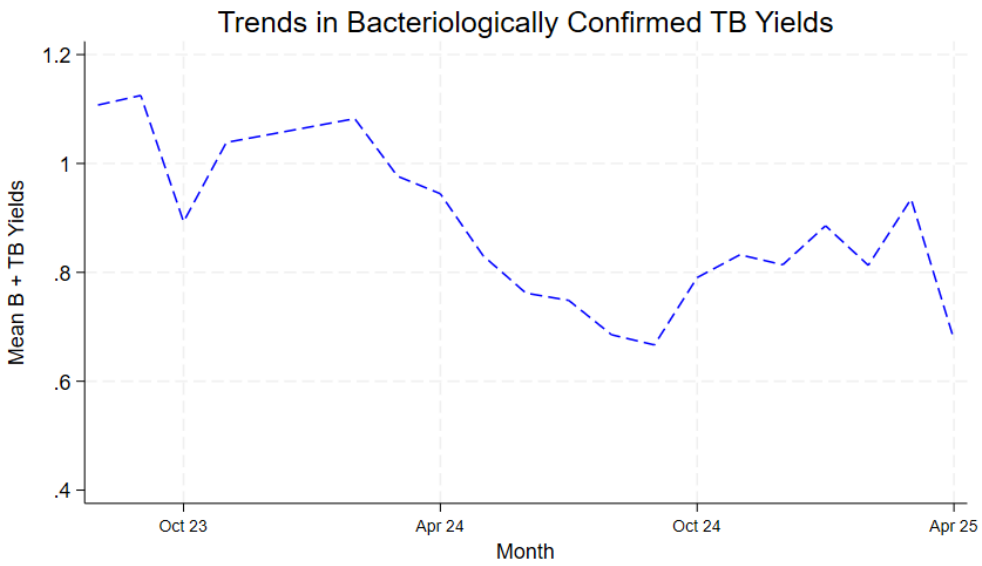


Figure 9: Distribution of All-Forms TB diagnosed from mobile X-ray camps



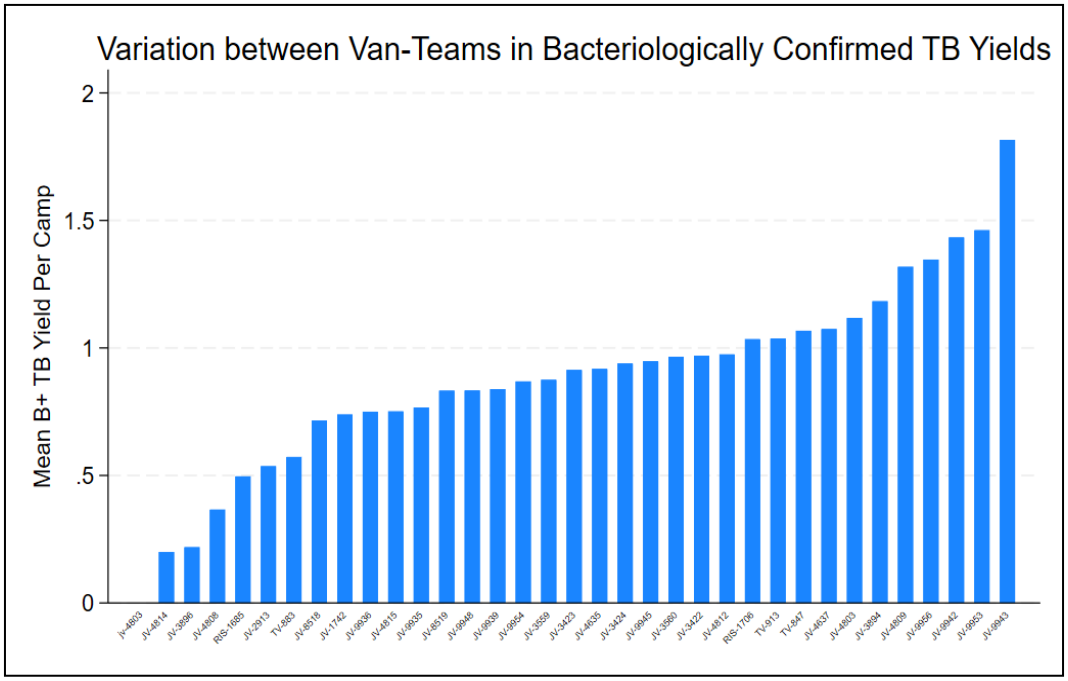
There were two major adjustments made in the regression models. These adjustments were pre-specified in the trial protocol. To account for secular trends (seasonal variation) in the incidence of TB in Pakistan, a fixed effect for time was added into the model. This was carried out by including each month of the intervention (Figure 10) as a categorical variable in the model. TB notifications were typically higher during the summer months in Pakistan; however, it remains unclear whether this pattern is due to health-seeking behavior, disease pathophysiology, or local contextual factors (such as the non-flooding or non-festival season).. Including each month as a categorical variable, allowed for comparisons between the same time-period for the intervention and control groups. This approach is commonly utilized in stepped-wedge trials and inclusion of fixed effect has been recommended previously by trial methodologists. A fixed effect assumes that the seasonal trends were the same across all randomized units. This was a reasonable assumption since the seasonal effect is unlikely to different between different districts of Pakistan.

Figure 10: Monthly trend of bacteriologically confirmed TB yield per camp



There was also significant clustering of outcomes within the randomization unit (van-teams), i.e., the outcomes varied significantly depending on the van-team (Figure 11). This was likely due to variation in underlying TB prevalence. It may also be related to variation in team performance and capabilities. To account for this clustering, van-units were added as a random effect in the model. This assumes that each van-unit has an independent association (slope and intercept of regression lines) between intervention and control arms. The final model combines information across various clusters. Adjustments were carried out at the van-level and not at the district level since the unit of randomization was the van-team. This was pre-specified in the protocol.

Figure 11: Variation in bacteriologically confirmed TB yields by van-teams



Camps were categorized into intervention or control based on their exposure to the MATCH-AI intervention and added as a binary explanatory variable into the model to calculate the incidence rate ratio between the two arms. The primary outcome measure in this study was Camp Positivity Yield, defined as counts of bacteriologically confirmed TB (B+) cases diagnosed in each camp. Secondary outcome measures included, Camp Positivity Rate, defined as, B+ cases per population screened, Camp All-Forms Yield, defined as, counts of All-Forms TB (AF-TB) cases diagnosed in each camp and Camp All-Forms TB Rate, defined as AF-TB cases per population screened.

Clinical practice for enrolling individuals with bacteriologically unconfirmed TB vary significantly within the country. The primary outcome was therefore based on B+ TB that was considered a more reliable measure, relative to AF-TB that included individuals with bacteriologically unconfirmed TB, among whom there may be concern of over-diagnosis and clinical practitioner subjectivity. Per camp yield instead of yield per population screened was the used at the primary outcome since this is the main metric for measuring programmatic efficiency. There was also strong interest from Mercy Corps and the funders to utilize this metric since, the overall aim of the project was to increase the number of cases identified from resources invested per activity.

Table 2: Outcome measures for SPOT-TB

Outcome	Measures	Time of measurement
Primary outcome		
Camp positivity yield	Counts of bacteriologically confirmed TB (B+) diagnosed in each camp	1–2 days after each camp
Secondary outcomes		
Camp positivity rate	Bacteriologically confirmed TB (B+)/camp participants	1–2 days after each camp
Camp All-Forms TB yield	Counts of All-Forms TB (AF-TB) diagnosed in each camp	1–2 days after each camp
Camp All-Forms TB rate	All-Forms TB (AF-TB) diagnosed/camp participants	1–2 days after each camp

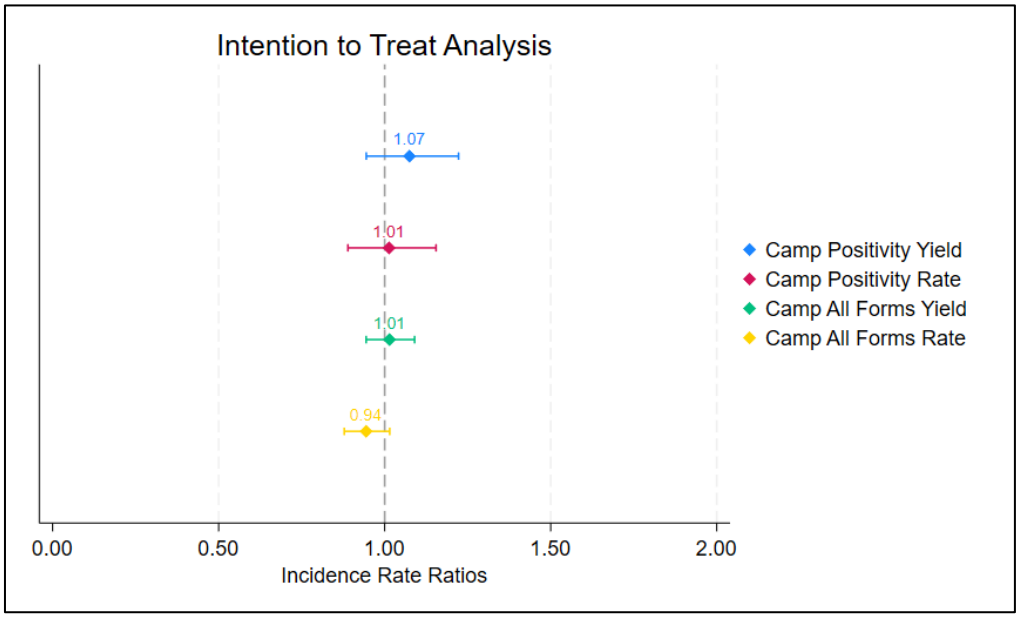
TB, tuberculosis.

The regression coefficient from the models is an "Incidence Rate Ratio (IRR)" comparing the study outcomes between intervention and control arms, after adjustments. An IRR>1 indicates that the intervention was superior whereas an IRR<1 indicates favors the control arm. An IRR=1 indicates no intervention effect.

In the intention to treat analysis as shown in figure 12, the Camps Positivity Yield was 7% higher in the intervention group relative to the control, however, this difference did not cross the superiority margin since the confidence intervals overlap 1 (IRR: 1.07, 95% CI: 0.94 – 1.22). The Camp Positivity Rate was 1% higher in the intervention group relative to the control however, this difference was not statistically significant as well (IRR: 1.01, 95% CI: 0.89 – 1.15).

The Camp All Forms Yield was 1% higher in the intervention group (IRR: 1.01, 95% CI: 0.94 – 1.09) and Camp All-Forms Rate was 6% lower in the intervention group (IRR: 0.94, 95% CI: 0.85 – 1.03) relative to the control. These differences were also not statistically significant.

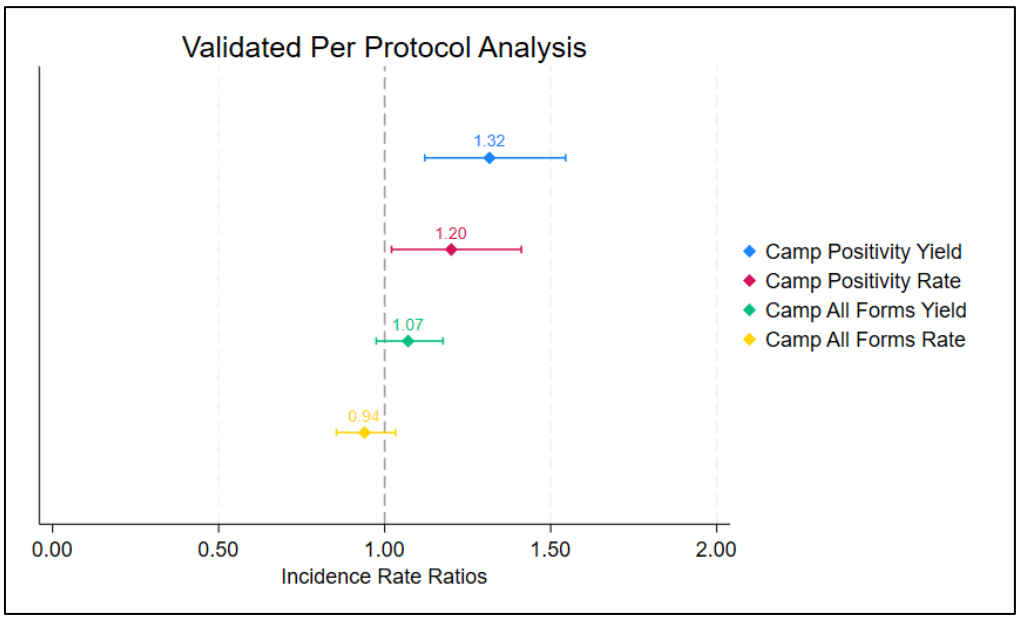
Figure 12: Intention to treat analysis of study outcomes



In the validated per-protocol analysis (where camps in the intervention arm are restricted to those verified to be within 5-km radius of the recommended location), the Camps Positivity Yield was 32% higher in the intervention group relative to the control and this difference was statistically significant (IRR: 1.32, 95% CI: 1.12 – 1.54). The Camp Positivity Rate was 20% higher in the intervention group relative to the control and this difference was statistically significant as well (IRR: 1.08, 95% CI: 1.02 – 1.41).

The Camp All Forms Yield was 7% higher in the intervention group relative to the control group, however, this difference did not cross the superiority margin and was not statistically significant (IRR: 1.07, 95% CI: 0.97 – 1.18). The Camp All-Forms Rate was 6% lower in the intervention group relative to the control and this difference was also not statistically significant (IRR: 0.94, 95% CI: 0.85 – 1.03). (Figure 13)

Figure 13: Primary Analysis (Validated Per Protocol)

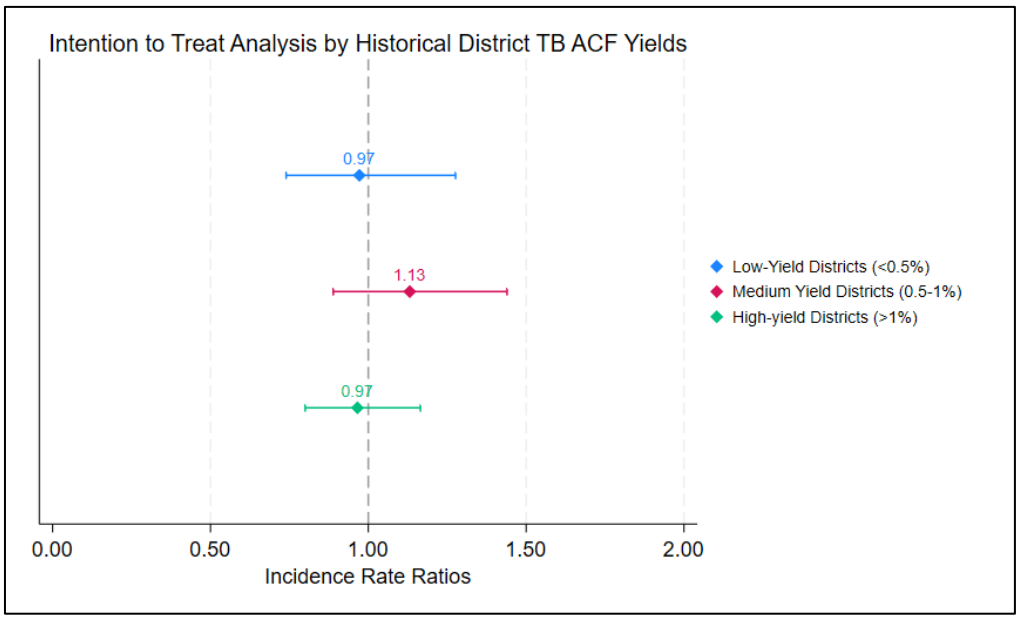


8.5. Subgroup analysis

A subgroup analysis (figure 14) was carried out by stratifying the effect of the intervention over district prevalence (yield) estimates of TB (high, medium, low) identified through a retrospective analysis of mobile X-ray-based camps conducted in the year preceding the trial using which the sample size calculations were performed and a recently published district-level analysis of ACF yields in Pakistan. Districts with yields below 0.5% were classified as "Low-Yield", those with 0.5% - 1% as "Medium-Yield" and those with >1% as "High-Yield." Districts where screening data was not available were classified as Medium Yield. This sub-group analysis was pre-specified in the protocol.

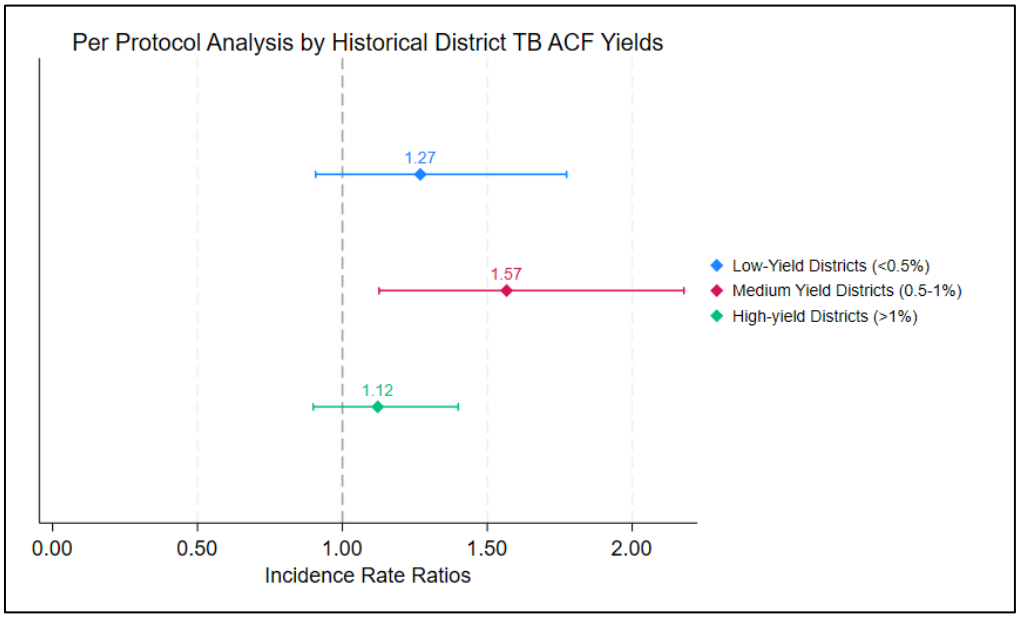
Subgroup analyses have been presented for only the primary outcome, i.e., Camp Positivity Yield outcome. For the intention-to-treat analysis, yields were 3% lower in the intervention group, relative to the control for Low (IRR: 0.97, 95% CI: 0.76 – 1.29) and High Yield (IRR: 0.97, 95% CI: 0.81 – 1.17) districts and 13% higher for the Medium Yield districts (IRR: 1.13, 95% CI: 0.86 – 1.39). However, this difference was not statistically significant for any of the subgroups (Figure 12).

Figure 14: Subgroup Analysis (Intention to Treat)



In the validated per-protocol analysis (Figure 15), yields were 57% higher for the Medium Yield districts and this difference was statistically significant (IRR: 1.57, 95% CI: 1.13 – 2.18). The yields were 27% higher for the Low Yield districts (IRR: 1.27, 95% CI: 0.91 – 1.77) and 12% higher in the High-Yield districts (IRR: 1.12, 95% CI: 0.90 – 1.40), however this difference was not statistically significant.

Figure 15: Subgroup Analysis (Validated Per Protocol)



An additional subgroup analysis by province was also conducted that was not previously specified. In the provincial-level subgroups intention-to-treat analysis (Figure 16, next page), yields were 18% (IRR: 1.28, 95% CI: 0.97 – 1.43) higher in Punjab province, however, this difference was not statistically significant. Yields were lower by 6% for Sindh (IRR: 0.95, 95% CI: 0.77 – 1.17) and 16% for KPK (IRR: 0.84, 95% CI: 0.56 – 1.26) in the intervention arm relative to the control however this difference was also not statistically significant. No differences in yields were observed for Baluchistan between intervention and control groups (IRR: 1.00, 95% CI: 0.51 – 1.97).

In the provincial-level subgroups validated per protocol analysis Figure 17, next page), yields were 52% (IRR: 1.52, 95% CI: 1.18 – 1.95) higher in Punjab province and this difference was statistically significant. Yields were higher by 26% for Sindh (IRR: 1.27, 95% CI: 0.98 – 1.61), however, this did not fully cross the superiority margin. The yields were 1% higher for KPK (IRR: 1.01, 95% CI: 0.62 – 1.64) in the intervention arm relative to the control however this difference was also not statistically significant. No differences in yields were observed for Baluchistan between intervention and control groups (IRR: 1.07, 95% CI: 0.39 – 2.97). Due a limited number of camps with validated GPS coordinates in Baluchistan, this result had wide confidence intervals.

Figure 16: Provincial Subgroup Analysis (Intention to Treat)

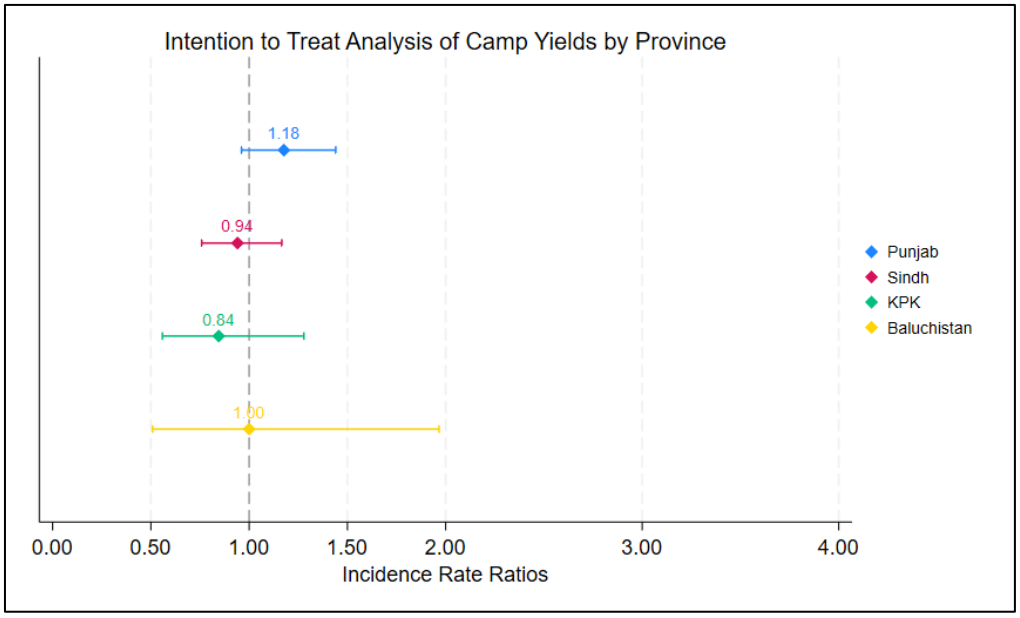
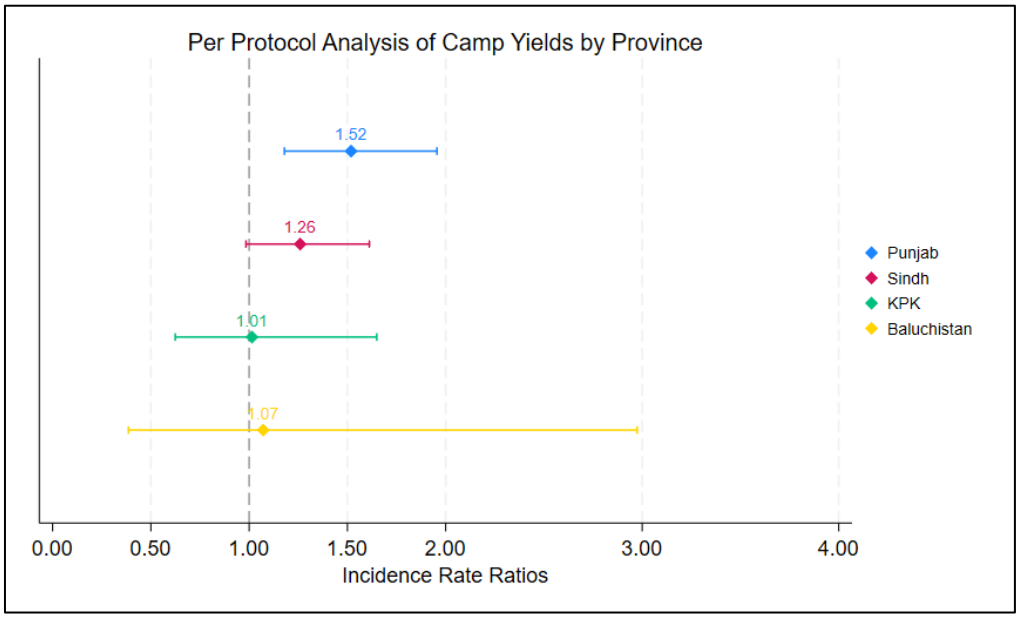


Figure 17: Provincial Subgroup Analysis (Validated Per Protocol)

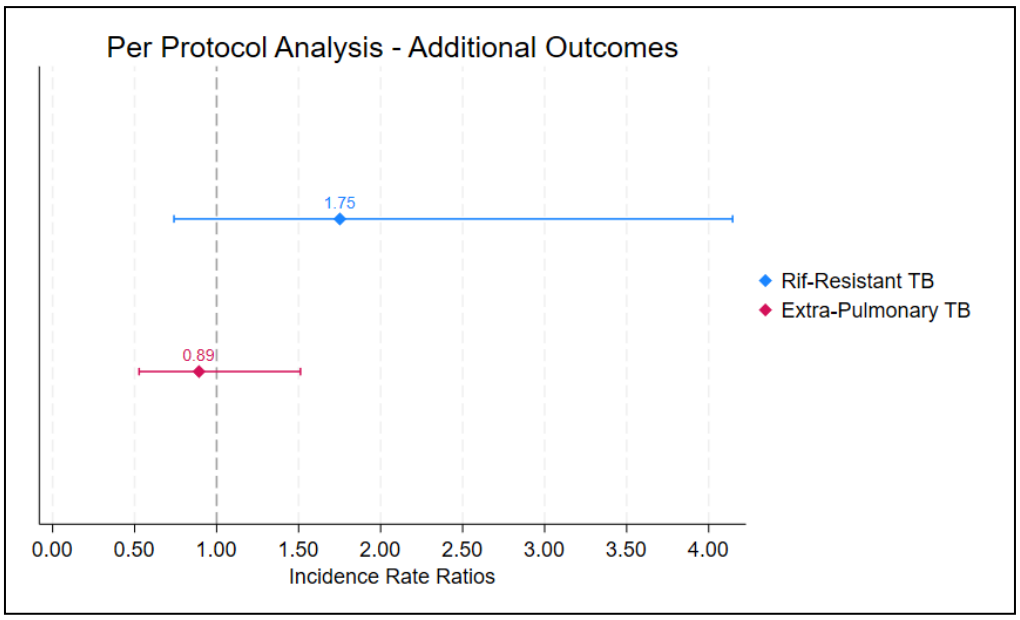


8.6. Post hoc analysis for additional outcomes

As displayed in figure 18, additional analyses were conducted after the conducted after the conclusion of the study that were not specified in the protocol. Using the approach outlined above, the effect of the intervention was evaluated on the detection of Rif-Resistant TB cases and on Extra-pulmonary TB cases. In the validated per-protocol analysis, there was a 75% increase in the detected of Rif-Resistant TB (IRR: 1.75, 95% CI: 0.73 – 4.14) in the intervention relative to the control group. While the size of the effect was quite large, the difference was not statistically significant. The study was not powered to detect differences in Rif-Resistant TB cases and this would have required a very large sample size.

There was an 11% decrease in the yield of EP-TB cases in the intervention group (IRR: 0.89, 95% CI: 0.52 – 1.51), however, this difference was also not statistically significant.

Figure 18: Post hoc analysis of additional outcomes (validated per protocol analysis)



9. CHALLENGES

The implementation of the project encountered a range of challenges both at the data and operational levels, which impacted verification, monitoring, and adherence to the MATCH-AI recommendations.

9.1. Data-related challenges:

- The secondary analysis relied on aggregate data collected from January to December 2022, which lacked individual-level details on persons screened and camp-specific records. This limited the depth of analysis and restricted the ability to assess outcomes at finer levels.
- Data quality issues were observed, including missing values, miscoding, and implausible entries. These gaps complicated the process of generating reliable baseline information.
- Data submission by SRs was decentralized, with variations in file structures and reporting frequencies, making harmonization resource-intensive
- Data entry into the eCBS system started late in the project, which meant that verification of camps held under MATCH-AI recommendations could not be done in real time. Back-entry of missing data for earlier camps also added delays.
- Camp verification relied heavily on unstructured WhatsApp logs shared by SR coordinators, which required manual review. In several cases, broken or incomplete records made it difficult to confirm locations and dates.

9.2. Operational and implementation challenges:

- Uncertainty remained around the adherence of SRs to MATCH-AI recommendations. Initial analysis of the first three months indicated that a low proportion of camps were being conducted as per AI guidance. In some cases, camps were incorrectly marked as MATCH-AI locations without actual alignment with the recommendations.
- Protocol deviations occurred due to logistical hurdles, community-level permissions, staffing constraints, and at times, staff preferences for familiar sites.
- GPS tracking of vans provided useful but incomplete information. While coordinates indicated where vans were parked, they did not always confirm that activities aligned with MATCH-AI recommendations.
- Variability in diagnostic resources influenced outcomes. Vans equipped with on-site GeneXpert provided quicker confirmatory testing and potentially higher yields, whereas locations relying on specimen transport faced delays and inconsistencies.
- Although Thiessen-level saturation prevented the system from recommending the same spots repeatedly, the slow data entry meant that the model was not updated in time and therefore produced the same recommendations again, which was

consistently reported by the SRs. This was also one of the reasons for not following the model's recommendations and instead conducting camps through their conventional methods.

9.3. Verification challenges

- For every camp labeled as a MATCH-AI camp, it became necessary to verify both the camp location and its alignment with AI-generated recommendations. This required substantial manual effort.
- All SRs were required to complete standardized camp activity sheets and ensure entry of all camps into eCBS, including retrospective entry for camps previously missed, which added further burden to field teams.
- Collectively, these challenges highlighted the complexities of implementing an AI-driven intervention in real-world settings. They also underscored the need for stronger data systems, better integration of technology into field operations, and continuous engagement with SRs to improve adherence and reporting fidelity.

10. CONCLUSION AND RECOMMENDATIONS

10.1. Discussion

This study was the first prospective trial of a geographical targeting approach for a tuberculosis case-finding intervention in a high TB burden country. The results suggest that geographical targeting of ACF using MATCH-AI software resulted in a 32% increase in the yield of bacteriologically confirmed TB in the intervention arm relative to the control arm in the validated per-protocol analysis. The software demonstrated a 7% benefit in the intention-to-treat analysis, however, this difference was not statistically significant.

It is important to note that the control arm in this study also included a form of targeting strategy based on previous knowledge and experience of healthcare workers in Pakistan. These teams have had several years of experience in conducting ACF and may have already been aware of potential TB hotspots. As a result, the yields from ACF were higher than expected in the control group and relative to earlier reported studies from Pakistan. The software therefore demonstrated superiority in the validated per-protocol analysis against a standard of care that was likely more effective than in other settings. The use of the software demonstrated non-inferiority in the ITT analysis and this suggests that it may be particularly useful in settings with limited or no prior ACF experience, by potentially augmenting decision-making of healthcare workers and resulting in higher yields.

The subgroup analysis suggests that the software performed better in districts with moderate pre-existing yields of TB while it was less effective in districts where yields were already high or very low. In the provincial subgroup analysis the software appeared to have improved performance in Punjab and Sindh that have relatively higher prevalence of TB relative to other provinces. These findings require further exploration and the contexts in which the software can maximize yields are an important area for further research.

The post-hoc analyses identified further interesting findings. While the study was not powered to detect a difference in identification of drug-resistant TB, the large coefficient effect suggests that the software may help identify spots of DR-TB transmission. This should potentially be further investigated.

Although the implementation teams did report that the camps were conducted at locations recommended by MATCH-AI, spatial data and follow-up verification revealed that, in many instances, the camps were held at randomly chosen locations which did not follow the protocol. This deviation from the protocol was significantly further from the allocated locations to fall beyond the 5-kilometer buffer zone, hence diluting the potential benefit of MATCH-AI guidance. This discrepancy highlights a key limitation in the practical application of MATCH-AI recommendations and led to a dilution of the true effect as shown in the ITT analysis. More specific analyses i.e., when limited to camps that were conducted within a 5-

kilometer radius of the locations recommended by MATCH-AI, the yield of TB cases detected was significantly higher than in camps organized using traditional site selection methods.

The finding of this study suggests that the AI-based approach has a higher predictive value in identifying areas with a greater burden of undiagnosed TB and focusing on such geographies significantly improves the yield of TB cases. MATCH-AI leverages on a large set of key predictors such as TB notifications, population density, disease patterns, and other geospatial indicators. This information is scientifically analyzed by the model to develop complex algorithms that determine optimal locations which should be targeted for a higher yield of TB cases. On the other hand, the conventional approach does incorporate past experiences of the field teams, generally depending on human judgment and localized knowledge, which, although important, may introduce bias and subjectivity.

The improvement in TB case detection even within a 5-kilometer buffer of the AI-recommended locations also suggests the geo-spatial continuity of TB transmission risk suggesting that disease hotspots are not confined to a localized geographic location but extend across adjacent areas. This is an extremely important consideration for the selection of camp sites and field implementation as the exact coordinates recommended by AI may not always be practically feasible for camp deployment due to physical or logistical constraints. This also supports the need of a flexible approach in selecting the exact site of a chest camps i.e., within a few kilometers around the geo-coordinates provided by MATCH-AI.

While the findings of this study suggest that the use of MATCH-AI software to guide chest camp locations can significantly improve TB case detection, the qualitative study where the perception and experiences of the implementing team were inquired has highlighted several key considerations.

- The staff perceived MATCH-AI to reduce personal bias and political pressure in camp site selection. They appreciated the introduction of new technology and saw potential in using AI for more transparent planning. They also acknowledged that it made their work easier as they no longer needed to look at registers or ask doctors to decide a suitable location for camp. Many expressed interests in its continued use with improvements.
- At the same time, the understanding of MATCH-AI was uneven across staff levels. Many field staff, especially DFSs, were unclear about how the software functioned, leading to lower trust and limited ownership.
- While AI helped identify new hot spots and hence cases that would have otherwise gone missing, staff consistently relied on local insights to meet yield targets. Gender norms, safety, accessibility, terrain, and distance were seen as factors the AI did not fully account for.
- The transition to electronic case-based surveillance (eCBS) was hampered by poor connectivity in provinces like KPK and Balochistan. The application had a slow interface

and not fully integrated with the system. Instead of improving efficiency, digital tools created frustration among users. As DFS were assigned clusters, some reported increased workload while some remained free if the AI location didn't fall in their designated district resulting in a disbalance of workload. DFSs faced increased travel requirements due to AI-recommended sites being farther and unfamiliar. Limited compensation, coupled with the dual burden of digital and paper-based reporting, added to their workload.

- Pre-camp activities especially conducting notable meetings remained the single most critical determinant of camp success. Engagement with local leaders (religious figures, teachers, tribal elders) was key for trust-building, turnout, and stigma reduction, regardless of whether the camp was AI-assisted or manual.

Overall, some of the staff members liked MATCH-AI because it helped them decide where to set up camps more quickly and reduced personal and political bias in site selection. They felt it saved time and made things more systematic. On the other hand, some staff members preferred manual methods because they trusted their local knowledge and felt the technology missed ground realities like distance, weather, security, or community needs. These staff members felt that MATCH-AI is useful for TB screening and selection of spots, but it works best when combined with local insight and involvement rather than used alone.

While the use of artificial intelligence can be helpful, the adoption of AI tools like MATCH does not negate the value of human expertise. Its use should complement the ACF approach rather than replace it. Feld staff who are operationalizing chest camps and field activities possess invaluable contextual knowledge that can refine and enhance AI-generated recommendations. Thus, a hybrid approach—where AI provides data-driven insights and local teams interpret and apply them in a context-specific manner would be the most effective model for field implementation.

This trial had a number of strengths:

- The scale and pragmatic nature of this study greatly supports its generalizability. By leveraging ongoing implementation, the study was able to screen over 250,000 individuals and conduct 44,000 Xpert tests, making it one of the largest TB trials ever to be conducted. This enabled the study to evaluate the impact of the intervention with high precision for inference including for per-protocol and subgroup analyses.
- By embedding the study within ongoing implementation, it is directly generalizable to the TB program in Pakistan and potentially to other countries with medium to high TB burdens.
- All cases identified during the study were reported to the NTP and this synergistic approach allowed greater ownership of the intervention enabled the implementation of the trial with a relatively modest research budget.

- The lessons learned through this trial are planned to be documented in a separate methodology manuscript related to the design and conduct of future TB active case-finding trials.

The limitations of the trial also need to be mentioned.

- Field-staff and researchers were not masked to the intervention status. Individuals visiting the camps were masked from the intervention status, however, they were informed that screening data may be used for research purposes during consent for CXR.
- To maximize external validity, the trial was embedded with routine programmatic activities of the NTP. As a result, the trial data was based on the indicators reported during routine surveillance and this limits accuracy and reliability to a certain degree, relative to trials conducted in a controlled setting. The concurrent rollout of the eCBS system during the trial was designed to address this limitation however, the system could not be fully implemented. This also limited the number of adjustments that could be made during the analysis, particularly for the indicators related to health access. The research team conducted independent monitoring visits, data-quality audits as well as data-cleaning to improve the internal validity of the trial. Due to the scale of the ACF program, we did not conduct independent monitoring or quality assurance audits of the laboratory testing for the microbiological outcomes. However, laboratory technicians involved in ACF receive regular trainings on Xpert testing and these are reviewed externally by the NTP. A review of individuals initiated on treatment empirically with bacteriologically unconfirmed TB was not conducted and this may have affected the results of the All-Forms TB outcomes.

This trial further demonstrates the need for standardization of management of bacteriologically unconfirmed TB in the programmatic setting and research into novel, shorter course regimens for individuals with early TB. It is important to note that while this trial evaluated the use of a targeted approach towards ACF, it was reliant on a commercial software to guide decision-making. It is likely that alternative platforms or prediction models will lead to different location recommendations. Since the trial results were contingent on the effectiveness of the software, they may not necessarily be generalizable to all geographically targeted approaches towards ACF. An analysis of the variables used in the MATCH-AI model that best predicted higher yields is planned by the developers using the trial data.

10.2. Recommendations

Based on the findings of this study, following are some of the recommendations:

- MATCH-AI is a potentially useful tool to identify high risk areas for outreach ACF activities, especially for mobile chest X-ray screening camps, and the National and Provincial TB programs should consider integrating MATCH-AI into their routine planning processes. However, in operational settings, it may not always be feasible to conduct camps exactly at the coordinates suggested by AI. Operational guidance needs to be developed where a specific buffer zone around AI-identified locations should be agreed upon involving field implementors to allow for practical implementation while maintaining high yield. This could be either a buffer zone of 5 kms or less, leveraging upon evidence available from this study as well as program data and based on local suggestions and knowledge of the field teams.
- The recommendations made by MATCH would be most effective when combined with the local knowledge and active involvement of field staff especially at the level of a district. We strongly recommend engaging field staff in the planning process to ensure that MATCH-AI outputs are not only scientifically sound but also context specific and operationally realistic. While staff should be involved, it is equally important to regularly provide orientation and training to the field staff on how MATCH-AI works, what data is needed and what could be the limitations of its use. This will empower the field staff to critically assess AI recommendations and adapt them to local contexts and effectively use them in their routine ACF activities.
- Further research should be conducted to investigate the impact of the software on DR-TB identification through ACF or other outreach activities.
- The cost-effectiveness of MATCH-AI should be systematically measured by comparing it to traditional methods of decision-making in selecting spots for ACF. While MATCH-AI has the potential to optimize resource allocation and enhance operational efficiency, these benefits must be evaluated against the costs of development, implementation and maintenance with the perspective of long-term sustainability.
- While this study has shown improved performance of MATCH-AI in comparison to conventional spot finding approach, the approach also needs to be tested by operationalizing it at the field implementer's level with the SRs given charge of operations and the national team provides technical support and monitors the whole process. A key component of this deployment will be comprehensive staff orientation and training on how to use MATCH-AI in their day-to-day workflows. This training will cover both the technical aspects (how to use the AI tool, interpret results, and enter data) and to ensure that the chest camps are conducted as per protocol and within a reasonable radius of the MATCH-AI provided geo-coordinates. The whole process needs regular inputs from the technical team (PR) to reinforce knowledge, update staff on any changes, and maintain high-quality data capture and interpretation. Designated AI champions could be identified at each SR level to lead the process within their

organization and provide ongoing support to their peers. Research should also be conducted on whether yields from the software are sustained overtime in the medium-term through routine programmatic use to determine the field-level efficiency, outside of a research setting.

- Over the longer term, the impact of the use of the software on targeting spots for better case finding and the broader implementing districts needs to be investigated.
- While the study provides ample evidence about the effectiveness of MATCH-AI for better case detection in chest camps, further explorations and research might be needed. Further studies are needed to understand whether this translates into better treatment outcomes, reduced transmission, or more efficient use of resources at scale.
- Finally, the model's accuracy may vary depending on data quality and regional factors. Ensuring that MATCH-AI is trained and updated with high-quality, locally relevant data is essential for maintaining its predictive power. This includes inclusion of ongoing ACF data to update location recommendations.

11. CONCLUSION

To conclude, this study demonstrated that geographical targeting of active case-finding using MATCH-AI may improve the efficiency of the intervention. This is particularly relevant in a country like Pakistan with limited resources. With identification of camp locations with higher underlying prevalence, resources could be deployed where they are most likely to generate high yields, thereby improving cost-effectiveness and public health impact. By improving the efficiency of screening interventions, MATCH-AI has the potential to support national TB programs in accelerating progress toward TB elimination goals, particularly by identifying and addressing gaps in detection among underserved populations.

12. REFERENCES

Ahn, E., & Kang, H. (2023). Intention-to-treat versus as-treated versus per-protocol approaches to analysis. *Korean Journal of Anesthesiology*, *76*(6), 531. https://doi.org/10.4097/KJA.23278

Burke, R. M., Nliwasa, M., Feasey, H. R. A., Chaisson, L. H., Golub, J. E., Naufal, F., Shapiro, A. E., Ruperez, M., Telisinghe, L., Ayles, H., Corbett, E. L., & MacPherson, P. (2021). Community-based active case-finding interventions for tuberculosis: a systematic review. *The Lancet Public Health*, *6*(5), e283–e299. https://doi.org/10.1016/S2468-2667(21)00033-5

Corbett, E., Corbett, E. L., Bandason, T., Duong, T., Dauya, E., Makamure, B., Churchyard, G. J., Williams, B. G., Munyati, S. S., Butterworth, A. E., Mason, P. R., Mungofa, S., & Hayes, R. J. (2010). Articles Comparison of two active case-fi nding strategies for community-based diagnosis of symptomatic smear-positive tuberculosis and control of infectious tuberculosis in Harare, Zimbabwe (DETECTB): a cluster-randomised trial. *The Lancet*, *376*, 1244–1253. https://doi.org/10.1016/S0140

Deya, R. W., Masese, L. N., Jaoko, W., Muhwa, J. C., Mbugua, L., Horne, D. J., & Graham, S. M. (2022). Yield and Coverage of Active Case Finding Interventions for Tuberculosis Control:A Systematic Review and Meta-analysis. *Tuberculosis Research and Treatment*, *2022*, 1–12. https://doi.org/10.1155/2022/9947068

EPCON. (2022). *Finding TB Cases in Pakistan: Using AI for Active Case Finding*. https://www.epcon.ai/post/finding-tb-cases-in-pakistan-using-ai-for-active-case-finding

JPRM. (2024). 2022 TB programme review in Pakistan: strengthening governance, with better patient diagnosis and treatment. *IJTLD OPEN*, *1*(3), 136–143. https://doi.org/10.5588/ijtldopen.23.0587

KIT. (2025a). *Active TB Case Finding in Pakistan - KIT*. https://www.kit.nl/institute/project/electronic-case-based-surveillance-using-predictive-modelling-and-real-time-data-to-plan-active-tb-case-finding-in-pakistan/

KIT. (2025b). *The KIT MATCH Approach for Enhancing TB Care Coverage - KIT*. https://www.kit.nl/institute/project/the-kit-match-approach-for-enhancing-tb-care-coverage/

Kranzer, K., Afnan-Holmes, H., Tomlin, K., Golub, J. E., Shapiro, A. E., Schaap, A., Corbett, E. L., Lönnroth, K., & Glynn, J. R. (2013). The benefits to communities and individuals of screening for active tuberculosis disease: A systematic review. *International Journal of Tuberculosis and Lung Disease*, *17*(4), 432–446. https://doi.org/10.5588/IJTLD.12.0743,

McClean, M., Panciu, T. C., Lange, C., Duarte, R., & Theis, F. (2024). Artificial intelligence in tuberculosis: a new ally in disease control. *Breathe*, *20*(3). https://doi.org/10.1183/20734735.0056-2024

Ortiz-Brizuela, E., & Menzies, D. (2021). Tuberculosis active case-finding: looking for cases in all the right places? In *The Lancet Public Health* (Vol. 6, Issue 5, pp. e261–e262). Elsevier Ltd. https://doi.org/10.1016/S2468-2667(21)00048-7

Rajpurkar, P., O'Connell, C., Schechter, A., Asnani, N., Li, J., Kiani, A., Ball, R. L., Mendelson, M., Maartens, G., van Hoving, D. J., Griesel, R., Ng, A. Y., Boyles, T. H., & Lungren, M. P. (2020). CheXaid_ deep learning assistance for physician diagnosis of tuberculosis using chest x-rays in patients with HIV. *Npj Digital Medicine*, *3*(1). https://doi.org/10.1038/s41746-020-00322-2

WHO. (2021a). *Module 2: Screening*.

WHO. (2021b). *Optimizing active case-finding for tuberculosis Implementation lessons from South-East Asia*.

WHO. (2024). Global TB report 2024. In *2024*. https://iris.who.int/bitstream/handle/10665/379339/9789240101531-eng.pdf?sequence=1

WHO. (2025). *Global Programme on Tuberculosis & Lung Health*. https://www.who.int/teams/global-tuberculosis-programme/data

Yuen, C. M., Amanullah, F., Dharmadhikari, A., Nardell, E. A., Seddon, J. A., Vasilyeva, I., Zhao, Y., Keshavjee, S., & Becerra, M. C. (2015). Turning off the tap: Stopping tuberculosis transmission through active case-finding and prompt effective treatment. In *The Lancet* (Vol. 386, Issue 10010, pp. 2334–2343). Lancet Publishing Group. https://doi.org/10.1016/S0140-6736(15)00322-0

Zaidi, S. M. A., Mahfooz, A., Latif, A., Nawaz, N., Fatima, R., Rehman, F. U., Reza, T. E., & Emmanuel, F. (2024). Geographical targeting of active case finding for tuberculosis in Pakistan using hotspots identified by artificial intelligence software (SPOT-TB): Study protocol for a pragmatic stepped wedge cluster randomised control trial. *BMJ Open Respiratory Research*, *11*(1). https://doi.org/10.1136/bmjresp-2023-002079

13. ANNEXURE

13.1. Consent Form

Information Sheet

Instructions for chest camp in-charge or District Field Supervisor (DFS):

- *Use this form for participants who are over 18 years of age.*
- *The DFS or camp in-charge should read out the consent form to all attending participants at the timethe participants receive their OPD token slip for the chest camp.*
- *The DFS or camp in-charge would mark the consent and sign it on behalf of the participant as theconsent is obtained.*

An evaluation study is being conducted by the Centre for Global Public Health Pakistan in partnership withthe Mercy Corps to assess the effectiveness of Artificial Intelligence to identify spots for TB cases compared to the conventional method of selecting the hotspots. This evaluation is being done in all chestcamps organized by SR organizations across all 90 districts of Pakistan. There will be no separate interviewfor this, you will be registered and tested as per the usual practice, but your information other than yourname will be used in the evaluation study.

The results of this study will lead us to identify the method which is better in identification of locations forfinding missing TB cases which in turn will decrease the overall burden of TB from the country. We will also develop a report, but YOU WILL NOT BE NAMED OR IDENTIFIED IN ANY WAY in this report because WE DO NOT NEED TO KNOW YOUR NAME or any other relevant information which can be used to identifyyou. If you do not want us to use your information for the study, you are completely allowed to refrain us from using your information. By refusing to use your information for the evaluation study will not affect provision of any services within this camp. All diagnosis and treatment will be provided to you as per routine.

There is no coercion of any type for you to participate in this evaluation. If you have any questions aboutyour rights as a subject participating in a research survey, or if you wish to discuss your participation in the study, you can contact Ms. Amna Mahfooz, Co-Investigator at CGPH-Pakistan on 051-8357603.

As this study is confidential, we do not want you to sign anything. If you agree to participate, I will mark yes on the consent form indicating your consent for participation.

DO YOU AGREE TO PARTICIPATE?	**Yes**		**No**		(Clarify/Discuss)

if yes*, sign and date below to indicate that informed consent was given by the participant)*

Name and signature of DSF/camp in-charge: _____

Date: _____

رضامندی نامہ

<u>ہدایات برائے چیسٹ کیمپ انچارج یا ڈسٹرکٹ فیلڈ سپروائزر (ڈی ایف ایس):</u>

- یہ فارم 18 سال سے زیادہ عمر کے شرکاء کے لیے استعمال کریں
- کیمپ انچارج یا ڈی ایف ایس یہ رضامندی نامہ تمام شرکاء کو او پی ڈی کی ٹوکن سلپ حاصل کرتے وقت پڑھ کر سنائیں
- کیمپ انچارج یا ڈی ایف ایس شرکت کرنے والے کی رضامندی حاصل کر کے اس کی جانب سے رضامندی نامے پر دستخط کرے

سینٹر فار گلوبل پبلک ہیلتھ پاکستان (CGPH-Pakistan) مرسی کور (Mercy Corp) کے اشتراک سے ایک تشخیصی سٹڈی کر رہا ہے جس میں ٹی بی کیسز کے ہاٹ سپاٹ کی شناخت کے لیے روائتی طریقہ کار کے مقابلے میں آرٹیفیشل انٹیلیجنس (AI) کی افادیت کو جانچا جائے گا۔ یہ تشخیصی سٹڈی پاکستان کے 90 اضلاع میں ایس آر (SR) تنظیموں کے زیر اہتمام ہونے والے تمام چیسٹ کیمپوں میں کی جا رہی ہے۔ اس کے لیے کوئی علیحدہ انٹرویو نہیں ہوگا، آپ کا معمول کے مطابق اندراج (رجسٹر) اور ٹیسٹ کیا جائے گا، اس تمام عمل میں آپ کا نام کہیں بھی استعمال نہیں کیا جائے گا۔

اس سٹڈی کے نتائج ہمیں اس طریقہ کار کی نشاندہی کرنے میں مدد کریں گے جو ٹی بی کے مسنگ کیسز کو تلاش کرنے کے لیے ہاٹ سپاٹ کی شناخت میں بہتر ہے، جس کے نتیجے میں ملک میں ٹی بی کے مجموعی بوجھ میں کمی آئے گی۔ ہم ایک رپورٹ بھی تیار کریں گے، لیکن اس رپورٹ میں کسی بھی طرح سے آپ کا نام یا شناخت نہیں دی جائے گی کیونکہ ہمیں آپ کے نام یا کسی دوسری متعلقہ معلومات کو جاننے کی ضرورت نہیں ہے جس کے ذریعے آپ کی شناخت ہوسکے۔ اگر آپ نہیں چاہتے / چاہتی کہ ہم آپ کی معلومات کو اس سٹڈی کے لیے استعمال کریں، تو آپ کو مکمل طور پر اجازت ہے کہ آپ ہمیں اپنی معلومات استعمال کرنے سے روک دیں۔ تشخیصی سٹڈی کے لیے آپ کی معلومات استعمال کرنے سے انکار پر کیمپ کے اندر کسی طرح کی بھی خدمات کی فراہمی پر کوئی اثر نہیں پڑے گا۔ تمام تر تشخیص اور علاج آپ کو معمول کے مطابق فراہم کیا جائے گا۔

اس تشخیصی سٹڈی میں حصہ لینے کے لیے آپ پر کسی قسم کا کوئی جبر (زبردستی) نہیں ہے۔ اگر آپ اس سٹڈی میں حصہ لینے کے حوالے سے اپنے حقوق کے بارے میں کوئی سوال پوچھنا چاہتے ہیں، یا اگر آپ اس سٹڈی میں اپنی شرکت کے بارے میں کوئی بات کرنا چاہتے ہیں، تو آپ CGPH-Pakistan میں سٹڈی کی کو-انویسٹیگیٹر محترمہ آمنہ محفوظ سے 051-8357603 پر رابطہ کر سکتے ہیں۔

چونکہ یہ سٹڈی کانفیڈنشل (رازدارانہ) ہے، ہم نہیں چاہتے کہ آپ کسی چیز پر دستخط کریں۔ اگر آپ شرکت کرنے کے لیے راضی ہیں، تو میں رضامندی کے فارم پر ہاں کا نشان لگا دوں گا / گی جس میں شرکت کے لیے آپ کی رضامندی کی نشاندہی ہوگی۔

کیا آپ شرکت کے لیے رضامند ہیں؟	ہاں ☐	نہیں ☐	(وضاحت / ڈسکس کریں)

ہاں کی صورت میں نیچے دی گئی جگہ میں نام لکھ کر دستخط کریں اور تاریخ لکھیں:

کیمپ انچارج یا ڈی ایف ایس کا نام _____ دستخط _____ تاریخ _____

Center for Global Public Health (CGPH), Pakistan
(Collaborative Partner of IGPH – University of Manitoba, Canada)
Office 10-11, 2nd Floor, Pakland Square
Sector G-8 Markaz, Islamabad
Email: info@cgphpk.org
Website: www.cgphpk.org