

Clustering of pulmonary tuberculosis in Ethiopia: repeated population-based symptom screening

Authors

Abiot Bezabeh Banti^{1,2}, Daniel Gemechu Datiko³, Brita Askeland Winje⁴, Sven Gudmund Hinderaker¹, Einar Haldal⁵, Mesay Hailu Dangisso⁶

¹University of Bergen, Bergen, Norway

²REACH Ethiopia, Addis Ababa, Ethiopia

³Management Sciences for Health, Addis Ababa, Ethiopia

⁴Faculty of Health Science, Oslo Metropolitan University, Oslo, Norway

⁵Norwegian Institute of Public Health, Oslo, Norway

⁶Ethiopian Public Health Institute, Addis Ababa, Ethiopia.

*Corresponding author: Abiot.Banti@uib.no

Word count abstract: 296

Word count of the main body: 3084

Table: 3

Figure: 3

Running title: Clustering of pulmonary tuberculosis from repeated population-based screening

Abstract

Objective: A “Cluster” is an area with a higher occurrence of tuberculosis than would be expected in an average random distribution of that area. Tuberculosis clustering is commonly reported in Ethiopia, but most studies rely on registered data, which may miss patients who do not visit health facilities or those who attend but are not identified as having tuberculosis. This makes the detection of actual clusters challenging. This study analysed the clustering of pulmonary tuberculosis and associated risk factors using symptom-based population screening in Dale, Ethiopia.

Design: A prospective population-based cohort study.

Setting: All households in 383 enumeration areas were visited three times over 1 year period, at four-month intervals.

Participants: Individuals with pulmonary tuberculosis aged ≥ 15 years with demographic, socioeconomic, clinical, and geographic data residing in 383 enumeration areas (i.e., the lowest unit/village in the kebele, each with approximately 600 residents).

Outcome measures: Pulmonary tuberculosis (i.e., bacteriologically confirmed by sputum microscopy, GeneXpert or culture plus clinically diagnosed pulmonary tuberculosis) and pulmonary tuberculosis clustering.

Results: We identified pulmonary tuberculosis clustering in 45 out of the 383 enumeration areas. During the first round of screening, 39 enumeration areas showed pulmonary tuberculosis clustering, compared to only three enumeration areas in the second and third rounds. Our multilevel analysis found that enumeration areas with clusters were located farther from the health centres than other enumeration areas. No other determinants examined were associated with clustering.

Conclusions: The distribution of pulmonary tuberculosis was clustered in enumeration areas distant from the health centres. Routine systematic community screening using existing health infrastructure with Health extension workers may be costly but through targeted screening they can identify and refer persons with TB symptoms more quickly for diagnosis and treatment, thereby decreasing the duration of disease transmission and contributing to the reduction of TB burden.

Keywords: Tuberculosis; Screening; Clustering; Spatial analysis; Sidama

Strengths and limitations of this study

- The study applied a three-round total household symptom screening strategy to identify undiagnosed as well as diagnosed tuberculosis cases, and hence identify real clusters.
- Health extension workers actively involved the entire population in screening, benefitting from their trust and familiarity with the community.
- The study also sought for risk factors for clustering that may require attention from public health practices at the lowest community level.
- Smear microscopy will due to relatively low sensitivity always miss some cases of tuberculosis.

Background

Although tuberculosis (TB) is preventable and curable, globally an estimated 1.3 million people died from TB in 2022. Modelling shows that 3.1 million TB cases were not detected or reported globally, which sustains disease transmission and suffering in the community (1,

2). According to the World Health Organization, Ethiopia has a high-burden of TB, with estimated incidence rates of 126 per 100,000 population. The TB burden is particularly acute in certain areas, creating a "TB belt" in the southern part of the country. Higher notification rates in the study area than the national average disproportionately affect these communities.

Efforts to improve services to the most vulnerable and disadvantaged populations using existing health systems are among strategic priorities (3, 4). Clustering of TB in an area means there are more cases than expected from random distribution; this represents an increased risk of transmission (5), but can also provide valuable information for the national TB programme. Clusters can be measured using surveillance data, genotypic data, and statistical models (6). Many studies, including some Ethiopian, have reported TB clustering using surveillance data (6-9). These studies relied on notified data from passive case-finding, which may underestimate the actual TB rate because of undetected cases (10, 11). Population-based studies may help better substantiate the existence of clusters, but they are relatively few (3, 6).

A suggested prevention and care strategy is to target the population in clusters (12). Repeated population screening may help identify these clusters, facilitating the design of focused, less expensive and easily accessible interventions. In our study of pulmonary tuberculosis (PTB) prevalence and incidence, we conducted three rounds of household screening in the Dale district in Ethiopia. We discovered that two-thirds of symptomatic PTB patients were undetected at the time of the household visit (13). This study is based on active case-finding data, complementing information from TB registries, resulting in a more comprehensive dataset on PTB. Our objectives in the Dale district of Ethiopia were to: 1) identify symptomatic PTB clustering through population screening, and 2) explore the association between PTB clustering and demographic, socioeconomic, clinical and environmental risk factors.

Methods

Study design and setting

This study is based on data from a cohort study with repeated household screening between October 2016 to September 2017, combined with a follow-up survey to collect geographical data in August and September 2018. The target population included everyone residing in the Dale district, which is located in the Sidama region of southern Ethiopia. According to the 2017 census, the estimated adult population aged 15 years and older in Dale was 136,181. It's worth noting that the majority of the population in this region resides in rural areas. Dale has 36 kebeles, which are the lowest administrative units. A kebele is divided into 383 smaller villages called enumeration areas (EAs). EA has 600 persons on average. TB health care services are delivered based on the End TB strategy across 10 health centres, 2 clinics, and 36 health posts. However, neither GeneXpert nor X-ray services are available within Dale. Access to these services requires a referral to the hospital in neighbouring Yirgalem town. In Ethiopia, public local community health providers, Health Extension Workers (HEWs), offer cost-effective health services, which include patient referrals and regular household visits. Approximately 33% of rural residents turn to HEWs as their initial source of care when they become ill (14). We followed the Strengthening the Reporting of Observational Studies in Epidemiology cohort reporting guidelines (Supplemental file 1) (15).

Study population

In the three rounds of household visits between 2016 and 2017, the initial visits occurred from October 2016 to January 2017, followed by visits from February to May 2017, and then again from June to September 2017. The visits occurred every 4-month intervals and

lasted 1-2 months each. During these visits, trained HEWs went door-to-door, asking all family members if they had any symptoms compatible with TB, they also asked about symptoms concerning household members who were not at home during the screening and revisited households that were missed. We used the term "presumptive TB" for individuals having a persistent cough for 14 days or more with or without haemoptysis, weight loss, fever, night sweats, chest pain or difficulty breathing. During these visits, demographic, socioeconomic, clinical, and environmental data to assess associated risk factors were collected (13). Those with symptoms were closely followed and asked to come to the health post to provide sputum. Sputum samples were collected and transported to the health facility for smear microscopy. GeneXpert testing was performed on smear-positive samples to validate the results. In total, the screening identified 442 PTB cases, with 90 clinically diagnosed and 352 bacteriologically confirmed cases. In 2018, we conducted a follow-up survey in the same population collecting geographic information on the identified PTB patients. This allowed us to link the previously collected data on risk factors to the geographic data. The addresses of 17 PTB patients from two kebeles (Hida Kaliti and Kaliti Simita) were not found due to industrial zone construction. Further details on household screening and follow-up can be found elsewhere (13,16).

Operational definitions

Bacteriologically confirmed PTB was defined as any TB diagnosed based on sputum smear-microscopy, culture or GeneXpert laboratory results. Clinically diagnosed PTB was defined by a clinical decision to start TB treatment in those with a persistent cough, usually supported by radiological findings and evidence from other tests according to the national TB and leprosy guidelines (16). An individual diagnosed with bacteriologically confirmed or clinically diagnosed PTB was considered a PTB case. The most likely and secondary clusters were reported by scan statistics (17). Assuming constant risk, the "most likely PTB cluster" within the EAs has the highest probability of having the most cases (18,19). The

“secondary PTB cluster” does not overlap with the most likely cluster but has an increased cluster probability (20). Depending on the risk in the cluster secondary clusters are further divided into secondary cluster 1 and 2. Clustering is a combination of the most likely and secondary clusters. To estimate distance we calculated the straight line from the health centres to the patient’s household.

Patient and public involvement

No patients were involved in setting the research question or the outcomes, interpretation, writing or dissemination of the results.

Data collection

The geographical information was collected using a handheld Global Positioning System (GPS). We used GPS to measure the location and altitude, computed population density from census data of the EA, and linked it to the geographic data (GPS coordinates) corresponding enumeration areas. We also used GPS points in the enumeration areas to calculate the straight-line distance between each household and the nearby health center. Locations of patients were collected by trained personnel. HEWs who were familiar with the villages provided support to the data collectors. We used principal component analysis to calculate a wealth index (Supplemental file 2). Socioeconomic covariates obtained directly from the study population were assigned wealth-index scores, which were derived from the original variables measuring household socioeconomic status. We used the wealth index to indicate household income inequality. We determined a Body Mass Index (BMI) cut-off point using national TB guidelines (<16 kg/m² refers to severe acute malnutrition; 16-17 kg/m² refers to moderate acute malnutrition, 17-18.5 kg/m² refers to mild malnutrition, and ≥18.5 kg/m² refers to normal) while the Guinea-Bissau study for the Mid-upper Arm Circumference (MUAC) <22cm refers to malnutrition; MUAC ≥ 22 refers to normal (21,22).

We calculated the median values of the dataset. We used the median value when we divided our dataset into two groups for distance, population, population density and altitude. EA served as the spatial unit of analysis. We used Scan statistics to identify clusters. The cluster was the outcome of interest. We classified EAs with PTB clusters as "yes" while those without as "no" to facilitate further analysis using the multilevel technique.

Statistical analysis

Spatial statistics are commonly used to detect risk factors, hotspots (areas with increased risk transmission or burden), and transmission patterns (23). We performed purely spatial analysis using poisson model to identify the most likely and secondary PTB clusters, using Kulldroff's scan statistics, with the freely available software (SaTScan 10.1) (17). We analysed data after exploration to avoid systematic errors that could occur during data collection and entry that impact our study's results. For spatial analysis, we used corrected shape file data linked with EAs. We used the population size for each EA projected from the 2007 census and calculated the expected incidence rate of PTB for Ethiopia. We used ArcGIS 10.4 (ESRI 2016) to calculate centroids for the polygons and construct output maps and maps created using the UTM Zone 37 North projected coordinate system. We used case, population, and coordinate as input for the purely spatial analysis. This analysis distributes the number of events in an area according to the at-risk population. Spatial clustering analysis provides the data with a multilevel structure in which individuals are nested in cluster locations, which allows for an exact estimate of the effect. A between-class coefficient of zero means no clustering effects within the data, that is, all people within the same group have identical responses to the outcome variable (24,25). When dealing with multilevel analyses, multiple groups were involved (24). The analysis included 425 individuals with PTB in two-level regression (patient and EA levels). We evaluated random effects at the EA level by using independent covariance and examined independent factors.

We used three models: In Model 1, we evaluated a null model with no independent covariates. In Model 2, we evaluated individual-level covariates. In Model 3, we added EAs-level covariates to Model 2 and allowed observation days to vary across both levels. We used Jamovi (Version 2.3) software for multilevel analysis (26).

Ethics clearance

Ethics approval was obtained from the Armauer Hansen Research Institute-Alert Ethics Review Committee, Ethiopia (PO12/15), the National Research Ethics Committee, Ethiopia (no 104/2016), and the Regional Committees for Medical and Health Research Ethics in Norway (2015/1006). Written consent was obtained from participants and parent(s)/Guardian(s) for participants aged 15-17 years.

Results

Patient characteristics and distribution

From October 2016 to September 2017, 442 PTB patients were identified. Among them, 425 patients were included in this report, while 17 (3.8%) PTB patients were excluded due to missing residential addresses (see Figure 1). The median age of the patients was 28 years, with an age range of 15 to 90 years. The patients' mean number of observation days was 34 days with a standard deviation of 2.6 days. Of the patients, 52% were male, 72% were ever married, and only 15% had BCG scars. Of 383 EAs in the district, 197 reported at least one PTB patient, and the geographic distribution of PTB patients exhibited variations (Figure 2).

[Inser Figure 1 here]

Figure1. Study flow chart for a household investigation for tuberculosis in Dale district Ethiopia, 2018.

[Inser Figure 2 here]

Figure 2. Rate per 100,000 population of pulmonary tuberculosis by Enumeration area in the three rounds of household visits from October 2016 to September 2017, Dale Ethiopia.

Table 1 shows the characteristics of patients inside and outside clusters. Most variables such as age, sex, marital status, BMI, MUAC, Wealth index, and BCG scar, were comparable. Differences appear in schooling (74% vs 59%), distance from health facilities (84% vs 43% in more than 3km), and population density (73% vs 53% in less dense areas <665/square km).

Table 1. Characteristics of symptomatic pulmonary tuberculosis patients inside and outside clusters, Dale, Ethiopia 2018.

Covariate	Cluster (n, 425)	
	Yes, n (%)	No, n (%)
All patients	81	344
Sex		
Men	39 (48)	180 (52)
Women	42 (52)	164 (48)
Age in years		
15-34	60 (74)	224 (65)
≥35	21 (26)	125 (35)
Marital status		

	Single	25 (31)	93 (27)
	Ever married	56 (69)	251 (73)
Schooling			
	Yes	60 (74)	204 (59)
	No	21 (26)	140 (41)
BMI kg/m ²			
	<16	22 (27)	99 (29)
	16-17	12 (15)	79 (23)
	17-18.5	18 (35)	80 (23)
	≥18.5	19 (23)	86 (25)
MUAC			
	<22 cm	65 (80)	267 (78)
	≥22 cm	16 (20)	77 (22)
Wealth index			
	Poor	35 (43)	145 (42)
	Middle	19 (24)	106 (31)
	Rich	27 (33)	93 (27)
BCG scar			
	Yes	13 (16)	51(15)
	No	68 (84)	293 (85)
Distance from the health centre			
	<3 km	13 (16)	197 (57)
	≥3 km	68 (84)	147 (43)
*Population density per square km			
	<665	59 (73)	183 (53)
	≥665	22 (27)	161 (47)
*Population			

	<644	57(70)	196(57)
	≥644	24(30)	148(43)
*Altitude (m)			
	<1793	57 (70)	178 (52)
	≥1793	24 (30)	166 (48)

BMI, Body Mass Index; MUAC, Mid Upper Arm Circumference; BCG, Bacillus and Calmette

*These variables used cut-off for groups using the mean values of dataset.

Clustering of symptomatic PTB

In purely spatial analysis, we found that 45 out of the 383 EAs visited in each round had PTB clustering during the study period, as shown in Table 2. One EA exhibits the most likely clustering (RR 11.9, $p < 0.0001$), while 41 EAs had secondary cluster 1 (RR 2.09, $p < 0.001$), and 3 EAs had secondary cluster 2 (RR 3.9, $p < 0.044$). PTB clusters were common in the southeast, northwest, and southcentral parts of the district, as shown in Figure 3a-c. These figures show PTB clustering during (a) the entire study period, (b) the first round of household screening, and (c) the second and third rounds of household screening. In the first round, 39 EAs (10%) had PTB clusters, but only three PTB clusters (1%) were detected in two kebeles during the second and third visits between February and September 2017.

Table 2. Results of cluster analysis of patients with pulmonary tuberculosis in Dale, Ethiopia, by scan statistics, in Dale, Ethiopia 2018.

Spatial Clusters	Location	Observed cases (n)	Expected cases (n)	RR	Likelihood ratio	P
Most likely cluster	1	17	1.48	11.9	26.3	<0.001
Secondary clusters 1	41	50	23.7	2.09	15.4	<0.001

Secondary clusters 2	3	14	3.7	3.92	8.6	0.044
----------------------	---	----	-----	------	-----	-------

n, number; RR, relative risk. P, P-value

[Insert Figure 3 here]

Figure 3a-c. (a) Pulmonary tuberculosis clustering in the three rounds of household visits from October 2016 to September 2017; (b) Pulmonary tuberculosis clustering in the first round of household visits (prevalent cases) from October 2016 to January 2017; (c) Pulmonary tuberculosis clustering in the second and third rounds (incident cases) from February 2017 to September 2017.

Risk factors for PTB clustering

Table 3 shows the results of a multilevel analysis of the patients and EAs. EAs in PTB clusters were more likely far from health facilities than EAs not in PTB clusters. In our analysis, the PTB cluster was not associated with other factors examined in our regression model.

Table 3. Results of multilevel analysis: association between pulmonary tuberculosis clusters and risk factors in Dale, Ethiopia 2018.

Covariates	Effect estimate		SE	z	p
	Coefficient	95% CI			
(Intercept)	-25.5	(-38.2, -12.6)	6.5	-3.9	<0.001
Women	-2.8	(-8.0, 2.5)	2.7	-1	0.303
Age ≥35 years	-0.5	(-5.76, 4.82)	2.7	-0.2	0.862
Ever married	-1.1	(-5.8, 3.6)	2.4	-0.5	0.649
Schooling (No)	-4.9	(-10.47, 0.76)	2.9	-1.7	0.09
BMI kg/m ²	1.6	(-0.57, 3.7)	1.1	1.4	0.152

<16	1,0				
16-17	-1	(-3.59, 1.6)	1.3	-0.8	0.453
17-18.5	1.2	(-0.73, 3.09)	1	1.2	0.225
≥18.5	-0.2	(-2.76, 2.34)	1.3	-0.2	0.871
MUAC in cm (>22)	-5.1	(-13.36, 3.14)	4.2	-1.2	0.225
Wealth index	-0.6	(-1.6, 0.45)	0.5	-1.1	0.275
Poor	1				
Middle	-0.7	(-5.92, 4.54)	2.7	-0.3	0.795
Rich	-8.4	(-16.96, 0.19)	4.4	-1.9	0.055
BCG Scar (no)	7.2	(-5.75, 20.1)	6.6	1.1	0.276
*Number population ≥644	-3.8	(-9.36, 1.74)	2.8	-1.4	0.178
*Distance in km from health centre ≥3	16.2	(4.9, 27.5)	5.8	2.8	0.005
*Population density per square km ≥655	0.7	(-4.15, 5.63)	2.5	0.3	0.767
*Altitude in metres ≥1793	-3.8	(-9.69, 2.18)	3	-1.2	0.215
**Number of observation days	0.1	(-0.03, 0.05)	0	0.7	0.506

cm, centimetre; kg, kilogram; km, kilometre; m², meter square; SE, standard error; p, p-value *These variables used cut-off above the mean values.**continuous variable.

Discussion

Based on active case-finding data, this study revealed clusters of symptomatic PTB cases in the Dale district in Ethiopia. In the southeast, northwest, and south-central parts of the district there was more PTB clustering, and the EAs with PTB clusters were farther from health centres. These areas may play an important role in transmission and suggest the need for targeted screening for early detection and treatment in areas with limited TB care. The study highlights the benefits of more active TB case-findings in these areas as majority of patients had not sought health services despite having symptoms (13, 27). Repeated

active case-finding made TB care more accessible to everyone in the district. The strategy involves several rounds of screening, which may identify missing and subclinical TB cases in rural locations and improve our ability to pinpoint clusters since this technique basically removed most clusters in subsequent visits.

In program settings, the locations of cases are often linked to the diagnostic centres and not the patients' households. Since many patients must travel to health centres for diagnosis, the notification data can, therefore, not always identify "real geographic" clusters, and units with "clusters" may include many visitors from other areas. Hence, the study findings from previous clustering studies using notification data, often reported in urban or semi-urban areas, where access to diagnosing and treating TB is less challenging (28,29).

Contrary to passive case-finding, active screening can be more costly but indicates clusters more likely to occur in areas farther from health centres. Such screening can help decentralize TB services, which may offer an effective means of improving access to early diagnosis, treatment and preventive therapy that helps long-term efforts to reduce TB clusters and infectious sources in the community (30, 31). Our case identification strategy did not depend on patient care seeking behaviour. Unlike previous studies that relied solely on passive case finding without accounting for undetected cases, we used trained HEWs who conducted regular screening visits on various health issues in their target population. These HEWs actively involve the entire population in the process, benefitting from their trust and familiarity with community screening. Expanding on this infrastructure, our study incorporated PTB symptom screening in all households within the district. The proportion of undetected symptomatic PTB cases is expected to be lower compared to the notification data from passive case-finding. While there may be a latency period between TB infection and symptom onset, repeated screening of the entire population over an extended period could help identify new cases previously asymptomatic who later become symptomatic as a disease progresses (13, 16). We used TB unit registrations to validate PTB cases and ensured that the identified patients received care and treatment.

In the multilevel analysis, we found that EAs in cluster areas were more distant from health facilities than other EAs; low TB notification and distance from the health facility were positively correlated (8). Distance from TB services affects access to TB services since it is interrelated with cost and time (32). Still, in rural settings, symptomatic people receive care late and with less sensitive diagnostic (13). TB prevention and care efforts should address patient delays and stigma affecting healthcare-seeking (33). TB screening targeted remote areas could detect clusters more reliably and reduce community transmission. There has been no previous research linking socioeconomic position and population clustering. In our analysis at the EA level, we sought not only to identify clustering but also risk factors for clustering that may require attention from public health practices. However, our findings did not show associations between clustering and wealth index and a multisite study would better identify socioeconomic-based differences in PTB clustering.

Our study has some limitations. We focused on patients with TB symptoms, missing the asymptomatic. While this may introduce bias for most individuals, this is usually a temporary condition, as symptoms in most will develop later (31, 34). Moreover, our study observed the populations for a year, with repeated screening of symptoms, so the number of undetected cases and asymptomatic individuals should be low. It's important to note that misdiagnosis can occur despite the implementation of quality assurance measures, primarily due to the low sensitivity of smear microscopy. We did not include 17 PTB with missing location information in the study, which may influence identifying true clusters.

Populated urban areas have both more TB services and more TB contacts and transmission than rural areas, even though infection is often within close relatives in rural areas. Therefore, clusters of PTB cases may be more prevalent in urban areas where access to care is greater while being less reported in rural areas due to limited access to care (22). On the contrary, TB incidence has shown an inverse association with

urbanization in developed, accessible health services and improved living standards.

Identifying TB patients and clusters based only on passive case-finding misses undiagnosed and un-notified patients. GIS studies based on notifications identifying clusters cannot distinguish between "real" TB incidence differences and the role of health services (35).

In program context, many areas may perform well, but some areas with many cases and less developed systems may need more attention and focused strategies. GIS has the potential to be helpful in close contact and neighbourhood TB screening in areas where TB cases are concentrated (30) and indicate areas with low notification but with many cases as well. Active TB screening, on the other hand, can enhance TB programme outcomes in these areas but is costly and usually regarded almost impossible as a routine TB prevention and care strategy in limited-resource settings, and hence requires a thorough evaluation before implementation, particularly with X-ray and molecular diagnostic tests, which are very expensive. Therefore, a cheap regular HEW-based symptom screening, with routine monitoring of presumptive TB patients identified and confirmed TB among referred could help identify catchment facilities with lower than expected presumptive TB rates as low cluster areas may be more alarming than high cluster areas (36).

Conclusions

The distribution of PTB was not uniform but concentrated in particular regions distant from health centres, indicating the need to strengthen decentralized TB services in remote settings. Routine systematic community screening using existing health infrastructure with HEWs may be costly but through targeted screening they can identify and refer persons with TB symptoms more quickly for diagnosis and treatment, thereby decreasing the transmission and contributing to the reduction of TB burden.

Abbreviations

BCG, Bacillus-Calmette Guerin; BMI, Body mass index; EA, Enumeration area; GPS, Global positioning systems; GIS, Geographic information systems; HEW, Health extension worker; MUAC, Mid-upper arm circumference; PTB, pulmonary tuberculosis; TB, tuberculosis

Ethics approval and consent to participate

Ethics approval was obtained from the AHRI-Alert Ethics Review Committee, Ethiopia (PO12/15), the National Research Ethics Committee, Ethiopia (no 104/2016), and the Regional Committees for Medical and Health Research Ethics in Norway (2015/1006). Written consent was obtained from participants and parent(s)/Guardian(s) for participants aged 15-17 years.

Data sharing statement

Data cannot be shared publicly because of confidentiality to protect the study participants since data sets generated and/or analysed in this study are not publicly available. Data are available upon reasonable request from the corresponding author, ABB. To receive access to the data, the applicant will need to provide ethical approval from IRB or an equivalent body and approval from the AHRI-ALERT and National Research Ethics Committees in Ethiopia.

Competing interests

The authors have declared that no competing interests exist.

Funding information

The Norwegian Health Association (NHA) and Norwegian Institute of Public Health (NIPH) funded the project. The University of Bergen funded the publication fee. ABB received funding. Grant number N/A. Funders have no role in the study design, data collection, analysis, decision to publish, or preparation of the manuscript.

Author contributions

Conceptualization: ABB, BAW, DGD, EH; Data analysis: ABB, MHD; Methodology: ABB, BAW, MHD, SGH; Visualization: ABB, MHD; Writing original draft: ABB, MHD, BAW, SGH; Writing review and editing: ABB, MHD, DGD, EH, BAW, SGH. All authors reviewed and approved final manuscript.

Acknowledgements

We want to acknowledge, data collectors, healthcare workers and patients. We would like to thank the Sidama Regional Health Bureau for their support during the project period. The University of Bergen funded the publication fee.

ORCID

Abiot Bezabeh Banti: <https://orcid.org/0000-0003-0441-8845>

Brita Askeland Winje: <https://orcid.org/0000-0003-2858-7248>

Sven Gudmund Hinderaker: <https://orcid.org/0000-0003-2725-0248>

Einar Heldal: <https://orcid.org/0000-0002-8741-792X>

References

1. Reid MJA, Arinaminpathy N, Bloom A, Bloom BR, Boehme C, Chaisson R, et al. Building a tuberculosis-free world: The Lancet Commission on tuberculosis. *Lancet* (London, England). 2019;393(10178):1331-84. Epub 2019/03/25. doi: 10.1016/s0140-6736(19)30024-8. PubMed PMID: 30904263.
2. World Health Organization. Global Tuberculosis Report, 2020.
3. World Health Organization. Systematic screening for active tuberculosis: principles and recommendations: WHO; 2013.
4. Bolstad P. GIS Fundamentals: a first text on geographic information systems. 6th edition. ed. Ann Arbor, MI: XanEdu; 2019.
5. Alene KA, Python A, Weiss DJ, Elagali A, Wagaw ZA, Kumsa A, et al. Mapping tuberculosis prevalence in Ethiopia using geospatial meta-analysis. 2023:dyad052.
6. Shaweno D, Karmakar M, Alene KA, Ragonnet R, Clements ACA, Trauer JM, et al. Methods used in the spatial analysis of tuberculosis epidemiology: a systematic review. *BMC Medicine*. 2018;16(1):193. doi: 10.1186/s12916-018-1178-4.
7. Tadesse T, Demissie M, Berhane Y, Kebede Y, Abebe M. The clustering of smear-positive tuberculosis in Dabat, Ethiopia: a population-based cross-sectional study. 2013;8(5):e65022.
8. Dangisso MH, Datiko DG, Lindtjørn B. Identifying geographical heterogeneity of pulmonary tuberculosis in southern Ethiopia: a method to identify clustering for targeted interventions. 2020;13(1):1785737.
9. Dangisso MH, Datiko DG, Lindtjørn B. Spatio-Temporal Analysis of Smear-Positive Tuberculosis in the Sidama Zone, Southern Ethiopia. *PLOS ONE*. 2015;10(6):e0126369. doi: 10.1371/journal.pone.0126369.
10. Shaweno D, Trauer JM, Denholm JT, McBryde ES. The role of geospatial hotspots in the spatial spread of tuberculosis in rural Ethiopia: a mathematical model. 2018;5(9):180887. doi: doi:10.1098/rsos.180887.

11. Tadesse S, Enqueselassie F, Hagos S. Spatial and space-time clustering of tuberculosis in Gurage Zone, Southern Ethiopia. PLOS ONE. 2018;13(6):e0198353. doi: 10.1371/journal.pone.0198353.
12. Shaweno D, Trauer JM, Doan TN, Denholm JT, McBryde ES. Geospatial clustering and modelling provide policy guidance to distribute funding for active TB case finding in Ethiopia. Epidemics. 2021;36:100470. doi: <https://doi.org/10.1016/j.epidem.2021.100470>.
13. Banti AB, Winje BA, Hinderaker SG, Heldal E, Abebe M, Dangisso MH, et al. Prevalence and incidence of symptomatic pulmonary tuberculosis based on repeated population screening in a district in Ethiopia: a prospective cohort study. 2023;13(7):e070594.
14. Fekadu L, Hanson C, Osberg M, Makayova J, Mingkwan P, Chin D. Increasing access to tuberculosis services in Ethiopia: findings from a patient-pathway analysis. 2017;216(suppl_7):S696-S701.
15. Erik von Elm, Douglas G. Altman, Matthias Egger, Stuart J. Pocock, Peter C. Gøtzsche, Jan P. Vandembroucke. The Strengthening the Reporting of Observational Studies in Epidemiology (STROBE) statement: guidelines for reporting observational studies. Journal of Clinical Epidemiology, Volume 61, Issue 4, 2008, Pages 344-349.
16. Banti AB, Datiko DG, Hinderaker SG, Heldal E, Dangisso MH, Mitiku GA, et al. How many of persistent coughers have pulmonary tuberculosis? Population-based cohort study in Ethiopia. 2022;12(5):e058466.
17. Kulldorff M. A spatial scan statistic. 1997;26(6):1481-96.
18. Zhang Z, Assunção R, Kulldorff M. Spatial Scan Statistics Adjusted for Multiple Clusters. Journal of Probability and Statistics. 2010;2010:642379. doi: 10.1155/2010/642379.

19. Roth D, Otterstatter M, Wong J, Cook V, Johnston J, Mak S. Identification of spatial and cohort clustering of tuberculosis using surveillance data from British Columbia, Canada, 1990–2013. *Social Science & Medicine*. 2016;168:214-22. doi: <https://doi.org/10.1016/j.socscimed.2016.06.047>.
20. Kulldorff M, Athas WF, Feurer EJ, Miller BA, Key CRJAjoph. Evaluating cluster alarms: a space-time scan statistic and brain cancer in Los Alamos, New Mexico. 1998;88(9):1377-80.
21. Federal Ministry of Health. Guidelines for Clinical and Programmatic Management of TB, TB/HIV, DR-TB and Leprosy in Ethiopia. 2021.
22. Porskrog A, Bjerregaard-Andersen M, Oliveira I, Joaquím LC, Camara C, Andersen PL, et al. Enhanced tuberculosis identification through 1-month follow-up of smear-negative tuberculosis suspects. *The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease*. 2011;15(4):459-64. Epub 2011/03/15. doi: 10.5588/ijtld.10.0353. PubMed PMID: 21396203.
23. Trauer JM, Dodd PJ, Gomes MGM, Gomez GB, Houben RMGJ, McBryde ES, et al. The Importance of Heterogeneity to the Epidemiology of Tuberculosis. *Clinical Infectious Diseases*. 2019;69(1):159-66. doi: 10.1093/cid/ciy938.
24. Holodinsky JK, Austin PC, Williamson TS. An introduction to clustered data and multilevel analyses. *Family Practice*. 2020;37(5):719-22. doi: 10.1093/fampra/cmaa017 J Family Practice.
25. Hox JJ, Moerbeek M, Van de Schoot R. *Multilevel analysis: Techniques and applications*: Routledge; 2017.
26. Şahin M, Aybek E. Jamovi: an easy to use statistical software for the social scientists. 2019;6(4):670-92.
27. Zelner JL, Murray MB, Becerra MC, Galea J, Lecca L, Calderon R, et al. Identifying Hotspots of Multidrug-Resistant Tuberculosis Transmission Using Spatial and

- Molecular Genetic Data. *The Journal of Infectious Diseases*. 2016;213(2):287-94.
doi: 10.1093/infdis/jiv387.
28. Tafess K, Beyen TK, Girma S, Girma A, Siu G. Spatial clustering and genetic diversity of *Mycobacterium tuberculosis* isolate among pulmonary tuberculosis suspected patients, Arsi Zone, Ethiopia. *BMC Pulmonary Medicine*. 2021;21(1):206.
doi: 10.1186/s12890-021-01567-7.
29. Taye H, Alemu K, Mihret A, Ayalew S, Hailu E, Wood JLN, et al. Epidemiology of *Mycobacterium tuberculosis* lineages and strain clustering within urban and peri-urban settings in Ethiopia. *PLOS ONE*. 2021;16(7):e0253480. doi: 10.1371/journal.pone.0253480.
30. Marks GB, Nguyen NV, Nguyen PT, Nguyen T-A, Nguyen HB, Tran KH, et al. Community-wide screening for tuberculosis in a high-prevalence setting. *PLoS ONE*. 2019;381(14):1347-57.
31. Wong EB. It Is Time to Focus on Asymptomatic Tuberculosis. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America*. 2021;72(12):e1044-e6. Epub 2020/12/08. doi: 10.1093/cid/ciaa1827. PubMed PMID: 33283223; PubMed Central PMCID: PMC8204778.
32. Hoa NP, Chuc NTK, Thorson A. Knowledge, attitudes, and practices about tuberculosis and choice of communication channels in a rural community in Vietnam. *Health Policy*. 2009;90(1):8-12. doi: <https://doi.org/10.1016/j.healthpol.2008.08.006>.
33. Ochom E, Robsky K, Gupta A, Tamale A, Kungu J, Turimumahoro P, et al. Geographic distribution and predictors of diagnostic delays among possible TB patients in Uganda. *BMC Public Health*. 2023;13(3):70-6.
34. Frascella B, Richards AS, Sossen B, Emery JC, Odone A, Law I, et al. Subclinical Tuberculosis Disease-A Review and Analysis of Prevalence Surveys to Inform Definitions, Burden, Associations, and Screening Methodology. *Clinical infectious diseases: an official publication of the Infectious Diseases Society of America*.

2021;73(3):e830-e41. Epub 2020/09/17. doi: 10.1093/cid/ciaa1402. PubMed PMID: 32936877; PubMed Central PMCID: PMC8326537.

35. Deng LL, Zhao F, Li ZW, Zhang WW, He GX, Ren X. Epidemiological characteristics of tuberculosis incidence and its macro-influence factors in Chinese mainland during 2014-2021. *Infect Dis Poverty*. 2024 May 21;13(1):34. doi: 10.1186/s40249-024-01203-6. PMID: 38773558; PMCID: PMC11107005.
36. Dlodlo RA, Brigden G, Heldal E, Allwood B, Chiang C-Y, Fujiwara PI, et al. *Management of Tuberculosis: a guide to essential practice*. 7th ed. Paris, France: International Union Against Tuberculosis and Lung Disease, 2019.
<https://www.theunion.org/what-we-do/publications/technical/management-of-tuberculosis-a-guide-to-essential-practice>.

Supplemental appendix

Supplemental file 1.

Supplemental file 2.

Target adult population in 36 kebeles in Dale

Round I
screening

Presumptive TB=2052

PTB=122

Round II
screening

Presumptive TB=749

PTB=159

Round III
screening

Presumptive TB=945

PTB=161

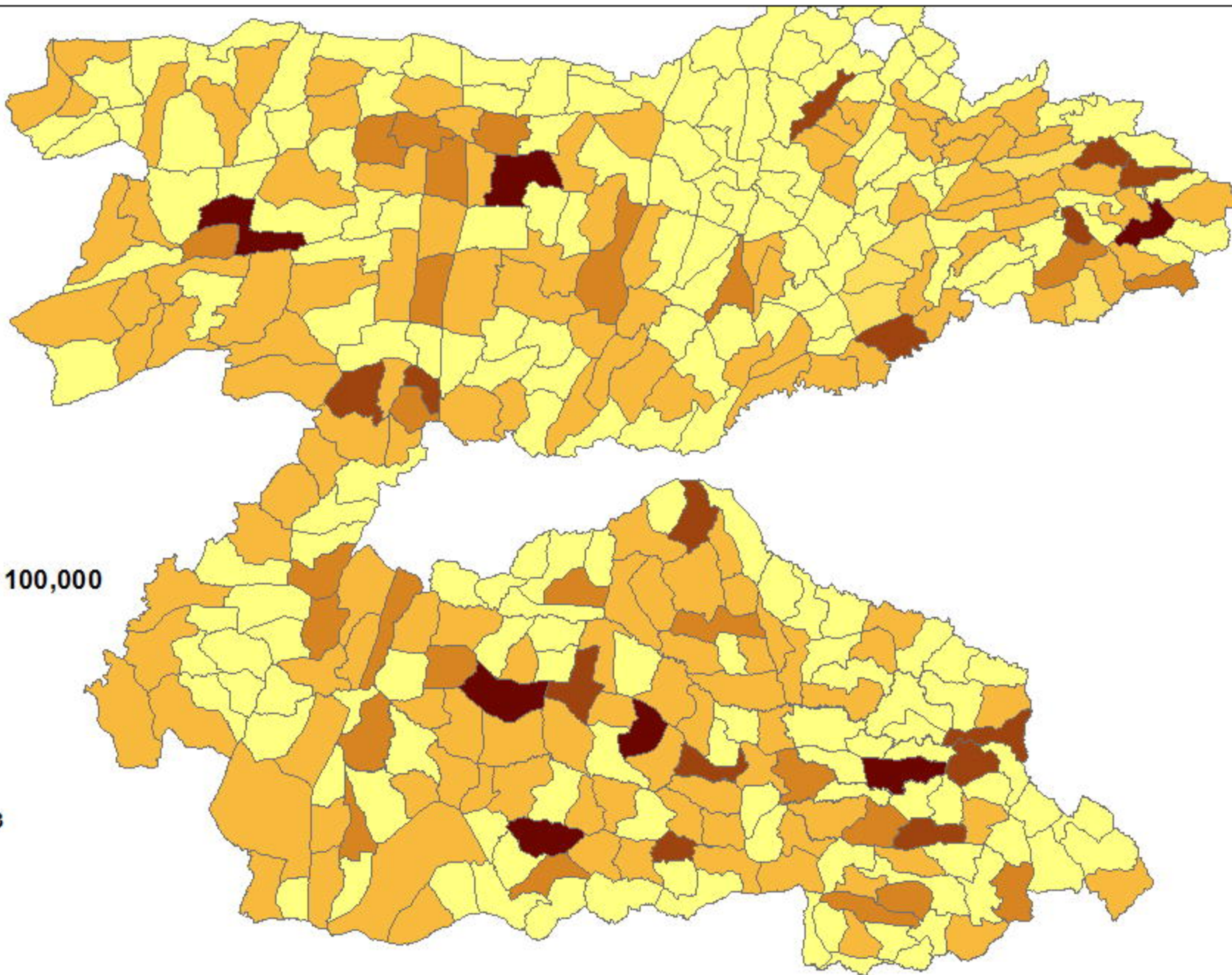
All round
screening

Presumptive TB=3746

PTB=442

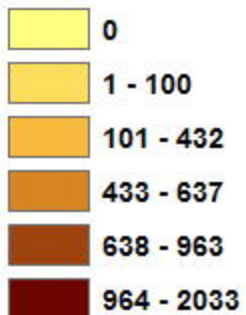
PTB without
geographic
information were
excluded=17

PTB eligible for the
study=425



Legend

PTB rate per 100,000



0 1 2 4 6 8 Kilometers

