Kapwata *et al. BMC Research Notes* (2024) 17:364
https://doi.org/10.1186/s13104-024-07015-z

BMC Research Notes

RESEARCH NOTE **Open Access**

The identification and spatial distribution of hotspots of tuberculosis occurrence in South Africa

Thandi Kapwata[1,2*], Gregory Breetzke[3], Caradee Y. Wright[3,4] and Tessa S. Marcus[5]

Abstract

Objective Prior research has shown strong evidence of spatial clustering of tuberculosis across a range of contexts. Identifying the spatial patterning of infectious diseases such as tuberculosis is crucial as it allows for targeted intervention strategies, directing healthcare resources efficiently to areas where tuberculosis incidence is concentrated. This is especially true for low- and middle-income countries that typically experience greater resource constraints relative to their Global North counterparts. In this study, we extend existing literature by investigating the spatial patterning of tuberculosis among vulnerable communities in South Africa, notably in the relatively under-researched provinces of the North-West and Gauteng.

Results Data for this study were collected from several locations implementing community-oriented primary care in the country. Community health workers used AitaHealth™, a custom-built mobile information management application, to obtain data on tuberculosis status and environmental conditions of households. We find notable clusters of tuberculosis in these provinces which we speculate could be associated with urban formal and informal settlement densification and overcrowding, the incidence of mining activities prevalent in sampled locations and poor access to healthcare.

Keywords GIS, Infectious disease, Spatial scan statistics

Background

Tuberculosis (TB) is an infectious disease caused by Mycobacterium tuberculosis [6]. Globally, an estimated two billion people contract TB each year with roughly 10.6 million becoming ill as a result of the disease [2]. Despite being preventable and treatable, TB remains deadly leading to an estimated 1.3 million deaths annually, worldwide [2]. In 2022 alone, 2.5 million people across the African continent fell ill with TB, contributing to a quarter of new TB cases worldwide [21]. Although TB incidence and deaths have been declining in South Africa in recent years, this disease remains among the leading causes of death in the country [12]. According to the World Health Organization's (WHO) 2023 Global TB report, approximately 54 200 people died of TB in South Africa in 2022 [20]. Moreover, the country's first national TB prevalence survey conducted in 2018 confirmed that TB poses a significant health challenge in the country with an estimated prevalence of 852 cases per 100,000 population [1].

---

*Correspondence:
Thandi Kapwata
Thandi.Kapwata@mrc.ac.za
[1] Environment and Health Research Unit, South African Medical Research Council, Johannesburg 2193, South Africa
[2] Environmental Health Department, Faculty of Health Sciences, University of Johannesburg, Johannesburg, South Africa
[3] Department of Geography, Geoinformatics and Meteorology, University of Pretoria, Pretoria 0002, South Africa
[4] Environment and Health Research Unit, South African Medical Research Council, Pretoria 0002, South Africa
[5] COPC Research Unit, Faculty of Health Sciences, University of Pretoria, Pretoria 0002, South Africa

In order to reduce and ultimately eliminate TB, the World Health Assembly recently adopted the WHO's post-2015 End TB Strategy [19] This strategy aims to reduce the incidence of TB globally by 90% by 2030 (compared with 2015 levels) [19]. In South Africa, the National Strategic Plan (NSP) which guides the country's approach and commitment to addressing Human immunodeficiency virus (HIV), TB and sexually transmitted infections (STIs) in the country, has also adopted this target [14]. TB control programmes are an important component of tackling the health burden of TB through disease surveillance and identifying communities at risk in order to facilitate the prioritization of intervention strategies and other control activities.

Recent evidence suggests that control strategies that target TB incidence hotspots could improve local control and reduce the risk of transmission spillover into surrounding areas [4]. To provide evidence for a spatially-targeted intervention approach this study used spatial epidemiology to investigate the spatial distribution of TB to determine local clustering of the disease in selected vulnerable communities in South Africa to provide evidence for spatially-targeted interventions.

Methods

Study sites and participant data

The data used in the study were collected as part of a community-oriented primary care (COPC) programme. The initiative was implemented across five provinces in South Africa including the Eastern Cape, Free State, Gauteng, KwaZulu Natal and North-West. Site selection was guided by high disease burden and poor healthcare coverage; therefore, the majority of the sampled communities were located in informal and densified settlements in these provinces [8]. Health status for TB was collected as part of the COPC programme between 2014 and 2019. The majority of households in the data set were located in Gauteng Province (76.6%), followed by the North-West (19%). The Free State, Mpumalanga, Eastern Cape and KwaZulu-Natal accounted for 4% of the study sampled (3%, 0.5%, 0.3% and 0.2%, respectively). Individual TB cases for each household were extracted from the database using the definitions provided in Fig. 1. To calculate prevalence, total population data was obtained from the most recent national census conducted in the country [15].

Analysis

Spatial scan statistics were conducted using SaTScan™ [11] to detect statistically significant clusters for high risk of TB in the regions where participants were recruited. Thereafter, a purely spatial analysis using a Poisson-based model was employed to determine if there were more TB cases than would be expected by any regions. SaTScan imposes circles of different sizes on the geographic area and computes a likelihood ratio statistic (known as the scan statistic) based on the number of observed and expected cases within and outside the circle and compared with the likelihood under the null hypothesis. The analysis was undertaken at the ward administrative level because that is the resolution at which population data was available. The prevalence of TB calculated per 100,000 population for each ward as shown below:

$$\text{TB Prevalence} = \frac{\text{Number of TB cases per ward}}{\text{Total population per ward}} \times 100{,}000.$$

The maximum cluster size was set to 50% of the population at risk because it is recommended that clusters should contain at most 50% of the total population at risk. Relative risk (RR) was calculated in each cluster to evaluate the risk of TB within the identified circular boundaries. A p-value <0.05 indicated 95% statistical significance. Clusters were mapped using ArcMap version 10.5.1 [5].

Results and discussion

TB prevalence

The North-West province had the highest TB prevalence in the study followed by Gauteng with averages of 19.82 cases and 11.51 per 100,000 population, respectively (see Fig. 2). This is particularly concerning because the North-West province had about four times fewer households enrolled in the study compared to Gauteng. An ongoing challenge for TB programmes globally is patients who are diagnosed with TB, but not started on treatment or are on delayed treatment. A report by the Health System Trust found that less than half the patients diagnosed with TB in the North-West province in South Africa had started treatment for TB [12]. In addition, the report found that the province had amongst the lowest treatment success rates with 6.3% lost to follow-up, having interrupted treatment for two consecutive months or more. The high TB burden in the North-West could be indicative of a need for more targeted TB monitoring and control policies.

TB cluster detection

High-rate TB clusters detected by purely spatial analysis are described in Table 1. The most likely cluster, which was least likely due to chance, covered the northern

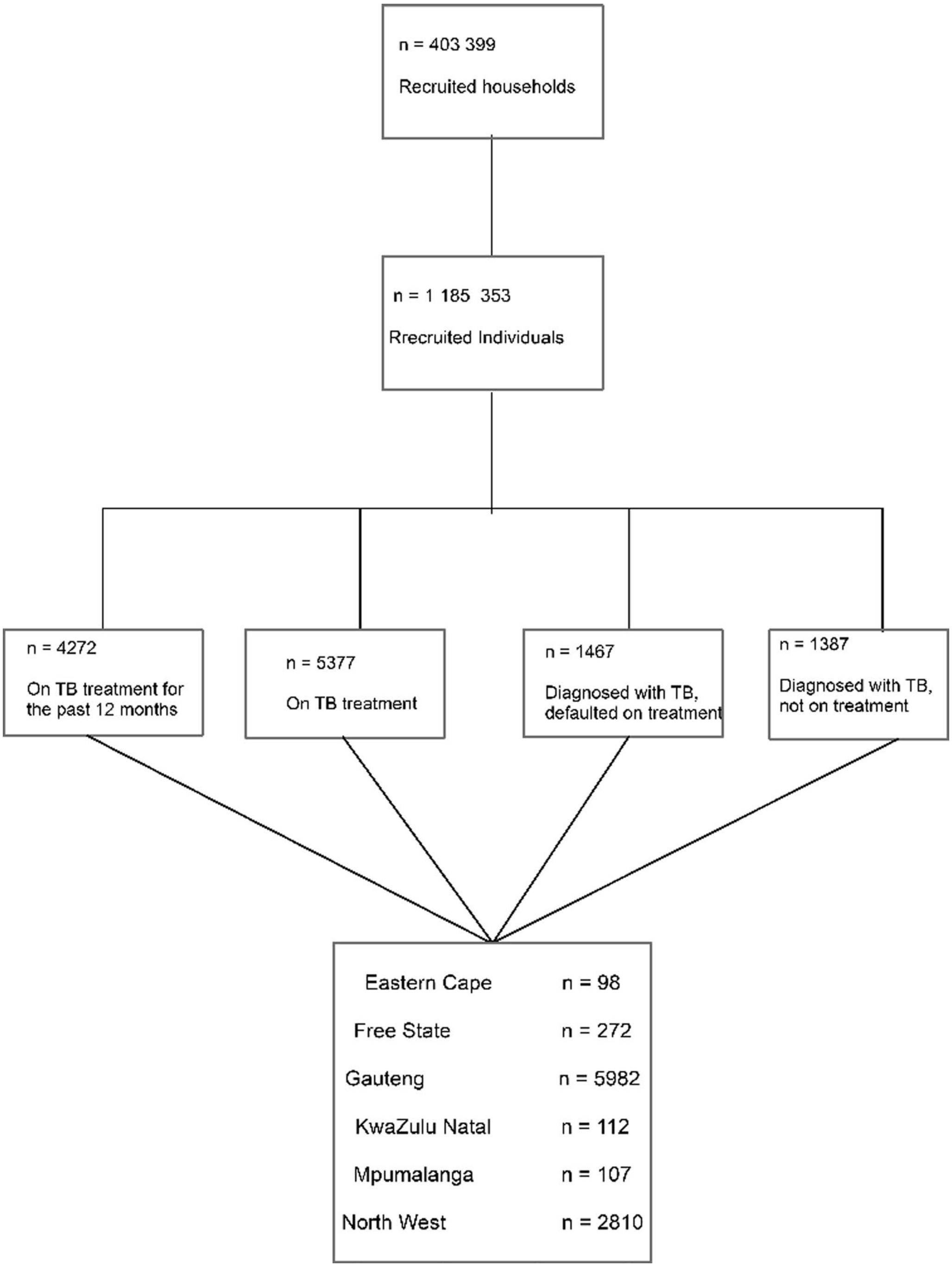

Fig. 1 Overview of study population by treatment/diagnosis and province

part of the Gauteng province and the north-eastern part of the North-West province (see Fig. 3). This cluster, with a radius of 106.21 km, had 5660 cases of TB compared to the expected number of cases of 761 (based on the population if the risk of TB was evenly distributed). People within this cluster had 20 times higher risk of TB infection than those outside the cluster. The first significant secondary cluster includes regions within the borders of Gauteng and North-West province. Individuals inside this cluster had 18 times more risk of TB infection

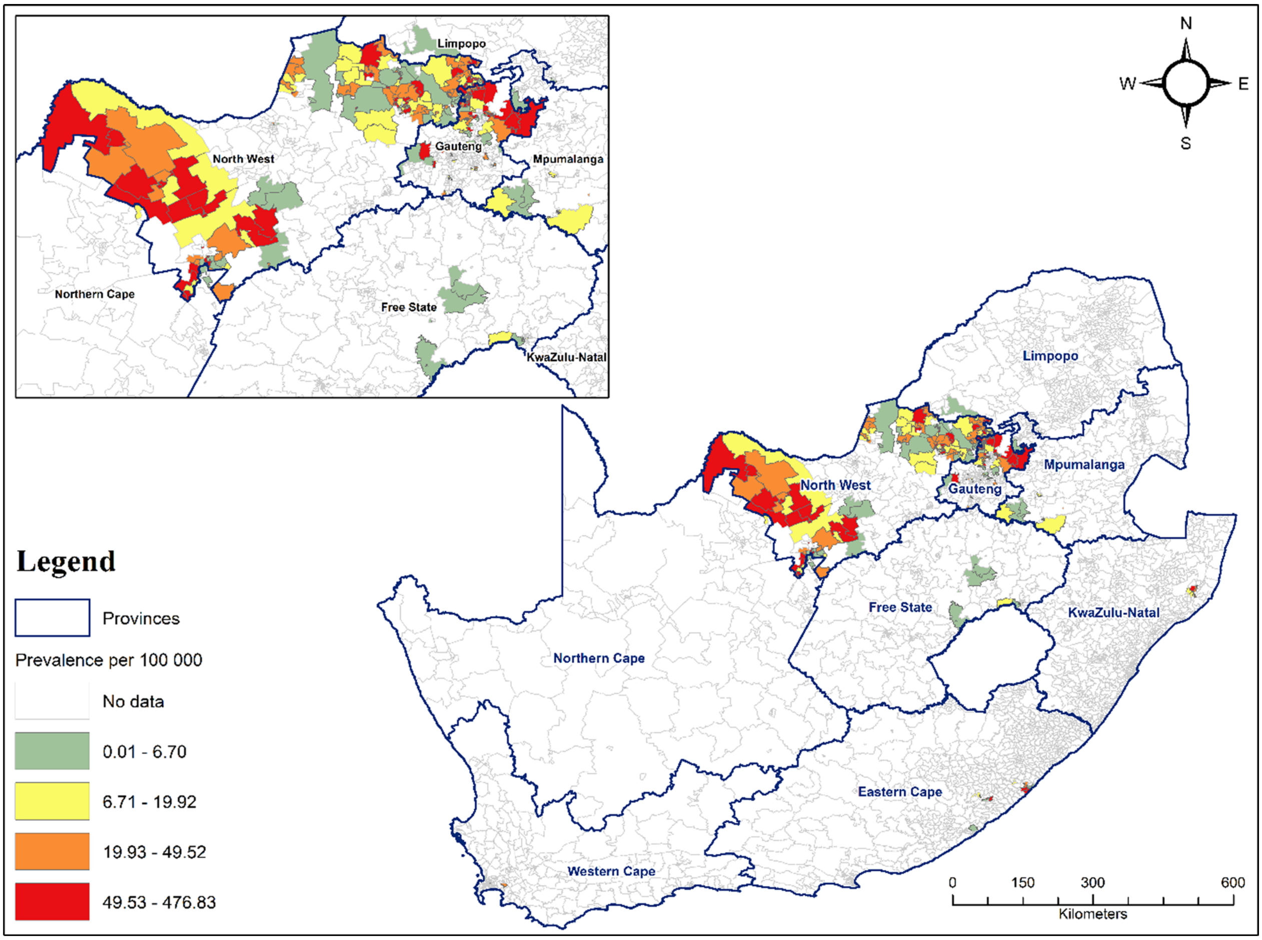

Fig. 2 Spatial distribution of TB prevalence per ward

compared to other regions outside the cluster. The remaining 16 statistically secondary clusters were distributed across Gauteng, North-West and Mpumalanga provinces with smaller clusters being detected in Free State and KwaZulu-Natal provinces.

Our study used the Kulldorff's scan statistic to detect statistically significant clusters for high occurrence of TB in vulnerable communities in South Africa. This method has been used globally to detect purely spatial and space–time clusters of infectious diseases, including TB [17, 22]. Several clusters with high-risk ratios for TB were detected in regions where TB data were collected with the most likely clusters being located around border areas. The occurrence of the most likely cluster in Gauteng and North-West provinces was unlikely to be due to higher COPC recruitment numbers in those regions as service delivery coverage using AitaHealth™ has only been ubiquitous in Mamelodi in the City of Tshwane. Whilst speculative, the most plausible explanation for this clustering could be settlement growth, population densification, household crowding as evidenced by previous studies [3, 7, 13]. Significant socio-economic internal and international in-migration could also play a role as people move closer to livelihood opportunities, especially around mining. Several studies have found that mining activities were associated with higher population TB incidence rates [16] as mine workers are at a higher risk of contracting and transmitting TB due to prolonged exposure to pollutants, poor living conditions, and high HIV prevalence in mining communities [10, 18]. In addition, regions with statistically significant clusters for high occurrence of TB in the North West province are known to have low healthcare density [9]. Therefore, poor access to healthcare facilities could also be a contributing factor to the high burden of TB disease in these areas although more research is required to substantiate this assertion.

Table 1 Purely spatial clusters for high occurrence of TB in South Africa, 2014–2018

Cluster type	Coordinates (South, East)/ radius (km)	P value	Population	Observed cases	Expected cases	RR
Most likely cluster	(27.96, – 25.01)/106.21	< 0.001	20,302,305	5660	761.38	20.54
Secondary cluster 1	(28.27, – 25.39)/40.98	< 0.001	8,944,443	3700	335.44	18.86
Secondary cluster 2	(24.67, – 27.31)/100.77	< 0.001	2,844,495	986	106.68	10.33
Secondary cluster 3	(27.21, – 25.66)/68.9	< 0.001	4,866,324	1061	182.50	6.51
Secondary cluster 4	(28.01, – 25.94)/22.85	< 0.001	6,862,365	699	257.35	2.87
Secondary cluster 5	(27.66, – 26.36)/4.36	< 0.001	56,973	80	2.14	37.79
Secondary cluster 6	(28.3, – 26.21)/0.99	< 0.001	108,945	95	4.09	23.51
Secondary cluster 7	(29.24, – 25.96)/74.69	< 0.001	6,285,507	550	235.72	2.43
Secondary cluster 8	(32.1, – 28.28)/0	< 0.001	41,715	59	1.56	37.97
Secondary cluster 9	(27.76, – 26.2)/2.5	< 0.001	168,885	80	6.33	12.74
Secondary cluster 10	(28.23, – 32.33)/0	< 0.001	24,918	27	0.93	28.98
Secondary cluster 11	(28.19, – 26.4)/0	< 0.001	182,805	48	6.86	7.04
Secondary cluster 12	(28.87, – 32.15)/6.69	< 0.001	27,735	23	1.04	22.17
Secondary cluster 13	(28.42, – 26.36)/0	< 0.001	59,112	28	2.22	12.67
Secondary cluster 14	(27.57, – 28.9)/1.33	< 0.001	31,470	19	1.18	16.13
Secondary cluster 15	(27.78, – 26.74)/4.68	< 0.001	122,385	30	4.59	6.56
Secondary cluster 16	(27.88, – 26.55)/0	< 0.001	40,830	17	1.53	11.12
Secondary cluster 17	(27.84, – 26.5)/0	< 0.001	153,117	27	5.74	4.71
Secondary cluster 18	(27.88, – 26.24)/0	< 0.001	111,966	20	4.20	4.77

Conclusion

The main aim of this research was to identify spatial clustering of increased TB risk in vulnerable regions of South Africa. We did this using data captured by a new mobile community healthcare management application which allowed direct and personal contact with participants in their households. The results of the study indicated that TB was spatially concentrated in communities made vulnerable by settlement and dwelling densification, population movement and livelihood seeking and occupation. In addition, the lack of health care facilities could contribute to clusters of TB transmission due to poor access to diagnosis and treatment. As one of the first of its kind in southern Africa, these findings are highly significant, not least of all because they support the urgent need to use spatial analysis of clusters of TB to ramp up and make TB disease control more effective. This is especially important in South Africa where TB remains the number one cause of death in the country. We believe that the results of this study can help in the identification of high-risk populations, such as those in overcrowded or impoverished areas, who may have limited access to healthcare services, and assist in informing future equitable health planning and resource allocation.

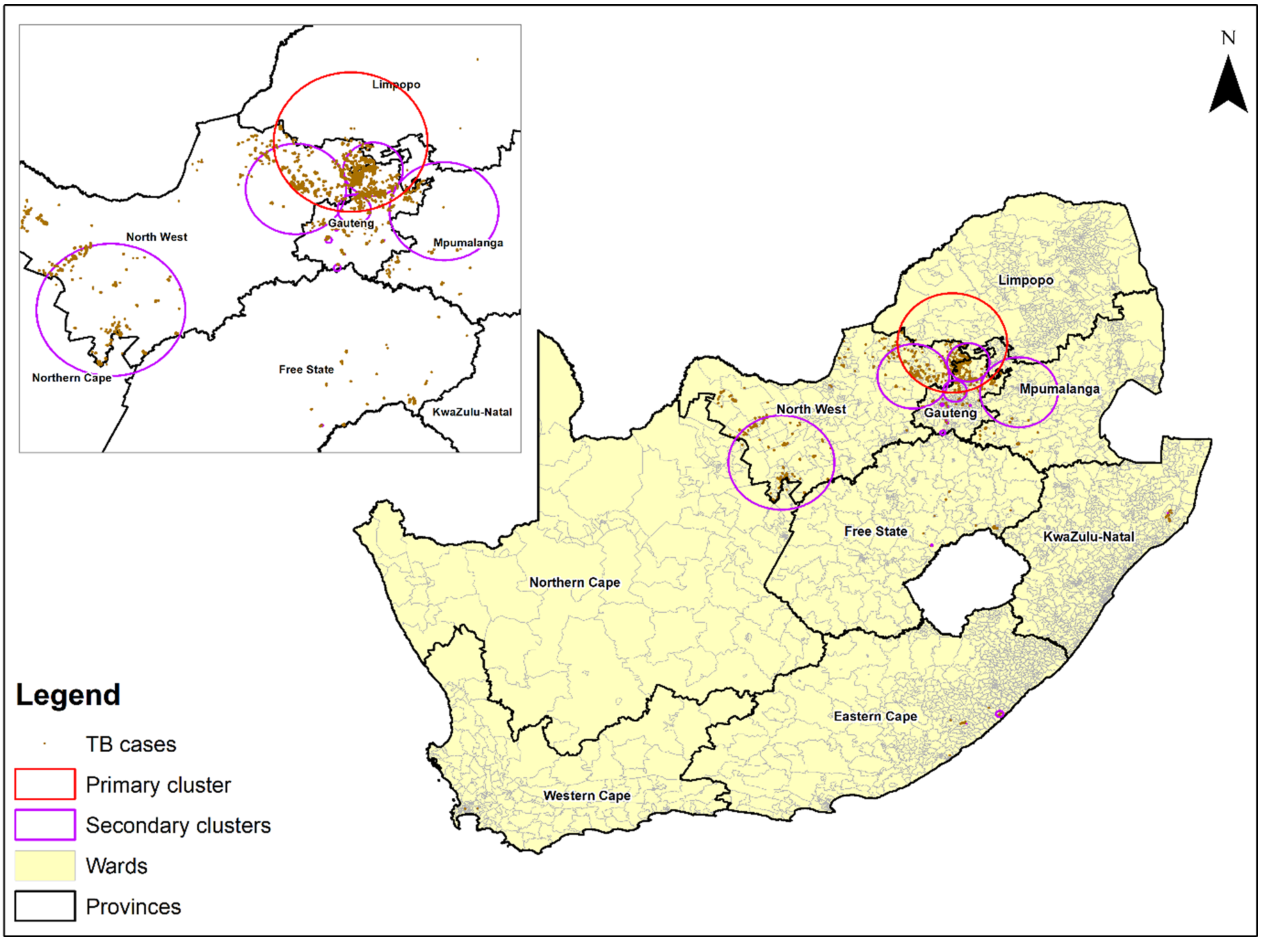

Fig. 3 Statistically significant purely spatial clusters for high occurrence of TB in South Africa

Limitations

One of the limitations of our findings is that it is possible that the clusters of TB cases were underestimated due to the rapid growth of informal settlements whose population may not have been included in the population count.

Abbreviations

COPC	Community-oriented primary care
TB	Tuberculosis
HIV	Human immunodeficiency virus
STIs	Sexually transmitted infections

Acknowledgements

We acknowledge the University of Pretoria Department of Family Medicine and the City of Tshwane for access to the data.

Author contributions

Conceptualization: TK, GB, C.Y.W, TSM; acquisition of data: GB, TSM; formal analysis and interpretation of data: TK; writing—original draft: TK, GB, C.Y.W.; writing—review and editing: All authors. All authors have read and agreed to the published version of the manuscript.

Funding

TK and CYW receive funding from the South African Medical Research Council.

Availability of data and materials

All data analysed for this study are available upon reasonable request and should be directed to: Prof JFM Hugo, University of Pretoria COPC Research Unit, South Africa.

Declarations

Ethics approval and consent to participate

The ethics for this study was approved by the University of Pretoria's Faculty of Health Science's Research Ethics Committee, Pretoria, South Africa (Reference number—UP 102/2011). This study used third party data however, informed consent was provided by all study participants in the primary study.

Consent for publication

Study participants were informed that collected data will be anonymised for use in publications and they provided signed consent as part of the informed consent process. No individual/identifying data is presented.

Competing interests

The authors declare no competing interests.

Received: 4 June 2024 Accepted: 26 November 2024

References

1. Ayles H, Mureithi L, Simwinga M. The state of tuberculosis in South Africa: what does the first national tuberculosis prevalence survey teach us? Lancet Infect Dis. 2022;22:1094–6.
2. CDC 2024. Centers for Disease Control and Prevention (CDC): Global TB Overview. Available at https://www.cdc.gov/globalhivtb/who-we-are/about-us/globaltb/globaltb.html.
3. Cramm JM, Koolman X, Møller V, Nieboer AP. Socio-economic status and self-reported tuberculosis: a multilevel analysis in a low-income township in the Eastern Cape, South Africa. J Public Health Afr. 2011;2.
4. Cudahy PGT, Andrews JR, Bilinski A, Dowdy DW, Mathema B, Menzies NA, Salomon JA, Shrestha S, Cohen T. Spatially targeted screening to reduce tuberculosis transmission in high-incidence settings. Lancet Infect Dis. 2019;19:e89–95.
5. ESRI 2010. Environmental Systems Research Institute (ESRI). ArcGIS 10.0. Redlands, CA.
6. Furin J, Cox H, Pai M. Tuberculosis. The Lancet. 2019;393:1642–56.
7. Harling G, Ehrlich R, Myer L. The social epidemiology of tuberculosis in South Africa: a multilevel analysis. Soc Sci Med. 2008;66:492–505.
8. Kapwata T, Breetzke G, Wright CY, Marcus TS, Eales O. Demographic and socio-economic risk factors associated with self-reported TB. Int J Tuberc Lung Dis. 2022;26:33–7.
9. Kapwata T, Manda S. Geographic assessment of access to health care in patients with cardiovascular disease in South Africa. BMC Health Serv Res. 2018;18:197.
10. Kootbodien T, Iyaloo S, Wilson K, Naicker N, Kgalamono S, Haman T, Mathee A, Rees D. Environmental silica dust exposure and pulmonary tuberculosis in Johannesburg, South Africa. Int J Environ Res Public Health. 2019;16:1867.
11. Kulldorff M, Heffernan R, Hartman J, Assunçao R, Mostashari F. A space–time permutation scan statistic for disease outbreak detection. PLoS Med. 2005;2.
12. Loveday M, Mzobe Y, Pillay Y, Barron P. Figures of the dead: a decade of tuberculosis mortality registrations in South Africa. S Afr Med J. 2019;109:728–32.
13. Marcus SM, Marcus TS. Infrastructural inequality and household COVID-19 vulnerability in a south African urban settlement. J Urban Health. 2022;99:571–81.
14. NSP 2016. South Africa's National Strategic Plan (NSP) for HIV, TB and STIs 2017–2022. Available at https://sanac.org.za//wp-content/uploads/2017/06/NSP_FullDocument_FINAL.pdf.
15. STATS SA 2011. Census 2011 Statistical release. Stats SA Library Cataloguing-in-Publication (CIP) Data Statistics South Africa. Pretoria: Statistics South Africa.
16. Stuckler D, Basu S, McKee M, Lurie M. Mining and risk of tuberculosis in sub-Saharan Africa. Am J Public Health. 2011;101:524–30.
17. TADESSE, T., DEMISSIE, M., BERHANE, Y., KEBEDE, Y. & ABEBE, M. 2013. The clustering of smear-positive tuberculosis in Dabat, Ethiopia: a population based cross sectional study. *PloS one*, 8.
18. THE WORLD BANK 2020. The World Bank: The Southern Africa TB in the Mining Sector Initiative. Available at https://www.worldbank.org/en/programs/the-southern-africa-tb-in-the-mining-sector-initiative#:~:text=In%20South%20Africa%20alone%2C%20TB,3%2C000%20cases%20per%20100%2C000%20individuals.&text=Mine%20workers%20are%20at%20a,HIV%20prevalence%20in%20mining%20communities. Accessed on 6 Nov 2020.
19. WHO 2015. World Health Organization (WHO): The End TB Strategy. Available at https://www.who.int/publications/i/item/WHO-HTM-TB-2015.19.
20. WHO 2023. Global tuberculosis report 2023. Geneva: World Health Organization (WHO); 2023. Licence: CC BY-NC-SA 3.0 IGO. Available at https://iris.who.int/bitstream/handle/10665/373828/9789240083851-eng.pdf?sequence=1.
21. WHO. 2024. World Health Organization (WHO): WHO Health Topics—Tubercolosis (TB). Available at https://www.afro.who.int/health-topics/tuberculosis-tb.
22. Zhao F, Cheng S, He G, Huang F, Zhang H, Xu B, Murimwa TC, Cheng J, Hu D, Wang L. Space-time clustering characteristics of tuberculosis in China, 2005–2011. PLoS One. 2013;8.

Publisher's Note