Title

Spatial clustering of drug-resistant tuberculosis in Hlabisa sub-district, KwaZulu-Natal, 2011-2015

Catherine M Smith

Department of Public Health Informatics, Institute of Health Informatics, UCL

Richard Lessells

Department of Clinical Research, London School of Hygiene and Tropical Medicine

Africa Health Research Institute, Somkhele, South Africa

Alison Grant

Department of Clinical Research, London School of Hygiene and Tropical Medicine

Africa Health Research Institute, School of Nursing and Public Health, University of KwaZulu-Natal, South Africa

School of Public Health, University of the Witwatersrand, Johannesburg, South Africa

Kobus Herbst

Africa Health Research Institute, Somkhele, South Africa

Frank Tanser

Africa Health Research Institute, Somkhele, South Africa

School of Nursing and Public Health, University of KwaZulu-Natal, Durban, South Africa

Centre for the AIDS Programme of Research in South Africa – CAPRISA, University of KwaZulu-Natal, Congella, South Africa.

Running head

Spatial clusters of tuberculosis in KwaZulu-Natal

Summary 199 words

Body: 2499

References: 34

Tables: 1

Figures: 3

Keywords

Geographical information systems, disease clustering, HIV infections

SUMMARY

Setting

Incidence rates of tuberculosis in South Africa are amongst the highest in the world, and drug-resistance is a major concern. Understanding geographic variation in disease may guide targeted interventions.

Objective

We aimed to characterise the spatial distribution of drug-resistant tuberculosis (DRTB) in a rural area of KwaZulu-Natal, South Africa, and test for clustering.

Design

This was a cross-sectional analysis of DRTB patients managed at a rural district hospital from 2011-2015. We mapped all patients in hospital data to local areas; and linked to a population-based demographic surveillance system to map patients to individual homesteads. We used kernel density estimation to visualise the distribution of disease and tested for clustering using spatial scan statistics.

Results

There were 489 patients with DRTB in the sub-district; 111 lived in the smaller demographic surveillance area. Spatial clustering analysis identified a high-risk cluster (relative risk of DRTB within cluster compared to outside: 3.0, $p<0.001$) in the south-east, a region characterised by high population density and high HIV prevalence.

Conclusion

We have demonstrated evidence of a geographic high risk cluster of DRTB. This suggests that targeting interventions to spatial areas of highest risk, where transmission may be ongoing, could be effective.

INTRODUCTION

Incidence rates of tuberculosis (TB) in South Africa are amongst the highest in the world.[1] In 2015 there were an estimated 454,000 new diagnoses, a rate of 834 per 100,000 population, and it is the leading natural cause of death in the country.[1, 2] Rates of TB are particularly high in the province of KwaZulu-Natal, largely driven by the high prevalence of human immunodeficiency virus (HIV) and complicated by TB drug resistance.[3, 4]

Understanding the spatial distribution of disease is important for effective control. Spatial analyses can be used to identify the worst affected areas, generate hypotheses about transmission, and guide interventions.[5] Tests of spatial clustering can be used to identify groups of patients that occur closer together in space than would be expected by chance. These analyses have been used to identify areas of likely TB transmission.[6-14] Visualisation of spatial data on maps also provides a powerful means of communicating information about the disease to policy makers and the public.

The Africa Health Research Institute (AHRI) in the Hlabisa sub-district of KwaZulu-Natal, South Africa, maintains a large health and demographic surveillance system. This includes individual residential locations mapped to an accuracy of less than two metres, and routine linkage to public sector records.[15, 16] The aim of this study was to characterise the spatial distribution of drug-resistant TB (DRTB) in the sub-district, test for spatial clustering, and discuss implications for prevention and care.

STUDY POPULATION AND METHODS

Study area

Hlabisa health sub-district is an area of approximately 1,450km^2 and 247,350 residents in uMkhanyakude district, northern KwaZulu-Natal, South Africa (Figure 1A). It is characterised by high prevalence of HIV with high rates of associated TB (577 recorded TB cases per 100,000 population in uMkhanyakude in 2015; 64.3% HIV positive).[3, 17] The AHRI demographic surveillance area is located within the Hlabisa sub-district (Figure 1B). This a region of 435km^2, with approximately 11,000 homesteads and 60,000 residents, in which AHRI has undertaken population-based demographic surveillance since 2000.[15]

Data sources and identification of DRTB patients

This was a cross-sectional analysis of patients diagnosed with DRTB between 2011 and 2015 who were resident in the Hlabisa health sub-district. Since 2011, all patients with DRTB aged 12 years and older were admitted to the TB ward at Hlabisa hospital for at least the

first month of treatment. Since July 2013, individuals in the sub-district with DRTB have been identified through Xpert MTB/RIF testing at one of 17 primary health care clinics. Prior to that, most patients were identified through culture-based methods, apart from a small number of patients diagnosed using Xpert MTB/RIF tests at a clinical trial site.[18]

We identified patients with DRTB using ICD-10 discharge codes in the Hlabisa hospital information system. We defined DRTB based on South African coding standards, incorporating codes for rifampicin mono-resistant, multidrug-resistant (MDR) and extensively drug-resistant (XDR) TB.[19] We calculated the proportion of patients in the hospital admissions data who had DRTB, and described characteristics of patients. Data from the hospital information system are routinely linked to the AHRI demographic surveillance data using the South African identification number or through a standard probabilistic matching algorithm.[15] We used linked data to identify individual homestead locations.

The research was approved by the Biomedical Research Ethics Committee of the University of KwaZulu-Natal (ref. BE290/16), the Ethics Committee of the London School of Hygiene and Tropical Medicine (ref. 11814), and the Health Research Committee of the KwaZulu-Natal Department of Health (ref. 378/16). These committees waived the requirement for individual informed consent to use the hospital admissions data, as the data were routinely collected from hospital records and there was no direct interaction with individual patients.

Spatial analysis

We conducted two spatial analyses which derived patient geographic locations using different methods: a local area analysis which covered the entire Hlabisa sub-district; and a micro-geographic analysis using the precise locations of patient homesteads within the smaller AHRI demographic surveillance area.

In the local area analysis, we compared the spatial distribution of DRTB patients with all other hospital admissions. Local areas are informal regions used by local populations to describe the sub-district, and have been mapped by AHRI (315 local areas in the Hlabisa sub-district in total, Figure 1B). We extracted patient-reported local areas of residence from a free text field in the hospital data, and matched them to mapped names of local areas.

In the micro-geographic analysis, we compared point residential locations for patients with DRTB to the spatial distribution of the general population. We used residential locations from the AHRI demographic surveillance data. For DRTB patients, we identified the exact homestead of residence which was recorded in the surveillance system closest in time to the patient's date of admission to Hlabisa hospital. The distribution of the general population was

derived by calculating total person years of residence in each homestead over the study period.

We tested for spatial clustering of DRTB in both the local area and micro-geographic analyses. We used spatial scan statistics, implemented in the SaTScan software,[20] to test the hypothesis that DRTB patients were closer together in space than the underlying population distribution. Scan statistics are used to compare the observed number of cases within spatial windows of various sizes with those that would be expected, in this case, under a random Poisson distribution. A likelihood ratio is calculated for each window which compares the observed and expected numbers of cases inside and outside the window. Monte Carlo simulations are then used to generate random distributions of cases under the Poisson distribution, which are compared to the observed data to calculate a p-value. We set the maximum cluster size to 3 km, because spatial dependencies have previously been reported for HIV within this distance in this study area.[21]

We also plotted the locations of clusters on a smoothed map of the relative proportion of DRTB patients compared to the underlying distribution in continuous geographical space. These maps were produced using kernel density estimation with a standard Gaussian kernel of 3km radius.

Analyses were performed using R v 3.2.3, using the packages *spatstat* and *rsatscan*.[22, 23]

RESULTS

Between 2011 and 2015, there were 19,408 individuals admitted to Hlabisa hospital who could be allocated to a local area in the Hlabisa sub-district. Of these, 489 (2.5%) had an ICD-10 hospital discharge code indicating DRTB, among whom the majority (478, 98%) had MDR disease.

Characteristics of patients with DRTB are shown in Table 1. Approximately half (250, 51%) the patients were female, and the modal age group was 25-34 years. There were 340 (70%) HIV positive DRTB patients, among whom 202 (60%) were on antiretroviral therapy at the time of admission. One in six (78, 16%) DRTB patients died prior to discharge, and five absconded from hospital.

Local area analysis of DRTB in the Hlabisa sub-district

We used the distribution of all 19,408 patients admitted to Hlabisa hospital across the local areas in the sub-district as a denominator for analyses of spatial clustering amongst the 489 DRTB patients.

There was one high relative risk cluster, located in the south-east of the sub-district (p<0.001). This cluster had a radius of 1.9 km, comprised four local areas with a total of 79 DRTB patients compared to 29 that would be expected by chance, and had a relative risk of 3.0. There was some evidence of a low relative risk cluster in the west of the sub-district, close to the hospital (p=0.08). This cluster had a radius of 2.1 km; comprised four local areas with six patients compared to 19 expected, and a relative risk of 0.3. Locations of these two clusters, overlaid on a smoothed map of the relative proportion of DRTB patients compared to all hospital admissions are displayed in Figure 2.

Micro-geographic analysis of DRTB in the AHRI demographic surveillance area

There were 144 DRTB patients whose hospital data could be linked to the AHRI population surveillance data. Of these, 111 had a recorded homestead location during the study period and were included in the analysis (94 patients had a residential location recorded in the same year as their admission date; 16 of the remaining patients had a residence recorded before their admission date, and one had a residence location recorded the year after their admission date). The remaining 33 patients were excluded from this analysis because they did not have a homestead of residence recorded during the period of this study (2011-2015).

The 111 patients with DRTB resided at 106 unique homestead locations; 10 patients shared homesteads with another patient. The most likely high relative risk cluster (p=0.057) had a radius of 2.8 km. The cluster comprised 55 patients compared to 31 that would be expected by chance, had a relative risk of 2.5, and all the homesteads with more than one patient were in this area. It was in a similar region to the high-risk cluster resulting from the local area-level analysis of the entire Hlabisa sub-district, in the south-east of the demographic surveillance area. This is the area around a township and is characterised by high population density and HIV prevalence compared to the rest of the demographic surveillance area.[21] There were no low relative risk clusters identified in this analysis.

DISCUSSION

In this study, we have described the spatial distribution of DRTB in the Hlabisa sub-district of KwaZulu-Natal for the first time. DRTB was highly prevalent in this region, with 489 (2.5%) of inpatients at Hlabisa hospital affected over a five-year period. Almost all had MDR disease and 16% died in hospital. There was clear evidence of a geographic high risk cluster of DRTB in the south-east of the region. This area is characterised by relatively high population density and high incidence and prevalence of HIV.[21] This spatial heterogeneity of DRTB in a high burden, predominantly rural area was consistent with findings from lower HIV

prevalence settings, although our analysis was at a more granular level than most previous studies.[6, 10, 11]

Establishing the spatial distribution of disease in rural areas such as the Hlabisa sub-district is challenging. This is because residential addresses are not recorded routinely in hospital systems, many people live in informal settlements which are not accurately mapped, and the population is highly mobile.[24] A strength of our study is that we used precise residential locations collected in the AHRI demographic surveillance system. We were therefore able to derive the geographic distribution of DRTB from two different data sources: the self-reported local area of residence from hospital data, and, for residents of the AHRI demographic surveillance area, the homestead of residence. High risk clusters of disease were indicated in the same approximate area using both methods, suggesting that the observed clustering is genuine.

The area of spatial clustering was characterised by high population density, and only 10 (18%) patients in the cluster shared residences with other patients. This implies that transmission of DRTB in this community may have occurred in public places as well as within households. Other studies have also indicated the importance of community-based transmission of TB in similar settings.[12, 25-30] Indoor venues with poor ventilation in which people come into close contact including healthcare facilities, public transport, churches and bars have been implicated as possible areas of transmission. An important component of DRTB prevention is therefore to identify such venues in the community and implement interventions including active case finding by regular screening; contact tracing; improving access to treatment, and airborne infection control measures in health facilities.

The distinct spatial clustering of the disease suggests that targeting these interventions to suspected high transmission areas could be effective. However, our findings only reveal where people with DRTB reside, and uncovering precisely where transmission is occurring will require more detailed clinical and molecular epidemiology. A prospective cohort of people with DRTB is now operational in the study area where information is collected about social contact patterns and use of shared public spaces. This will be integrated with whole genome sequence data to provide better understanding of transmission.

The results of our study also highlight the importance of the interaction between HIV and TB in this population. Almost three quarters of the patients with DRTB were HIV positive, compared to a prevalence in the population of approximately one quarter. The area of spatial clustering of DRTB was characterised by high HIV prevalence, and is in a similar region to a geographic cluster of HIV positive individuals identified previously.[21] In this study population, approximately 60% of HIV positive patients were on antiretroviral therapy. Previous studies

have suggested that improved coverage of antiretroviral therapy at both the individual and community levels can contribute to reducing the incidence of TB.[31-33]

Our study had several limitations. The analysis was restricted to DRTB patients, because we were only able to ascertain cases through hospital admissions data. Drug-sensitive TB patients are only admitted when clinically essential whereas policy at the time of the study was for all drug-resistant patients to be admitted for at least one month. We were therefore unable to determine whether the distribution of drug-resistant disease is similar to drug-sensitive disease. Future studies in this area will integrate additional data from the electronic TB register (ETR.net) to characterise the distribution of drug-sensitive TB.

The hospital information system used in this study also did not contain clinical information about history of treatment for TB. With the data available we therefore could not make any inference about the balance of primary and secondary drug resistance in the population. However, most recent data from high HIV prevalence settings suggest that, regardless of treatment history, the majority of DRTB cases arise from transmission.[34]

Another limitation was the use of hospital discharge codes to define DRTB patients, which is an underestimate of the true number. We will have missed patients that were not coded in the hospital data as DRTB; children younger than 12 years who may have been managed elsewhere; individuals who had DRTB detected but did not go to hospital, and those with undetected disease. However, we have no reason to suspect that these factors would operate in a geographically heterogeneous way that would lead to spurious spatial clusters.

The spatial analysis of DRTB in the wider Hlabisa sub-district was limited by use of local areas as opposed to individual addresses. However, results were similar to the analysis of precise point locations in the demographic surveillance area, which suggests that this method may be useful for similar analyses in future. We also used a hospital-based denominator as a proxy for the underlying population in this analysis. This will therefore be influenced by spatial factors which govern the distribution of conditions relating to other admissions. Finally, we described the characteristics of individuals using hospital data, but further information on patients would allow a more detailed analysis of risk factors in this population.

CONCLUSIONS

Our study shows concerning evidence of possible ongoing transmission of DRTB in this area of high prevalence. This suggests that targeting interventions to spatial areas of highest risk could be effective in supporting progress towards the WHO's End TB strategy for a 90% reduction in new cases by 2035.[35]

ACKNOWLEDGEMENTS

CMS received funding from the South African MRC Flagship program at the AHRI and University of KwaZulu-Natal, and from the Farr Institute of Health Informatics Research. We thank Petros Khambule for collection of the hospital data, and Kathy Baisley for technical support with data manipulation and analysis. CMS performed the analysis and drafted the manuscript. FT, RL and CMS designed the study. KH oversaw collection of hospital data. All authors contributed to and approved the final manuscript. We have no conflicts of interest to declare.

REFERENCES

1. World Health Organization. Global Tuberculosis Report 2016. Geneva: 2016.
2. Statistics South Africa. Mortality and causes of death in South Africa, 2011: Findings from death notification. 2014.
3. Massyn N, Peer N, English R, Padarath A, Barron P, Day C. District Health Barometer 2015/16. Durban: Health Systems Trust, 2016.
4. Nanoo A, Izu A, Ismail NA, Ihekweazu C, Abubakar I, Mametja D, et al. Nationwide and regional incidence of microbiologically confirmed pulmonary tuberculosis in South Africa, 2004-2012: a time series analysis. Lancet Infec Dis. 2016;15:1066-76.
5. Smith CM, Le Comber SC, Fry H, Bull M, Leach S, Hayward AC. Spatial methods for infecitous disease outbreak investigations: systematic literature review. Euro Surveill. 2015;20(39).
6. 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. J Infect Dis. 2016;213(2):287-94.
7. Saavedra-Campos M, Welfare W, Cleary P, Sails A, Burkitt A, Hungerford D, et al. Identifying areas and risk groups with localised Mycobacterium tuberculosis transmission in northern England from 2010 to 2012: spatiotemporal analysis incorporating highly discriminatory genotyping data. Thorax. 2015;71(8):742-8.
8. Izumi K, Ohkado A, Uchimura K, Murase Y, Tatsumi Y, Kayebeta A, et al. Detection of Tuberculosis Infection Hotspots Using Activity Spaces Based Spatial Approach in an Urban Tokyo, from 2003 to 2011. PLOS ONE. 2015;10(9):e0138831.
9. Althomsons SP, Kammerer JS, Shang N, Navin TR. Using Routinely Reported Tuberculosis Genotyping and Surveillance Data to Predict Tuberculosis Outbreaks. PLOS ONE. 2012;7(11):e48754.
10. Jenkins HE, Plesca V, Ciobanu A, Crudu V, Galusca I, Soltan V, et al. Assessing spatial heterogeneity of MDR-TB in a high burden country. Eur Respir J. 2013;42(5):10.1183/09031936.0111812.
11. Lin H, Shin S, Blaya JA, Zhang Z, Cegielski P, Contreras C, et al. Assessing spatiotemporal patterns of multidrug-resistant and drug-sensitive tuberculosis in a South American setting. Epidemiol Infect. 2011;139(11):1784-93.
12. Dissou A, Frank F, N'Dira S, Gladys A, Isdore Chola S, Leen R, et al. Possible Outbreak of Streptomycin-Resistant Mycobacterium tuberculosis Beijing in Benin. Emerg Infect Dis. 2009;15(7):1123.
13. Smith CM, Maguire H, Anderson C, Macdonald N, Hayward AC. Multiple large clusters of tuberculosis in London: a cross-sectional analysis of molecular and spatial data. ERJ Open Res. 2017;3(1).
14. Smith CM, Trienekens SC, Anderson C, Lalor MK, Brown T, Story A, et al. Twenty years and counting: epidemiology of an outbreak of isoniazid-resistant tuberculosis in England and Wales, 1995 to 2014. Euro Surveill. 2017;22(8).
15. Tanser F, Hosegood V, Bärnighausen T, Herbst K, Nyirenda M, Muhwava W, et al. Cohort Profile: Africa Centre Demographic Information System (ACDIS) and population-based HIV survey. Int J Epidemiol. 2008;37(5):956-62.
16. Herbst K, Law M, Geldsetzer P, Tanser F, Harling G, Barnighausen T. Innovations in health and demographic surveillance systems to establish the causal impacts of HIV policies. Curr Opin HIV AIDS. 2015;10(6):483-94.
17. Zaidi J, Grapsa E, Tanser F, Newell ML, Barnighausen T. Dramatic increase in HIV prevalence after scale-up of antiretroviral treatment. AIDS. 2013;27(14):2301-5.
18. Lessells RJ, Cooke GS, McGrath N, Nicol MP, Newell ML, Godfrey-Faussett P. Impact of a novel molecular TB diagnostic system in patients at high risk of TB mortality in rural South Africa (Uchwepheshe): study protocol for a cluster randomised trial. Trials. 2013;14:170.

19. National Task Team for the Implementation of ICD-10. South African ICD-10 Coding Standards, Version 1.06. 2007.
20. Kulldorff M, Information Management Services Inc. SaTScan™ v9.4.2: Software for the spatial and space-time scan statistics. 2015.
21. Tanser F, Bärnighausen T, Cooke GS, Newell M-L. Localized spatial clustering of HIV infections in a widely disseminated rural South African epidemic. Int J Epidemiol. 2009;38(4):1008-16.
22. Baddeley A, Turner R. spatstat: An R Package for Analyzing Spatial Point Patterns. J Stat Softw. 2005;12(6):1-42.
23. Kleinman K. rsatscan: Tools, Classes, and Methods for Interfacing with SaTScan Stand-Alone Software. 0.3.9200 ed2015.
24. Muhwava W, Hosegood V, Nyirenda M, Herbst K, Newell ML. Levels and determinants of migration in rural KwaZulu-Natal, South Africa. Afr Popul Stud. 2010;24(3):259-80.
25. Yates TA, Tanser F, Abubakar I. Plan Beta for tuberculosis: it's time to think seriously about poorly ventilated congregate settings. Int J Tuberc Lung D. 2016;20(1):5-10.
26. Andrews JR, Morrow C, Walensky RP, Wood R. Integrating social contact and environmental data in evaluating tuberculosis transmission in a South African township. J Infect Dis. 2014;210:597-603.
27. Chamie G, Wandera B, Marquez C, Kato-Maeda M, Kamya MR, Havlir DV, et al. Identifying locations of recent TB transmission in rural Uganda: a multidisciplinary approach. Trop Med Int Health. 2015;20:537-45.
28. Wood R, Racow K, Bekker L-G, Morrow C, Middelkoop K, Mark D, et al. Indoor social networks in a South African township: potential contribution of location to tuberculosis transmission. PLOS ONE. 2012;7:e39246.
29. Andrews JR, Morrow C, Wood R. Modeling the role of public transportation in sustaining tuberculosis transmission in South Africa. Am J Epidemiol. 2013;177:556-61.
30. Shah NS, Auld SC, Brust JCM, Mathema B, Ismail N, Moodley P, et al. Transmission of Extensively Drug-Resistant Tuberculosis in South Africa. New Engl J Med. 2017;376(3):243-53.
31. Hermans S, Boulle A, Caldwell J, Pienaar D, Wood R. Temporal trends in TB notification rates during ART scale-up in Cape Town: an ecological analysis. J Int AIDS Soc. 2015;18:20240.
32. Lawn SD, Badri M, Wood R. Tuberculosis among HIV-infected patients receiving HAART: long term incidence and risk factors in a South African cohort. AIDS. 2005;19:2109-16.
33. Middelkoop K, Bekker L-G, Myer L, Whitelaw A, Grant A, Kaplan G, et al. Antiretroviral program associated with reduction in untreated prevalent tuberculosis in a South African township. Am J Respir Crit Care Med. 2010;182:1080-5.
34. Dheda K, Gumbo T, Maartens G, Dooley KE, McNerney R, Murray M, et al. The epidemiology, pathogenesis, transmission, diagnosis, and management of multidrug-resistant, extensively drug-resistant, and incurable tuberculosis. The Lancet Respiratory medicine. 2017.
35. World Health Organization. The End TB Strategy 2015 [cited 2016 30 August]. Available from: http://www.who.int/tb/strategy/en/.

TABLES

Table 1: Characteristics of drug-resistant tuberculosis patients, Hlabisa hospital (2011-2015)

Characteristic	N	%
Sex		
Male	239	48.9
Female	250	51.1
Age group (years)		
5-14	10	2.0
15-24	59	12.1
25-34	166	33.9
35-44	143	29.2
45-54	66	13.5
55-64	27	5.5
65-74	10	2.0
75+	8	1.6
Year of admission		
2011	77	15.7
2012	98	20.0
2013	103	21.1
2014	115	23.5
2015	96	19.6
Drug-resistant tuberculosis type		
MDR	478	97.8
Rifampicin monoresistant	5	1.0
XDR	6	1.2
Site of disease		
Pulmonary	421	86.1
Extrapulmonary	5	1.0
Missing	63	12.9
HIV/ ART status*		
HIV positive, on ART	202	41.3
HIV positive, not on ART	133	27.2
HIV positive, ART missing	5	1.0
HIV negative	67	13.7
Missing	82	16.8

Discharge status		
Discharged	394	80.6
Transferred	12	2.5
Died	78	16.0
Absconded	5	1.0

MDR, multidrug-resistant; XDR, extensively drug-resistant; HIV, human immunodeficiency virus; ART, antiretroviral therapy.

*at time of hospital admission

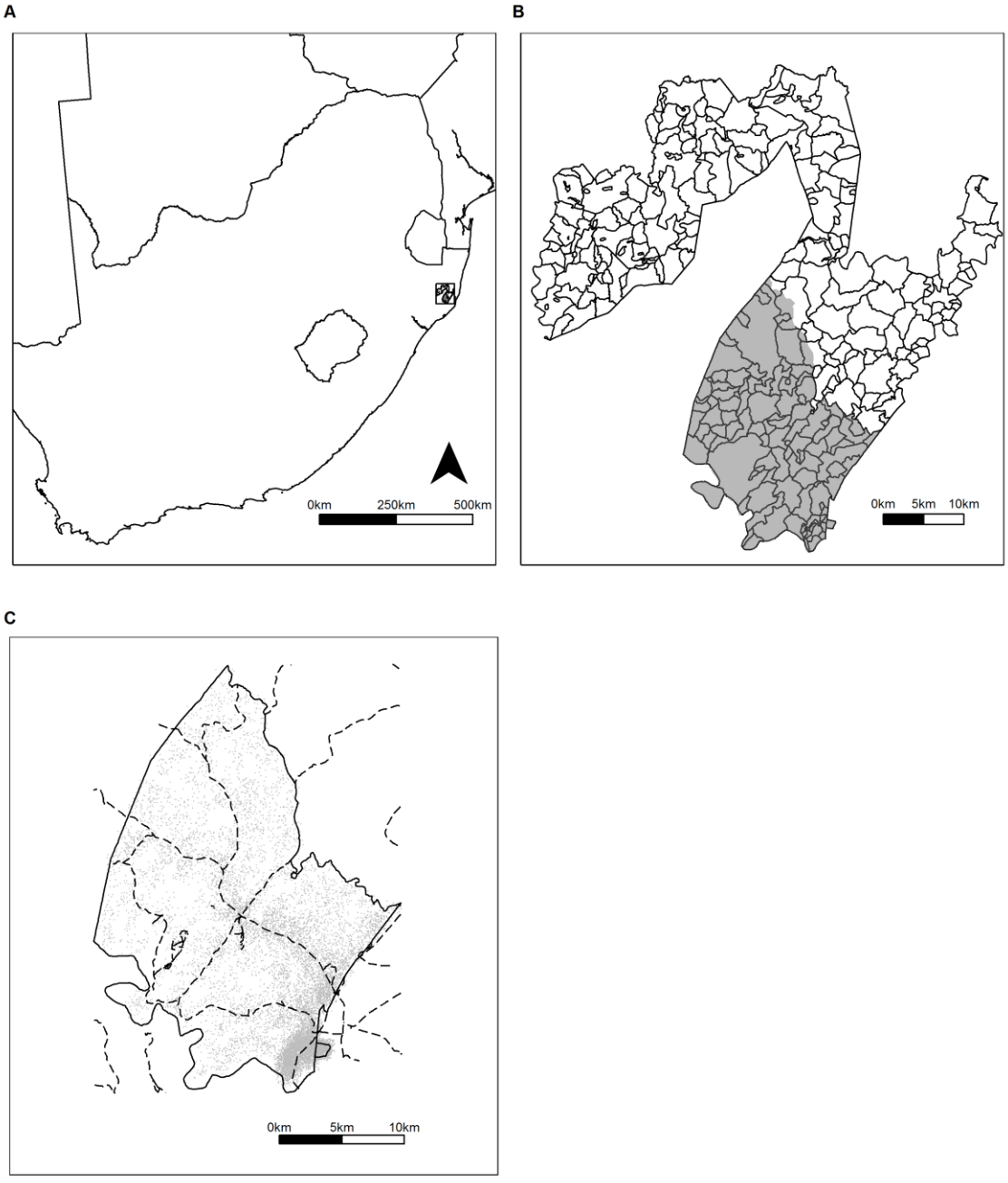
FIGURES

Figure 1: Study site

A: Location of Hlabisa sub-district within South Africa

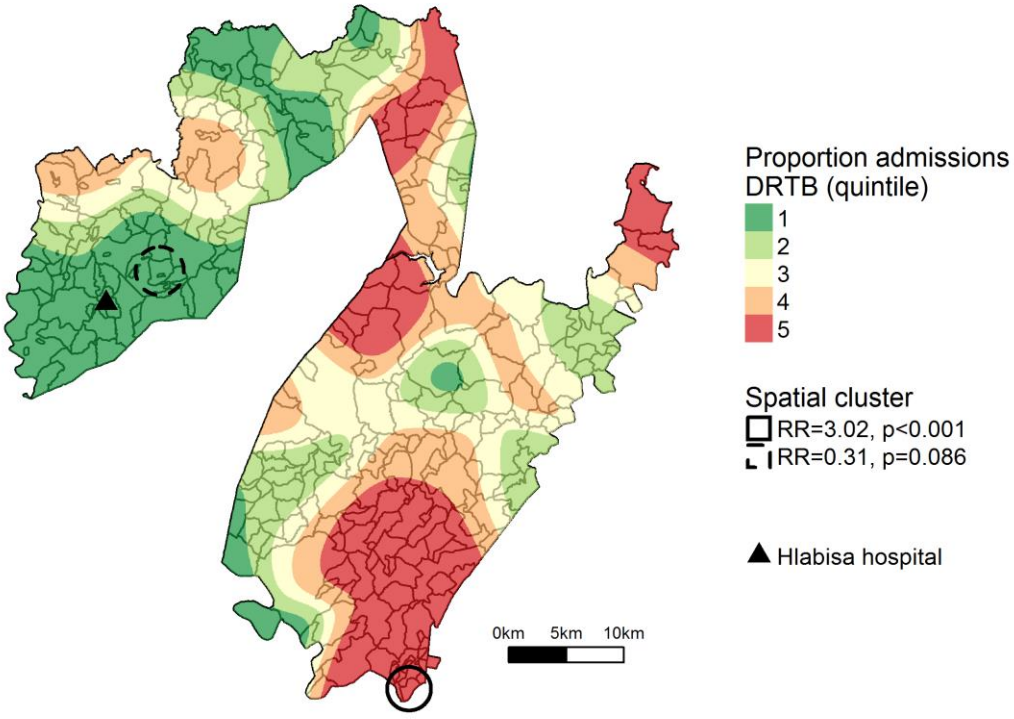
B: Hlabisa sub-district, showing local areas and Africa Health Research Institute demographic surveillance area (shaded)

C: Africa Health Research Institute demographic surveillance area, with roads and approximate locations of homesteads (incorporating intentional random error)

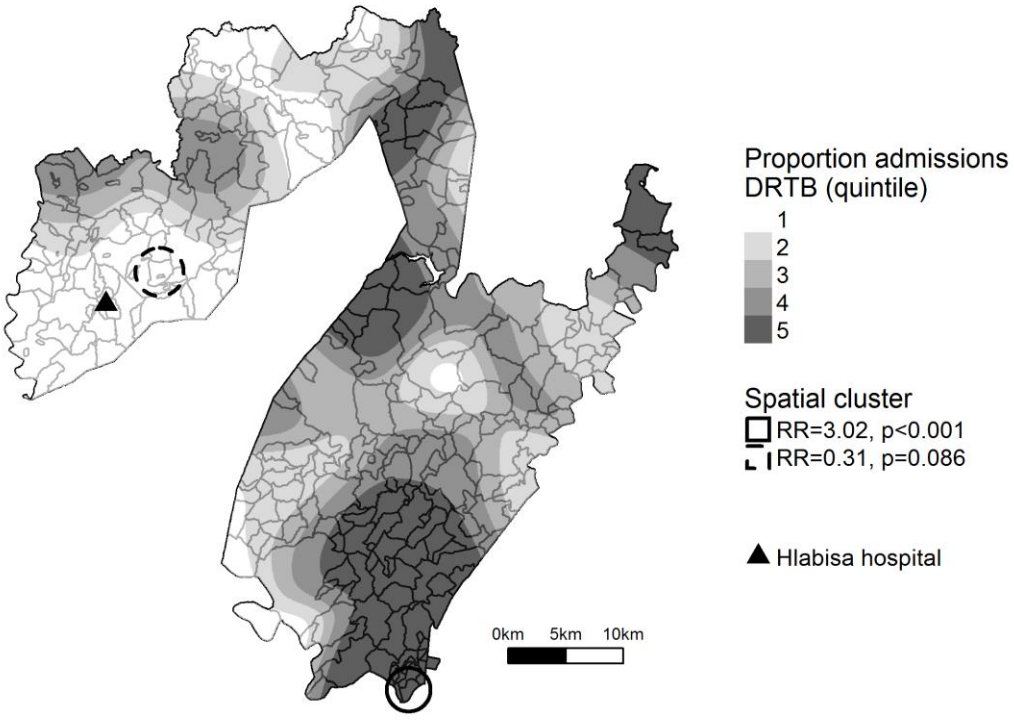


392 *Figure 2: Spatial clustering of drug-resistant tuberculosis in Hlabisa sub-district, 2011-2015.*

393 *Locations determined using patient-reported local areas in hospital information system.*



394



395

396

Figure 3: Spatial clustering of drug-resistant tuberculosis in Africa Health Research Institute demographic surveillance area, 2011-2015. Locations determined by linking hospital data to individual homesteads in demographic surveillance system.