UNIVERSITY OF CAPE TOWN
IYUNIVESITHI YASEKAPA • UNIVERSITEIT VAN KAAPSTAD

The spatial analysis of the associations between built environment and lung health outcomes in children from Drakenstein, Western Cape: a nested case-control study.

Demi Almaz Meyer (MRYDEM001)

BSc (UWC) BSc (hons) (UWC)

Dissertation submitted in partial fulfilment of the requirement for the degree.

Master of Public Health (Epidemiology and Biostatistics)

School of Public Health

Faculty of Health Sciences, University of Cape Town

Supervisor: Dr Maia Lesosky

February 2023

PREAMBLE

Declaration

I, Demi Meyer (MYRDEM001), hereby declare that the work on which this dissertation/thesis is based is my original work (except where acknowledgements indicate otherwise) and that neither the whole work nor any part of it has been, is being, or is to be submitted for another degree in this or any other university.

I empower the university to reproduce for the purpose of research either the whole or any portion of the contents in any manner whatsoever.

Signature: Signed by candidate

Date: 13 February 2023

Acknowledgements

I am thankful to God Almighty for good mental and physical health until this time.

I would like to acknowledge my supervisor, Professor Maia Lesosky. Thank you for the support and feedback provided throughout the MPH process. I would like to acknowledge the contribution of Professor Heather Zar, the co-supervisor of this project.

To my sponsors, National Research Foundation (NRF), who funded my MPH degree; I am eternally grateful.

I would also like to thank the Drakenstein Child Health Study teams at both Mbekweni Clinic and TC Newman Clinic for the opportunity to work with the project data.

To Miss Phindi Zwane, Mr Eke Arua and all my fellow colleagues from the School of Public Health, University of Cape Town (UCT), I am eternally grateful for your support, assistance with R code and for providing a safe space for me to brainstorm.

Lastly, I would like to dedicate this thesis to my family, your love and encouragement was more than I could ever ask for. I would like to thank my parents, Mr Anthony Meyer, and Mrs Glynnis Abrahams-Meyer for continually supporting my career choices. I am forever grateful. To my grandparents (Arthur and Winnie Abrahams) and aunty (Anthea Abrahams) for always keeping me in your prayers. To my one and only sister, Miss Yentl Meyer, sissie, thank you for keeping me sane, always knowing how to cheer me up. My best friends, Chelsea Phillips, and Caron Johannes for always checking up on me and making sure I'm okay. Last, but definitely not least; to my partner Qiran Ewers, thank you for always being there for me, wanting to help no matter the situation. I appreciate all of you.

Dissertation Abstract

In South Africa, childhood lung diseases continue to be a serious public health issue. Lower respiratory tract infections (LRTI) and tuberculosis (TB) are among the childhood lung disorders that have a diverse geographic distribution. However, information on the condition of children's lungs and related risk factors is typically only provided at the province, district, or subdistrict levels and is lacking at a local level in South Africa. Child lung health is associated with a complex combination of social and environmental risk factors. The built environment impacts public health by influencing human exposure to airborne pollutants. Overcrowding, which is often a result of poor infrastructure, is directly linked to the built environment and contributes to the spread of childhood lung diseases. The geographical analysis of the epidemiology and risk factors for child lung diseases can guide the targeting of health programs that work to address this issue. This study aims to assess if there is any relationship between the social and environmental risk factors (built environment) and spatial distribution of child lung health outcomes in Drakenstein.

To achieve this aim, this study employs a secondary analysis of data from the Drakenstein Child Health Study (DCHS). The parent study is an ongoing birth cohort of approximately 1000 mother and child pairs however, only 844 mother and child pairs consented to have their household addresses collected and geocoded. Data regarding demographics, household and environmental exposures as well as information on child lung health is periodically collected. The longitudinal nature of this data offers a unique opportunity to define children's lung health through in space, allowing for insightful inferences about the factors influencing children's health in South Africa.

The 844 consented cohort analysed in this study was further spatially subsampled into distance matched case control groups. We examined the relationships between individual-, household-, and community-level risk factors and child lung diseases (namely LRTIs and TB). To do this, we created and mapped the spatial data using the geo-location of study participants' households, known community-level risk factors, and key built environments. The built environment was categorized into distance between cases and controls and built environment in kilometres and number of built environment types within a 500-meter radius. We then used multivariable logistic regression to determine the association between lung health (namely LRTI and LTBI) cases and controls and the built environment.

The cohort included 408 LRTI cases (237 male children: 58%) and 408 LRTI controls (185 male children: 44%) (1:1 ratio) and 75 LTBI cases (49 male children: 65%) and 375 LTBI controls (185 male children: 49%) (1:5 ratio), subsampled and matched by distance. The LRTI population was equally spread between Mbekweni (50%) and TC Newman (50%) neighbourhoods whereas the LTBI population was higher in the TC Newman neighbourhoods at 65% compared to 35% in Mbekweni neighbourhoods. In the LRTI and LTBI models, the number of significant differences between cases and controls vary. The main built environment types that were associated with LRTI were spaza shops [OR = 1.05, 95%CI (1.02; 1.08), p = 0.01], formal liquor stores [OR = 2. 63, 95%CI (1.20; 5.90); p = 0.02] and agriculture (farming places) [aOR = 0.34, 95%CI, 0.11; 0.82 p = 0.03]. The main built environment types that were associated with LTBI were formal [OR = 0.26, 95%CI (0.09; 0.68), p = 0.01] and informal liquor stores [OR = 0.27, 95%CI (0.08; 0.85), p = 0.01] and religious (places of worship) [OR = 11.81, 95%CI (1.51; 101.88), p = 0.02]. Thus, creation of more green spaces or designing of buildings with better air flow would be useful and actionable by the government local authorities as potential strategies to reduce the burden of childhood LRTI and LTBI in LMICs.

To conclude, in Drakenstein, the associations between the built environment and the spatial distribution of child lung health outcomes are present after distance-matching cases and controls and adjusting for covariates. The observed associations are indicative of the SES and community circumstances. However, the relationship warrants further investigation in similar settings where LTBI and LRTI in children remains highly prevalent. Thus, creation of more green spaces or designing of buildings with better air flow would be useful and actionable by the government local authorities as potential strategies to reduce the burden of childhood LRTI and LTBI in LMICs.

List of Abbreviations

AIC	Akaike Information Criteria
BE	Built Environment
BES	Built Environment and Socio-demographic risk factors
CI	Confidence Interval
COPD	Chronic Obstructive Pulmonary Diseases
CRF	Case Report Form
DCHS	Drakenstein Child Health Study
EDHS	Ethiopian Demographic and Health Survey
GIS	Geographic Information Systems
GPS	Geographic Positioning System
HIB	Haemophilus Influenzae Type B
HIV	Human Immunodeficiency Virus
HREC	Health Sciences Research Ethics Committee
IJERPH	International Journal of Environmental Research and Public Health
Km	Kilometres
LMIC	Low and Middle Income Country
LTBI	Latent Tuberculosis Infection
LRTI	Lower Respiratory Tract Infections
M	Metres
OR	Odds Ratio

PID	Personal Identifier
QGIS	Quantum Geographic Information System
TB	Tuberculosis
TST	Tuberculin Skin Tests
SES	Socio-Economic Status
SA	South Africa
SDG	Sustainable Development Goals
USA	United States of America
WHO	World Health Organization
ZAR	South African Rand

Thesis Organization

This thesis is divided into three sections: Part A to Part C.

Part A details the research protocol which includes a background to the study, the study rationale, aim and objectives, methods, and ethical considerations.

The findings of the study are presented in a journal-ready manuscript format in Part B. This manuscript has been formatted according to the author submission guidelines of the targeted journal, *International Journal of Environmental Research and Public Health* journal (IJERPH) (*see Appendix D)*. In Part B, the findings, strengths, and limitations of the study are discussed in detail in the context of relevant published literature. In addition, conclusions and recommendations are presented here.

Lastly, all supplementary materials relevant to the study are provided as an appendix section in Part C. These supplementary materials include supporting data and analyses, the ethics clearance certificates to conduct this study and IJERPH submission guidelines. The Vancouver referencing style has been used throughout this thesis in accordance with the instructions for authors submission guidelines.

Table of Contents

List of Tables

Part A: Research Protocol

Part B: Manuscript

Part C: Appendices

List of Figures

Part B: Manuscript

Part C: Appendices

A. Supplementary tables and figures included in the manuscript

Part A: Research Protocol

1. Introduction

1.1 Rationale

The amount of data on spatial clustering of lung diseases at a sub-district level is scarce in South Africa as this data is collected at provincial level and district level. Routine data collected annually in the Western Cape since 2011 has shown a decrease in pneumonia fatalities however, the rate of hospital admittance is still increasing (1). Even though there is a readily accessible pneumonia conjugate vaccine against *Streptococcus pneumoniae* and *Haemophilus influenzae type b* (Hib), this has not reduced the mortality rate since 18% of child mortality worldwide is caused by pneumonia, among those, 42% of deaths are located within the African continent (2). Due to the geospatial data collected in this study, lung disease hotspots in children around Drakenstein area was identified. Furthermore, the availability of spatiotemporal clusters in Drakenstein could assist in determining if the built environment and other risk factors play a significant role in childhood lung health(1). Analysis of geospatial data at local level is useful in informing policy makers where to target health interventions and resource allocation. There is an existing birth cohort study being conducted within the community of Drakenstein, Western Cape province, South Africa in which demographics, household, environmental exposures, child lung health outcomes are being collected from approximately 1000 mother-child pairs living in the Drakenstein region. Nesting this study within the birth cohort study named the Drakenstein Child Health Study (DCHS) offers a unique opportunity to characterise child lung health in both space and time, and thus novel inferences can be made about the determinants of child health in South Africa.

1.2 Background

The leading causes of child mortality both in South Africa and globally are childhood respiratory conditions, such as lower respiratory tract infections (LRTIs) including pneumonia and tuberculosis (TB)(3). In addition, a primary reason for visits to hospital emergency room, hospitalizations, and death in children are due to acute LRTIs(4). Africa bears a disproportionate portion of the burden, with 131 million cases of lower respiratory infections occurring each year (5). Furthermore, South Africa has the largest burden of TB globally and data suggests that children contract TB from older persons as child-to-child transmission is rare(6). Chronic Obstructive Pulmonary Diseases (COPD) which is characterized by coughing, among other diseases are usually caused by inflammation due to smoking, exposure to fumes and burning(7). LRTIs which are also influenced by several aspects such as climatic conditions and environmental exposures in addition to living in crowded areas contribute largely to the

respiratory diseases in South Africa (6). However, it is evident in the DCHS, just like in Low and Middle Income Country (LMIC) settings they battle with the burden of the effects of childhood pneumonia, despite having strong health programmes (9).

Over the last ten years, the built environment has gained interest due to its impact on health (10). The region's characteristics may impact how infectious diseases are related to the built environment. Characteristics such as insufficient healthcare resources, crowded housing, and poor sanitation may affect how the built environment and infectious respiratory diseases interact (11). In previous studies, associations were observed between LRTI clusters and environmental risk factors such as socio-economic status, housing environment and air pollution(12). The built environment consists of healthcare facilities, educational facilities, bus and taxi ranks, food outlets and taverns or shebeens (informal establishments for alcohol consumption). There is research that suggests that households that are in close proximity to busy traffic have an increased risk of contracting respiratory diseases such as wheezing, dry cough and asthma(13). Gases such as nitrous dioxide (NO_2) which pose hazardous health risks are released from motor vehicles (13)(14). However, the evidence gathered is insufficient to determine a connection between transportation related pollution and asthma. Further studies suggest that an increase in traffic pollution density is linked to reduced lung function, but Gehring et al., only discovered a link among high levels of traffic density and dry cough in children under the age of one(13).

Geographic Information Systems (GIS) is an emerging powerful tool utilised to estimate spatial patterns of diseases. Spatial analysis aids in the detection of geographic regions increased disease burden or a cluster of diseased persons in close proximity with a low likelihood that it is due to chance(15). The earliest used of GIS was in 1854 by John Snow, who mapped cholera cases in Soho, London, and subsequently linked the outbreak to a contaminated water pump (16). GIS has come to be the essential resource for surveillance for the planning and implementation of public health interventions(15,17). Thus, the spatial epidemiology of child lung health outcomes is increasing in popularity due to its efficiency in identifying areas which can be targeted for interventions.

The Western Cape province has lower levels of child mortality compared to other provinces. However, there are highly concentrated areas within the Western Cape province, including the Drakenstein region, that have the highest levels of child mortality in South Africa(1). The reasons for these

inequalities are inadequately understood mainly due to lack of data. In South Africa, several studies have shown spatial clustering of child mortality (18). However, these studies were limited to child mortality in provinces, sub-districts or rural communities with high Human Immunodeficiency Virus (HIV) prevalence. Of these studies, only one study accounted for spatial dependence of the risk factors in the analysis(19). There are no studies on spatiotemporal epidemiology of child lung health in a South African peri-urban area.

Evidence has shown that there are associations between spatial and temporal variations of childhood lung health outcomes and complex relationships among demographic, environmental, behavioural, socio-economic, genetic, and infectious risk factors(20). Furthermore, these risk factors have also been shown to be spatially dependent (21). For examples air pollution, and poor housing conditions were found to associated with spatial clusters of LRTIs in the United States of America (USA)(12). Another study from north-western Ethiopia found significant spatial clustering of childhood TB within communities. There was also an association found between spatio-temporal transmission of childhood TB and district level socio-climatic factors such as urbanisation(22). Another Ethiopian study making use of the Ethiopian Demographic and Health Survey (EDHS) data stated that acute respiratory infection in children under the age of five varied across the country and found significant spatial clustering of childhood acute respiratory infection (23). A study in the US by Beck et.al. showed spatial clustering of bronchiolitis and pneumonia in children and an association between LRTI hospitalization and SES, as hotspots were in the central town and coldspots were in remote suburbs (24). Similarly, in Vietnam, spatial clustering of childhood acute LRTIs was observed. This clustering was influenced by climatic factors such as temperature and rainfall patterns(25). It is clear that despite being a more thorough method, spatiotemporal analysis of child lung health outcomes is gaining popularity due to its capacity to pinpoint specific locations that can benefit from interventions.

In South Africa, a study used a descriptive spatial analysis to show an association between latent tuberculosis infection (LTBI) in children and exposure to adult TB cases in a residential plot. According to the study, a residential plot is a geographical space that contains multiple households where communal water and sanitation services are shared(6). However, a limitation of this study is that it did not account for spatial dependency. Accounting for spatial dependence is becoming important when analysing risk factors that influence health outcomes since these factors often vary in space and time(25).

Geospatial analysis of child lung health outcomes will help identify areas with unusually high (hotspots) or low (coldspots) childhood lung diseases in the Drakenstein area. The presence of spatiotemporal clusters of childhood lung diseases in Drakenstein could provide evidence about local demographic, environmental, behavioural, socio-economic, genetic, and infectious risk factors. Moreover, understanding the geospatial patterns of child lung diseases and the impacts of the built environment in the Drakenstein community can help to target interventions and resources allocation.

2. Purpose of the study

2.1 Aim

To assess if there is any relationship between the social and environmental risk factors (built environment) and spatial distribution of child lung health outcomes in Drakenstein.

2.2 Research question

In children residing in Drakenstein, what are the risk factors associated with spatial distribution of respiratory disease outcomes?

2.3 Hypothesis

We hypothesize that the observed spatial distribution of child lung health outcomes is not associated with the social and environmental risk factors (built environment).

2.4 Objectives

1. To describe spatial patterns of childhood lung health outcomes in terms of geographical areas in Drakenstein.

2. To describe built environment in terms of geographical areas in Drakenstein.

3. To determine the relationship between spatial distribution of childhood lung health outcomes and social and environmental risk factors (built environment) using spatial analysis.

3. Methodology

3.1 Study design

This is a nested case-control study of children included in the Drakenstein Child Health Study. This study is a secondary geospatial data analysis of the health outcomes from the Drakenstein Child Health Study (DCHS) cohort. The details of DCHS cohort and study procedures have been published(26). Comprehensive clinical, socio-economic, behavioural, demographic and geospatial data have already been collected in this cohort of approximately 1000 mother-child pairs, with measures of child lung health (including the outcomes to be examined here). This study seeks to add geospatial data to the parent study and to conduct a comprehensive analysis of the geospatial and temporal determinants of child lung health outcomes.

3.2 Study Setting

Drakenstein is one of the five local municipalities located within the Cape Winelands District of Western Cape Province, South Africa. The study area is 1 538 sq.km. with an estimated population of 200 000 people. Unemployment in the area is high, with many people of low socio-economic status. The DCHS cohort was recruited from two primary healthcare facilities (sites): Mbekweni Clinic, which mainly provides services to the Black community, and TC Newman Clinic, which mainly provides services to those of coloured (a South African race classification) community. Levels of unemployment and poverty are high in these settings and most people in the area live in informal or crowded housing conditions. The burden of infectious diseases such as TB and HIV is high(27). This setting is ideal for the proposed study because it is typical of many peri-urban disadvantaged communities in South Africa and has a stable population with low immigration and emigration.

3.3 Characteristics of study population

The participants for this study will be approximately 1000 mother-child pairs in DCHS study who will consent to have their household addresses collected and geocoded. Additional data will be collected on mothers, families and other caregivers of these children. The parent study's initial inclusion criteria were women aged 18 or older, at 20-28 gestational weeks, having attended the clinic, and staying in the area. However, only those who agreed to have their household addresses collected and geocoded

are included in this analysis. Participant enrolment in the cohort was entirely voluntary. All of the births took place at Paarl Hospital. The inclusion of participants in the Geospatial study were conducted by a qualified research team. The participant (the child) was included if they resided in the Drakenstein community and were excluded if the child lived outside of the Drakenstein community or if the trained study staff could not locate them or contact their parents.

3.4 Recruitment and enrolment

Trained study staff located the study participants (approximately 1000 mother-child pairs) enrolled in the Drakenstein Child Health Study (DCHS). The participants contact details are stored in the REDCap (Research Electronic Data Capture) database of the study and they were contacted to arrange the visit time and verify the correct location. Participants were visited according to their next birthday at the year 6 timepoint, therefore the participants geolocation had to be collected by their 7th birthday.

3.5 Data collection and confidentiality

Two trained study staff visited all the households of the DCHS participants and using a hand-held geographic positioning system (GPS) device (Garmin 010-00970-00 eTrex 10 Worldwide Handheld GPS Navigator), recorded latitude and longitude coordinates of each household. The documentation and download of data were conducted on REDCap in the appropriate case report form (CRF). Data was also stored on Google MyMaps and on the DCHS Google Drive. The GPS device eTrex 10, Garmin is calibrated at the first recording of the day. Each household visit's duration was roughly 5 minutes long and was conducted twice a week. Participants information was stored in unique patient identifier (PID) to maintain patient confidentiality. Furthermore, the geospatial data collected from each household was anonymised to ensure the addresses are confidential. Research assistants and trained study staff also collected latitude and longitude coordinates of built environment such as health facilities, schools, creches or day care centres, bus and taxi ranks, food outlets, taverns or shebeens. The built environment data was collected on REDCap in the appropriate CRF. All other geodatabases including shapefiles and maps of Drakenstein study areas were obtained from Drakenstein Municipality Geographic Information Systems (GIS) Unit in Paarl.

3.6 Quality control

Once geocodes were downloaded and documented on the CRFs, Google MyMaps and on the DCHS Google Drive, the addresses are verified by a 3rd trained study staff. Geocoordinates were recollected if there were any discrepancies. This quality control was conducted monthly. Access to the study data on REDCap, MyMaps and the Google drive is only granted to study researchers and trained study staff working on specific CRF's. Every night, a thorough REDCap database export was performed; and is encrypted and kept in accordance with the database retention schedule(28).

3.7 Data on Childhood Lung health

Data on childhood LTBI in DCHS database Tuberculin skin tests (TST) are collected at study visits, annually, and at the time of a LRTI episode. LTBI is defined as an induration reaction greater than or equal to 10 mm. Data on childhood TB is in the DCHS database. Childhood TB is diagnosed by experienced physicians and nurses' chest radiographs, TB microbiological culture and Xpert MTB/RIF(29). Data on LRTIs will be obtained from active surveillance system for LRTIs in the DCHS. LRTIs are measured using World Health Organization (WHO) case definition criteria. LRTIs are categorised into LRTIs and severe LRTIs, with additional descriptors based on hospitalization requirements(30).

3.8 Data Analysis

The geocoordinates will be imported into QGIS software after they are converted into shapefiles. The shapefiles will be projected into the WGS84 (EPSG:4326) using QGIS. These coordinates will be layered in the Drakenstein basemap and colour coded according to the different childhood lung diseases. Based on the geospatial coordinates on the map, the clusters of the lung diseases will be visually interpreted to describe spatial patterns of childhood lung health outcomes in terms of geographical areas in Drakenstein.

The coordinates of the built environment will also be imported into QGIS after they will be converted into shapefiles. These coordinates will be layered in the Drakenstein basemap and colour coded according to the built environment classification; agriculture (farming place), commercial or business site, educational facility, healthcare facility or provider, religious (places of worship) and safety and security. Furthermore, the built environment will be tabulated by their classification.

To visualize the spatial distribution of LTBI and LRTIs relative to the built environment, maps of hotspots and coldspots will be generated using QGIS. Furthermore, the association between childhood lung health outcomes and the built environment, along with other social risk factors, will be estimated in R using multivariable logistic regression. For each independent variable, we present univariable models and multivariable models (direct effects, after adjustment for the effects of potential confounding variables).

3.9 Ethical clearance

Ethical clearance was given by faculty of health science University of Cape Town and Stellenbosch university ethics committee and written informed consent was obtained from parents and care givers in the primary DCHS study. Informed consent was obtained after recruitment when mothers were still pregnant, and compensation for time is offered after home visits. The informed consent is renewed annually, and participants are free to discontinue from the study any time.

3.10 Risks and Benefits

In the DCHS the parent received compensation for their time and the children received health benefits, free healthcare and regular check-ups. The communities were under an ongoing surveillance for emerging diseases in children and prevention interventions are implemented. Parents receive education on risk factors for lung diseases and ways to prevent them. With regards to the protection of the study participants, there are no direct risks or benefits with the data collection however, there are indirect risks. These risks are the potential loss of confidentiality of their private data including home addresses. The data used in this analysis does not contain any personal identifiers such as name, surname or identity numbers. In addition, the potential loss of confidentiality of home addresses will be minimised using donuts geomasking.

Donut geomasking involves moving a given address randomly from its original location. There is a minimum and maximum distance the location of these address can be randomly moved by so that patient confidentiality is maintained while minimizing the loss of spatial information(31). The distance a location is moved from its origin is inversely proportional to the underlying population density

(32,33). The minimum and maximum distances used in this study were 5 and 20 meters respectively. To ensure security of geospatial data, all the data was stored in a password protected computer.

3.11 Study timeline

Study Timeline Table 1: The anticipated timeline for the study

	2021		2022			2023
Activities	Mar - Jun	July - Dec	Jan -Apr	Jun - Sep	Oct - Dec	Jan - Mar
Protocol Development	X	X				
Ethics Approval		X	X			
Data Cleaning and Exploration			X	X		
Data Analysis			X	X	X	
Manuscript Write-Up				X	X	X
Mini-Dissertation Submission						X

4. References

1. Bamford L, Barron P, Kauchali S, Dlamini N, Paeds M. Inpatient case fatality rates improvements in children under 5: Diarrhoeal disease, pneumonia and severe acute malnutrition. 2018;108(March):33–7.

2. Rudan I, O'Brien KL, Nair H, Liu L, Theodoratou E, Qazi S, et al. Epidemiology and etiology of childhood pneumonia in 2010: Estimates of incidence, severe morbidity, mortality, underlying risk factors and causative pathogens for 192 countries. J Glob Health. 2013;3(1).

3. Troeger C, Forouzanfar M, Rao PC, Khalil I, Brown A, Swartz S, et al. Estimates of the global, regional, and national morbidity, mortality, and aetiologies of lower respiratory tract infections in 195 countries: a systematic analysis for the Global Burden of Disease Study 2015. Lancet Infect Dis. 2017 Nov 1;17(11):1133–61.

4. Mansbach JM, Emond JA, Camargo CA. Epidemiology and Practice Variation [Internet]. 1992. Available from: http://www2.

5. Mehta S, Shin H, Burnett R, North T, Cohen AJ. Ambient particulate air pollution and acute lower respiratory infections: A systematic review and implications for estimating the global burden of disease. Air Qual Atmos Health. 2013;6(1):69–83.

6. Middelkoop K, Bekker LG, Morrow C, Zwane E, Wood R. Childhood tuberculosis infection and disease: A spatial and temporal transmission analysis in a South African township. South African Medical Journal. 2009;99(10):738–43.

7. Tian L, Yang C, Zhou Z, Wu Z, Pan X, Clements ACA. Spatial patterns and effects of air pollution and meteorological factors on hospitalization for chronic lung diseases in Beijing, China. Sci China Life Sci. 2019;62(10):1381–8.

8. Bhutta ZA, Das JK, Walker N, Rizvi A, Campbell H, Rudan I, et al. Interventions to address deaths from childhood pneumonia and diarrhoea equitably: what works and at what cost? The Lancet. 2013 Apr 20;381(9875):1417–29.

9. Bamford L, Barron P, Kauchali S, Dlamini N, Paeds M, O'SULLIVAN D, et al. Investigating the early-life determinants of illness in Africa: The Drakenstein Child Health Study. BMJ Open [Internet]. 2017;14(1):1–6. Available from: http://dx.doi.org/10.1016/S2352-4642(17)30149-9

10. Villanueva K, Pereira G, Knuiman M, Bull F, Wood L, Christian H, et al. The impact of the built environment on health across the life course: Design of a cross-sectional data linkage study. BMJ Open. 2013;3(1).

11. Indriyani W, Yudhistira MH, Sastiono P, Hartono D. The relationship between the built environment and respiratory health: Evidence from a longitudinal study in Indonesia. SSM Popul Health. 2022 Sep 1;19.

12. Beamer PI, Lothrop N, Lu Z, Ascher R, Ernst K, Stern DA, et al. Spatial clusters of child lower respiratory illnesses associated with community-level risk factors. Pediatr Pulmonol. 2016;51(6):633–42.

13. Gehring U, Cyrys J, Sedlmeir G, Brunekreef B, Bellander T, Fischer P, et al. Traffic-related air pollution and respiratory health during the first 2 yrs of life. European Respiratory Journal. 2002;19(4):690–8.

14. Morgenstern V, Zutavern A, Cyrys J, Brockow I, Gehring U, Koletzko S, et al. Respiratory health and individual estimated exposure to traffic-related air pollutants in a cohort of young children. Occup Environ Med. 2007 Jan;64(1):8–16.

15. Andrade ALSS de, Silva SA e., Martelli CMT, Oliveira RM de, Morais Neto OL de, Siqueira Júnior JB, et al. Population-based surveillance of pediatric pneumonia: use of spatial analysis in an urban area of Central Brazil. Cadernos de saúde pública / Ministério da Saúde, Fundação Oswaldo Cruz, Escola Nacional de Saúde Pública. 2004;20(2):411–21.

16. Mukhopadhyay AK. Mapping of cholera cases using satellite based recording systems to investigate the outbreak. Vol. 142, Indian Journal of Medical Research. Indian Council of Medical Research; 2015. p. 509–11.

17. Nuckols JR, Ward MH, Jarup L. Using geographic information systems for exposure assessment in environmental epidemiology studies. Vol. 112, Environmental Health Perspectives. Public Health Services, US Dept of Health and Human Services; 2004. p. 1007–15.

18. Sartorius BKD, Sartorius K, Chirwa TF, Fonn S. Infant mortality in South Africa - distribution, associations and policy implications, 2007: an ecological spatial analysis. Int J Health Geogr [Internet]. 2011;10(1):61. Available from: http://www.ij-healthgeographics.com/content/10/1/61

19. Tlou B, Sartorius B, Tanser F. Space-time variations in child mortality in a rural South African population with high HIV prevalence (2000–2014). PLoS One. 2017;12(8):1–13.

20. Auchincloss AH, Gebreab SY, Mair C, Diez Roux A v. A review of spatial methods in epidemiology, 2000-2010. Vol. 33, Annual Review of Public Health. 2012. p. 107–22.

21. Congdon P. Spatiotemporal frameworks for infectious disease diffusion and epidemiology. Int J Environ Res Public Health. 2016;13(12):15–8.

22. Alene KA, Viney K, McBryde ES, Clements ACA. Spatiotemporal transmission and socio-climatic factors related to paediatric tuberculosis in north-Western Ethiopia. Geospat Health. 2017;12(2):342–50.

23. Amsalu ET, Akalu TY, Gelaye KA. Spatial distribution and determinants of acute respiratory infection among under-five children in Ethiopia: Ethiopian demographic Health Survey 2016. PLoS One. 2019;14(4):1–14.

24. Beck AF, Florin TA, Campanella S, Shah SS. Geographic variation in hospitalization for lower respiratory tract infections across one county. JAMA Pediatr. 2015;169(9):846–54.

25. Ho NT, Thompson C, Nhan LNT, van Hoang MT, Dung NT, Tran My P, et al. Retrospective analysis assessing the spatial and temporal distribution of paediatric acute respiratory tract infections in Ho Chi Minh City, Vietnam. BMJ Open. 2018;8(1):1–9.

26. Zar HJ, Barnett W, Myer L, Stein DJ, Nicol MP. Investigating the early-life determinants of illness in Africa: The Drakenstein Child Health Study. Thorax. 2015;70(6):592–4.

27. Zar HJ, Barnett W, Myer L, Stein DJ, Nicol MP. Investigating the early-life determinants of illness in Africa: The Drakenstein Child Health Study. Thorax. 2015;70(6):592–4.

28. REDCap Terms of Use ARC-REDCap Terms of Use REDCap Terms of Use [Internet]. 2018. Available from: https://activate.id.ubc.ca/iamweb/

29. Martinez L, le Roux DM, Barnett W, Stadler A, Nicol MP, Zar HJ. Tuberculin skin test conversion and primary progressive tuberculosis disease in the first 5 years of life: a birth cohort study from Cape Town, South Africa. Lancet Child Adolesc Health [Internet]. 2018;2(1):46–55. Available from: http://dx.doi.org/10.1016/S2352-4642(17)30149-9

30. MacGinty RP, Lesosky M, Barnett W, Stein DJ, Zar HJ. Associations between maternal mental health and early child wheezing in a South African birth cohort. Pediatr Pulmonol. 2018 Jun 1;53(6):741–54.

31. Stinchcomb D. Procedures for Geomasking to Protect Patient Confidentiality Overview Motivations Survey of geomasking methods Random perturbation method: General geomasking procedures Issues and potential problems Example of geomasking within a GIS environment Automation and extensions Conclusions. 2004.

32. Armstrong MP, Rushton G, Zimmerman DL. Geographically Masking Health Data To Preserve Confidentiality. Statistics In Medicine Statist Med [Internet]. 1998;18:497–525. Available from: https://onlinelibrary.wiley.com/doi/10.1002/

33. Hampton KH, Fitch MK, Allshouse WB, Doherty IA, Gesink DC, Leone PA, et al. Mapping health data: Improved privacy protection with donut method geomasking. Am J Epidemiol. 2010 Nov 1;172(9):1062–9.

Part B: MANUSCRIPT

The spatial analysis of the associations between built environment and lung health outcomes in children from Drakenstein, Western Cape: a nested case-control study.

Author: Demi A Meyer 1*

Affiliations

1. Division of Epidemiology and Biostatistics, School of Public Health and Family Medicine, of Cape Town, South Africa

Corresponding Author (*)

Demi Meyer

Mailing Address: Division of Epidemiology & Biostatistics

School of Public Health & Family Medicine

University of Cape Town Faculty of Health Sciences

Anzio Road, Observatory

Cape Town, South Africa

Email Address: myrdem001@myuct.ac.za

Contact number: +27 74 881 8586

Keywords

Lower Respiratory Tract Infections; Tuberculosis; Built Environment; Spatial Analysis.

Word Count

Manuscript: 5463

Abstract: 200

§ This manuscript is written according to the requirements in the Instructions for Authors for International Journal of Environmental Research and Public Health (outlined in Appendix D). Deviations have been made to accommodate the University of Cape Town MPH dissertation requirements. As a result, co-authors are not listed on the journal ready manuscript.

ABSTRACT

Despite the wide geographic distribution of childhood lung disease, data on geospatial risk factors at the local level in South Africa is scarce. This study aims to assess if there is any relationship between the social and environmental risk factors (built environment) and spatial distribution of child lung health outcomes in Drakenstein. An informed consent to collect household addresses was obtained from 884 mother-child pairs. The average distance in kilometres between participant's address and the built environment categories, as well as the average number of each built environment category within a 500-metre radius of the participant's address, were calculated. To evaluate the relationship between lung health and the built environment, multivariable logistic regression was used. Using this logistic regression outputs, density hotspots and coldspots were mapped. The subsampled distance matched cohort included 408 LRTI cases and controls respectively (1:1 ratio) and 75 LTBI cases and 375 LTBI controls (1:5 ratio). In the LRTI and LTBI models, the number of significant differences between cases and controls vary. The main built environment types that were associated with LRTI were spaza shops [OR = 1.05, 95%CI (1.02; 1.08), p = 0.01], formal liquor stores [OR = 2. 63, 95%CI (1.20; 5.90); p = 0.02] and agriculture (farming places) [aOR = 0.34, 95%CI, 0.11; 0.82 p = 0.03]. The main built environment types that were associated with LTBI were formal [OR = 0.26, 95%CI (0.09; 0.68), p = 0.01] and informal liquor stores [OR = 0.27, 95%CI (0.08; 0.85), p = 0.01] and religious (places of worship) [OR = 11.81, 95%CI (1.51; 101.88), p = 0.02]. In Drakenstein, the associations between the built environment and the spatial distribution of child lung health outcomes are present after distance-matching cases and controls and adjusting for covariates. Thus, creation of more green spaces or designing of buildings with better air flow would be useful and actionable by the government local authorities as potential strategies to reduce the burden of childhood LRTI and LTBI in LMICs.

Introduction

Common childhood lung conditions, such as lower respiratory tract infections (LRTIs) including pneumonia and tuberculosis (TB), are some of the major sources of child mortality in South Africa and globally (1). According to recent data, overall child mortality has been gradually decreasing at an annual rate of 1.1%, and as a result, South Africa may fail to meet the post-2015 Sustainable Development Goals (SDGs) child health target of 25 deaths per 1000 livebirths by 2030 (2). The data also show that there are significant geographic disparities in child mortality. The Western Cape province, for example, has lower levels of child mortality than other provinces, but this masks significant heterogeneity within the province, with specific regions of the Western Cape, including the Drakenstein region, having some of the highest levels of child morality in South Africa (3). Although the risk factors associated with child mortality are well studied, the drivers of these geographical disparities are poorly understood. Understanding the drivers of these disparities can inform the development of targeted public health interventions.

A substantial proportion of child deaths can be attributed to lung conditions. About 60 000 in-hospital deaths are caused by LRTI's in children under the age of five with just under half in infants under the age of six months (4). Evidence has demonstrated that complex interactions of demographic, environmental, behavioral, socio-economic, genetic, and viral factors have a substantial effect on the spatial variation of children respiratory health outcomes [4, 5]. Furthermore, it has been demonstrated that some of these risk factors may be spatially dependent (7). For example, it was discovered that geographical clusters of LRTIs in the United States of America (USA) were associated with air pollution and substandard living conditions (5). In addition, the type of housing materials had a significant impact on acute respiratory infections symptoms in a study conducted in Pakistan(8).

There are studies in South Africa have revealed spatial clustering of child mortality. A South African study employed a descriptive geographical analysis to demonstrate a link between adult TB cases exposure in a suburban setting and latent tuberculosis infection (LTBI) in children (6). A study by Middlekoop et al. indicates that in an immediate social network, the frequency of infectious adult TB was statistically correlated with childhood illness and infection (6). In Cape Town, a study by Wood et al. has reported that Human Immunodeficiency Virus (HIV) infection is affecting the increasing LTBI epidemic in an urban–rural community (9). There are additional risk factors to consider for childhood respiratory diseases. According to Aftab et al., children of mothers with primary or middle education have significantly higher odds of developing acute respiratory infections than children of mothers with

no education (8). The findings are similar to earlier studies carried out in Bangladesh (10) and Nigeria (11). Smoking is a known risk factor of childhood respiratory infections and studies have proven that it may cause or be worsened by household passive smoking (12). Because these risk factors are spatially and temporally varied, it is increasingly important to consider spatial dependence impacting spatial patterns of health outcomes (13).

Identifying exposures of public health significance can be done by mapping the regional distributions of illness incidence (14). Exploratory spatial analysis combines a number of statistical methods with the goal of defining and visualizing spatial distributions, locating spatial outliers, finding clusters or hotspots. Cases of disease have been shown to cluster at specific sites by modelling the spatial aspect of epidemiological data (15). Drakenstein's spatial clusters of childhood lung diseases may hold information about the region's socio-economic, behavioral, environmental, genetic, and viral risk factors. Furthermore, identifying areas that are more affected by LRTI and LTBI would lead to more efficient allocation of resources in the Drakenstein neighbourhood.

There have been few studies exploring the geospatial patterns of childhood lung diseases outside of South Africa. For example, one study from north-western Ethiopia found significant spatial clustering of childhood TB within local communities (16). According to an American study, children's bronchiolitis and LRTI cases were also spatially clustered (17). A study in Vietnam found spatio-temporal clustering of paediatric acute LRTIs (13). The ability to identify specific regions for interventions makes spatio-temporal analysis of child lung health outcomes popular. Thomas et al. found that although children from urban areas made up the majority of those with pneumonia, rural areas had an excessively higher prevalence of severe and very severe pneumonia (18). Pneumonia patients present with a cough or difficulty breathing and either have severe pneumonia (lower chest wall indrawing) or very severe pneumonia (central cyanosis, difficulties breastfeeding/drinking, vomiting everything, convulsions, lethargy, coma, or head nodding) (19). In an effort to reduce the burden of disease and child mortality caused by pneumonia, the pneumococcus vaccine has been rolled out globally. A study in the Philippines examined how the spatiotemporal clustering of pediatric LRTI affects the local level of pneumococcal vaccination effectiveness. Compared to vaccination with the placebo, pneumococcal vaccination was linked with a lower hazard ratio for pneumonia in children who lived further from the hospital. However, there was no significant difference in the likelihood of pneumonia diagnosis between children who lived closer to the hospital and those who received the placebo vaccine (20).

This demonstrated the effectiveness of the vaccination. Although vaccination is of vital importance, there is a link between where you stay and respiratory disease.

It is widely recognized that the built environment influences public health (21). The term "built environment", as it is referred to in urban planning, includes elements of municipal layout, the use of land, the transport network, and what the human movement patterns are in the physical environment (22). The study by Xu et al. shows a clear association between respiratory disease and the environment in which individuals reside, supporting the premise that urban redevelopment could perform a vital role in lowering respiratory disease risk (23). The effect of built environments on childhood lung health was also demonstrated in a study conducted in the Philippines which highlighted how proximity to hospitals affected the prognosis of children diagnosed with pneumonia (20). In addition, the built environment impacts public health by influencing human exposure to airborne pollutants. When airborne particles from a symptomatic person leave their body (via coughing, speaking, or singing) and come into contact with a person who does not have the disease, TB is transmitted (24). The spread of TB is facilitated by overcrowding and poor ventilation in households, hospitals, and public transit. This is often a result of poor infrastructure and is directly linked to built environments (25)**.** Research studies have investigated the link between air pollution and the lung health disease incidence and prevalence. Field studies have demonstrated that the development of urban land environments, transport networks, green space, and many other built environmental factors impacts the dissemination and distribution of air contaminants (26). According to a study conducted in rural South Africa, opening windows and doors to allow more air to circulate reduced the probability of TB transmission from 55.4% to 9.6% (21). Improving the built environment provides advantages beyond those that directly help health. Better roads, for instance, enhance accessibility to educational institutions, healthcare facilities, and food outlets, all of which can significantly lower the burden of LRTI (25)**.** The burden of childhood lower respiratory tract infections such as pneumonia varies greatly within and between provinces, districts, and sub-districts (27). However, the data is usually presented only at provincial, district, or subdistrict levels. To date, there has been no research on the spatial epidemiology of child lung health outcomes and determinants at local level in South Africa.

This study aims to assess if there is any relationship between the social and environmental risk factors (built environment) and spatial distribution of child lung health outcomes in Drakenstein. This setting is ideal for the proposed study because it is typical of many peri-urban disadvantaged communities in South Africa and has a stable population with minimum immigration and emigration**.**

Methods

Study design

This is a nested case-control study of children included in the Drakenstein Child Health Study. The current study is a secondary analysis of data collected from a birth cohort, the Drakenstein Child Health Study (DCHS). The characteristics of the birth cohort and study procedures have been detailed elsewhere (28). Briefly, comprehensive clinical, socio-economic, behavioural, and demographic data was collected annually by trained personnel on roughly 1000 mother-child pairs. Clinical data included measures of child lung health such as the presence or absence of LRTI and LTBI. The study design was a case-control where cases were children with LRTI and LTBI.

Study Setting

The study is set in Drakenstein which is a peri-urban area within the Cape Winelands District Municipality, South Africa (figure 1). Drakenstein is one of the five local municipalities located within the Cape Winelands District of Western Cape Province, South Africa. The study area covers 1538 km², with a population of approximately 200 000 people. Unemployment in the area is high, with many people of low socio-economic status. The DCHS cohort was recruited from two primary healthcare facilities (sites): Mbekweni Clinic, which mainly provides services to the Black community, and TC Newman Clinic, which mainly provides services to those of coloured (a South African race classification) community. This setting is ideal for the proposed study because it is typical of many peri-urban disadvantaged communities in South Africa and has a stable population with low immigration and emigration.

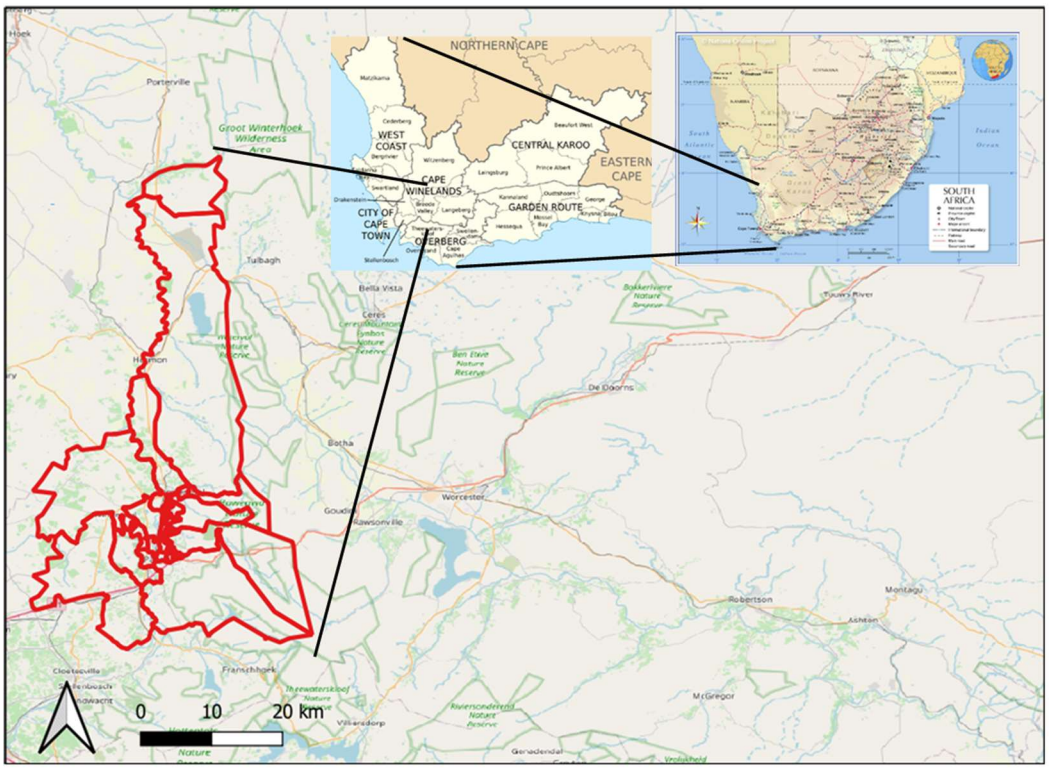


Figure 1: The map of the study areas within the Drakenstein region, in the Western Cape, South Africa.

Study participants

The parent study recruited women who were 20-28 weeks pregnant, 18 years or older, attending antenatal care at either Mbekweni or TC Newman Clinic. The women were then followed through pregnancy to childbirth, where their infants joined the study. Of the approximately 1000 mother-infant pairs in the parent study, only 844 were included in this analysis as they agreed to have the of their household addresses collected and geocoded.

Data on Childhood Lung health

Childhood lung health was determined using the Tuberculin skin tests (TST) chest radiographs, TB microbiological culture, and Xpert MTB/RIF(29). These tests were conducted annually and whenever the child presented with a LRTI episode. A child was considered to be a LTBI case if the TST test that resulted in an induration reaction with a diameter of 10mm or more or was on TB treatment. Data on LRTIs which include ambulatory and hospitalized LRTI cases was obtained from active surveillance

system for LRTIs in the DCHS. LRTIs are characterised using World Health Organization (WHO) case definition criteria [28, 29]. The sociodemographic covariates included in the analysis were known and include maternal HIV exposure, maternal smoking, house type, number of people in the household, household income, site (Mbekweni and TC Newman) and day care attendance (31).

Matched case-control sampling

To account for possible spatial dependence, cases were matched to controls that were within a 1-kilometre radius of each other. The rationale for using a 1-kilometre radius was because TC Newman and TC Mbekweni are relatively small regions within Drakenstein. This was done using an iterative algorithm that randomly sampled controls that fell within the 1-kilometre radius for each case. The distance matched subsampled populations included 408 LRTI cases and controls (1:1 ratio) and 75 LTBI cases and 375 LTBI controls (1:5 ratio). The matched case control ratio was chosen based on the overall case control ratio of the total population.

Data Analysis

Statistical methods

Multivariable logistic regression was used to estimate the association between the built environment and the presence of LTBI and LRTI. For each independent variable, we present univariable models and multivariable models (direct effects, after adjustment for the effects of potential confounding variables). Two methods were used for this analysis. The first method used logistic regression on the whole sample, without any prior matching of cases to control. The second method used the same logistic regression models on distance matched cases and controls. Both models were included in this analysis for comparison. The matched case-control sampling was explained above under the same heading. The significance level was set to < 0.05.

Socio-demographic covariates considered for the analyses included: HIV exposure; type of housing; number of people in the household; household income; maternal smoking; day-care attendance; Site of visit (Mbekweni or TC Newman) and socio-economic status (SES). This was determined by a total score of four socio-economic variables: level of education, employment status, household income,

and number of assets (32). Standardised scores were divided into quartiles: 'low,' 'low-moderate,' 'high-moderate,' and 'high'.

Data Visualisation

Several maps were used to illustrate the relationship between the built environment and lung health outcomes. These include hotspot maps that show the spatial distribution of LTBI and LRTI cases relative to their distance from certain built environments. These hotspots were defined as areas (500 metre^2 grids) where cases were found to be within a certain distance of built environment types that included healthcare facilities or providers, educational facilities, formal liquor stores, informal liquor stores (bar, tavern, shebeen - an unlicensed establishment selling alcohol), spaza shops (informal convenient stores) and street vendors, malls or supermarkets, agriculture (farming places), safety and security (police stations), and places of worship. Similar maps were used to show the relationship between lung health outcomes and the frequency of the abovementioned built environment types. All maps were created using QGIS (33).

Hotspot Analysis

Hotspot analysis was done to visualize the proximity of various built environment types to LRTI and LTBI cases, as well as to visualize the average number of built environments within a 500-metre radius to the cases using QGIS. A hotspot was when a built environment type was within a 500-metre radius of a case. A hotspot was when the average number of built environment types (agricultural or farming places and formal liquor stores) within a 500-metre radius was higher than 1. A hotspot was when the average number of built environment (spaza shop or street vendor) within a 500-metre radius was higher than 25.

Ethical clearance

Ethical clearance was given by University of Cape Town Faculty of Health Sciences Research Ethics Committee for the parent study (HREC Ref number: 401/2009) (Appendix C1) and the current study (HREC Ref number: 758/2021) (Appendix C2)

Results

Descriptive statistics and exploratory analysis

The method of distance matching LRTI cases to controls resulted in a subsample of 816 mother-child pairs, 408 cases and 408 controls (1:1 ratio). Whereas distance matching of LTBI cases to controls resulted in a subsample of 450 mother-child pairs, 75 cases and 375 controls (1:5 ratio). The LRTI and LTBI cases were mostly male children, with 58% and 65% respectively (Table 1).

Socio-demographics and clinical factors by case control sampling

The socio-demographic and clinical factors described are based on the distance matched sub-sample. Of those participants who are LRTI cases, 25% are exposed to HIV-positive mothers, compared to 17% LRTI controls are exposed to HIV-positive mothers. Additionally, 20% of LTBI cases compared to 14% LTBI controls are exposed to HIV-positive mothers. The LRTI cases and controls are closer to the total number of participants in the study. The average number of people in a household was 3 among cases and controls in both LRTI and LTBI populations. In LRTI and LTBI populations, the participants are of all socio-economic status backgrounds however, most of the participants have an average household income of 1000-5000 ZAR per month at 50% and 53%, respectively. Of those participants who are LRTI cases, 39% are exposed to mothers who are active smokers, compared to 30% from the participants who are LRTI controls. Additionally, 48% of LTBI cases compared to 39% LTBI controls are exposed to mothers who are active smokers. The LRTI population is equally spread between the two sites of Mbekweni (50%) and TC Newman (50%) whereas the LTBI population is higher in the TC Newman neighbourhoods at 65% compared to 35% in Mbekweni neighbourhoods (Table 1).

Table 1: Descriptive table of sociodemographic characteristics of child- mother pairs matched by distance to nearest neighbour within a 1-kilometre radius and total population child- mother pairs included in this study by case control groups within disease status.

Variable: n(%)	Matched case controls sample						Total study sample					
	LRTI			LTBI			LRTI			LTBI		
	CASES	CONTROLS	TOTAL	CASES	CONTROLS	TOTAL	CASES	CONTROLS	TOTAL	CASES	CONTROLS	TOTAL
Number of Participants (%)	408 (50)	408 (50)	816	75 (16.7)	375 (83.3)	450	409 (48.5)	435 (51.5)	844 (100)	75 (8.9)	769 (91.1)	844 (100)
HIV Status (Positive Mothers)	100 (25)	68 (17)	168 (21)	15 (20)	54 (14)	69 (15)	100 (24)	73 (17)	173 (20)	15 (20)	158 (21)	173 (20)
Sex of child (% male)	237 (58)	185 (45)	422 (52)	49 (65)	185 (49)	234 (52)	238 (58)	194 (45)	432 (51)	49 (65)	383 (50)	432 (51)
Currently employed mothers	101 (25)	110 (27)	211 (26)	15 (20)	94 (25)	109 (24)	101 (25)	118 (27)	219 (26)	15 (20)	204 (27)	219 (26)
Highest Education Level attained												
Primary	34 (8)	27 (7)	61 (7)	8 (11)	28 (7)	36 (8)	34 (8)	30 (7)	64 (8)	8 (11)	56 (7)	64 (8)
Some secondary	226 (56)	229 (56)	455 (56)	45 (60)	206 (55)	251 (56)	226 (55)	237 (54)	463 (55)	45 (60)	418 (54)	463 (55)
Completed secondary	127 (31)	125 (31)	252 (31)	20 (27)	117 (31)	137 (30)	128 (31)	139 (32)	267 (32)	20 (27)	247 (32)	267 (32)
Any tertiary	19 (5)	26 (6)	45 (6)	1 (1)	23 (6)	24 (5)	19 (5)	28 (6)	47 (6)	1 (1)	46 (6)	47 (6)
Average household income per month												
<1000ZAR	146 (36)	160 (39)	306 (38)	34 (45)	131 (35)	165 (37)	147 (36)	170 (39)	317 (38)	34 (45)	283 (37)	317 (38)
1000-5000ZAR	205 (50)	204 (50)	409 (50)	35 (47)	202 (54)	237 (53)	205 (50)	215 (49)	420 (50)	35 (47)	385 (50)	420 (50)
>5000ZAR	55 (13)	43 (11)	98 (12)	1 (1)	41 (11)	46 (10)	55 (13)	49 (11)	104 (12)	1 (1)	99 (13)	104 (12)
SES Quartiles												
Lowest SES	93 (23)	96 (24)	189 (23)	20 (27)	84 (22)	104 (23)	93 (23)	102 (23)	195 (23)	20 (27)	175 (23)	195 (23)
Low-moderate SES	118 (29)	105 (26)	223 (27)	24 (32)	101 (27)	125 (28)	119 (29)	109 (25)	228 (27)	24 (32)	204 (27)	228 (27)
Moderate-high SES	105 (26)	109 (27)	214 (26)	18 (24)	106 (28)	124 (28)	105 (26)	115 (26)	220 (26)	18 (24)	202 (26)	220 (26)
High SES	92 (22)	98 (26)	190 (23)	13 (17)	84 (22)	97 (22)	92 (22)	109 (25)	201 (24)	13 (17)	188 (24)	201 (24)
Creche / day-care at birth	102 (25)	89 (22)	191 (23)	17 (23)	85 (23)	102 (23)	102 (25)	92 (21)	194 (23)	17 (23)	177 (23)	194 (23)
Antenatal smoking												
Nonsmoker	80 (20)	90 (22)	170 (21)	15 (20)	67 (18)	82 (18)	80 (20)	99 (23)	179 (21)	15 (20)	164 (21)	179 (21)
Passive smoker	155 (38)	182 (45)	337 (41)	23 (31)	152 (41)	175 (39)	155 (38)	191 (44)	346 (41)	23 (31)	323 (42)	346 (41)
Active smoker	160 (39)	123 (30)	283 (35)	36 (48)	145 (39)	181 (40)	161 (39)	131 (30)	292 (35)	36 (48)	256 (33)	292 (35)
Average nr of people in household. Mean (SD)	3 (1.75)	3 (1.67)	3 (1.71)	3 (2.38)	3 (1.69)	3 (1.82)	3 (1.75)	3 (1.64)	3 (1.7)	3 (2.38)	3 (1.61)	3 (1.7)
Housing type												
House	190 (47)	202 (50)	392 (48)	31 (41)	155 (41)	186 (41)	190 (46)	212 (49)	402 (48)	31 (41)	371 (48)	402 (48)
Flat	78 (19)	72 (18)	150 (18)	20 (27)	95 (25)	115 (26)	79 (19)	75 (17)	154 (18)	20 (27)	134 (17)	154 (18)
Shack	85 (21)	72 (18)	157 (19)	9 (12)	62 (17)	71 (16)	85 (21)	81 (19)	166 (20)	9 (12)	157 (20)	166 (20)
Wendy house/ backyard dwelling	52 (13)	59 (14)	111 (14)	14 (19)	61 (16)	75 (17)	52 (13)	64 (15)	116 (14)	14 (19)	102 (13)	116 (14)
SITE												
Mbekweni	208 (51)	204 (50)	412 (50)	24(32)	132 (35)	156 (35)	208 (51)	213 (49)	421 (50)	24(32)	397 (52)	421 (50)
TC Newman	200 (49)	204 (50)	404 (50)	51 (68)	243 (65)	294 (65)	201 (49)	222 (51)	423 (50)	51 (68)	372 (48)	423 (50)

Notes: All results show frequencies (percentages) or mean (Standard deviation). Self-reported smoking status at the baseline study visit. LRTI = lower respiratory tract infection; LTBI = latent tuberculosis infection; SES = socioeconomic status; ZAR= South African Rand.

Socio-demographics and clinical factors by site

In the overall cohort of 844 mother- child pairs between the TC Newman and Mbekweni neighbourhoods (Table S1) is the proportion of mothers living with HIV: it is 3% and 28% respectively. Furthermore, the percentage of children exposed to mothers who are active smokers in the TC Newman and Mbekweni neighbourhoods are 54% and 15% respectively. The type of housing is also different between the two neighbourhoods. The percentage of people living in shacks in the TC Newman and Mbekweni neighbourhoods are 5% and 34% respectively. In addition, percentage of people living in Wendy houses or backyard dwellings in the TC Newman and Mbekweni neighbourhoods are 25% and 2% respectively.

Type of Built environment

The frequency of the type of built environment was recorded and split by site (Figure S2). Spaza shops and Street vendors make up more than half the built environments at 160 and 152 for Mbekweni and TC Newman respectively. They are followed by educational facilities: 34 and 53 for Mbekweni and TC Newman respectively. Furthermore, there were 32 and 46 religious (places of worship) captured for Mbekweni and TC Newman respectively. Lastly, there were 19 and 18 shopping centre or mall or supermarkets that have been recorded for Mbekweni and TC Newman, respectively. The rest of the built environment types made up 12 % of the total recorded frequency namely: agriculture (farming site), informal and formal alcohol sale point, healthcare provider or facility, place of safety and security.

Logistic Regression Results

Firstly, the distance in kilometres between built environment and participants' addresses (independent variable) was used to quantify the spatial differences between LRTI and LTBI cases and controls (dependent variable). Similarly, the number of built environment facilities within a 500-metre radius of participants' addresses (independent variable) was used to quantify the spatial differences between LRTI and LTBI cases and controls (dependent variable).

Table 2**:** Multivariable logistic regression models of LTBI within Drakenstein using the distance to the nearest built environment structure in kilometres.

	TOTAL UNMATCHED MODELS				DISTANCE MATCHED MODELS			
	BE		BES		BE		BES	
	AIC (509.68)		AIC (493.84)		AIC (418.63)		AIC (412.78)	
Covariates	**OR (95% CI)**	**P-value**	**aOR (95% CI)**	**P-value**	**OR (95% CI)**	**P-value**	**aOR (95% CI)**	**P-value**
Built Environment								
Distance to nearest agriculture/ farming area (km)	1.12 (0.34; 3.55)	0.85	0.56 (0.14; 2.11)	0.40	0.43 (0.12; 1.56)	0.20	0.24 (0.05; 1.04)	0.06
Distance to nearest educational facility (km)	0.09 (0.01; 1.02)	0.05	0.08 (0.01; 1.02)	0.05	0.24 (0.02; 2.96)	0.27	0.21 (0.01; 3.33)	0.27
Distance to nearest formal liquor store (km)	**0.36 (0.13; 0.91)**	**0.04***	**0.26 (0.09; 0.68)**	**0.01***	0.38 (0.12; 1.13)	0.09	**0.27 (0.08; 0.85)**	**0.03***
Distance to nearest informal liquor store (e.g., shebeen) (km)	1.40 (0.63; 3.01)	0.40	0.46 (0.14; 1.48)	0.20	0.67 (0.28; 1.58)	0.36	**0.23 (0.06; 0.90)**	**0.04***
Distance to nearest healthcare facility/provider (km)	**2.45 (1.04; 5.81)**	**0.04***	0.56 (0.15; 2.16)	0.40	0.89 (0.35; 2.30)	0.81	0.25 (0.05; 1.06)	0.06
Distance to nearest religious / places of worship (km)	3.05 (0.68;14.22)	0.15	**8.76 (1.45; 57.01)**	**0.02***	3.99 (0.72; 22.78)	0.12	**11.81 (1.51; 101.88)**	**0.02***
Distance to nearest safety & security (km)	0.82 (0.40; 1.70)	0.59	1.80 (0.73; 4.38)	0.20	1.59 (0.74; 3.48)	0.24	**3.03 (1.12; 8.44)**	**0.03***
Distance to nearest spaza shop (km)	0.14 (0.00; 1.74)	0.34	0.56 (0.01; 18.57)	0.80	0.48 (0.00; 41.19)	0.75	1.20 (0.01; 137.25)	0.94
Distance to nearest mall/ supermarket (km)	2.12 (0.65; 6.96)	0.21	1.65 (0.47; 5.78)	0.43	1.46 (0.38; 5.62)	0.58	1.19 (0.28; 4.94)	0.80
HIV exposed (YES)			1.75 (0.80; 3.76)	0.15			1.69 (0.73; 3.82)	0.21
Type of house								
Flat			reference category				reference category	
House			0.81 (0.38; 1.69)	0.57			0.92 (0.43; 1.96)	0.83
Shack			0.72 (0.25; 1.98)	0.53			0.84 (0.27; 2.43)	0.75
Wendy house/ backyard dwelling			1.12 (0.50; 2.48)	0.78			1.30 (0.56; 2.99)	0.53
Number of people in household			1.06 (0.96; 1.16)	0.23			1.07 (0.96; 1.18)	0.18
SES quartile								
Moderately-high			reference category				reference category	
Lowest			1.44 (0.50; 4.26)	0.50			1.22 (0.39; 3.87)	0.74
Moderately-low			1.30 (0.53; 3.34)	0.58			1.13 (0.43; 3;12)	0.80
Highest			1.09 (0.46; 2.71)	0.84			1.01 (0.41; 2.61)	0.99
Household income								
<1000ZAR			reference category				reference category	
1000-5000ZAR			0.65 (0.38; 1.09)	0.10			0.58 (0.33; 1.00)	0.05
>5000ZAR			**0.34 (0.11; 0.87)**	**0.04***			0.45 (0.14; 1.19)	0.13
Smoking								
Active smoker			reference category				reference category	
Non-smoker			1.32 (0.61; 2.78)	0.46			1.24 (0.53; 2.28)	0.62
Passive smoker			0.81 (0.44; 1.46)	0.48			0.83 (0.44; 1.56)	0.57
Day care (YES)			1.09 (0.58; 1.94)	0.78			1.09 (0.57; 2.03)	0.79
Site (TC Newman)			**5.55 (1.56; 20.37)**	**0.01***			**4.67 (1.19; 19.75)**	**0.03***

Notes: LTBI = Latent Tuberculosis Infection; BE = Built Environment; BES = Built Environment and Socio-demographic risk factors; AIC = Akaike Information Criteria; OR = Odds Ratio; aOR = adjusted Odds Ratio; CI = confidence interval; * = Statistically Significant

In the total unmatched BE and BES model, an increased distance from the nearest formal liquor stores is associated with a 64% and 74% reduction in the odds of being an LTBI case [OR = 0.36, 95%CI (0.13; 0.91), p = 0.04]; [OR = 0.26, 95%CI (0.09; 0.68), p = 0.01]. This association persisted in the distance matched BE [OR = 0.38, 95%CI (0.12; 1.13), p = 0.09] and BES [OR = 0.27, 95%CI (0.08; 0.85), p = 0.01] models, however this association in the BE model was not statistically significant. In the distance matched BES model, an increased distance from the nearest informal liquor store is associated with 0.23 times the odds (a reduction of 77%) of being a LTBI case [OR = 0.23, 95%CI (0.06; 0.90), p = 0.04]. However, this association in the total unmatched BES model was not statistically significant.

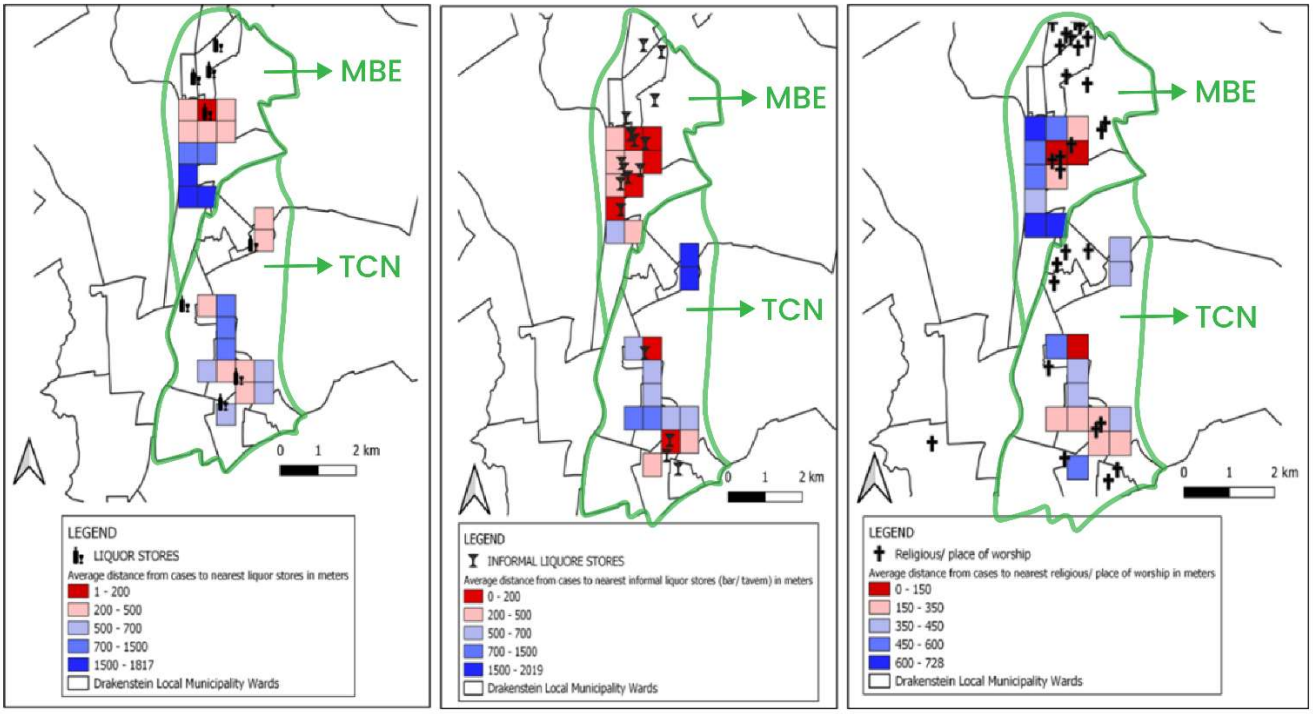
In the distance matched BES model, an increased distance from the nearest healthcare facility is associated with 0.25 times the odds of being a LTBI case [OR = 0.25, 95%CI (0.05; 1.06), p = 0.06]. In the total unmatched BES model, an increased distance from the nearest healthcare facility is associated with 0.56 times the odds of being a LTBI case [OR = 0.56, 95%CI (0.15; 2.16), p = 0.40]. These association were not significant. Interestingly, the only time this association was significant was in the unmatched BE model with an odds ratio of 2.45 [OR = 2.45, 95%CI (1.04; 5.81), p = 0.04].

In the total unmatched BES and distance matched BES models, an increased distance from the nearest religious (places of worship is associated with 8.76 and 11.81 times the odds of being a LTBI case respectively [OR = 8.76, 95%CI (1.45; 57.01), p = 0.02]; [OR = 11.81, 95%CI (1.51; 101.88), p = 0.02].

In the distance matched BES model, an increased distance from the nearest place of safety and security and the address location is associated with 3.03 times the odds of being a LTBI case [OR = 3.03, 95%CI (1.12; 8.44), p = 0.03]. However, this association in the total unmatched BES model was not statistically significant [OR = 1.80, 95%CI (0.73; 4.38), p = 0.20].

Of the sociodemographic risk factors included in the models, only household income and site were statistically significant. In the total unmatched BES model, the odds of LTBI are 66% lower if the participant lives in a household where the income is >5000ZAR compared with households with an income of <1000ZAR [OR = 0.34, 95%CI (0.11; 0.87), p = 0.04]. However, this is no longer statistically significant when matching cases and controls by distance. In the total unmatched BES model, the odds of LTBI are 5.55 times higher in TC Newman compared to Mbekweni [OR = 5.55, 95%CI (1.56; 20.37) p = 0.01]. In the distance matched BES model, the odds of LTBI are 4.67 times higher in TC Newman compared to Mbekweni [OR = 4.67, 95%CI (1.19; 19.75) p = 0.03] (Table 2).

The map in figure 2A reveals that, on average, there is a density hotspot of LTBI cases in the Mbekweni region that have a distance of between 1 - 200 metres to the nearest formal liquor store. The highest density of LTBI cases in the TC Newman region have a distance of between 200 - 500 metres to the nearest formal liquor store, on average. The pattern disperses from the hotspot to the surrounding areas, expanding to the outskirts of the study area. The map in figure 2B shows that, on average, there is a high-density hotspot cluster of LTBI cases in the Mbekweni region that have a distance of between 0 - 500 metres to the nearest informal liquor store. Similarly, two density hotspots of LTBI cases in the TC Newman region are 0 - 200 metres from the nearest informal liquor store, on average. The map in figure 2C shows that, on average, there is a density hotspot of LTBI cases in the Mbekweni region that have a distance of between 0 - 150 metres to the nearest religious (place of worship). The highest density of LTBI cases in the TC Newman region have a distance of between 150 - 350 metres to the nearest religious (place of worship), on average.



A **B** **C**

Figure 2: Maps of the spatial distribution of hotspots and coldspots of the average distance from residential addresses of LTBI cases to the nearest formal liquor store in meters (A), informal liquor store (bar or tavern) in meters (B) and religious (place of worship) in meters (C).

Table 3: Multivariable logistic regression models of LRTI within Drakenstein using the distance to the nearest built environment structure in kilometres.

	TOTAL UNMATCHED MODELS				DISTANCE MATCHED MODELS			
	BE AIC (1173.7)		BES AIC (1124)		BE AIC (1144.5)		BES AIC (1095.6)	
Covariates	**OR (95% CI)**	**P-value**	**aOR (95% CI)**	**P-value**	**OR (95% CI)**	**P-value**	**aOR (95% CI)**	**P-value**
Built Environment								
Distance to the nearest agriculture / farming area (km)	1.80 (0.97; 3.36)	0.06	**2.00 (1.02; 3.92)**	**0.04***	1.73 (0.91; 3.33)	0.10	1.94 (0.96; 3.98)	0.07
Distance to the nearest educational facility (km)	0.40 (0.10; 1.53)	0.18	0.32 (0.08; 1.32)	0.12	0.47 (0.12; 1.82)	0.27	0.37 (0.09; 1.56)	0.18
Distance to the nearest formal liquor store (km)	1.28 (0.76; 2.14)	0.35	1.20 (0.70; 2.03)	0.50	1.31 (0.76; 2.26)	0.34	1.25 (0.71; 2.21)	0.44
Distance to the nearest informal liquor store (e.g., shebeen) (km)	1.14 (0.74; 1.76)	0.54	1.35 (0.72; 2.52)	0.34	1.30 (0.82; 2.06)	0.27	1.52 (0.78; 2.98)	0.21
Distance to the nearest healthcare facility/provider (km)	1.45 (0.92; 2.30)	0.11	1.75 (0.83; 3.71)	0.14	1.27 (0.79; 2.02)	0.32	1.52 (0.70; 3.30)	0.29
Distance to the nearest religious / places of worship (km)	1.79 (0.79; 4.10)	0.17	1.58 (0.63; 4.01)	0.33	1.50 (0.64; 3.51)	0.35	1.24 (0.47; 3.25)	0.67
Distance to the nearest safety & security (km)	**0.63 (0.43; 0.91)**	**0.01***	**0.57 (0.35; 0.92)**	**0.02***	0.68 (0.46; 1.00)	0.05	0.62 (0.37; 1.03)	0.07
Distance to the nearest spaza shop (m)	**0.08 (0.01; 0.55)**	**0.01***	**0.07 (0.01; 0.57)**	**0.02***	0.17 (0.02; 1.75)	0.14	0.18 (0.01; 2.13)	0.18
Distance to the nearest mall/ supermarket (km)	1.16 (0.61; 2.19)	0.66	1.23 (0.63; 2.41)	0.55	1.08 (0.56; 2.08)	0.81	1.13 (0.57; 2.25)	0.72
HIV exposed (YES)			**1.73 (1.16; 2.59)**	**0.01***			**1.78 (1.19; 2.67)**	**0.01***
Type of house								
Flat			reference category				reference category	
House			0.79 (0.50; 1.26)	0.33			0.82 (0.51; 1.32)	0.41
shack			0.96 (0.53; 1.73)	0.88			1.08 (0.59; 1.98)	0.80
Wendy house/ backyard dwelling			0.76 (0.45; 1.28)	0.31			0.77 (0.45; 1.30)	0.32
Number of people in household			1.01 (0.95; 1.07)	0.76			1.01 (0.95; 1.07)	0.73
SES quartile								
Moderately-high			reference category				reference category	
Lowest			1.31 (0.72; 2.40)	0.37			1.24 (0.68; 2.28)	0.49
Moderately-low			1.40 (0.85; 2.32)	0.19			1.29 (0.78; 2.16)	0.32
Highest			1.18 (0.74; 1.90)	0.48			1.15 (0.71; 1.85)	0.57
Household income								
<1000ZAR			reference category				reference category	
1000-5000ZAR			1.14 (0.84; 1.55)	0.40			1.15 (0.84; 1.57)	0.39
>5000ZAR			1.29 (0.80; 2.09)	0.30			1.41 (0.86; 2.31)	0.17
Smoking								
Active smoker			reference category				reference category	

Non-smoker	0.67 (0.43; 1.04)	0.07	0.67 (0.43; 1.04)	0.08
Passive smoker	0.74 (0.52; 1.05)	0.09	0.73 (0.51; 1.03)	0.07
Day care (YES)	1.32 (0.94; 1.86)	0.11	1.25 (0.89; 1.77)	0.20
Site (TC Newman)	0.81 (0.41; 1.59)	0.54	0.86 (0.43; 1.71)	0.67

Notes: LRTI = Lower Respiratory Tract Infection; BE = Built Environment; BES = Built Environment and Socio-demographic risk factors; AIC = Akaike Information Criteria; OR = Odds Ratio; aOR = adjusted Odds Ratio, CI = confidence interval; * = Statistically Significant

In the total unmatched BES model, an increased distance from the nearest agriculture (farming place) is associated with 2 times the odds of being a LRTI case [OR = 2.00, 95%CI (1.02; 3.92) p = 0.04]. In the same model, an increased distance from the nearest safety and security facility is associated with 0.57 times (a reduction of 43%) the odds of being a LRTI case; [OR = 0.57, 95%CI (0.35; 0.92), p = 0.02]. Additionally, an increased distance from the nearest spaza shop is associated with 0.07 times (a reduction 93%) the odds of being a LRTI case [OR = 0.07, 95%CI (0.01; 0.57) p = 0.02]. It is worth noting that none of the built environment types were statistically significant in the distance matched BE and BES model.

Of the sociodemographic risk factors included, only maternal HIV status was statistically significant in the total unmatched and distance matched models. In the total unmatched BES model, the odds of having LRTI are 1.7 times higher for a child exposed to a mother living with HIV compared to a mother living without HIV [OR = 1.73, 95%CI (1.16; 2.59), p = 0.01]. In the distance matched BES model, the odds of having LRTI are 1.8 times higher for a child exposed to a mother living with HIV compared to a mother living without HIV [OR = 1.78, 95%CI (1.19; 2.67) p = 0.01] (Table 3).

Table 4: Multivariable logistic regression models of LTBI within Drakenstein using the number of built environment structure within a 500-meter radius.

	TOTAL UNMATCHED MODELS				DISTANCE MATCHED MODELS			
	BE AIC (504.52)		BES AIC (493.13)		BE AIC (415.24)		BES AIC (414.17)	
Covariates	**OR (95% CI)**	**P-value**	**aOR (95% CI)**	**P-value**	**OR (95% CI)**	**P-value**	**aOR (95% CI)**	**P-value**
Built Environment								
Number of agriculture/ farming within 500 m radius	1.12 (0.25; 3.36)	0.86	1.02 (0.22; 3.25)	0.97	1.24 (0.26; 4.53)	0.76	1.13 (0.23; 4.33)	0.87
Number of educational facilities within 500 m radius	1.20 (0.98; 1.46)	0.07	1.22 (0.99; 1.50)	0.06	1.17 (0.95; 1.44)	0.13	1.15 (0.92; 1.43)	0.21
Number of formal liquor stores within 500 m radius	**2.48 (1.28; 4.83)**	**0.01***	**2.71 (1.32; 5.61)**	**0.01***	**2.42 (1.18; 5.00)**	**0.02 ***	**2.63 (1.20; 5.90)**	**0.02***
Number of informal liquor stores (shebeens) within 500 m radius	0.95 (0.78; 1.14)	0.58	1.09 (0.86; 1.40)	0.47	1.03 (0.83; 1.27)	0.76	1.15 (0.87; 1.52)	0.32
Number of healthcare facilities/providers within 500 m radius	**0.34 (0.16; 0.69)**	**0.01***	**0.41 (0.18; 0.85)**	**0.02***	0.60 (0.27; 1.24)	0.18	0.65 (0.28; 1.45)	0.31
Number of religious/ places of worship within 500 m radius	0.91 (0.77; 1.08)	0.28	0.85 (0.70; 1.03)	0.10.	0.83 (0.68; 0.99)	0.05	0.82 (0.66; 1.00)	0.05
Number of Safety & security within 500 m radius	1.66 (0.91; 2.80)	0.07	1.50 (0.8; 2.58)	0.16	1.59 (0.83; 2.84)	0.14	1.48 (0.75; 2.76)	0.23
Number of Spaza shops within 500 m radius	1.00 (0.96; 1.04)	0.87	1.00 (0.96; 1.05)	0.94	1.00 (0.95; 1.04)	0.85	0.99 (0.94; 1.04)	0.69
Number of Mall/ supermarkets within 500 m radius	0.95 (0.65; 1.32)	0.77	0.95 (0.62; 1.35)	0.78	0.92 (0.63; 1.28)	0.65	0.92 (0.63; 1.32)	0.68
HIV exposed (YES)			1.63 (0.74; 3.51)	0.22			1.63 (0.70; 3.70)	0.24
Type of house								
Flat			reference category				reference category	
House			1.01 (0.48; 2.10)	0.99			1.13 (0.53; 2.43)	0.75
shack			0.86 (0.30; 2.35)	0.77			0.94 (0.31; 2.73)	0.91
Wendy house/ backyard dwelling			1.24 (0.56; 2.72)	0.59			1.46 (0.63; 3.32)	0.37
Number of people in household			1.07 (0.97; 1.17)	0.16			1.06 (0.96; 1.17)	0.24
SES quartile								
Moderately-high			reference category				reference category	
Lowest			1.43 (0.49; 4.26)	0.51			1.39 (0.45; 4.45)	0.57
Moderately-low			1.30 (0.53; 3.32)	0.58			1.19 (0.46; 3.26)	0.72
Highest			1.11 (0.47; 2.76)	0.82			1.06 (0.42; 2.76)	0.90
Household income								
<1000ZAR			reference category				reference category	
1000-5000ZAR			0.63 (0.37; 1.07)	0.09			0.61 (0.35; 1.05)	0.07
>5000ZAR			**0.34 (0.11; 0.86)**	**0.04***			0.39 (0.12; 1.05)	0.08
Smoking								

Active smoker	reference category		reference category	
Non-smoker	1.29 (0.60; 2.69)	0.51	1.17 (0.51; 2.63)	0.70
Passive smoker	0.83 (0.45; 1.51)	0.54	0.82 (0.43; 1.55)	0.55
Day care (YES)	1.08 (0.58; 1.93)	0.80	1.17 (0.61; 2.17)	0.62
Site (TC Newman)	**3.06 (1.06; 9.33)**	**0.04***	1.91 (0.61; 6.27)	0.27

Notes: LTBI = Latent Tuberculosis Infection; BE = Built Environment; BES = Built Environment and Socio-demographic risk factors; AIC = Akaike Information Criteria; OR = Odds Ratio; aOR = adjusted Odds Ratio; CI = confidence interval; * = Statistically Significant

In the total unmatched BES model, the increased number of formal liquor stores is associated with 2.70 times the odds of being a LTBI case [OR = 2. 71, 95%CI (1.32; 5.61), p = 0.01]. In the distance matched BES model, the increased number of formal liquor stores is associated with 2.60 times the odds of being a LTBI case [OR = 2. 63, 95%CI (1.20; 5.90); p = 0.02]. In the distance matched BES model, the increased number of healthcare facilities is associated with 0.41 times the odds of being a LTBI case [OR = 0.41, 95%CI (0.18; 0.85), p = 0.02]. However, this was not statistically significant in the distance matched BES model.

Of the sociodemographic risk factors included, only household income and site were statistically significant in the total unmatched model. The odds of LTBI are 66% if the mother-child pair lives in a household where the income is >5000ZAR compared with households with an income of <1000ZAR [OR = 0.34, 95%CI (0.11; 0.86) p = 0.04]. The odds of LTBI are 3 times higher in TC Newman compared to Mbekweni [OR = 3.06, 95%CI (1.06; 9.33), p = 0.04]. However, these are no longer statistically significant when matching cases and controls by distance (Table 4).

In figure 3 there are density hotspots of LTBI cases in both Mbekweni and TC Newman regions that have 1 liquor store within a 500-meter radius, on average.

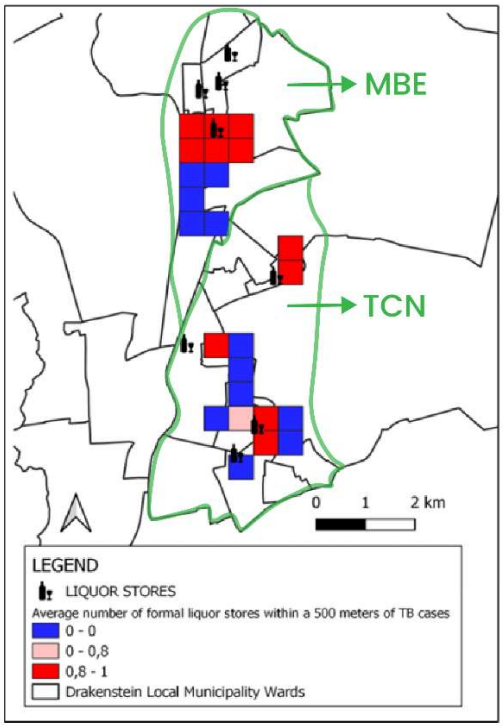


Figure 3: Map of spatial distribution of hotspots and coldspots of the average number of formal liquor stores within 500-metres of the residential addresses of LTBI cases.

Table 5: Multivariable logistic regression models of LRTI within Drakenstein using the number of built environment structure within a 500-meter radius.

	TOTAL UNMATCHED MODELS				DISTANCE MATCHED MODELS			
	BE AIC (1173)		BES AIC (1120)		BE AIC (1133.9)		BES AIC (1083.2)	
Covariates	**OR (95% CI)**	**P-value**	**aOR (95% CI)**	**P-value**	**OR (95% CI)**	**P-value**	**aOR (95% CI)**	**P-value**
Built Environment								
Number of agriculture / farming within 500 m radius	**0.39 (0.15; 0.89)**	**0.04***	0.40 (0.15; 0.93)	0.05	**0.33 (0.11; 0.82)**	**0.02***	**0.34 (0.11; 0.87)**	**0.03***
Number of educational facilities within 500 m radius	0.98 (0.88; 1.09)	0.67	0.96 (0.86; 1.07)	0.475	0.97 (0.87; 1.08)	0.54	0.96 (0.86; 1.07)	0.43
Number of formal liquor stores within 500 m radius	0.76 (0.52; 1.12)	0.17	0.74 (0.49; 1.11)	0.15	0.81 (0.55; 1.18)	0.27	0.78 (0.51; 1.17)	0.23
Number of informal liquor stores (shebeens) within 500 m radius	1.04 (0.94; 1.16)	0.46	1.08 (0.95; 1.22)	0.25	1.031 (0.90; 1.12)	0.91	1.05 (0.92; 1.19)	0.50
Number of healthcare facilities/providers within 500 m radius	**0.65 (0.45; 0.91)**	**0.01***	**0.60 (0.41; 0.87)**	**0.01***	**0.67 (0.46; 1.95)**	**0.03***	**0.61 (0.42; 0.90)**	**0.01***
Number of religious / places of worship within 500 m radius	0.98 (0.90; 1.07)	0.66	0.99 (0.90; 1.09)	0.90	0.97 (0.89; 1.06)	0.47	0.98 (0.89; 1.08)	0.67
Number of safety & security within 500 m radius	1.18 (0.81; 1.72)	0.38	1.18 (0.80; 1.73)	0.40	1.16 (0.80; 1.70)	0.43	1.17 (0.79; 1.73)	0.42
Number of spaza shops within 500 m radius	**1.03 (1.00; 1.05)**	**0.03***	**1.04 (1.01; 1.06)**	**0.01***	**1.04 (1.01; 1.07)**	**0.01***	**1.05 (1.02; 1.08)**	**0.01***
Number of malls/ supermarkets shops within 500 m radius	1.04 (0.85; 1.27)	0.70	1.05 (0.85; 1.29)	0.67	1.05 (0.86; 1.29)	0.63	1.05 (0.85; 1.30)	0.64
HIV exposed (YES)			**1.70 (1.14; 2.54)**	**0.01***			**1.74 (1.16; 2.63)**	**0.01***
Type of house								
Flat			reference category				reference category	
House			0.75 (0.47; 1.20)	0.24			0.81 (0.50; 1.31)	0.39
shack			0.86 (0.48; 1.55)	0.62			1.02 (0.56; 1.87)	0.94
Wendy house/ backyard dwelling			0.70 (0.41; 1.18)	0.18			0.73 (0.43; 1.25)	0.26
Number of people in household			1.01 (0.96; 1.08)	0.62			1.02 (0.96; 1.08)	0.61
SES quartile								
Moderately-high			reference category				reference category	
Lowest			0.12 (0.66; 2.20)	0.54			1.13 (0.61; 2.08)	0.69
Moderately-low			1.32 (0.80; 2.18)	0.28			1.18 (0.71; 1.98)	0.52
Highest			1.12 (0.70; 1.80)	0.64			1.06 (0.65; 1.71)	0.82
Household income								
<1000ZAR			reference category				reference category	
1000-5000ZAR			1.12 (0.82; 1.52)	0.49			1.12 (0.82; 1.54)	0.48

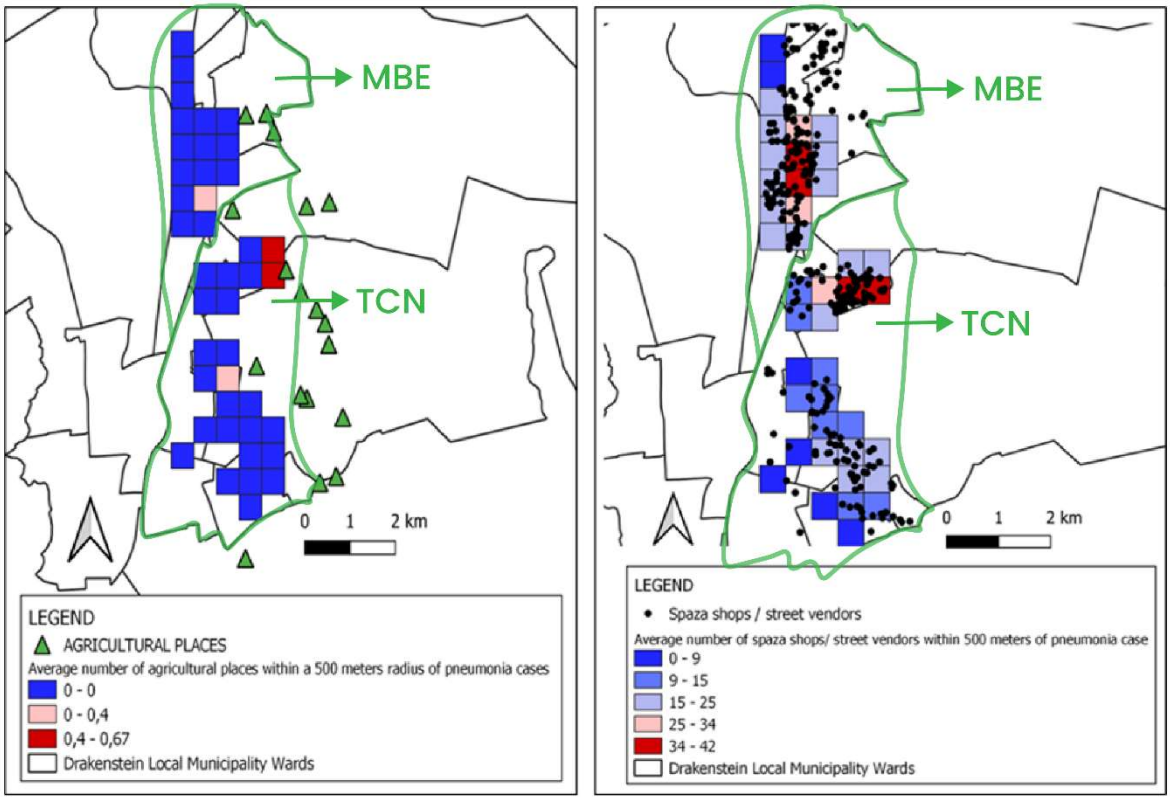
	>5000ZAR	1.31 (0.81; 2.12)	0.27	1.44 (0.88; 2.38)	0.15
Smoking					
	Active smoker	reference category		reference category	
	Non-smoker	**0.62 (0.40; 0.96)**	**0.03***	0.68 (0.44; 1.06)	0.09
	Passive smoker	0.76 (0.53; 1.07)	0.12	0.76 (0.53; 1.08)	0.12
Day care (YES)		1.35 (0.96; 1.90)	0.09	1.28 (0.90; 1.81)	0.17
Site (TC Newman)		1.00 (0.58; 1.73)	0.99	1.17 (0.67; 2.05)	0.59

Notes: LRTI = Lower Respiratory Tract Infection; BEM = Built Environment; BES = Built Environment and Socio-demographic risk factors; AIC = Akaike Information Criteria; OR = Odds Ratio; aOR = adjusted Odds Ratio; CI = confidence interval; * = Statistically Significant

In the distance matched BES model, the increased number of agriculture (farming places) is associated with 0.34 times the odds of being a LRTI case [aOR = 0.34, 95%CI, 0.11; 0.82 p = 0.03]. However, this is not statistically significant in the total unmatched model. In the total unmatched model, the increased number of healthcare facilities or providers is associated with 0.60 times the odds of being a LRTI [OR = 0.60, 95%CI (0.41; 0.87), p = 0.01]. Similarly, in the distance matched BES models, the increased healthcare facilities or providers is associated with 0.61 times the odds of being a LRTI case [OR = 0.61, 95%CI, 0.42; 0.90 p = 0.01]. In the total unmatched BES model, the increased number of spaza shops is associated with 4% increase in the odds of being a LRTI case [OR = 1.04, 95%CI (1.01; 1.06), p = 0.01]. Similarly, in the distance matched BES model, the increased number of spaza shops is associated with 5% increase in the odds of being a LRTI case [OR = 1.05, 95%CI (1.02; 1.08), p = 0.01].

Maternal HIV status and maternal smoking were statistically significant among the sociodemographic risk factors included in the models. The odds of an LRTI are 1.7 times higher if you are exposed to a mother living with HIV compared to a mother living without HIV in both the total unmatched BES model and the distance matched BES model [OR = 1.70, 95%CI (1.14; 2.54), p = 0.01], [OR = 1.74, 95%CI (1.16; 2.63), p = 0.01]. The risks of an LRTI are 1.6 times greater in participants whose mothers are active smokers compared to participants whose mothers are non-smokers in the total unmatched BES model [OR = 0.62, 95%CI (0.40; 0.96), p = 0.03]. When cases and controls are matched by distance, maternal smoking is no longer significant (Table 5).

In figure 4A, the number of agricultural (farming places) within a 500-meter radius of an LRTI cases is >1, on average and the cases are further from agricultural (farming places). In figure 4B, on average, there are density hotspots of LRTI cases in Mbekweni that have 25-42 spaza shops or street vendors within a 500-meter radius. The highest density in the TC Newman region is 15-25 spaza shops or street vendors within a 500-meter radius.



A **B**

Figure 4: Maps of spatial distribution of hotspots and coldspots of the average number of agricultural places (A) and spaza shops or street vendors (B) within 500-metres of the residential addresses of LRTI cases.

Discussion

Child lung health remains a public health problem. Research has suggested that assessing child lung health outcomes and their risk factors geospatially is invaluable as it allows for better, more precise targeting of interventions. The aim of this study was to assess if there is any relationship between the social and environmental risk factors (built environment) and spatial distribution of child lung health outcomes in Drakenstein. This was done through a distance-matched case-control sampling method to account for spatial dependence. When comparing the unmatched and distance matched models for both outcomes, the distance matched models had the lowest Akaike Information Criteria (AIC) values compared to the unmatched models (34). This indicates that the distance matched models are superior to the unmatched models suggesting that accounting for spatial dependence was the best modelling approach.

In this peri-urban, low-income area of South Africa, LRTI and LTBI were associated with a few of the built environment types, after adjusting for clinical and socio-demographic risk factors. Some of the social risk factors that are commonly found to be associated with lung diseases (number of people living in a household, house type and maternal smoking) had no effect in this analysis after adjusting for spatial distribution. This is inconsistent with existing literature which states that type of housing is often associated with risk of respiratory infections (35). This is as the study by Aftab et al. associated the risk with the type of building material used which was not included in our study. Similarly, with regards to number of people living in a household, overcrowded households put young children at risk of developing an acute lower respiratory infection (36). A number of well-known exposures, including maternal HIV infection and household income was also found to be linked to LRTI and LTBI (37). Children who are exposed to HIV have an increased chance of LRTI. This is supported by literature which states that the HIV-exposed infants would be at a higher risk of LRTI compared to the unexposed. Furthermore, a South African study found that HIV exposed and uninfected infants were 40% more likely to be hospitalized with LRTI, have longer hospital stays and needed mechanical ventilation than unexposed infants (38). Household income, an indicator of SES, is no longer significant when conducting the distance matched model. A possible explanation for this observation is that cases and controls were matched within 1km of each other, and often people residing in the same neighbourhood are of the same SES. When comparing the two study sites, TC Newman and Mbekweni, we observe a difference in the odds of developing LTBI as there is twice as many cases in TC Newman compared to Mbekweni (Table S1).

Public health is impacted by the built environment (39). The physical environment of a child can either cause or prevent illness. Thus, the environment is critical for children's optimal health and development. Much less focus has been given to the potential effects of the built environment, such as substandard housing and poor land use, as well as community planning. In reality, children spend more time indoors and in neighbourhoods than they do in the natural environment (40). The built environment can also be utilised as a marker of a set of social conditions and community circumstances that often go hand in hand. An article by Chuang et al., found that lower community SES and higher convenience store concentration, as assessed by density and distance, were both significantly linked to greater levels of individual smoking (41). In addition, fast food restaurants and convenience stores may outnumber supermarkets where people can purchase more nutritious food (39). Our findings are consistent with this statement as spaza shop density in Drakenstein is much higher in comparison to the malls and supermarkets. Furthermore, fast food restaurants and convenience stores may be an avenue for faster spread of LRTI and LTBI, especially if they are congested. The density of formal and informal liquor stores reflects a community's SES in the same manner as the density of spaza shops does.

Xu et al. found that the built environment was associated with lung health and posit that urban redevelopment can help to reduce the risk of respiratory disease (23). There are several interconnected aspects that contribute to the transmission of LTBI or LRTI, such as overcrowding and unemployment. Furthermore, Munch et.al has suggested that shebeens were clustered in areas of high unemployment (14). In our study setting of high unemployment, informal liquors stores may drive TB transmission. This is as our study found that close proximity to informal and formal liquor stores is associated with LTBI. Prior literature has found that TB can be connected with socio-economic status, even more so in urban areas, and among the homeless and substance abusers (42). This suggests that TB reflects the fundamental socio-economic problems of inequality and poverty (14). Low socioeconomic status residents appear to face a twofold disadvantage since they lack both private and public resources, such as green areas (43).

As such, increasing the number of agriculture (farming places) decreased the odds of LRTI. This is consistent with existing literature. Recent studies showed that schoolchildren with higher levels of green space nearby had improved lung function (44). However, these improvements could be related to a decrease in air pollution when in the green spaces. One study has demonstrated that the more time one spends in green spaces, the better their lung function (45). While increasing the number of agriculture (farming places) would be a challenge in the Drakenstein region due to space, increasing

the number of urban greenspaces could be plausible. Urban greenspaces are defined as greenness or green infrastructure and can include parks and trees, street gardens, and a variety of other vegetation within an urban setting (46). Like agricultural spaces, health care facilities also impacted the odds of LRTI and LTBI. Our results indicated that an increase in the number of healthcare facilities decreased the odds of LRTI. According to a study in West Africa, children under 5 visited primary healthcare facilities less frequently the further they travelled (47). In contrast, Robsky et al. did not find an association between poor treatment outcomes once treatment has begun as they were prepared to go a greater distance be more motivated to continue with treatment (48). However, distance to healthcare facilities should not be a barrier to access. This can be mitigated by expanding the number of facilities or concentrating on community engagement in more remote communities, which may strengthen early childhood healthcare in similar settings (47).

A strength of this study is that the DCHS study team surveyed respiratory disease symptoms, including tuberculosis, prospectively and actively, which may enhance the precision of episode prevalence (31). In addition, the high sample size of comparable lung function acquired by the same investigators using the same methods of testing is one of the study's strengths. Furthermore, our study took spatial dependence into account which was the drawback Middlekoop et al's study and could bias their results (6). Similar to the parent study, this study's limitation includes the possibility of a lack of generalizability for different settings with varied exposures; however, most of these exposures are prevalent in LMIC contexts (49). The primary disadvantage of nested case-control studies is that not all relevant risk factors are likely to have been captured (50). The main limitation of this study was the lack of built environment features. A study conducted previously included moisture and dust inside the homes, housing materials used, closeness to contamination sources, level of nearby traffic, bedroom sizes and space per occupant as additional variables (23). Further research into the built environment should be considered with these variables to assess the effect of possible risk factors on LRTI and LTBI. In addition, the exposure was assessed based on location of residence and likelihood of transmission occurring in the home, while exposure to infected individuals may occur elsewhere (15). While the relationships between lung health and socio-economic status and built environment variables were consistent with literature, it was insufficient in demonstrating how these variables could justify variations in disease distributions.

Conclusions

To conclude, in Drakenstein, the associations between some of the social and environmental risk factors (built environment) and the spatial distribution of child lung health outcomes are present after distance-matching cases and controls and adjusting for covariates. The observed associations are indicative of the SES and community circumstances. Thus, creation of more green spaces or designing of buildings with better air flow would be useful and actionable by the government and local authorities as potential strategies to reduce the burden of childhood LRTI and LTBI in LMICs.

References

1. Zar HJ, Barnett W, Stadler A, Gardner-Lubbe S, Myer L, Nicol MP. Aetiology of childhood pneumonia in a well vaccinated South African birth cohort: A nested case-control study of the Drakenstein Child Health Study. Lancet Respir Med. 2016;4(6):463–72.
2. Pillay Y, Barron P. On the path to reach the SDG targets: Decreasing maternal and child mortality in South Africa. *South African Medical Journal*. 2018 May 7;108(3):2-3.
3. Bamford L, Barron P, Kauchali S, Dlamini N, Paeds M. Inpatient case fatality rates improvements in children under 5 : Diarrhoeal disease , pneumonia and severe acute malnutrition. 2018;108(March):33–7.
4. Shi T, McAllister DA, O'Brien KL, Simoes EAF, Madhi SA, Gessner BD, et al. Global, regional, and national disease burden estimates of acute lower respiratory infections due to respiratory syncytial virus in young children in 2015: a systematic review and modelling study. The Lancet. 2017 Sep 2;390(10098):946–58.
5. Beamer PI, Lothrop N, Lu Z, Ascher R, Ernst K, Stern DA, et al. Spatial clusters of child lower respiratory illnesses associated with community-level risk factors. Pediatr Pulmonol. 2016;51(6):633–42.
6. Middelkoop K, Bekker LG, Morrow C, Zwane E, Wood R. Childhood tuberculosis infection and disease: A spatial and temporal transmission analysis in a South African township. South African Medical Journal. 2009;99(10):738–43.
7. Congdon P. Spatiotemporal frameworks for infectious disease diffusion and epidemiology. Int J Environ Res Public Health. 2016;13(12):15–8.
8. Aftab A, Noor A, Aslam M. Housing quality and its impact on Acute Respiratory Infection (ARI) symptoms among children in Punjab, Pakistan. PLOS Global Public Health. 2022 Sep 21;2(9):e0000949.
9. Wood R, Middelkoop K, Myer L, Grant AD, Whitelaw A, Lawn SD, et al. Undiagnosed tuberculosis in a community with high HIV prevalence: Implications for tuberculosis control. Am J Respir Crit Care Med. 2007;175(1):87–93.
10. Hasan M, Tasfina S, Haque SMR, Saif-Ur-Rahman KM, Khalequzzaman M, Bari W, et al. Association of biomass fuel smoke with respiratory symptoms among children under 5 years of age in urban areas: Results from Bangladesh Urban Health Survey, 2013. Environ Health Prev Med. 2019 Nov 27;24(1).

11. Adesanya OA, Chiao C. A multilevel analysis of lifestyle variations in symptoms of acute respiratory infection among young children under five in Nigeria. BMC Public Health. 2016 Aug 25;16(1).

12. El Sawy IH, Nasr FM, Mowafy EW, Sharaki OA, Abdel Bakey AM. Passive smoking and lower respiratory tract illnesses in children. *EMHJ-Eastern Mediterranean Health Journal*, 3 (3), 425-434, 1997. 1997.

13. Ho NT, Thompson C, Nhan LNT, Van Hoang MT, Dung NT, Tran My P, et al. Retrospective analysis assessing the spatial and temporal distribution of paediatric acute respiratory tract infections in Ho Chi Minh City, Vietnam. BMJ Open. 2018;8(1):1–9.

14. Webster T, Vieira V, Weinberg J, Aschengrau A. Method for mapping population-based case-control studies: An application using generalized additive models. Int J Health Geogr. 2006;5:1–10.

15. Munch Z, Van Lill SW, Booysen CN, Zietsman HL, Enarson DA, Beyers N. Tuberculosis transmission patterns in a high-incidence area: a spatial analysis. *International journal of tuberculosis and lung disease*. 2003 Mar 7;7(3):271-7.

16. Alene KA, Viney K, McBryde ES, Clements ACA. Spatiotemporal transmission and socio-climatic factors related to paediatric tuberculosis in north-Western Ethiopia. Geospat Health. 2017;12(2):342–50.

17. Beck AF, Florin TA, Campanella S, Shah SS. Geographic variation in hospitalization for lower respiratory tract infections across one county. JAMA Pediatr. 2015;169(9):846–54.

18. Thomas DSK, Anthamatten P, Root ED, Lucero M, Nohynek H, Tallo V, et al. Disease mapping for informing targeted health interventions: childhood pneumonia in Bohol, Philippines. Tropical Medicine and International Health. 2015;20(11):1525–33.

19. Scott JAG, Wonodi C, Moïsi JC, Deloria-Knoll M, Deluca AN, Karron RA, et al. The definition of pneumonia, the assessment of severity, and clinical standardization in the pneumonia etiology research for child health study. Clinical Infectious Diseases. 2012 Apr 1;54(SUPPL. 2).

20. Root ED, Lucero M, Nohynek H, Stubbs R, Tallo V, Lupisan SP, et al. Distance to health services modifies the effect of an 11-valent pneumococcal vaccine on pneumonia risk among children less than 2 years of age in Bohol, Philippines. Int J Epidemiol. 2017;46(2):706–16.

21. Shenoi S V, Brooks RP, Bhushan A, Brust JCM, Zelterman D, Deng Y, et al. Natural ventilation reduces high TB transmission risk in traditional homes in rural KwaZulu-Natal, South Africa. BMC Infect Dis. 2013;13:300.

22. Handy SL, Boarnet MG, Ewing R, Killingsworth RE. How the built environment affects physical activity: Views from urban planning. Am J Prev Med. 2002;23(2 SUPPL. 1):64–73.

23. Xu W, Zhao X, Wang L. Impact of Built Environment on Respiratory Health: An Empirical Study. Nano Life [Internet]. 2018 Feb 13;08(02):1840001. Available from: https://doi.org/10.1142/S1793984418400019

24. Sterling TR, Njie G, Zenner D, Cohn DL, Reves R, Ahmed A, Menzies D, Horsburgh Jr CR, Crane CM, Burgos M, LoBue P. Guidelines for the treatment of latent tuberculosis infection: recommendations from the National Tuberculosis Controllers Association and CDC, 2020. *American Journal of Transplantation*. 2020 Apr 1;20(4):1196-206.

25. Ortblad KF, Salomon JA, Bärnighausen T, Atun R. Stopping tuberculosis: A biosocial model for sustainable development. The Lancet. 2015;386(10010):2354–62.

26. Wang L, Sun W, Zhou K, Zhang M, Bao P. Spatial analysis of built environment risk for respiratory health and its implication for urban planning: A case study of shanghai. Int J Environ Res Public Health. 2019;16(8).

27. Massyn N, Padarath A, Peer N. District health barometer 2016/17. Health Systems Trust; 2017.

28. le Roux DM, Myer L, Nicol MP, Zar HJ. Incidence of childhood pneumonia: Facility-based surveillance estimate compared to measured incidence in a South African birth cohort study. BMJ Open. 2015;5(12):1–6.

29. Martinez L, le Roux DM, Barnett W, Stadler A, Nicol MP, Zar HJ. Tuberculin skin test conversion and primary progressive tuberculosis disease in the first 5 years of life: a birth cohort study from Cape Town, South Africa. Lancet Child Adolesc Health [Internet]. 2018;2(1):46–55. Available from: http://dx.doi.org/10.1016/S2352-4642(17)30149-9

30. World Health Organization. "Integrated Management of Childhood Illness: distance learning course." (2014). Available from: https://apps.who.int/iris/handle/10665/104772

31. MacGinty RP, Lesosky M, Barnett W, Stein DJ, Zar HJ. Associations between maternal mental health and early child wheezing in a South African birth cohort. Pediatr Pulmonol. 2018 Jun 1;53(6):741–54.

32. Myer L, Stein DJ, Grimsrud A, Seedat S, Williams DR. Social determinants of psychological distress in a nationally-representative sample of South African adults. Soc Sci Med. 2008 Apr;66(8):1828–40.

33. Fletcher RS, Fisher DK. Spatial Analysis of Soybean Plant Height and Plant Canopy Temperature Measured with On-the-Go Tractor Mounted Sensors. Agricultural Sciences. 2019;10(11):1486–96.

34. Cavanaugh JE, Neath AA. The Akaike information criterion: Background, derivation, properties, application, interpretation, and refinements. Vol. 11, Wiley Interdisciplinary Reviews: Computational Statistics. Wiley-Blackwell; 2019.

35. Aftab A, Noor A, Aslam M. Housing quality and its impact on Acute Respiratory Infection (ARI) symptoms among children in Punjab, Pakistan. PLOS Global Public Health. 2022 Sep 21;2(9):e0000949.

36. Regina M, Cardoso A, Cousens SN, Fernando De Góes Siqueira L, Alves FM, Antônio V D'angelo L. Crowding: risk factor or protective factor for lower respiratory disease in young children? [Internet]. 2004. Available from: http://www.biomedcentral.com/1471-2458/4/19

37. Jackson S, Mathews KH, Pulanić D, Falconer R, Rudan I, Campbell H, et al. Risk factors for severe acute lower respiratory infections in children - a systematic review and meta-analysis. Croat Med J. 2013 Apr;54(2):110–21.

38. Slogrove AL, Frigati L, Gray DM. Maternal HIV and Paediatric Lung Health. Vol. 21, Paediatric Respiratory Reviews. W.B. Saunders Ltd; 2017. p. 47–53.

39. Perdue WC, Stone LA, Gostin LO. The Legal Perspective The Built Environment and Its Relationship to the Public's Health: The Legal Framework. Vol. 93, American Journal of Public Health. 2003.

40. Cummins SK, Jackson RJ. The Built Environment and Children's Health. Pediatr Clin North Am [Internet]. 2001;48(5):1241–52. Available from: https://www.sciencedirect.com/science/article/pii/S0031395505703722

41. Chuang YC, Cubbin C, Ahn D, Winkleby MA. Effects of neighbourhood socioeconomic status and convenience store concentration on individual level smoking. J Epidemiol Community Health (1978). 2005 Jul;59(7):568–73.

42. Hurtig AK, Porter JDH, Ogden JA. Tuberculosis control and directly observed therapy from the public health/human rights perspective. Vol. 3, INT J TUBERC LUNG DIS. 1999.

43. Hoffimann E, Barros H, Ribeiro AI. Socioeconomic inequalities in green space quality and Accessibility—Evidence from a Southern European city. Int J Environ Res Public Health. 2017 Aug 15;14(8).

44. Yu H, Hu LW, Zhou Y, Qian Z, Schootman M, LeBaige MH, et al. Association between eye-level greenness and lung function in urban Chinese children. Environ Res. 2021 Nov 1;202:111641.

45. Mueller W, Milner J, Loh M, Vardoulakis S, Wilkinson P. Exposure to urban greenspace and pathways to respiratory health: An exploratory systematic review. Science of The Total Environment [Internet]. 2022;829:154447. Available from: https://www.sciencedirect.com/science/article/pii/S0048969722015406

46. Mueller W, Milner J, Loh M, Vardoulakis S, Wilkinson P. Exposure to urban greenspace and pathways to respiratory health: An exploratory systematic review. Vol. 829, Science of the Total Environment. Elsevier B.V.; 2022.

47. Oldenburg CE, Sié A, Ouattara M, Bountogo M, Boudo V, Kouanda I, et al. Distance to primary care facilities and healthcare utilization for preschool children in rural northwestern Burkina Faso: results from a surveillance cohort. BMC Health Serv Res [Internet]. 2021 Mar;21(1):212. Available from: https://europepmc.org/articles/PMC7941928

48. Robsky KO, Robsky KO, Hughes S, Kityamuwesi A, Kendall EA, Kitonsa PJ, et al. Is distance associated with tuberculosis treatment outcomes? A retrospective cohort study in Kampala, Uganda. BMC Infect Dis. 2020 Jun 11;20(1).

49. Gray DM, Turkovic L, Willemse L, Visagie A, Vanker A, Stein DJ, et al. Lung function in African infants in the Drakenstein child health study impact of lower respiratory tract illness. Am J Respir Crit Care Med. 2017 Jan 15;195(2):212–20.

50. Sedgwick P. Nested case-control studies: Advantages and disadvantages. Vol. 348, BMJ (Online). 2014.

Funding statement:

This work was supported by The National Research Foundation (NRF) grant number [MND190729460466].

Informed Consent Statement:

Informed consent was obtained from all subjects involved in the study.

Institutional Review Board Statement:

The study was conducted according to the guidelines of the Declaration of Helsinki, and approved by the Institutional Review Board (or Ethics Committee) of The University of Cape Town for the parent study (HREC Ref number: 401/2009) on 11/10/2021 (Appendix C1) and the current study (HREC Ref number: 758/2021) on 16/11/2021.

Conflicts of Interest:

The authors declare no conflict of interest.

Part C: Appendices

A. Supplementary tables and figures included in the manuscript

Table S 1: Descriptive table of sociodemographic characteristics of child- mother pairs enrolled in the Drakenstein Child Health Study (DCHS) cohort by clinic, Mbekweni and TC Newman.

Variable: n (%)	N	Black African (Mbekweni)	Coloured (TC Newman)	Total
Number of Participants (%)		421 (49.9)	423 (50.1)	844 (100)
HIV Status	844	161 (38)	12 (3)	173 (20)
Sex of child (male)		202 (48)	230 (54)	432 (51)
Currently employed	844	99 (24)	120 (28)	219 (26)
Highest Education Level attained				
Primary	844	31 (7)	33 (8)	64 (8)
Some secondary		231 (55)	232 (55)	463 (55)
Completed secondary		123 (29)	144 (34)	267 (32)
Any tertiary		33 (8)	14 (3)	47 (6)
Average household income per month				
<1000ZAR	844	175 (42)	142 (34)	317 (38)
1000-5000ZAR		203 (48)	217 (51)	420 (50)
>5000ZAR		40 (10)	64 (15)	104 (12)
SES Quartiles				
Lowest SES	844	115 (27)	80 (19)	195 (23)
Low-moderate SES		121 (29)	107 (25)	228 (27)
Moderate-high SES		97 (23)	123 (29)	220 (26)
High SES		88 (21)	113 (27)	201 (24)
Creche / day-care at birth	844	89 (21)	105(25)	194 (23)
Antenatal smoking				
Nonsmoker	844	140 (33)	39 (9)	179 (21)
Passive smoker		200 (47)	146 (35)	346 (41)
Active smoker		63 (15)	229 (54)	292 (35)
Average nr of people in household Mean (SD)	844	3 (1.53)	3 (1.82)	3 (1.7)
Housing type				
house	844	263 (62)	139 (33)	402 (48)
Flat		1 (0)	153 (36)	154 (18)
shack		144 (34)	22 (5)	166 (20)
Wendy house/ backyard dwelling		9 (2)	107 (25)	116 (14)
Lung disease	844			
LTBI		24 (6)	51 (12)	75 (9)
LRTI		208 (49)	201 (48)	409 (48)

Notes: All results show frequencies (percentages) or mean (Standard deviation). Self-reported smoking status at the baseline study visit. SES = socio-economic status; ZAR= South African Rand.

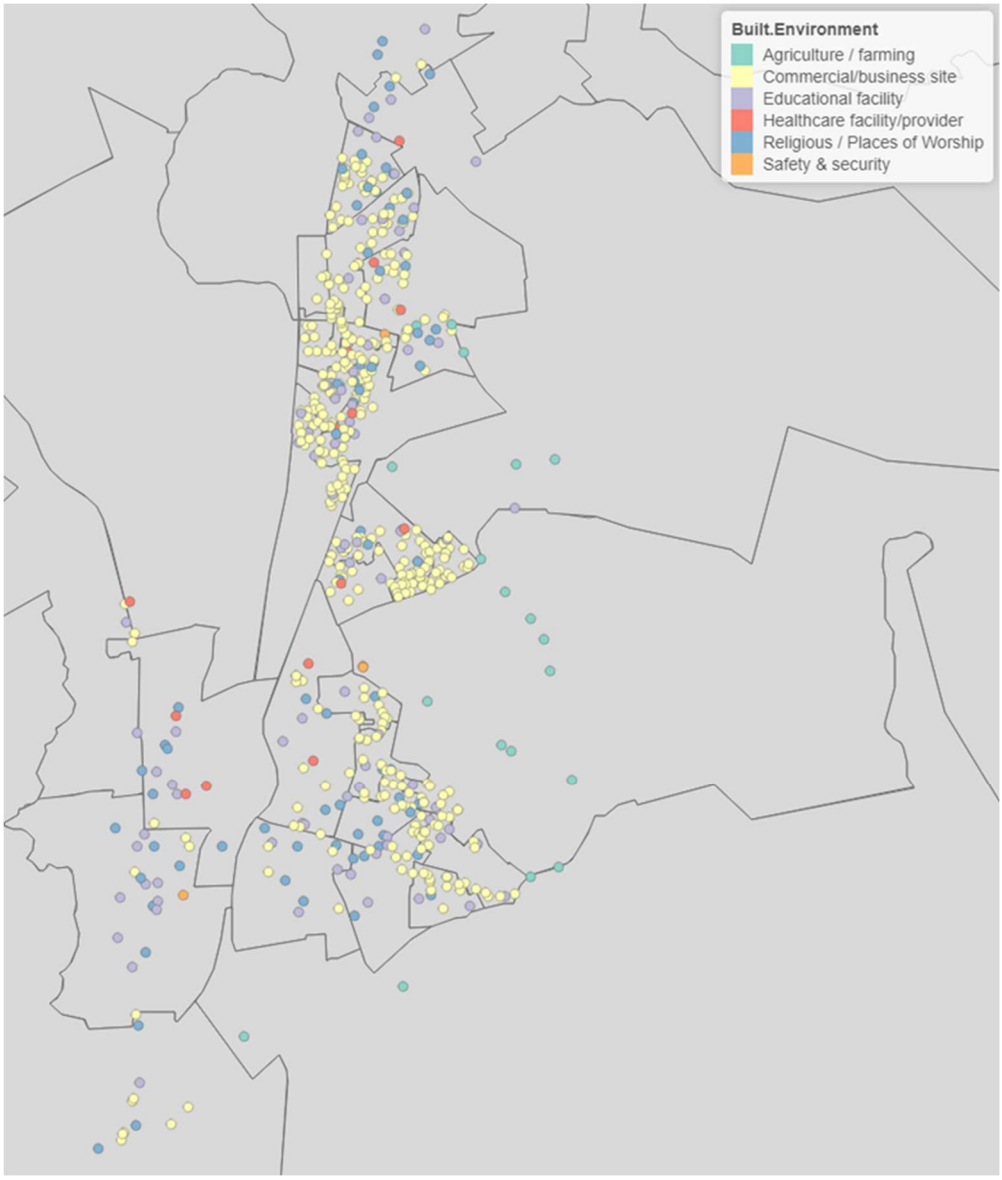


Figure S 1: Leaflet plot of the type of built environment within Drakenstein, Western Cape, South Africa.

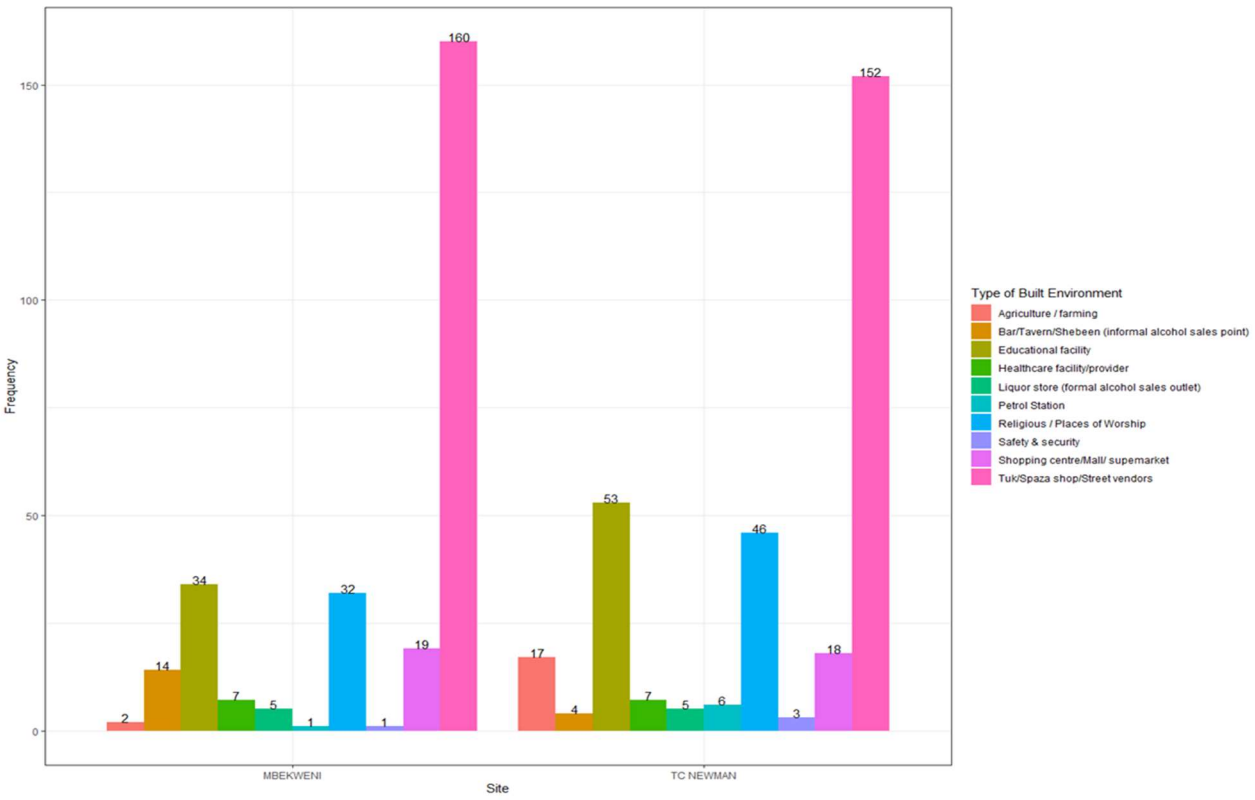


Figure S 2: Bar graph of the frequency of the type of built environment obtained via GPS in Drakenstein by site: Mbekweni and TC Newman.

Figure S 3: Density Heatmap of LTBI (A) and LRTI (B) cases and matched controls within the Mbekweni, Drakenstein subsetted region.

Figure S 4: Density Heatmap of LTBI (A) and LRTI (B) cases and controls within the TC Newman, Drakenstein subsetted region.

B. Data Collection Tool used in Parent Study

B1. DRAKENSTEIN CHILD HEALTH STUDY

Standard Operating Procedure (SOP)

SOP for Environmental and Geospatial Home Visits

First Review Date:

Version	Reviewed by/date	Description of changes
1.1	January 2020	Removed instructions on phone triangulation, update Garmin instructions, data export and QC procedures
1.2	August 2020	Add metadata collection and remove household data collection

Drafted by: Muhammed Adil Khan, edited Maresa Botha, edited Demi Meyer

Author Signature:

Reviewed by: Date: June 2019

Approved by: Tiffany Burd

PI Signature: Date of Approval:

Training Implications

Admin Assistant		Lab Assistant	
Study Nurse Coordinator	**X**	Lab Manager	
Data Administrator		Medical Officer	**X**
Data Capturer		Nursing Assistant	
Data Manager		Professional Nurse	
Driver	**X**	Project Manager/Coordinator	**X**
Fieldworker	**X**	Research Assistant	**X**

TABLE OF CONTENTS:

6.0 Documentation

7.0 Training

8.0 References

1. Purpose

The purpose of this SOP is to describe the Home Visits to be performed and the procedures surrounding the devices to be used for measuring the outside of the participants' homes and recording the GPS coordinates.

2. Background

One home visit is planned for all participants enrolled on DCHS at the year 6 timepoint.

During the home visit, a case report form (CRF) will be administered, measurements of the home will be taken and the GPS co-ordinates of the participant's home will be recorded.

3. Scope

This SOP will be used by DCHS staff members (SM) involved in performing or assisting with home visits as delegated by the Project Manager (PM) / Medical Officer (MO).

4. Equipment

a. CRF

b. Laser measuring device

c. GPS device

5. Procedure

a. Pre-arrange a suitable date and time to perform the home visit.

b. Ask permission to visit the home. Explain the purpose of the home visit and what you will be doing while in the home. Any questions about the home visit should be answered prior to beginning.
c. Confirm that the child lives at this address. If the child does not live at this address**, do not continue with the CRF.**

Identify the main person who will be available to answer questions regarding the home environment and their relationship to the child. This person needs to be someone that lives in the home.

Field worker does telephonic interview

To measure GPS coordinates:

Stand outside the participants home closest to the front door while remaining in the car. If there is a covering over the front door, stand further away so that there is no obstruction between the GPS device and the sky.

Switch on the Garmin and go to "Mark Waypoint". Enter participant PID (e.g. 1234). Click "Done" to start measurement

Wait 60 seconds in the same position for the Garmin GPS to triangulate before tagging GPS coordinates on the Garmin GPS device. Click "Done" to end measurement. Go to "Waypoint Manager". Check that PID is less than 3m away from current point. If the PID is more than 3m, repeat and rename the PID (e.g. 1234NEW)

Write down the coordinates of the printed CRF and if there are any delays or issues with the Garmin, write them down too.

E.g. weather overcast therefore, it took longer for the GPS to triangulate or it's the first recording of the day.

6. Documentation and download of data

Home visits will be documented on REDCap in the appropriate CRF.

Garmin data upload: Plug in the Garmin to laptop. Open MyMaps in Google Drive. Click on "Add layer" on MyMaps. Upload Garmin data according to date. Copy coordinates from MyMaps according to PID. Go to CRF of participant on REDCap. Paste coordinates from MyMaps into REDCap. Click complete and Save CRF. On MyMaps, drag and drop PID into TC or MBE layers. Delete the newly added "layer" once all pins have been dragged and dropped.

7. Training

Training in this SOP will be documented per the DCHS Study Training Plan.

8. References

- SOP for Environmental home visit, DCHS, version 1.3, July 2013
- SOP for Location Data Collection, SS4ME project, version 1.1, 19 Nov 2018

Environmental Postnatal Home Visit Year 6

Child PID	_____

Only complete this CRF if you are at the home of the child in the study

Is this the home of the child in the study?	○ Yes ○ No
Date of visit	_____
Time at the start of the home visit	_____ (Please use 24 hour clock)
Latitude	_____ (Using GPS device)
Longitude	_____ (Using GPS device)

B2. Built environment/Amenities GPS coordinates collection Form

Confidential

Built Environment

Record ID	_____ (Display)
Form_ID	_____
Zone	◯ Zone 1 ◯ Zone 2 ◯ Zone 3 ◯ Zone 4 ◯ Zone 5 ◯ Zone 6 ◯ Zone 7 ◯ Zone 8 ◯ Zone 9 ◯ Zone 10 ◯ Zone 11 ◯ Zone 12 ◯ Zone 13 ◯ Zone 14 ◯ Zone 15 ◯ Zone 16 ◯ Zone 17
Built Environment	◯ Healthcare facility/provider ◯ Educational facility ◯ Commercial/business site ◯ Public transport ◯ Agriculture / farming ◯ Services & sanitation ◯ Safety & security ◯ Religious / Places of Worship
Type of Healthcare facility/provider:	◯ Clinic (Public) ◯ Private GP ◯ Hospital (Public) ◯ Hospital (Private) ◯ Traditional Healer
Type of Educational facility:	◯ School (Primary) ◯ School (Secondary) ◯ Day-care/Creche (formal) ◯ Tertiary Institute
Type of Commercial/business site:	◯ Shopping centre/Mall ◯ Supermarket (e.g. Pick n pay) ◯ Tuk/Spaza shop/Street vendors ◯ Liquor store (formal alcohol sales outlet) ◯ Bar/Tavern/Shebeen (informal alcohol sales point) ◯ Petrol Station
Type of Public transport:	◯ Taxi / Bus rank ◯ Train station
Type of Agriculture / farming:	◯ Commercial Farm ◯ Informal farming

Type of Services & sanitation:	◯ Communal water point/tap ◯ Communal toilets ◯ Rubbish / Refuse Dump
Type of Safety & Security:	◯ Police station ◯ Community safety kiosks
Type of Religious / Places of Worship:	◯ Formal buildings (Churches, Mosques etc) ◯ Informal & outdoor worship areas
Name of built environment (if applicable)	_____ (If not applicable, Type "N/A")
For Street Vendors: what variety of food is for sale? (Select all that apply)	☐ Fruits ☐ Vegetables ☐ Sugary drinks ☐ Sweets/Candy ☐ Amakipkip
Latitude	_____ (Using GPS device)
Longitude	_____ (Using GPS device)
Date CRF completed	_____

C3. Consent Form

DRAKENSTEIN CHILD HEALTH STUDY

CONSENT AND INFORMATION SHEET FOR MOTHERS – MAIN COHORT

July 2020

CONSENT FORM AT YEAR 5

You and your child are invited to continue to take part in a study that is being done in the Drakenstein sub-district, in collaboration with the Universities of Cape Town and Stellenbosch. We would like to thank you and your child for taking part in this study, we hope to impact child health and your participation will help us achieve that. The following information describes the study and you and your child's role for the next year. Please read this carefully and feel free to ask any questions.

Why is this study being done?

Lung infections and chest problems are common in young children. This study is being done to find out the effect of chest infections in the early years of life on the development of lung disease in children. The study will also look at a number of other factors that may affect your child's health.

You and your child will continue to attend occasional scheduled visits at your primary health care clinic and at Paarl Hospital. During these visits, we will assess the health of you and your child by using questionnaires and doing tests. Should your child get sick with a chest infection, then he/ she will be carefully investigated to try and find out the cause of this infection. This study will help us to better understand why children get chest illness and may help to improve child health.

Cognition and social emotional development are important parts of a child's overall health. By cognition we mean the way a child thinks, learns, plans and pays attention. Social emotional and social cognitive development refer to how a child learns to understand their own and others' emotions, and to reason about what other people think, believe and feel. We want to see how these factors develop over time. This study will help us understand how various other factors we

look at in the main study impact on this development, and also how this development affects overall health.

What must I do if I agree to continue in the study?

If you agree to continue in this study, we will follow you and your child regularly to assess his/ her health. We will see you and your child at Paarl hospital or your primary health care clinic when your child is about 5 years of age and again at about 6 years of age. We will also schedule a visit at 5 ½ years and a single visit between these at the primary health care clinic. We will ask you some questions about your child's health, nutrition, growth and development, and any chest illnesses. We will do regular tests to watch these.

At study visits in the next year, you will be asked some questions about you and your child's health. Your child will be examined. Tests will be done on you and your child to assess whether there is any chest problem. The tests that may be done on your child are:

1. Blood tests - these will be to test for allergies or blood problems.
2. Throat swab to test for germs (oropharyngeal swab).
3. Nasal swab to check for germs including the coronavirus
4. Saliva will be collected to check for germs which may cause pneumonia
5. A skin test for tuberculosis infection.
6. A urine test for smoke exposure.
7. A stool test to check what germs are in the stool.
8. At 5 years and 6 years of age a breathing test will be done, to measure the air moving in and out of his/her lungs.
9. A skin test if your child has a rash
10. A hearing test to see if your child may have any hearing problems
11. Teeth that fall out naturally or are extracted during normal dental care will be collected and analysed for environmental exposures.
12. Vision test to check your child's eyes.

The tests YOU may be asked to complete are:

1. Questionnaires about your physical activity, socio-economic status and your levels of emotional distress, stress, life events, social support, aggression, emotional regulation, exposure to community violence, assessment of maternal parenting styles, your and your child's emotional style and empathy, resilience as well as behaviours and drug and alcohol use. If a mental health condition or abuse is suspected, you will be referred to the appropriate local services.
2. Neurocognitive and socioemotional assessment at 5.5 and 5 years. We will ask you to answer some questions about your child's feelings and behaviors, and about your own feelings and behaviors. While you are doing this, we will be doing some tasks, games and puzzles with your child.
3. Blood tests – We may ask you to take a blood test for diabetes, fats in the blood or for COVID.

We will only share your test results with primary health care staff if it indicates that you or your child require treatment or further follow up. For some assessments, study staff may follow up with you and provide you with information on where you can seek help, if necessary.

Should your child get sick with a chest infection, wheeze or asthmathen additional tests will be done to try and find out the cause of your child's illness. The tests that will be done will depend on how sick your child is and what the illness is. These tests may include:

1. Blood tests to test for infections, at the time of the illness, and again 4-6 weeks afterwards
2. A test of the mucus from the nose or mouth (nasopharyngeal or oropharyngeal swab) to test for infection
3 A skin test for tuberculosis infection.
4 A test of the mucus from the lungs (induced sputum test) for chest infection.
5 A urine test for smoke exposure
6 Chest X-ray
7 Breathing test

8 A ultrasound test of the lungs

9 A stool test to check what germs are in the stool.

If your child is enrolled in the study and is admitted to hospital, he/she will be followed up in hospital by a member of the study team. The study member will ask you questions about your child's illness, and some tests may be done, including a nose swab and an induced sputum. All of these tests are usual for investigating the cause of pneumonia.

What are the benefits of my child being in the study?

You and your child will be closely followed for the first few years of your child's life. Any medical illness or problem should be found soon after it develops. Your child's growth and development will be carefully followed. If an illness or problem is found then your child will be promptly referred for treatment. If your child gets sick you will be able to take him/ her to your usual health facility, where additional tests to find out the cause of your child's illness may be done, depending on how sick your child is. If your child requires hospitalisation, then he/ she will be hospitalised at Paarl hospital as is usually done. If your child is hospitalised, then one of the study staff will see your child in hospital and additional investigations may be done to try and find out the cause of the illness. A specific focus of the study includes pneumonia, wheeze and asthma. If study staff believe that your child has any of these illnesses, your child will be referred to the clinic or hospital for treatment, depending on the severity of illness. In addition, study staff will ask questions about this illness and take specimens as detailed below. Study staff will continue to follow your child and the development of this illness as well as discuss where to seek appropriate medical care. Therefore the study offers an opportunity for your child to receive appropriate medical care. The study will also help us to better understand the causes of illness in children, and identify the things that may harm their health. We hope that this will lead to improvements in child health. When the study is finished or if you choose to withdraw from the study, you and your child will continue to go to your usual health facility for care and study staff will no longer be involved with you or your child.

What are the risks to my child?

There are no major risks to your child. Your child may also become tired during the tasks, games and puzzles. To minimize this risk, we will take breaks whenever it seems your child is getting tired.

There may also be some discomfort associated with some of the tests we will do. These tests are listed below:

(1) Blood tests

Your child may feel sore when blood samples are taken with a needle. Where possible an anaesthetic cream will be used to dull the pain from the needle. Some bruising may occur, but this is not harmful and will disappear. Only a small amount of blood (not more than 3 teaspoons) will be taken from your child at any time.

(2) Nasopharyngeal swab

A sample of mucus will be taken from your child's nose, to test for germs that can cause chest infections and to monitor which germs are usually in your child's nose. Your child may experience minor discomfort when the nasal swab is done. Occasionally it can cause bleeding from the nose, but this is not serious, and usually stops by itself.

(3) Oropharyngeal swab

A sample of mucus will be taken from your child's mouth, to test for germs that can cause chest infections and to monitor which germs are usually in your child's mouth. Your child may experience minor discomfort when the swab is done. Occasionally it can cause your child discomfort or to gag when taken.

(4) Nasal Swab

To do a nasal swab, a small, soft-tipped swab will be inserted into one of your child's nostrils and twirled a few times until it is covered in secretions. The swab will be inserted about 2cm into the nose. This may be a little uncomfortable but should not be painful. Rarely it can cause bleeding from the nose, but this is not serious and will usually stop by itself.

(5) TB skin test

A small injection is made on your child's arm. This is to test whether your child has TB or not, and will be done at regular visits. Your child will experience minor discomfort due to the needle, with the skin test. There may also be irritation of the skin if the test is positive (reactive). This test will need to be checked 2-3 days after the injection is given.

(6) Induced sputum

Your child will be given salt-water through a nebulizer to loosen the mucous in the lungs. Then a sample of that mucus will be suctioned, or your child will be asked to cough up the mucus. Your child may experience a little discomfort while the sputum test is done. He/ she may develop some coughing or have a small amount of bleeding from the nose after this. These are not serious. Occasionally this test can cause the airways of the lungs to close. If this occurs your child will be given medicine through an inhaler/nebulizer to open the airways.

(7) Breathing test

This test is done after a child recovers from pneumonia, and at the 5 year and 6 year visits at Paarl Hospital and should not cause any discomfort. A mask will be put on his/ her face and the air going in and out of his/ her lungs while breathing will be recorded.

(8) Stool test

This test may be done monthly on your child and then every 6 months after 1 year. Study staff will collect stool from your child's nappy if passed during a study visit. If there is no stool available, a small tube will be inserted into your child's bottom and some stool will be sucked out with a syringe. The tube is thin and bendable and is only put in 1-2 centimeters to reach stool. There is a very small chance of bleeding at the rectum right where the tube goes is.

(9) Ultrasound test of the lungs

This test will be done if your child develops pneumonia so as to better see how the infection is affecting your childs lungs. This is a very safe procedure and there are no side effects.

(10) Neurocognitive and socioemotional assessment at 5 & 5.5 years.

We will ask you to answer some questions about your child's feelings and behaviors, and about your own feelings. This will be very safe, though you or your child may become tired. We will take lots of breaks so your child can relax and play in between these tasks. We will provide you both with refreshments. In total, this visit will probably last about 1.5 hours.

(11) Hearing test

This test will assess whether your child may have hearing problems. This will be very short and there will be no pain or risk.

(12) Teeth Collection

We will collect teeth that fall out naturally or are extracted during normal dental care for analysis of exposure to environmental risks. There will be no pain or risk.

(13) Vision Test

We will perform a test to check the vision of your child's eyes. There will be no pain or risk.

What are the risks to you?

There are no major risks to you. You may feel sore when blood samples are taken with a needle. Some bruising may occur, but this is not harmful and will disappear.

Some of the questionnaires ask for sensitive information relating to mental health, and this may cause some emotional distress or discomfort. Where significant issues are identified, and if you agree, study staff will offer referral to mental health support. You may also choose not to answer certain questions and still remain in the study. You will be able to take breaks, if you need to, and you will be free to terminate or reschedule the interview should the need arise.

What happens if I get hurt taking part in this study?

This research study is covered by an insurance policy taken out by the University of Cape Town. If you suffer a bodily injury because you are taking part in the study, the insurer will pay for all reasonable medical costs required to treat your bodily injury, according to the SA Good Clinical Practice Guidelines 2006. The insurer will pay without you having to prove that the research was responsible for your bodily injury. You may ask the study doctor for a copy of these guidelines.

The insurer will *not* pay for harm if, during the study, you:

- Do not take reasonable care of yourself

If you are harmed and the insurer pays for the necessary medical costs, usually you will be asked to accept that insurance payment as full settlement of the claim for medical costs. However, accepting this offer of insurance cover does not mean you give up your right to make a separate claim for other losses based on negligence, in a South African court.

Will I be paid to participate in the study?

No, you will not be paid to participate in this study. If you agree to take part, we will reimburse your transport costs for visits that are not part of your well child clinic visits.

Will there be any cost to participate in the study?

No, there will be no cost to you.

How long will my child be in the study?

This consent form is for permission for you and your child to participate in the study from 5 to 6 years of age. We hope to continue the study for many years until your child reaches at least 10 years of age; each year we will ask you again to sign permission for you and your child to continue in the study for another year.

Will my child's participation in the study be confidential?

All information that you provide will be considered confidential, and no mention of you or your child's name will appear on the stored samples or in any publication in connection with this study. No persons other than the health care workers overseeing your child's care and the study nurses and doctors will have access to any information that identifies your child personally. All your test results will not be disclosed to anyone other than for the purpose of treating you if there is a problem.

The video material will be securely stored, and only researchers directly involved in this part of the study will have access to it. This material will also be treated as very strictly confidential. Your names will not be attached to the video. The video will give us information about how you and your child interact, and about your child's behavior.

Mandatory reporting of abuse and/or deliberate neglect

The researcher(s) may not be able to keep confidential, information about known or reasonably suspected incidents of deliberate neglect or physical, sexual or emotional abuse of an adult or a child. If a researcher is given such information, he or she may report it to the authorities such as child welfare.

Does my child have to be in the study?

You can choose not to take part in the study. This will not affect the quality of care your child receives. You will be able to decline to participate at any time should any part of the study be unacceptable to you, you may still take part in the rest of the study.

What do I do if I have any questions?

If you have any questions about this study, you can ask study staff, the Principal Investigator, Professor Heather Zar, or the lung study doctor at: 021 860 2802. The UCT's Faculty of Health Sciences Human Research Ethics Committee can be contacted on 021 406 6338 in case you have any ethical concerns or questions about your rights or welfare as a participant on this research study.

Informed Consent

1. **I, _____ understand the information contained in this consent form, as explained to me in a language that I understand. I am prepared to participate in this study and give consent for my child to participate in this study. I agree to allow study staff to access my medical and hospital records as well as those of my child during the course of the study.**

2. To be completed by mother:

Child's Name:_____

Mother's Name:_____

Mother's Signature: _____

Date: _____

3. Study staff providing information: **Study staff confirming consent:**

Name: _____ Name: _____

Role in Study:_____ Role in Study:_____

Signature: _____ Signature: _____

Date: _____ Date: _____

4. If the mother is unable to read or write the entire counselling process must be observed by an independent witness who can then confirm the procedure once the mother has given consent.

Fingerprint of mother:

Witness: I confirm that I am independent of the study and that I witnessed the entire enrolment counselling process in the home language of the mother.

Name:_____

Signature: _____

Date: _____

C. Ethics Approval Documents

C1. **Ethical Approval for the parent study**

UNIVERSITY OF CAPE TOWN
Faculty of Health Sciences
Human Research Ethics Committee

Room E53-46 Old Main Building
Groote Schuur Hospital
Observatory 7925
Telephone [021] 650 7260
Email: hrec-enquiries@uct.ac.za
Website: www.health.uct.ac.za/fhs/research/humanethics/forms

11th October 2021

HREC REF: 401/2009

Prof Heather Zar
(SCAH) School of Child & Adolescent Health
Red Cross War Memorial Children's Hospital
Kilpfontein Road,
Rondebosch
Email: Heather.zar@uct.ac.za

Dear Prof Zar

Project Title: DRAKENSTEIN CHILD HEALTH STUDY

Thank you for your email to the Faculty of Health Sciences Human Research Ethics Committee dated 7th October 2021.

As per the FHS006 submitted to the Human Research Ethics committee dated 22 August 2021.

This study is **approved** until the **30th August 2022.**

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Please quote the HREC REF number 401/2009 in all your correspondence.

Yours sincerely

Signed by candidate

PROFESSOR MARC BLOCKMAN
CHAIRPERSON, FACULTY OF HEALTH SCIENCE HUMAN RESEARCH ETHICS

C2: **Ethics approval for current study**

UNIVERSITY OF CAPE TOWN
Faculty of Health Sciences
Human Research Ethics Committee

Room 45, E-52- Old Main Building
Groote Schuur Hospital
Observatory 7925
Telephone [021] 406 6492
Email: hrec-enquiries@uct.ac.za
Website: www.health.uct.ac.za/fhs/research/humanethics/forms

16 November 2021

HREC REF: 758/2021

A/Prof M Lesosky
School of Public Health & Family Medicine
Falmouth Building-FHS
Email: Maia.lesosky@uct.ac.za
Student: MYRDEM001@myuct.ac.za

Dear A/Prof Lesosky

PROJECT TITLE: THE SPATIAL ANALYSIS OF THE POTENTIAL ASSOCIATION BETWEEN BUILT ENVIRONMENT AND LUNG HEALTH OUTCOMES IN CHILDREN FROM DRAKENSTEIN, WESTERN CAPE-MPH CANDIDATE-DEMI MEYER-SUB-STUDY LINKED TO 401/2009

Thank you for submitting your study to the Faculty of Health Sciences Human Research Ethics Committee (HREC) for review.

It is a pleasure to inform you that the HREC has **formally approved** the above-mentioned study.

This approval is subject to strict adherence to the HREC recommendations regarding research involving human participants during COVID -19, dated 17 March 2020; 06 July 2020 & 01 July 2021.

Approval is granted for one year until the 30 November 2022.

Please submit a progress form, using the standardised Annual Report Form if the study continues beyond the approval period. Please submit a Standard Closure form if the study is completed within the approval period.
(Forms can be found on our website: www.health.uct.ac.za/fhs/research/humanethics/forms)

The HREC acknowledge that the student: Ms Demi Meyer will also be involved in this study.

Please quote the HREC REF 758/2021 in all your correspondence.

Please note that the ongoing ethical conduct of the study remains the responsibility of the principal investigator.

Please note that for all studies approved by the HREC, the principal investigator **must** obtain appropriate institutional approval, where necessary, before the research may occur.

Yours sincerely

Signed by candidate

PROFESSOR M BLQCKMAN
CHAIRPERSON, FACULTY OF HEALTH SCIENCES H UMAN RESEARCH ETHICS COMMITTEE
Federal Wide Assurance Number: FWA00001637.
Institutional Review Board (IRB) number: IRB00001938
NHREC-registration number: REC-210208-007

This serves to confirm that the University of Cape Town Human Research Ethics Committee complies to the Ethics Standards for Clinical Research with a new drug in patients, based on the Medical Research Council (MRC-SA), Food and Drug Administration (FDA-USA), International Council for Harmonisation of Technical Requirements for Pharmaceuticals for Human Use: Good Clinical Practice (ICH GCP), South African Good Clinical Practice Guidelines (DoH 2006), based on the Association of the British Pharmaceutical Industry Guidelines (ABPI), and Declaration of Helsinki (2013) guidelines. The Human Research Ethics Committee granting this approval is in compliance with the ICH Harmonised Tripartite Guidelines E6: Note for Guidance on Good Clinical Practice (CPMP/ICH/135/95) and FDA Code Federal Regulation Part SO, 56 and 312.

D. Manuscript preparation for the International Journal of Environmental Research and Public Health

https://www.mdpi.com/journal/ijerph/instructions