Prediction of post-tuberculosis lung damage using CT lung imaging measures among adults in Malawi

Vester Poyamba Gunsaru

GNSVES001

Supervisor: Associate Professor Maia Lesosky

Dissertation submitted to
the School of Public Health and Family Medicine,
University of Cape Town,
in partial fulfilment of the requirements for the
Master of Public Health (Epidemiology & Biostatistics) degree

October 2018

PREAMBLE

DECLARATION

I, Vester Poyamba Gunsaru (GNSVES001), 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: 29th October 2018

DEDICATION

To my mother Kudzai, for always believing in me.

DISSERTATION ABSTRACT

The burden of tuberculosis (TB) is high globally and in sub-Saharan Africa (SSA). Although TB treatment regimens are now widely available, and with high success rates in most parts of the world, TB can lead to long-term health consequences. Evidence has shown that completion of TB treatment can mark the beginning of chronic lung damage and other sequelae. There is a need for early identification of individuals at risk of post-tuberculosis lung damage (PTLD) to necessitate its appropriate management. This study evaluated the significance of radiological features from computed tomography (CT) scan in describing PTLD patterns at treatment completion and predicting PTLD outcomes at 12 months post treatment, as a possible tool to detect risk of PTLD at an early stage.

This thesis comprises three parts. Part A is the study protocol, which outlines the purpose, methods and ethical considerations of the proposed study. Part B is the literature review, which summarizes the existing literature on PTLD, the risk factors for PTLD, and the relationship between imaging data and PTLD. The objective of the literature review was to identify gaps in literature which would provide context for further research. Part C is the journal-ready manuscript, which provides the results of the conducted study and a discussion on the implication of these results.

The results of the study showed that CT imaging features were important in determining concurrent abnormal spirometry at baseline in post-TB individuals. These features were not useful in predicting spirometry outcomes at 12 months after treatment completion. However, there is need for further validation of these results in other settings and the consideration of other lung damage outcomes.

ACKNOWLEDGEMENTS

I would like to thank my supervisor, Associate Professor Maia Lesosky, for your constant support and guidance throughout the development of this dissertation. Your door was ever open for me and I am grateful.

I extend my gratitude to Dr Jamilah Meghji, from the Liverpool School of Tropical Medicine, for granting me the opportunity to work with the LAT data and for your support with any queries I had.

Finally, a big thank you to all who reviewed any part of my work. Your feedback and sacrifice of your precious time is greatly appreciated.

LIST OF ABBREVIATIONS

ANOVA	Analysis of variance
ATS	American Thoracic Society
AUC	Area under the curve
BMI	Body mass index
BOLD	Burden of Obstructive Lung Disease
COPD	Chronic obstructive pulmonary disease
CRD	Chronic respiratory disease
CT	Computed tomography
cTB	Clinically confirmed tuberculosis
CXR	Chest X-ray
FEV_1	Forced expiratory volume in one second
FVC	Forced vital capacity
HIV	Human immunodeficiency virus
HRCT	High resolution computed tomography
HREC	Human Research Ethics Committee
IQR	Interquartile range
LAT	Life after pulmonary tuberculosis
LLN	Lower limit of normal

mTB	Microbiologically confirmed tuberculosis
NCD	Non-communicable disease
OR	Odds ratio
PC	Principal Component
PCA	Principal Component Analysis
PCR	Principal Component Regression
PFT	Pulmonary function test
PI	Principal Investigator
PR	Penalized regression
PTB	Pulmonary tuberculosis
PTLD	Post tuberculosis lung damage
R^2	R squared
RF	Random forest
RMSE	Root mean squared error
ROC	Receiver operating characteristic
SD	Standard deviation
SES	Socio-economic status
SSA	sub-Saharan Africa
TB	Tuberculosis

USA	United States of America
WHO	World Health Organization

TABLE OF CONTENTS

LIST OF TABLES

PART B: LITERATURE REVIEW

PART C: MANUSCRIPT

PART D: APPENDICES

LIST OF FIGURES

PART C: MANUSCRIPT

PART A: PROTOCOL

PURPOSE OF THE STUDY

Aim: The aim of this study is to describe the relationship between post-tuberculosis (TB) treatment lung imaging measures and an *a-priori* definition of post-TB lung damage (PTLD), and to develop models for prediction of clinical outcomes at 12 months post-TB treatment completion based on early imaging data and other characteristics.

- **Objective 1**: Describe the characteristics and lung disease outcomes in a cohort of post-TB treatment adults from an urban setting in Malawi.

- **Objective 2**: Carry out variable reduction/variable selection methods on lung imaging features measured on computed tomography (CT) to identify the features most useful for discrimination between clinical outcomes.

- **Objective 3**: Estimate the predictive performance of variables identified in Aim 2 with pre-labelled groups of individuals.

BACKGROUND

Tuberculosis (TB), an infectious disease usually affecting the lungs, is one of the leading causes of death worldwide.[1,2] In 2017 alone, 10.0 million incident cases were reported with more than 1.6 million deaths in both human immunodeficiency virus (HIV)-infected and uninfected people.[2] This burden of TB is spread unequally across the world regions. The World Health Organization (WHO) ranks the African region second in the number of incident cases per year after the South-East Asia region, but leading in terms of TB and HIV co-infection.[2]

TB infection is usually accompanied by acute complications such as acute respiratory symptoms which are often the focus of treatment and care.[3] However, TB disease and its treatment can also cause chronic complications which are rarely considered in many low-resource health-care settings.[3] These chronic complications include persistent respiratory symptoms, impaired pulmonary functions[4,5] and abnormal radiological findings.[5] A number of studies have provided evidence that these complications are usually associated with poor clinical outcomes such as fibrosis or lung scarring,[5] chronic obstructive pulmonary disease (COPD) which is defined as "persistent respiratory symptoms and airflow limitation"[6,7] and bronchiectasis, which is the "permanent dilatation and distortion of the airways".[5,7] These complications and outcomes have been grouped together as "post-tuberculosis lung damage" (PTLD).

Fibrotic changes in TB patients are present after treatment completion and signify treatment effectiveness.[8,9] A study in India, one of the countries with high TB incidence rates, looked at abnormal radiological findings in post-TB patients, finding that the majority of individuals (86.0%, 170/198) had abnormal radiological findings 14-18 years after long duration treatment for TB.[10] One of the most common abnormal radiological findings was fibrosis, which was present in 36.0% of those with abnormal radiological findings. Fibrosis was also reported in another study in India as the common (38.7%) pulmonary lesion in 120 cases of pulmonary TB (PTB).[11]

COPD is also common among individuals with a history of TB.[5,7] A systematic review by Byrne et al.,[7] which included studies that investigated the relationship between a history of TB and COPD, found an association between previous TB and spirometry confirmed COPD in adults of all ages. The association was higher for countries with higher TB incidence rates.[7] It is to be noted, however, that most of the studies in the review were cross-sectional studies which cannot fully assess temporality. Obstructive ventilatory defects were observed in 26 of 47 individuals (55.3%) previously treated for PTB in a case series study in Pakistan.[12] Similarly, an evaluation of the prevalence of chronic respiratory symptoms and pulmonary dysfunction in Brazil found 41.0% of the patients had pulmonary dysfunction with a majority of these being obstructive disorders.[4]

Furthermore, previous TB has been reported to be a common aetiology of bronchiectasis.[1] Qi et al.[13] studied the aetiology and clinical characteristics of bronchiectasis in a Chinese population (another country with high TB incidence rates). In this study, previous TB was the second most common risk factor of bronchiectasis in the study population.[13] Individuals with a history of TB had a statistically significant higher frequency of both left and right upper lobe involvement compared to those without a history of TB.[13] Byrne et al.[7] also reported a statistically significant association between bronchiectasis and a history of TB.

One of the factors that differentiates the risk of such outcomes in post-TB individuals is the severity of disease as seen on radiological imaging of TB patients before treatment.[14,15] Chung et al.[15] also reported that less improvement in this radiological damage after TB treatment was suggestive of even more poor clinical outcomes. A further exploration of this relationship has been done in a few studies mostly in Korea and India.[16,17] It has been observed that severe radiological damage after TB treatment is significantly associated or moderately correlated with other PTLD clinical outcomes.[17-19]

It is clear that radiological findings are useful in describing severity of lung function impairment as well as various clinical outcomes. These can be used as a possible tool for early detection of PTLD outcomes. However, as can be seen from the evidence above, this association has been studied mainly in the Asian world region using a majority of cross-sectional and retrospective studies. Furthermore, the statistical methods used to derive the suggested associations have been mainly tests of association or correlation coefficients. This proposed study aims to look at the relationship between lung imaging findings and PTLD with the use of computed tomography (CT) measures that have been derived by robust data-driven variable selection methods. Using prospective data from sub-Saharan Africa (SSA), results from this study might be beneficial to management of PTLD in similar settings.

METHODS

STUDY DESIGN

The proposed work will use secondary data from “The Life After Pulmonary TB” study, which is a prospective cohort study carried out in urban Blantyre, Malawi. The parent study had a primary aim to look at the prevalence, nature and severity of post-TB lung damage. Individuals finishing pulmonary TB treatment with a “Treatment success” outcome were followed for 12 months after the completion of TB treatment. Study visits occurred at TB treatment completion, as a baseline study visit, at six months post-TB treatment and at 12 months post-TB treatment. The data used in the study were gathered from a non-probability sample of 405 individuals.

STUDY POPULATION

A total of 405 individuals were included in the parent study. These individuals were between the ages of 15 and 84. During recruitment, individuals who were residents of urban Blantyre and on TB treatment were eligible for the study. The inclusion criteria for the baseline visit was individuals with TB "Treatment success" ("Cured" or "Treatment completed"). Both HIV-infected and uninfected individuals were included.

RECRUITMENT AND ENROLMENT

TB patients visiting hospitals/clinics in the vicinity of urban Blantyre for their TB treatment were approached. Those meeting eligibility criteria for recruitment were asked for and provided written informed consent for participation in the study, and were able to withdraw at any point during the study.

Inclusion and Exclusion Criteria

Recruitment

The inclusion criteria for recruitment covered individuals who were ≥ 15 years of age, who had PTB as defined by the National TB program, with the first episode of TB treatment and within two months of TB treatment success. All individuals with WHO danger signs or with positive TB smear or culture were not eligible for recruitment.

Baseline Visit

Individuals who were alive, with “Treatment success” (“Cured” or “Treatment completed”) at TB treatment completion and healthy enough to participate in the study were eligible for the baseline visit. Those who failed TB treatment, died, were acutely unwell, relocated from Blantyre or unable to be reached, were excluded from the baseline visit.

Six-month visit

At the six-month visit, all those who participated in the baseline visit were eligible for participation. In addition to this, individuals must have been alive and well at the time of the six-month visit. All individuals who had relocated, withdrew or were unable to be reached were excluded from participation.

12-month visit

Individuals who had been in the study from the baseline visit through to the six-month visit were eligible for the final follow-up visit at 12 months post-TB treatment. Participation in this visit was only possible if individuals were alive, well, still willing to continue in the study and still residing within the Blantyre area.

RESEARCH PROCEDURES AND DATA COLLECTION METHODS

Baseline characteristics

At the baseline visit, detailed questionnaires were administered to the individuals for the purpose of getting necessary patient characteristics. These patient characteristics included demographic data, HIV- and TB-related information, respiratory exposures and socio-economic status (SES). The questionnaires were administered by trained interviewers who were hired by the principal investigator (PI) of the parent study. All interviews were done in the local language.

Spirometry data

All study visits included spirometry tests, using the *ndd EasyOne* Spirometer. Spirometry was done to American Thoracic Society (ATS) standards. Spirometry abnormalities were defined according to the NHANES III, GLI and Local reference ranges.

Imaging data

CT scans were done at baseline only. These scans were done at two locations in the country, depending on availability and functionality of the equipment. The same imaging protocol was used for both locations to ensure comparability. The CT images were assessed by two experienced radiologists and discrepancies were resolved by a third experienced reader. The primary CT variables included parenchymal variables, bronchiectasis variables, airway variables, cavity variables, other lobar variables and whole lung variables.

DATA SAFETY AND MONITORING

Data with no patient-identifying information were stored on a central server at the Liverpool School of Tropical Medicine with regular back-ups. The data were password-protected and only approved personnel had access. Electronic devices used to collect the data were erased regularly. Any physical copies of patient information were stored in locked cupboards.

Data given by the participants were verified by examining patient health books and the TB National Treatment Program register. The PI of the parent study performed regular weekly reviews and queries of the questionnaire data collected to ensure completeness and quality. Quality control of the spirometry data was done on a daily basis by the PI and included retraining of the study team where necessary. Further quality assurance on the spirometry data was sourced externally. CT imaging data were reviewed by the PI as well, and in some instances, it was done together with collaborating radiologists.

DATA ANALYSIS

Descriptive Statistics

Descriptive statistics for patient characteristics, spirometry and CT variables will be used for exploratory analysis of the data. These will include frequency distributions (%) for categorical variables and mean (standard deviation)/median (inter-quartile range) for continuous variables.

Variable selection

Data-driven variable selection methods such as Principal Component Analysis (PCA), penalized regression and random forest will be used to develop reduced variable sets. These methods have been selected because of their capabilities in handling many predictor variables and collinearity. Results will be presented via variable importance scores.

Predictive models

The most important variables from above will be used for predictive modelling using principal components regression, penalized regression and random forest. Several PTLD outcomes will be used as outcome variables. Other covariates to be included in the prediction models are age, gender, body mass index (BMI), HIV status, smoking status and SES status. The full data set will be randomly split into a training data set and a testing data set. The training data set will be used for the initial formation of the regression models using cross-validation and the testing data set will be used for validation of these models. Receiver operating characteristic (ROC) curves will be used to assess the prediction performance of the models. Other results will be reported with 95% confidence intervals.

Software

All analyses will be performed using R software version 3.4.3.

ETHICAL CONSIDERATIONS

ETHICS APPROVAL

The parent study was carried out under ethics approval from the Malawi College of Medicine Research and Ethics Committee (P.10/15/1813, Appendix B). For the proposed study, ethics approval from the Human Research Ethics Committee (HREC) of the Faculty of Health Sciences at the University of Cape Town will be sought.

RISKS AND BENEFITS

This will be an analysis of secondary data, so there are no direct risks or benefits to the individuals.

INFORMED CONSENT PROCESS

Informed consent was sought from all individuals prior to study enrolment. Participants between the ages of 15 and 18 provided assent and informed consent was given by their guardians. The informed consent included information on the purpose of the study and the potential risks and benefits of the study. This information was given in the local language. Individuals were informed that their participation was voluntary and that they had the right to withdraw at any point of the study if they decided to do so.

PRIVACY AND CONFIDENTIALITY

This will be an analysis of secondary data that has already had identifiers removed. No patient identifying information would be available to the study investigators. Study data will be kept on a password-protected computer and only non-identifying summary measures will be reported and published.

REFERENCES

1. Jordan TS, Spencer EM, Davies P. Tuberculosis, bronchiectasis and chronic airflow obstruction. Respirology (Carlton, Vic). 2010;15(4):623-8.

2. World Health Organization. Global Tuberculosis Report. Geneva, Switzerland2018. Available from: http://apps.who.int/iris/bitstream/handle/10665/274453/9789241565646-eng.pdf?ua=1.

3. Shah M, Reed C. Complications of tuberculosis. Current Opinion in Infectious Diseases. 2014;27(5):403-10.

4. Nihues Sde S, Mancuzo EV, Sulmonetti N, Sacchi FP, Viana Vde S, Netto EM, et al. Chronic symptoms and pulmonary dysfunction in post-tuberculosis Brazilian patients. The Brazilian journal of infectious diseases : an official publication of the Brazilian Society of Infectious Diseases. 2015;19(5):492-7.

5. Chakaya J, Kirenga B, Getahun H. Long term complications after completion of pulmonary tuberculosis treatment: A quest for a public health approach. Journal of Clinical Tuberculosis and Other Mycobacterial Diseases. 2016;3:10-2.

6. Vogelmeier CF, Criner GJ, Martinez FJ, Anzueto A, Barnes PJ, Bourbeau J, et al. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease 2017 Report. GOLD Executive Summary. American journal of respiratory and critical care medicine. 2017;195(5):557-82.

7. Byrne AL, Marais BJ, Mitnick CD, Lecca L, Marks GB. Tuberculosis and chronic respiratory disease: a systematic review. International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases. 2015;32:138-46.

8. Lee JJ, Chong PY, Lin CB, Hsu AH, Lee CC. High resolution chest CT in patients with pulmonary tuberculosis: characteristic findings before and after antituberculous therapy. European journal of radiology. 2008;67(1):100-4.

9. Majmudar DK, Rajput DK. Role of HRCT in diagnosing disease activity in pulmonary tuberculosis. International Journal of Contemporary Medical Research. 2017;4(8):1724-7.

10. Banu Rekha VV, Ramachandran R, Kuppu Rao KV, Rahman F, Adhilakshmi AR, Kalaiselvi D, et al. Assessment of long term status of sputum positive pulmonary TB patients successfully treated with short course chemotherapy. The Indian journal of tuberculosis. 2009;56(3):132-40.

11. Menon B, Nima G, Dogra V, Jha S. Evaluation of the radiological sequelae after treatment completion in new cases of pulmonary, pleural, and mediastinal tuberculosis. Lung India : official organ of Indian Chest Society. 2015;32(3):241-5.

12. Baig IM, Saeed W, Khalil KF. Post-tuberculous chronic obstructive pulmonary disease. Journal of the College of Physicians and Surgeons--Pakistan : JCPSP. 2010;20(8):542-4.

13. Qi Q, Wang W, Li T, Zhang Y, Li Y. Aetiology and clinical characteristics of patients with bronchiectasis in a Chinese Han population: A prospective study. Respirology (Carlton, Vic). 2015;20(6):917-24.

14. Ralph AP, Kenangalem E, Waramori G, Pontororing GJ, Sandjaja, Tjitra E, et al. High morbidity during treatment and residual pulmonary disability in pulmonary tuberculosis: under-recognised phenomena. PloS one. 2013;8(11):e80302.

15. Chung K-P, Chen J-Y, Lee C-H, Wu H-D, Wang J-Y, Lee L-N, et al. Trends and predictors of changes in pulmonary function after treatment for pulmonary tuberculosis. Clinics. 2011;66(4):549-56.

16. Seon HJ, Kim YI, Lim SC, Kim YH, Kwon YS. Clinical significance of residual lesions in chest computed tomography after anti-tuberculosis treatment. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2014;18(3):341-6.

17. Santra A, Dutta P, Manjhi R, Pothal S. Clinico-Radiologic and Spirometric Profile of an Indian Population with Post-Tuberculous Obstructive Airway Disease. Journal of clinical and diagnostic research : JCDR. 2017;11(3):Oc35-oc8.

18. Panda A, Bhalla AS, Sharma R, Mohan A, Sreenivas V, Kalaimannan U, et al. Correlation of chest computed tomography findings with dyspnea and lung functions in post-tubercular sequelae. Lung India : official organ of Indian Chest Society. 2016;33(6):592-9.

19. Meyyappan D, Chockalingam P. Evaluation of respiratory impairment and health related quality of life in pulmonary tuberculosis sequelae patients. 2018. 2018;5(2):5.

PART B: LITERATURE REVIEW

INTRODUCTION

Tuberculosis (TB) remains one of the leading causes of mortality and morbidity in the world, with 1.6 million deaths and 10.0 million new infections in 2017 alone.[1] In Africa, which is the region with the highest TB incidence rates after South-East Asia, there were 0.6 million TB associated deaths and 2.4 million new TB infections in that same year.[1] These new infections in Africa represented 25.0% of the global TB incident cases.[1]

Amid this epidemic, effective TB treatment has been widely available with coverage rates of up to 52.0% in Africa.[1] Consequently, high treatment success rates have been reported (up to 82.0%).[1] This has seen many people successfully treated for TB and regarded as cured. Nevertheless, when TB patients have successfully completed their treatment, most TB health services do not provide for their life after treatment. In addition, the WHO 'end TB' strategy, which is a strategic plan for achieving a world free of TB, focuses only on mortality, morbidity and suffering associated with TB infection.[2,3]

Empirical evidence has shown that TB infection is usually associated with lung damage, acute or chronic, after completion of treatment.[3-5] Despite this, there has been insufficient focus on the management of post-tuberculosis lung damage (PTLD). One form of PTLD management that has been suggested in the literature is pulmonary rehabilitation for TB patients after successful treatment.[6] This involves exercise training, relaxation, breathing retraining,[6] resistance training[7] and education on TB,[6,7] chronic respiratory diseases (CRD) and their management.[7] This is done with the aim of improving pulmonary function, physical function and quality of life in post-TB individuals.[6,7] Although management of PTLD may be possible, there is a need for identification of individuals likely to develop PTLD at an early stage to facilitate appropriate and timely care.

The proposed research, aims to investigate the detection of PTLD by using imaging data to describe and predict patient outcomes after completion of TB treatment. The current literature review aims to inform the proposed research by evaluating previous work done on the relationship between lung imaging features and PTLD. The review will specifically address the following objectives:

- To describe PTLD characteristics and outcomes
- To synthesize the evidence on risk factors for PTLD
- To synthesize the evidence on the significance of imaging features in post-TB individuals.

The above objectives, once met, will guide the focus of the proposed study.

SEARCH STRATEGY

An electronic search of the PubMed, Google Scholar, Scopus and EBSCO host (Africa-Wide Information, CINAHL, Health Source-Nursing/Academic information) databases was done to find the relevant literature. The search was restricted to English language articles published in the last 10 years. No restrictions were placed on the geographic region of study. The search terms used were "post-tuberculosis OR post-TB", "radiology OR radiography", "computed tomography OR CT" and "lung damage". These were combined with the Boolean operator "AND", where appropriate and adapted for each database. Titles and abstracts were reviewed from the initial search results to obtain relevant literature. Furthermore, reference lists of the identified relevant literature were also reviewed to obtain additional articles. The detailed inclusion and exclusion criteria have been presented in Table B1 and a summary of the included studies is presented in Table B2.

Table B1: Inclusion and exclusion criteria of the selected relevant articles.

Category	Inclusion	Exclusion
Study population	• Adults with a history of pulmonary TB	• Adults with a history of other forms of TB • Adults without a history of TB • Children with or without a history of pulmonary TB
Study design	• Review (systematic or non-systematic) • Comparative studies (controlled trials, cohort and case-control) • Case series	• Case study or report
Study setting	• All world regions	
Outcome measures	• Lung damage or other related outcomes after TB treatment	
Language	• English	
Publication dates	• 2008 - 2018	
Text availability	• Full text	

OUTLINE OF THE REVIEW

The review will summarize the literature by first focusing on the description of PTLD characteristics and outcomes. A synthesis of the suggested risk factors for these PTLD characteristics and outcomes will follow. Subsequently, there will be a section on imaging features and lung damage, and then a conclusion on what has been discussed thus far. Finally, the review will end with a section on the gaps identified in the literature.

SUMMARY OF THE LITERATURE

PTLD CHARACTERISTICS AND OUTCOMES

Lung damage after the completion of TB treatment is evident in most patients,[3-5] and often leads to poor quality of life.[8,9] Patients who have completed treatment for pulmonary tuberculosis (PTB) usually move from the burden of an infectious disease (TB) to a new burden of non-communicable diseases (NCDs) that relate to long-term lung damage.[4] An example of such NCDs is chronic respiratory diseases (CRDs), for which PTB has been shown to be a risk factor.[3,4,10] CRDs are defined by the WHO as “chronic diseases affecting the airways and other structures of the lung”.[11] CRDs in post-TB individuals are manifested through several lung damage characteristics and outcomes.

PTLD characteristics

Lung damage characteristics, which are symptoms of compromised lung function, in post-TB individuals include persistent respiratory symptoms, abnormal spirometry patterns and abnormal radiological findings.

Respiratory symptoms

Persistent respiratory symptoms are common in post-TB individuals. These include haemoptysis,[12,13] cough,[6,14,15] sputum,[14] chronic bronchitis, asthma,[16] dyspnoea,[14-16] wheeze[14,16] and expectoration.[15] These symptoms have been reported with high prevalence

(45.0-62.1%) in countries such as South Korea, Cameroon,[6] South Africa,[16] Brazil[14] and India,[15] most of which are high TB burden areas. However, the majority of studies with these observations were cross-sectional studies, which make it difficult to determine direction of causation.

The most observed of these respiratory symptoms vary among post-TB populations. Santra et al.[15] reported dyspnoea as the most common respiratory symptom (95.7%), followed by cough (89.9%). This was in an Indian population of 138 patients (53.6% with respiratory symptoms) with a history of PTB (> 6 years prior) and symptoms of obstructive airway disease.[15] On the other hand, Nihues et al.[14] observed cough as the most common symptom (28%) and dyspnoea as the least observed symptom (8%) in 121 Brazilian adults who had been successfully cured of TB. In a review of epidemiological studies done in South Africa, chronic bronchitis was commonly reported (in five out of seven studies).[16]

Spirometry symptoms

Lung damage due to TB disease causes poor lung function in patients, as seen by spirometry tests (pulmonary function tests). This usually persists after treatment completion into spirometry patterns of obstruction, restriction or mixed, signifying affected lung function.[6,16] The dominance of these patterns also varies among post-TB populations. Allwood et al.[5] reported a prevalence of obstructive spirometry pattern ranging from 15.0-68.0% among the adults with previous TB in a systematic review. These individuals were either studied soon after TB treatment completion, or followed prospectively for 15 years and retrospectively for 5.6 years.[5] In contrast to the high prevalence of obstructive spirometry patterns, mixed spirometry patterns were observed in 9.3-34.0% of these individuals.[5] These results were from countries such as Pakistan, United States of America (USA), South Africa, Serbia, Brazil and Taiwan. In most of these studies, however, it was unclear whether airflow obstruction pre-dated TB treatment commencement. In their review, Munoz-Torrico et al.[6] also reported dominant obstructive spirometry patterns of 68.0%, 62.9%, 62.5% and 55.3% among study participants from the USA, Cameroon, India and Pakistan respectively, as compared to other spirometry patterns.

In other post-TB populations, mixed spirometry patterns have been reported to be dominant compared to the obstructive and restrictive spirometry patterns. In a study by Meyyappan et al.,[9] mixed spirometry patterns were the most reported (39.5%) in an Indian population with previous PTB. This was a population of 200 patients, of which 81% had spirometry abnormalities.[9] Obstructive spirometry pattern was the second most reported abnormality (23.5%) and 18% had restrictive spirometry abnormalities.[9] Santra et al.[15] also reported dominant mixed spirometry patterns (72.5%) in post-TB patients.

Restrictive spirometry patterns have not been reported as commonly as the other two spirometry patterns. Nevertheless, Nair et al.[8] reported restrictive abnormality as the most common spirometry abnormality in a cross-sectional study of 443 patients who had successfully completed TB treatment. The majority of these patients (65.7%) had abnormal spirometry patterns, of which 35.0%, 20.3% and 10.4% had restrictive, mixed and obstructive patterns respectively.[8]

Abnormal radiological findings

Abnormal radiological findings, suggestive of different lung pathologies, are another characteristic of PTLD. These have been observed in post-TB individuals on either chest X-rays (CXR) or computed tomography (CT) scan. Meghji et al.,[17] in a systematic review on the nature of post-TB lung disease, reported a prevalence of post-TB radiological features such as cavitation, fibrosis, pleural thickening, nodules, emphysema, consolidation and mosaicism. These features were observed with differing prevalence, depending on whether they were observed on CXR or CT scan. Pleural thickening, nodules, consolidation and mosaicism features were only observed on the CT scan.[17] In a retrospective study of 66 newly diagnosed PTB adults (>20 years) in Korea, 50.0% had possible active parenchymal lesions on CT after treatment while a further 8.0% had equivocal parenchymal lesions on CT after treatment.[18] However, none of these lesions persisted after a mean follow-up period of 15 months.[18] In another retrospective study of 441 new cases of pulmonary, pleural and mediastinal TB in New Delhi, India, 40.4% had residual CXR lesions after treatment.[19] Of these residual lesions, 67.4% were pulmonary, 23.6% were pleural and 9.0% were mediastinal in nature.[19] The pulmonary lesions consisted of fibrosis (38.7%), cavity (21.4%), and calcification (3.0%), among others.[19]

PTLD outcomes

The PTLD characteristics of persistent respiratory symptoms, abnormal spirometry patterns and radiological findings are usually associated with two main lung damage outcomes: bronchiectasis, and chronic obstructive pulmonary disease (COPD).

Bronchiectasis

Bronchiectasis, a “permanent dilatation and distortion of the airways”[20] characterised by a constant cough, is as a respiratory sequelae of PTB.[4,17,20] Meghji et al.[17] reported a prevalence of bronchiectasis ranging between 4.3-86.0%, depending on whether it was observed using CXR or CT scan. Higher prevalence of bronchiectasis was observed on CT imaging.[17] The results were found in mainly prospective studies done in Thailand, India, Brazil, Martinique, Canada and Taiwan.[17] In addition, two studies from China and Korea, included in a systematic review by Bryne et al.,[4] showed a significant positive association between previous TB and bronchiectasis in adults older than 18 years of age. Odds ratios (OR) of 3.1 and 4.6, respectively, were reported for this association.[4] However, these results were from large population-based cross-sectional studies with inadequate information, non-representative sampling methods, less than 50% participation rates, and assessment of exposure and outcome done in different samples.[4] More evidence of the relationship between bronchiectasis and a history of TB was presented in a systematic review by Gao et al.,[21] who looked at the aetiology of bronchiectasis in adults. In their review, TB was the widely reported post-infectious aetiology of bronchiectasis.[21] Studies that reported post-TB bronchiectasis were mainly from Europe and Asia, with only one from Africa.

Chronic obstructive pulmonary disease

Chronic obstructive pulmonary disease (COPD), "characterised by persistent respiratory symptoms and airflow limitation",[22] is strongly associated with a history of TB especially in areas with high TB incidence rates.[4-6,15,16,23] Bryne et al.[4] estimated ORs between 1.8 and 6.3 for the association between a history of TB and COPD. Study participants were adults (>15 years) from South Africa, Colombia, Korea, Philippines, Taiwan and other countries from Burden of Obstructive Lung Disease (BOLD) studies.[24] These results were adjusted for smoking and age, which are known risk factors of COPD.[4] Furthermore, Sarkar et al.[23] reported ORs ranging from 1.4 to 4.0 for the association between a history of TB and COPD after adjusting for sex, age, smoking, ethnicity and exposure to dust, and smoke.

From the description of the prevalence of respiratory diseases above, it is evident that chronic lung damage is common among post-TB individuals. The question that arises from this evidence, however, is whether there are factors that differentiate the risk of PTLD in post-TB individuals.

RISK FACTORS FOR PTLD

The risk of PTLD for patients who complete TB treatment is associated to a number of factors. Identified factors include: smear-positive disease,[25,26] delay in TB diagnosis and treatment,[27-29] longer treatment duration, regimen change,[25] human immunodeficiency virus (HIV) infection,[28] baseline radiological damage[25,26,28] and decreased improvement in radiological damage after treatment.[25] However, some studies have not found an association between PTLD and delay in TB diagnosis and treatment or HIV infection. In a prospective study done in Texas in patients treated for PTB, pulmonary impairment after treatment was not associated with delay in TB diagnosis and treatment.[10] While HIV infection was reported to be associated with PTLD by Ralph et al.,[28] this association was only stronger in univariate analysis but weakened in a multivariate analysis. In addition, Ross et al.[26] did not find any association between HIV infection and loss of lung function in a retrospective study of South African gold miners treated for PTB.

One of the commonly reported risk factors for PTLD is the extent of radiological damage before TB treatment. Radiological scores of TB patients before treatment could provide important information in understanding the nature of PTLD. In addition, smaller improvement in these scores after treatment characterises PTLD and is also suggestive of future abnormal lung function.[17,25] Establishing a relationship between imaging features and other PTLD characteristics and outcomes could then provide a method of early detection and proper management of PTLD.

IMAGING FEATURES AND LUNG DAMAGE

Radiological features are of clinical significance in TB patients as they provide information on disease activity as well as treatment response.[30,31] Lee et al.[30] looked at high resolution CT (HRCT) findings in 52 PTB patients before and after treatment to evaluate the usefulness of HRCT in assessing disease activity. In this study done in Taiwan, they concluded that HRCT findings were capable of aiding in diagnosis of TB and early initiation of TB treatment. Specifically, they reported that micronodules and tree-in-bud HRCT features, which cannot be seen on CXR, were important signs of disease activity and could thus be used as diagnostic signs.[30] In addition, the presence of fibrosis, the absence of pleural effusion and the disappearance of tree-in-bud features were reported to be signs of TB treatment effectiveness.[30] Similar findings were reported in a study by Majmudar et al.[31] in India. They conducted a retrospective study on 50 patients with suspected PTB to assess patterns of HRCT findings in active and inactive disease as well as to determine if HRCT findings can predict TB disease activity. More than half (28/50) of the patients had inactive disease and 22 patients had active disease.[31] Tree-in-bud appearance, ill-defined nodules and consolidation features were predictive of disease activity and these were observed in 77.0%, 72.0% and 68.0% of those with active disease, respectively.[31] Features predictive of inactive disease were traction bronchiectasis, scar formation and calcified granulomas. These findings were observed in 64.0%, 50.0% and 21.4% patients with inactive disease, respectively.[31] However, the details of their predictive analyses were not explicitly stated.

Besides relating radiological findings to TB disease activity, the significance of such findings has been further explored in relation to persistent disease activity or lung damage after TB treatment. Seon et al.[18] studied the significance of radiological lesions observed on a HRCT scan after TB treatment in a retrospective study in Korea. Residual active lesions after TB treatment were observed in half of the patients.[18] The active lesions were defined as either residual parenchymal lesions >30% on the initial lesions before treatment, or tree-in-bud or cavitation appearance, or internal fluid bronchogram, or non-reduced low attenuation area in pneumonic consolidation.[18] After a follow-up period of 15 months, these residual lesions did not predict any persistent disease activity or relapse of PTB.[18] Only three of the patients with active lesions after TB treatment had new or aggravating lesions after the follow-up period.[18]

On the other hand, several studies have evaluated the association between radiological findings and lung damage post-TB. Santra et al.[15] looked at the radiologic and spirometric profile of an Indian population with a history of PTB and with symptoms of obstructive airway disease. Based on CXR results of these patients, the extent of radiological lesions which included pulmonary fibrosis, cavitation, calcification and bronchiectasis was grouped into three severity groups. From the severity groups, more severe radiological damage was associated with severe airflow limitation,[15] This was characterised by severe involvement of radiological lesions in more than three zones of the lung, with or without cavitation.[15] This relationship was statistically significant using a Chi-square test of association.[15] According to Nair et al.,[8] a similar characterisation of severity of CXR radiological damage was associated with poor quality of life scores in post-TB patients.

Using radiological findings from a CT scan, Panda et al.[32] evaluated the correlation between such scores, dyspnoea and pulmonary function tests (PFT). They included CT measurements of fibrosis, cavitation, bronchiectasis, consolidation, nodule and aspergilloma. This was done in a sample of 101 individuals with a history of PTB, where 57.4% had dyspnoea and 77.2% had abnormal PFTs.[32] Significant differences in fibrosis, bronchiectasis and nodule scores were observed between patients with dyspnoea and patients without dyspnoea.[32] Negative correlation between PFTs (predicted FEV1 and FVC) and CT scores was observed, but this result was not statistically significant.[32] Bak et al.[33] also used CT findings to determine which features were associated with the severity of bronchiectasis in 127 Korean patients (40 post-TB patients) with moderate-to-severe COPD. They used quantitative assessments of emphysema, air trapping and bronchial wall thickening. Moderate-to-severe bronchiectasis, which was associated with a history of TB, was also associated with severe air trapping.[33] This association was determined using analysis of variance (ANOVA). The other quantitative measurements of CT features did not have a statistically significant relationship with severity of bronchiectasis.[33]

Meyyappan et al.[9] used radiological features observed on a chest skiagram and reported a significant moderate correlation (r = 0.3) between pulmonary function and radiological features in post-TB individuals. A moderate correlation (0.4) was also seen between radiological features and quality of life of these individuals.[9]

From the above evidence, radiological features are of significance in providing information on lung damage in both TB and post-TB patients. However, radiological features from a CT scan are of vital importance because of their high sensitivity to lung damage compared to other radiological features such as from CXR[17,30,31]

CONCLUSION

Despite successful treatment completion for many TB patients, most of these patients will suffer from long-term associated illness. There is evidence of a high prevalence of PTLD in TB patients after treatment completion. PTLD is characterised by persistent respiratory symptoms such as haemoptysis, coughing, chronic bronchitis, asthma, dyspnoea and wheeze.[6,12-16] In addition, lung function abnormalities are also characteristics of PTLD. These can be either obstructive, restrictive or mixed patterns.[5,6,8,9,15] Abnormal radiological findings on CXR or CT scan such as fibrosis, emphysema, cavitation and pleural thickening, among others, are also common in post-TB individuals.[17-19] These PTLD characteristics are associated with outcomes of bronchiectasis and COPD.[4,17,21,23] PTLD is likely in those patients with smear-positive disease, HIV infection, poor baseline radiological scores, decreased improvement in radiological scores after treatment, longer TB treatment duration and delayed TB diagnosis and treatment.[25-29] However, some studies have not found a relationship between PTLD and HIV status or delayed TB diagnosis and treatment.[10,26] The relationship between radiological scores and PTLD has further been explored to establish the significance of radiological findings in determining other PTLD characteristics and outcomes. Significant associations have been reported, but there is a need for further research.[8,9,15,32,33]

AREAS FOR FURTHER RESEARCH

This review has highlighted the large body of evidence present for PTLD, even though most the studies included were lacking control populations. Despite covering a variety of world regions, very few studies were done in Africa, especially sub-Saharan Africa (SSA) where the TB burden is high. This calls for further research in such a region, which could help identify different patterns of PTLD, if any exist. This information could help in informing targeted management of PTLD. The significance of radiological findings in post-TB individuals has been suggested as a method of determining who will experience PTLD. The review has revealed the methodological gap that is present in the suggested associations. As such, further research is required to investigate the significance of radiological findings, using robust methods, as a way of early detecting and managing PTLD.

With the revealed gaps in knowledge, the proposed research seeks to evaluate the relationship between early CT imaging data and other PLTD clinical outcomes and to develop a PTLD prediction model using these imaging data. The knowledge from this will potentially provide an option for early detection and proper management of PTLD.

Table B2: Summary of the literature included in the review.

Author, Year	Country	Study design	Population	Outcomes	Key Findings
Munoz-Torrico, 2016	Global	Non-systematic review	Adults treated for pulmonary tuberculosis (PTB)	TB sequelae	TB is responsible for functional sequelae, primarily obstruction spirometry pattern. There is a rationale for pulmonary rehabilitation after TB treatment
Meghji, 2016	Global	Systematic review	Adults post pleural, miliary or PTB	Post-TB lung disease	High prevalence of residual structural pathology and residual abnormalities on imaging following PTB treatment
Byrne, 2014	Global	Systematic review	General population	Chronic obstructive pulmonary disease (COPD) and bronchiectasis	TB associated with COPD and bronchiectasis especially in TB endemic areas
Santra, 2017	India	Cross-sectional	138 adults (≥18 years) with a history of PTB	Post-TB obstructive airway disease (OAD)	Most patients with post-TB OAD have associated restrictive component. Pulmonary function among such patients was nor different by smoking status.
Chakaya, 2016	Global	Review	Post-PTB adults	Long-term complications following PTB treatment.	Evidence of long-term complications after PTB treatment
Niheus, 2015	Brazil	Cross-sectional	120 adults (18-65 years) with a history of TB	Chronic symptoms Pulmonary dysfunction	High prevalence of chronic respiratory symptoms and pulmonary dysfunction in post-TB patients
Vecino, 2011	USA (Texas)	Prospective cohort	123 adults treated for culture-confirmed PTB	Pulmonary impairment after TB (PIAT)	PIAT was not related to duration of delay in tuberculosis diagnosis or treatment. No significant change in PIAT during follow-up period
Menon, 2015	India	Retrospective observational	446 adults diagnosed and treated for pulmonary, pleural and mediastinal TB	Radiological lesions	Residual radiological opacities are widely seen in TB treated patients. Pulmonary lesions show more residual lesions.
Pefura-Yone, 2014	Cameroon	Cross-sectional	177 post-PTB	Chronic respiratory signs	Forced expiratory flow 25-75%<65% is useful for the

		study	adults		assessment and diagnosis of post-TB distal airflow obstruction.
Seon, 2013	Korea	Retrospective	66 post-PTB adults (>20 years)	Residual lesions in computed tomography (CT)	Residual lesions did not predict persistent activity or relapse of PTB after 15-month follow-up
Allwood, 2013	Global	Systematic Review	Post-PTB adults (>19 years)	Chronic airflow obstruction (CAO)	Evidence of positive association between PTB and CAO
Chung, 2011	Taiwan	Retrospective	115 post-PTB adults	Pulmonary function	Risk factors for pulmonary function deterioration include smear-positive disease, extensive pulmonary involvement prior to anti-TB treatment, prolonged anti-TB treatment, and reduced radiographic improvement after TB
Lee, 2007	Taiwan	Prospective cohort	52 adults with active PTB	High resolution CT (HRCT) findings	HRCT may be helpful in the diagnosis of PTB and assessment of the efficacy of anti-TB treatment
Kumar, 2016	India	Prospective observational	168 adults with haemoptysis	Causes of haemoptysis Imaging modalities	TB the main cause of haemoptysis HRCT and Multidetector CT useful in identification of bleeding causes
Nair, 2018	India	Cross-sectional	443 post-TB adults (18-60 years)	Lung function Health related quality of life (HRQoL)	PTB associated with poor quality of life
Meyyappan, 2018	India	Observational	200 adults (>18 years) with PTB sequelae	Respiratory and cardiac impairment HRQoL	Significant functional impairment and poor quality of life associated with PTB sequelae Severe radiological scores correlated with respiratory and cardiac impairment, and HRQoL)
Majmudar, 2017	India	Retrospective co-relational	50 adults with suspected PTB	HRCT findings in active and inactive PTB patients	HRCT findings can be important in diagnosis and management of PTB
Gao, 2016	Global	Systematic review	Adults with bronchiectasis	Aetiologies of bronchiectasis	Most common aetiologies of bronchiectasis were idiopathic and post-infective.
Sarkar, 2017	Global	Review	TB and COPD patients	Epidemiology, characteristics, mechanisms and treatment of TB	History of TB a risk factor of COPD

				associated COPD	
Shankar, 2018	India	Retrospective observational	254 adults with haemoptysis	Aetiology, pathogenesis, radiological profile and outcomes	Active TB, TB sequelae, malignancy and bronchiectasis were common aetiologies of haemoptysis
Bak, 2018	Korea	Cohort	127 COPD patients	Quantitative CT features and clinical values determining extent of bronchiectasis	PTB history associated with moderate-to-severe bronchiectasis Air trapping associated with extent of bronchiectasis
Ehrlich, 2011	South Africa	Review	Adults and adolescents with or without TB history	Chronic symptoms Lung function loss	PTB associated with airflow obstruction and restriction
Ralph, 2013	Indonesia	Case-control	200 PTB adults (≥15 years) 40 healthy adult volunteers (≥18 years)	Morbidity during TB treatment Predictors of residual disability	High 6-month residual pulmonary disability Advanced disease at baseline and HIV infection predicted residual disability
Lee, 2012	Taiwan	Cohort	3176 TB adults 15880 non-TB adults	Risk factors of COPD	PTB and delay in TB treatment were risk factors of COPD
Ross, 2010	South Africa	Retrospective cohort	185 PTB gold miners 185 non-TB gold miners	Lung function loss Predictors of excess lung function decline	PTB associated with mean FEV_1 loss of 40.3 ml/year (95% CI: 25.4 – 55.1) and mean FVC loss of 42.7 ml/year (95% CI: 27.0 – 58.5) Severe clinical presentation of TB associated with greater lung function loss
Panda, 2016	India	Prospective cross-sectional	101 post-TB adults (18-65 years) with respiratory symptoms	Correlation coefficients between dyspnoea, radiological findings and pulmonary function tests (PFT)	Moderate correlation between radiological findings (fibrosis and bronchiectasis) and dyspnoea and not PFT

REFERENCES

1. World Health Organization. Global Tuberculosis Report. Geneva, Switzerland2018. Available from: http://apps.who.int/iris/bitstream/handle/10665/274453/9789241565646-eng.pdf?ua=1.

2. World Health Organization. The End TB Strategy. Geneva, Switzerland2015. Available from: http://www.who.int/tb/End_TB_brochure.pdf?ua=1.

3. van Kampen SC, Wanner A, Edwards M, Harries AD, Kirenga BJ, Chakaya J, et al. International research and guidelines on post-tuberculosis chronic lung disorders: a systematic scoping review. BMJ Global Health. 2018;3(4).

4. Byrne AL, Marais BJ, Mitnick CD, Lecca L, Marks GB. Tuberculosis and chronic respiratory disease: a systematic review. International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases. 2015;32:138-46.

5. Allwood BW, Myer L, Bateman ED. A systematic review of the association between pulmonary tuberculosis and the development of chronic airflow obstruction in adults. Respiration; international review of thoracic diseases. 2013;86(1):76-85.

6. Munoz-Torrico M, Rendon A, Centis R, D'Ambrosio L, Fuentes Z, Torres-Duque C, et al. Is there a rationale for pulmonary rehabilitation following successful chemotherapy for tuberculosis? Jornal brasileiro de pneumologia : publicacao oficial da Sociedade Brasileira de Pneumologia e Tisilogia. 2016;42(5):374-85.

7. Jones R, Kirenga BJ, Katagira W, Singh SJ, Pooler J, Okwera A, et al. A pre-post intervention study of pulmonary rehabilitation for adults with post-tuberculosis lung disease in Uganda. International journal of chronic obstructive pulmonary disease. 2017;12:3533-9.

8. Nair V, Patil S, Pratinidhi A, Pawar B, Jadhav A, Gaikwad R, et al. A cross-sectional study on the health related quality of life in patients who complete treatment for pulmonary tuberculosis. J Evolution Med Dent Sci. 2018;7(25):2939-43.

9. Meyyappan D, Chockalingam P. Evaluation of respiratory impairment and health related quality of life in pulmonary tuberculosis sequelae patients. 2018. 2018;5(2):5.

10. Vecino M, Pasipanodya JG, Slocum P, Bae S, Munguia G, Miller T, et al. Evidence for chronic lung impairment in patients treated for pulmonary tuberculosis. Journal of infection and public health. 2011;4(5-6):244-52.

11. World Health Organization. Chronic respiratory diseases. Available from: https://www.who.int/respiratory/en/.

12. Shankar M, Saha KK, Kumar P, Tiwari MK, Kumar S. Haemoptysis- aetiopathological, radiological profile and its outcome, our experience. J Evolution Med Dent Sci 2018;7(17):2078-83.

13. Kumar PVK, Gorantla R, Rachakonda R, Babu J, Venu M. Two years prospective observational study of haemoptysis at one of the tertiary health centres in Andhra Pradesh. Journal of Evolution of Research in Pulmonary Medicine. 2016;2(2, July-December 2016):1-5.

14. Nihues Sde S, Mancuzo EV, Sulmonetti N, Sacchi FP, Viana Vde S, Netto EM, et al. Chronic symptoms and pulmonary dysfunction in post-tuberculosis Brazilian patients. The Brazilian journal of infectious diseases : an official publication of the Brazilian Society of Infectious Diseases. 2015;19(5):492-7.

15. Santra A, Dutta P, Manjhi R, Pothal S. Clinico-Radiologic and Spirometric Profile of an Indian Population with Post-Tuberculous Obstructive Airway Disease. Journal of clinical and diagnostic research : JCDR. 2017;11(3):Oc35-oc8.

16. Ehrlich RI, Adams S, Baatjies R, Jeebhay MF. Chronic airflow obstruction and respiratory symptoms following tuberculosis: a review of South African studies. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2011;15(7):886-91.

17. Meghji J, Simpson H, Squire SB, Mortimer K. A Systematic Review of the Prevalence and Pattern of Imaging Defined Post-TB Lung Disease. PloS one. 2016;11(8):e0161176.

18. Seon HJ, Kim YI, Lim SC, Kim YH, Kwon YS. Clinical significance of residual lesions in chest computed tomography after anti-tuberculosis treatment. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2014;18(3):341-6.

19. Menon B, Nima G, Dogra V, Jha S. Evaluation of the radiological sequelae after treatment completion in new cases of pulmonary, pleural, and mediastinal tuberculosis. Lung India : official organ of Indian Chest Society. 2015;32(3):241-5.

20. Chakaya J, Kirenga B, Getahun H. Long term complications after completion of pulmonary tuberculosis treatment: A quest for a public health approach. Journal of Clinical Tuberculosis and Other Mycobacterial Diseases. 2016;3:10-2.

21. Gao YH, Guan WJ, Liu SX, Wang L, Cui JJ, Chen RC, et al. Aetiology of bronchiectasis in adults: A systematic literature review. Respirology (Carlton, Vic). 2016;21(8):1376-83.

22. Vogelmeier CF, Criner GJ, Martinez FJ, Anzueto A, Barnes PJ, Bourbeau J, et al. Global Strategy for the Diagnosis, Management, and Prevention of Chronic Obstructive Lung Disease 2017 Report. GOLD Executive Summary. American journal of respiratory and critical care medicine. 2017;195(5):557-82.

23. Sarkar M, Srinivasa, Madabhavi I, Kumar K. Tuberculosis associated chronic obstructive pulmonary disease. The clinical respiratory journal. 2017;11(3):285-95.

24. Buist AS, Vollmer WM, Sullivan SD, Weiss KB, Lee TA, Menezes AM, et al. The Burden of Obstructive Lung Disease Initiative (BOLD): rationale and design. Copd. 2005;2(2):277-83.

25. Chung K-P, Chen J-Y, Lee C-H, Wu H-D, Wang J-Y, Lee L-N, et al. Trends and predictors of changes in pulmonary function after treatment for pulmonary tuberculosis. Clinics. 2011;66(4):549-56.

26. Ross J, Ehrlich RI, Hnizdo E, White N, Churchyard GJ. Excess lung function decline in gold miners following pulmonary tuberculosis. Thorax. 2010;65(11):1010-5.

27. Pefura-Yone EW, Kengne AP, Tagne-Kamdem PE, Afane-Ze E. Clinical significance of low forced expiratory flow between 25% and 75% of vital capacity following treated pulmonary tuberculosis: a cross-sectional study. BMJ Open. 2014;4(7).

28. Ralph AP, Kenangalem E, Waramori G, Pontororing GJ, Sandjaja, Tjitra E, et al. High morbidity during treatment and residual pulmonary disability in pulmonary tuberculosis: under-recognised phenomena. PloS one. 2013;8(11):e80302.

29. Lee CH, Lee MC, Lin HH, Shu CC, Wang JY, Lee LN, et al. Pulmonary tuberculosis and delay in anti-tuberculous treatment are important risk factors for chronic obstructive pulmonary disease. PloS one. 2012;7(5):e37978.

30. Lee JJ, Chong PY, Lin CB, Hsu AH, Lee CC. High resolution chest CT in patients with pulmonary tuberculosis: characteristic findings before and after antituberculous therapy. European journal of radiology. 2008;67(1):100-4.

31. Majmudar DK, Rajput DK. Role of HRCT in diagnosing disease activity in pulmonary tuberculosis. International Journal of Contemporary Medical Research. 2017;4(8):1724-7.

32. Panda A, Bhalla AS, Sharma R, Mohan A, Sreenivas V, Kalaimannan U, et al. Correlation of chest computed tomography findings with dyspnea and lung functions in post-tubercular sequelae. Lung India : official organ of Indian Chest Society. 2016;33(6):592-9.

33. Bak SH, Kim S, Hong Y, Heo J, Lim MN, Kim WJ. Quantitative computed tomography features and clinical manifestations associated with the extent of bronchiectasis in patients with moderate-to-severe COPD. International journal of chronic obstructive pulmonary disease. 2018;13:1421-31.

PART C: MANUSCRIPT

Prediction of post-tuberculosis lung damage using computed tomography lung measures among adults in Malawi

Vester Gunsaru[1§]

[1]Division of Epidemiology and Biostatistics, School of Public Health and Family Medicine, University of Cape Town

Running head: PREDICTING PTLD USING CT MEASURES

Abstract: 197

Text: 4234

References: 33

Tables: 5

Figures: 2

Keywords: Tuberculosis, spirometry, machine learning, sub-Saharan Africa.

[§] Corresponding author:

Vester Gunsaru, Division of Epidemiology and Biostatistics, School of Public Health and Family Medicine, University of Cape Town, Level 5, Falmouth Building, Anzio Road, Observatory, Cape Town, South Africa. 0783841339 (gnsves001@myuct.ac.za)

[§]The Manuscript meets the requirements of the International Journal of Tuberculosis and Lung Disease (IJTLD) (Appendix D), except for an increased word count to meet the MPH mini-dissertation guidelines. In accordance with the MPH mini-dissertation guidelines, co-authors are not listed in the journal manuscript and the contribution of co-authors are listed in the acknowledgements section.

ABSTRACT

Setting: Tuberculosis (TB) burden is high in sub-Saharan Africa (SSA). Although TB treatment is widely available, and with high success rates, empirical evidence demonstrates that completion of treatment can mark the beginning of chronic post-TB lung damage (PTLD).

Objectives: To carry out feature selection and to evaluate the significance of computed tomography (CT) scan features in describing PTLD patterns at treatment completion and predicting PTLD outcomes at 12 months post-treatment.

Design: Data from a longitudinal study carried out in urban Malawi, on life post-TB treatment, was used. Machine learning techniques for cluster analysis and variable selection of important CT features were applied. The predictive performance of these features was evaluated in models using baseline abnormal spirometry and 12-month spirometry as outcomes.

Results: Among 346 adults, 24% had evidence of baseline abnormal spirometry. CT features were associated with baseline abnormal spirometry but not 12-month spirometry. For baseline abnormal spirometry groups, area under the curve ranged from 0.71 to 0.82 for three predictive models that were applied. The root mean squared error for 12-month spirometry ranged from 0.20 to 0.26.

Conclusion: The methods demonstrate promise in use of early imaging data to identify individuals at risk of PTLD.

INTRODUCTION

The disease Tuberculosis (TB) remains one of the major public health problems in the world. Recent statistics from the World Health Organization (WHO) show that nearly 10.0 million new TB infections and 1.6 million TB deaths were observed in 2017.[1] Strides towards treatment availability for the drug susceptible form of TB infection have been made over the years, with very high (82.0%) treatment success rates.[1] Unfortunately, patient clinical outcomes after treatment completion are rarely investigated, with much focus being only on the successful treatment completion.[2,3] Evidence has shown that TB treatment completion may mark the onset of chronic post-tuberculosis lung damage (PTLD) for many patients.[3-6]

Lung damage after TB treatment completion is manifested through various clinical characteristics and outcomes. Studies from South Korea, Cameroon, South Africa, India and Brazil have reported high prevalence (45.0 – 62.1%) of persistent respiratory symptoms such as coughing, bronchitis, asthma, dyspnoea, wheeze and expectoration in patients with a history of pulmonary TB (PTB) compared to patients without PTB.[7-10] In addition, obstructive, restrictive and mixed spirometry patterns suggesting compromised lung function have also been observed in these individuals in countries such as India, South Africa, Pakistan and the United States of America (USA).[4,8,10-12] Prevalence for such patterns of abnormal lung function are between 35.0-68.0% suggesting between a third to two-thirds of successfully treated TB patients suffer from long-term sequelae.[4,8,10-12] PTLD has also been characterised by parenchymal, airways and pleural lesions in lung-imaging features obtained from both chest X-rays (CXR) and computed tomography (CT).[6,13,14] These PTLD characteristics are associated with lung damage outcomes of chronic obstructive pulmonary disease (COPD) and bronchiectasis. These have been widely reported, with odds ratios (OR) as high as 6.3 and 4.6 for the association between a history of PTB and COPD, and bronchiectasis respectively.[5]

Post-TB individuals are at different risks of developing PTLD. One of the commonly reported risk factors of abnormal lung function after TB treatment is severe radiological damage, before or after TB treatment.[15-17] This suggests the importance that radiological features might have in evaluating lung damage and has been explored by several studies. Santra et al.[10] reported that the extent of radiological damage on CXR after TB treatment was associated with severe lung damage after TB treatment. A similar characterisation of radiological damage was seen to be associated with poor quality of life in post-TB individuals.[12] Furthermore, positive correlations and associations have been reported between CT radiological features and dyspnoea, pulmonary function, quality of life and bronchiectasis in post-TB individuals.[11,18,19]

The above associations were observed using correlation coefficients,[11,19] Chi-square tests of association[10,12] and Analysis of Variance (ANOVA).[12,18] Using prospectively collected data, we investigate the significance of radiological features from CT scan in describing PTLD patterns at treatment completion and predicting PTLD outcomes at 12 months post treatment, as a possible tool to detect risk of PTLD at an early stage. We used data-driven machine learning methods for variable reduction, selection and prediction. Radiological features on a CT scan were used because they provide more detail on lung damage compared to features on CXR.[6,12,20,21]

METHODS

STUDY SETTING AND PARTICIPANTS

This study is part of “The Life After Pulmonary TB” (LAT) study. The LAT study is a prospective cohort study of adults successfully completing TB treatment in urban Blantyre, Malawi. Individuals (≥15 years of age) with pulmonary TB, with a first episode of TB treatment and within two months of TB treatment success were eligible for recruitment from surrounding hospitals and clinics. Exclusions included no WHO danger signs or smear- or culture-positive TB. A total of 405 individuals were enrolled in the LAT study with all individuals followed for 12 months.

Informed consent was given by all participants prior to enrolment. Ethics approval for the LAT study was granted by the Malawi College of Medicine Research and Ethics committee (P.10/15/1813, Appendix B). This analysis received ethics approval from the Human Research Ethics Committee (HREC) of the Faculty of Health Sciences at the University of Cape Town (HREC REF: 177/2018, Appendix C).

STUDY MEASUREMENTS

Data for the LAT study were collected at three-time points; baseline, six and 12 months post-baseline. This analysis used a subset of the variables including baseline demographic characteristics, clinical outcomes, CT imaging features and spirometry data at baseline and 12 months.

Detailed questionnaires were administered by trained interviewers for demographic data. Spirometry was carried out using the *ndd EasyOne* spirometer and was done according to the American Thoracic Society (ATS) standards. Abnormal spirometry at baseline was defined by airway obstruction (FEV/FVC ratio< lower limit of normal (LLN) & FEV<LLN) or low FVC pattern (FVC<LLN but FEV1//FVC ratio>LLN). CT scans were carried out at baseline in two locations, with quality checked equipment. CT scans were read by qualified readers using a standardized form. CT variables were measured in percentages for parenchymal pathology, on a scoring of 0-18 for airways pathology and in frequencies for other pathology.

STATISTICAL ANALYSIS

Data processing

Observations with missing data in the characteristics of interest were dropped from all analyses. Data were missing mainly because other observations had no CT scans and no spirometry done. For variable reduction analyses, categorical CT imaging features were converted to numerical variables. Furthermore, all CT imaging features were cube root transformed to normality in these analyses. A cube root transformation was used because of the nature of the data, which had zeroes and negative values. For both variable reduction and predictive modelling analyses, data was centred and/or scaled where appropriate.

Patient characteristics

Descriptive statistics for the variables included in the analysis were computed as either mean and standard deviation (SD) or median and interquartile range (IQR) for continuous variables and as frequencies and percent for discrete variables.

Cluster analysis

Unsupervised clustering was used to find subgroups of individuals in the data. K-means and hierarchical clustering techniques were used.

K-means clustering works by first randomly choosing “K” user-defined observations which are used as initial cluster means.[22,23] Different K’s (2, 3 and 4) were initially specified. Each observation in the data is then assigned to their closest mean based on the distance between the observation and the mean.[22,23] The Euclidean distance was used for measuring this distance. New cluster mean values are then calculated and the clusters are updated by iteratively assigning observations to their closest cluster mean to minimize within-cluster variation.[22,23]

Agglomerative clustering was applied for hierarchical clustering.[24] With this method, each observation is initially treated as a cluster. Then two clusters that are close to each other are combined into a larger cluster.[24] This process is repeated until all observations form one cluster.[24] The Ward’s minimum variance was used as the criterion to select the clusters to combine at each stage. This minimizes within-cluster variation.[22]

For both techniques, the Elbow method[22] was applied to choose the optimal number of clusters.

Variable reduction and selection

Principal component analysis (PCA) was used for variable reduction. PCA is an unsupervised dimensionality reduction technique that creates components which are each a linear combination of the initial variables.[25,26] The principal components (PCs) produced, account for as much variation in the dataset as possible and are not correlated to each other.[26] The first PC accounts for the largest variation, the second PC accounts for the next largest remaining variation, and so on.[25,26]

Penalized regression (PR) and random forest (RF) were used for variable selection. These are both supervised learning techniques. PR uses penalties to regularize or control a model and improve the fit to the data.[26] This method works as a way of controlling for many predictors and collinearity among the predictors. The least absolute shrinkage and selection operator model (LASSO) was used. This has the advantage of achieving both regularization and variable selection.[26]

RF works by constructing several decision trees based on a random set of training data. The final prediction decision is made based on the decision of the majority of the trees.[27] RF has the advantage of improved predictive accuracy, variable selection and reduces correlation among predictors.[26]

Predictive modelling

Predictive models for baseline abnormal spirometry groups and spirometry at 12 months post-treatment were developed. The reduced variable set obtained from PCA was used in logistic and linear regression models. PR and RF models were also used as predictive models. The data sets used for predictive modelling were split into training (75%) and testing (25%) sets. A 10-fold cross validation method was used on the training data. Other covariates included in the prediction models were age, gender, body mass index (BMI), human immunodeficiency virus (HIV) status, smoking status and socioeconomic status (SES), to account for potential confounding. Predictive performance of classification models was assessed using receiver operating characteristic (ROC) curves and estimated area under the curve (AUC). For regression models, the predictive performance was assessed using the root mean squared error (RMSE). The level of statistical significance was set at α = 0.05. All analyses were performed using R software version 3.4.3.

RESULTS

PATIENT CHARACTERISTICS

Among the 405 individuals enrolled, 59 (14.6%) were excluded due to missing or incomplete CT data in the unsupervised learning analyses. Fifty-three (13.1%) more patients were excluded in the supervised learning analyses, for outcomes at baseline, due to missing data for demographic and clinical characteristics. For the outcomes at 12-month visit, the data set was further reduced because of missing outcome data. Overall, data from 346 (85.4%), 293 (72.3%) and 264 (65.2%) patients were used in the unsupervised, supervised learning for baseline outcomes and 12-month outcomes respectively. The patient characteristics of these reduced analysis sets were not different from the distribution of the patient characteristics in the total data set (Table D1 and D2).

Of the individuals enrolled, 313 (77.3%) had microbiologically confirmed TB (mTB) and 92 (22.7%) had clinically confirmed TB (cTB). The median age of the mTB group was 34.0 (IQR: 28.0, 39.0) compared to a median age of 38.0 (IQR: 32.6, 43.0) in the cTB group (Table D1). For both groups, there were more males than females. The majority of individuals in both groups had never smoked. A higher percentage (77.2%) of cTB individuals were HIV-infected compared to 55.3% for the mTB group.

SPIROMETRY CHARACTERISTICS

Spirometry characteristics at baseline, for both pre- and post-bronchodilator, were similar between the individual groups (Table D3). The mean pre- and post-bronchodilator forced expiratory volume in one second (FEV_1) across all groups was 2.6 litres (SD: 0.7) and the mean forced vital capacity (FVC) was 3.2 litres (SD: 0.8). The overall mean FEV_1/FVC ratio at baseline was above the LLN cut-off of 70%. Among all the individuals, 26.7% had abnormal spirometry at treatment completion.

CT IMAGING FEATURES

The majority of the individuals (385, 95.1%) had CT results available with a median time to scan after TB treatment completion of 4.1 weeks (IQR: 2.1, 7.0) (Table C1).

Abnormal parenchyma was common among the individuals (372, 96.6%), with a median percentage of abnormal parenchyma of 22.9 (IQR: 9.2, 39.2) (Table C1). In individuals with mTB, median percentage of abnormal parenchyma was higher (25.8 (IQR: 11.3, 40.2) versus 12.1 (IQR: 4.3, 34.0)). Atelectasis/banding was the most observed parenchymal pathology (12.1 (IQR: 5.0, 25.0)), followed by mosaicism (5.4 (IQR: 0.8, 14.2)).

The highest median airways pathology scores were observed for the tree-in-bud pathology (Table C1). The median tree-in-bud severity score was 3.5 (IQR: 1.5, 6.0), higher in the mTB group (4.0 (IQR: 1.5, 6.3)) than in the cTB patients (2.3, IQR: 0.6, 4.5). Cystic bronchiectasis was present in 12.7% of the total patients. In the microbiological and clinical diagnosis groups, 14.1% and 8.1% of the patients had cystic bronchiectasis respectively.

Few individuals (1.0%) had a miliary pattern at TB treatment completion (Table C1). Pleural pathology was present in 8.1% of the total individuals who had a CT scan.

VARIABLE REDUCTION

Variable reduction analyses used data from 346 individuals, who had a CT scan at TB treatment completion, and had complete CT imaging data.

PCA

Five PCs which had eigenvalues >1 were chosen from PCA results. These PCs explained 64.2% of the variance in the CT imaging data. The first PC was positively correlated (loadings > 0.3) with parenchymal variables (atelectasis/banding, abnormal parenchyma and abnormal parenchyma excluding mosaicism) and bronchiectasis (Table D4). The second PC was positively correlated with emphysema and pleural pathology and negatively correlated with mosaicism. As for the third PC, it was positively correlated with number of lobes with ≥90% destroyed parenchyma, pleural pathology and cystic bronchiectasis and negatively correlated with emphysema and miliary pattern. The fourth PC was positively correlated with time to CT scan after TB treatment completion and ground glass opacification. The fifth PC was positively correlated with ground glass opacification and negatively correlated with miliary pattern.

CLUSTER ANALYSIS

Three clusters by CT imaging features were identified for both clustering methods used. A cross-tabulation of the clusters by the two methods is presented in Table D7.

Hierarchical clustering

The clusters obtained from hierarchical clustering were plotted over the first two PCs (Figure C1-left). The patient and CT characteristics for each of these clusters have been presented in

Table C2. The cluster with the largest number of individuals was cluster 1 (175, 50.6%), followed by cluster 3 (102, 29.5%). Cluster 2 (69, 19.9%), had the highest median percentage of abnormal parenchyma (45.4 (IQR: 30.0, 64.2)), highest scores of airways pathology and other pathologies compared to the other two clusters. This group also had the lowest median BMI of 19.0 (IQR: 18.0, 21.1), a high percentage of males (88.4%) and a higher proportion of those who had ever smoked.

Cluster 1 was the group with fewer pathologies compared to the other two clusters. Patients in this cluster had the lowest median percentage of abnormal parenchyma (8.3, IQR: 3.8, 15.6) and lowest scores relating to airways pathology. These patients also had the highest median BMI of 20.9 (IQR: 19.2, 22.8) and a larger proportion of those who had never smoked.

Cluster 3 had characteristics that were between the other two clusters. Parenchymal pathology and airways pathology scores were less than those of cluster 2 but more than those of cluster 1. Patients in this cluster did not have other pathologies.

K-means clustering

Figure C1 (right) shows a plot of the clusters from K-means clustering over the first two PCs. Cluster 1 had 112 (32.4%) individuals, cluster 2 had 60 (17.3%) individuals and cluster 3 had 173 (50.3%) individuals (Table D6). Cluster 1 was the group with fewer pathologies, having the lowest median percentage of abnormal parenchyma (4.8 (IQR: 1.7, 8.9)) and lowest scores of airways pathology. Individuals in this group had a higher median BMI (21.1 (IQR: 19.4, 23.2)), a larger proportion of those who never smoked and those who were HIV-infected compared to the other clusters.

Cluster 2 was the group with more pathology. This group of patients had the highest median percentage of abnormal parenchyma (51.9 (IQR: 37.4, 67.7)) compared to the other two clusters. These individuals also had the highest scores of airways pathology and other pathologies except miliary pattern. They also had the lowest median BMI (18.9 (IQR: 18.0, 21.1)) and the largest proportion of males compared to the other two clusters

Cluster 3 had pathology scores that were between the other two groups, except for pleural pathology.

PREDICTIVE MODELS

A classification model was fit for the baseline abnormal spirometry outcome, and a regression model was fit for the 12-month spirometry outcomes. The models included other covariates, namely age, BMI, education, gender, SES, smoking status and HIV status.

Principal components regression

The five PCs obtained from PCA were used in principal components regression (PCR). Coefficients and odds ratios (OR) for the PCR models are presented in Table D5. The CT feature with the most important score in determining baseline abnormal spirometry was the first PC (Table C3). The third and fifth PCs also had higher scores. The other covariates which were important in the model were HIV-infected status and relative wealth. Performance was higher in the testing set compared to the training set (Figure C2-A). The AUC on the training set was 0.712 and the testing set had an AUC of 0.742.

The PCs were not important in the prediction of FEV_1 at 12 months (Table C3). Post-bronchodilator baseline FEV_1 was the most important variable, followed by HIV status.

RMSE associated with the model on the training set was 0.21 and the r-squared (R^2) was 0.89. The performance on the testing set was lower (RMSE = 0.25, R^2 = 0.87) than on the training set. In predicting 12-month FVC post-bronchodilator spirometry, the PCs were also of low importance. The most important variable was post-bronchodilator baseline FVC (Table C3). The RMSE on the training and testing set were 0.21 (R^2 = 0.92) and 0.22 (R^2 = 0.91) respectively.

PR and RF

For PR, the most important CT imaging features in determining baseline abnormal spirometry were pleural pathology and number of lobes with ≥90% destroyed parenchyma (Table C4). The other covariates that were important in classifying baseline abnormal spirometry were if an individual was HIV-infected, male or had ever smoked. The PR model on the training data had an AUC of 0.774 and an AUC of 0.809 on the testing data (Figure C2-B). Using RF, the most important CT imaging features in determining baseline abnormal spirometry were abnormal parenchyma (with or without mosaicism), mosaicism, atelectasis/banding, bronchiectasis, airway plugging, tree-in-bud, the time to CT scan and cavities (Table C5). The other important covariates were BMI, age and education years. The predictive performance of the RF model was 0.768 and 0.816 for the training and testing data respectively (Figure C2-C).

CT imaging features were not important in the prediction of 12-month post-bronchodilator FEV_1 for both PR and RF models. The most important variable for both methods was baseline post-bronchodilator FEV_1 (Table C4 and C5). The RMSEs from the PR model were 0.21 (R^2 = 0.91) and 0.26 (R^2 = 0.87) for the training and testing sets respectively. For the RF model, the RMSE was 0.23 (R^2 = 0.88) and 0.29 (R^2 = 0.83) for the training and testing set respectively. Baseline FVC was the most important variable in predicting 12-month FVC using both PR and RF. When using the PR model, the RMSE was 0.20 (R^2 = 0.93) for the training set and 0.26 (R^2 = 0.89) for the testing set. For the RF model, baseline FEV_1 was the second most important variable in predicting FVC (Table C5). CT imaging features had low importance scores for prediction of 12-month FVC in both models. The RMSE for the training and testing set was 0.24 (R^2 = 0.90) and 0.26 (R^2 = 0.89).

DISCUSSION

Radiological features, before and after TB treatment, have been shown to correlate with the severity of lung function abnormality after TB treatment.[10,15-17,19] This study used CT lung imaging features to describe patterns of PTLD at treatment completion and used those features to predict outcomes at 12 months post treatment. Three CT feature subgroups were identified, which on average demarcated patients with severe, moderate and low pathologies. Parenchymal pathology and bronchiectasis explained the largest variation in the data set and these were also important in predicting concurrent abnormal spirometry at TB treatment completion from. CT lung imaging features did not demonstrate utility in predicting spirometry at 12 months.

Cluster analysis has been recently used in epidemiological studies to identify important subgroups in a data driven approach.[28-32] To our knowledge, this study was the first to use cluster analysis on CT features to investigate post-TB lung function. Clusters with severe pathologies were associated with low BMI, male gender and low SES. The identified clusters could be distinguishing the characteristics of lung damage. On the other hand, these clusters could also be highlighting the internal biases of the CT reporters.

Guan et al.[32] used factor analysis for variable reduction and selection in a population of 148 patients with bronchiectasis for identification of bronchiectasis phenotypes. This method is similar to PCA. The initial variables included radiology, lung function and demographic characteristics, to mention just a few.[32] The first factor, which accounted for the greatest variation (63.9%) in the data, was associated with radiological and lung function characteristics.[32] From our results, a minimum of five PCs were required to reach that level of variance explanation, suggesting that there are other important variables that account for the variation in the data besides CT features.

All prediction models on abnormal spirometry at baseline performed better on testing data sets. Our results could mean that our models on abnormal spirometry at baseline are generalisable. However, because our testing data sets were small, these results could also suggest that our models were overfitting on the testing data sets. Nevertheless, overfitting is unlikely for the RF model because it combines multiple decision trees which help to create more accurate predictions.[27,33] For the outcomes at 12 months post treatment, the predictive performances of the models dropped on the testing data sets, as expected. Overall the predictive performances of the models were high, though the lack of a clear association between spirometry and patient outcomes limits the clinical utility of these results.

The relationship between imaging findings and spirometry outcomes after treatment is not clear. Studies that have looked at the relationship between radiological findings (CXR or CT) and PTLD have only identified correlations and associations.[10-12,18,19] Santra et al.[10] reported a significant association between CXR features of fibrosis, cavitation, calcification and bronchiectasis, and severe airflow limitation in a post-TB population with symptoms of airway obstruction disease. In a study by Panda et al.,[19] CT measurements of fibrosis, cavitation, bronchiectasis, consolidation, nodule and aspergilloma were negatively correlated with FEV_1 and FVC parameters. However, this correlation was not statistically significant. On the other hand, Meyyappan et al.[11] reported a significant correlation between skiagram radiological features and pulmonary function. In our study, bronchiectasis was an important CT feature in determining abnormal spirometry at baseline from PCR and RF models. In addition, parenchyma pathology variables were also important in determining abnormal spirometry at baseline from the same models. However, the important parenchymal variables in the present study were different from those as observed from the other studies. Pleural pathology was also observed as an important variable in determining abnormal spirometry at baseline from our PR model.

The significance of CT lung imaging findings in predicting long-term outcomes after TB treatment is also not well established. From our results, spirometry outcomes at 12 months post treatment were closely related with spirometry outcomes at baseline and not CT variables. This is not surprising because as time progresses after TB treatment completion, an individual's lung function status changes as well, either to recovery or more lung damage. This then would suggest that lung imaging findings soon after TB treatment completion would be of little or no use to explaining long-term lung function outcomes unless compared to long-term imaging findings.

The important CT variable sets for determining abnormal spirometry at baseline differed for each of the methods used. The differences observed could mean that there are different channels in the data of explaining the outcome of interest. In addition, the differences in the important variable sets could also be highlighting the inherent differences of the methods used. Firstly, compared to PR and RF, PCR used variables that were achieved using unsupervised learning. Secondly, even though PR and RF are both supervised learning techniques, they are different methods.

To our knowledge, this study is the first to investigate the significance of CT lung imaging features in predicting long-term PTLD outcomes. This kind of study is essential at a time when there is call for research on options for early detection and management of PTLD based on the overwhelming evidence of PTLD.[3] Furthermore, this study uses machine learning techniques that are data-driven and more applicable to high dimensional data. We also acknowledge that our study has some limitations. The sample size was reduced greatly due to missing data, which might affect some of our results. Some of the observations dropped were patients with missing data for bronchial wall thickening. These are probably those with the most severely damaged lung in whom the walls of the airways were not visible due to surrounding pathology. This then might have an impact on our results. However, a comparison of the other patient characteristics between those retained in the analysis and those dropped from the analysis did not show significant differences. The other limitation of our study is the availability of CT imaging data only at baseline. This did not allow us to evaluate how radiological lesions change over time.

Considering the high burden of TB in SSA, large numbers of individuals are at an increased risk of developing chronic lung damage even if successfully treated for TB. Early detection of PTLD may help in understanding the burden. Radiological features are a potential tool that can be used for this as they are associated with PTLD outcomes. In our findings, CT lung imaging features were able to determine abnormal spirometry soon after TB treatment but failed to predict spirometry outcomes at 12 months post treatment. These models might still be useful in early detection of PTLD, but further validation of the models and use of other PTLD outcomes is required.

CONCLUSION

There is need for early detection of PTLD in post-TB individuals to necessitate its appropriate management. We investigated the use of early imaging data to achieve this and our findings demonstrated this data could be very useful in identification of PTLD outcomes soon after TB treatment. More research is needed to investigate this same relationship in other settings as well as to focus on other PTLD outcomes.

REFERENCES

1. World Health Organization. Global Tuberculosis Report. Geneva, Switzerland2018. Available from: http://apps.who.int/iris/bitstream/handle/10665/274453/9789241565646-eng.pdf?ua=1.

2. World Health Organization. The End TB Strategy. Geneva, Switzerland2015. Available from: http://www.who.int/tb/End_TB_brochure.pdf?ua=1.

3. van Kampen SC, Wanner A, Edwards M, Harries AD, Kirenga BJ, Chakaya J, et al. International research and guidelines on post-tuberculosis chronic lung disorders: a systematic scoping review. BMJ Global Health. 2018;3(4).

4. Allwood BW, Myer L, Bateman ED. A systematic review of the association between pulmonary tuberculosis and the development of chronic airflow obstruction in adults. Respiration; international review of thoracic diseases. 2013;86(1):76-85.

5. Byrne AL, Marais BJ, Mitnick CD, Lecca L, Marks GB. Tuberculosis and chronic respiratory disease: a systematic review. International journal of infectious diseases : IJID : official publication of the International Society for Infectious Diseases. 2015;32:138-46.

6. Meghji J, Simpson H, Squire SB, Mortimer K. A Systematic Review of the Prevalence and Pattern of Imaging Defined Post-TB Lung Disease. PloS one. 2016;11(8):e0161176.

7. Ehrlich RI, Adams S, Baatjies R, Jeebhay MF. Chronic airflow obstruction and respiratory symptoms following tuberculosis: a review of South African studies. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2011;15(7):886-91.

8. Munoz-Torrico M, Rendon A, Centis R, D'Ambrosio L, Fuentes Z, Torres-Duque C, et al. Is there a rationale for pulmonary rehabilitation following successful chemotherapy for tuberculosis? Jornal brasileiro de pneumologia : publicacao oficial da Sociedade Brasileira de Pneumologia e Tisilogia. 2016;42(5):374-85.

9. Nihues Sde S, Mancuzo EV, Sulmonetti N, Sacchi FP, Viana Vde S, Netto EM, et al. Chronic symptoms and pulmonary dysfunction in post-tuberculosis Brazilian patients. The Brazilian journal of infectious diseases : an official publication of the Brazilian Society of Infectious Diseases. 2015;19(5):492-7.

10. Santra A, Dutta P, Manjhi R, Pothal S. Clinico-Radiologic and Spirometric Profile of an Indian Population with Post-Tuberculous Obstructive Airway Disease. Journal of clinical and diagnostic research : JCDR. 2017;11(3):Oc35-oc8.

11. Meyyappan D, Chockalingam P. Evaluation of respiratory impairment and health related quality of life in pulmonary tuberculosis sequelae patients. 2018. 2018;5(2):5.

12. Nair V, Patil S, Pratinidhi A, Pawar B, Jadhav A, Gaikwad R, et al. A cross-sectional study on the health related quality of life in patients who complete treatment for pulmonary tuberculosis. J Evolution Med Dent Sci. 2018;7(25):2939-43.

13. Menon B, Nima G, Dogra V, Jha S. Evaluation of the radiological sequelae after treatment completion in new cases of pulmonary, pleural, and mediastinal tuberculosis. Lung India : official organ of Indian Chest Society. 2015;32(3):241-5.

14. Seon HJ, Kim YI, Lim SC, Kim YH, Kwon YS. Clinical significance of residual lesions in chest computed tomography after anti-tuberculosis treatment. The international journal of tuberculosis

and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2014;18(3):341-6.

15. Chung K-P, Chen J-Y, Lee C-H, Wu H-D, Wang J-Y, Lee L-N, et al. Trends and predictors of changes in pulmonary function after treatment for pulmonary tuberculosis. Clinics. 2011;66(4):549-56.

16. Ralph AP, Kenangalem E, Waramori G, Pontororing GJ, Sandjaja, Tjitra E, et al. High morbidity during treatment and residual pulmonary disability in pulmonary tuberculosis: under-recognised phenomena. PloS one. 2013;8(11):e80302.

17. Ross J, Ehrlich RI, Hnizdo E, White N, Churchyard GJ. Excess lung function decline in gold miners following pulmonary tuberculosis. Thorax. 2010;65(11):1010-5.

18. Bak SH, Kim S, Hong Y, Heo J, Lim MN, Kim WJ. Quantitative computed tomography features and clinical manifestations associated with the extent of bronchiectasis in patients with moderate-to-severe COPD. International journal of chronic obstructive pulmonary disease. 2018;13:1421-31.

19. Panda A, Bhalla AS, Sharma R, Mohan A, Sreenivas V, Kalaimannan U, et al. Correlation of chest computed tomography findings with dyspnea and lung functions in post-tubercular sequelae. Lung India : official organ of Indian Chest Society. 2016;33(6):592-9.

20. Lee JJ, Chong PY, Lin CB, Hsu AH, Lee CC. High resolution chest CT in patients with pulmonary tuberculosis: characteristic findings before and after antituberculous therapy. European journal of radiology. 2008;67(1):100-4.

21. Majmudar DK, Rajput DK. Role of HRCT in diagnosing disease activity in pulmonary tuberculosis. International Journal of Contemporary Medical Research. 2017;4(8):1724-7.

22. Tan P-NAS, Michael; A Kumar, Vipin. Cluster Analysis: Basic Concepts and Algorithms. 2005 [cited 13 September, 2018]. In: Introduction to data mining [Internet]. US ed. [cited 13 September, 2018]; [487-568]. Available from: https://www-users.cs.umn.edu/~kumar001/dmbook/ch8.pdf.

23. Trevino A. Introduction to K-means Clustering 2016. Available from: https://www.datascience.com/blog/k-means-clustering.

24. Blei DM. Hierarchical Clustering 2008. Available from: http://www.cs.princeton.edu/courses/archive/spr08/cos424/slides/clustering-2.pdf.

25. Kassambara A. PCA - Principal Component Analysis Essentials 2017. Available from: http://www.sthda.com/english/wiki/practical-guide-to-principal-component-methods-in-r.

26. Kuhn M, Johnson K. Applied Predictive Modeling: Springer; 2013.

27. Polamuri S. How the random forest algorithm works in machine learning 2017. Available from: http://dataaspirant.com/2017/05/22/random-forest-algorithm-machine-learing/.

28. Lynch CM, van Berkel VH, Frieboes HB. Application of unsupervised analysis techniques to lung cancer patient data. PloS one. 2017;12(9):e0184370.

29. Kruse C, Eiken P, Vestergaard P. Clinical fracture risk evaluated by hierarchical agglomerative clustering. Osteoporosis international : a journal established as result of cooperation between the European Foundation for Osteoporosis and the National Osteoporosis Foundation of the USA. 2017;28(3):819-32.

30. Yang J, Angelini ED, Smith BM, Austin JHM, Hoffman EA, Bluemke DA, et al. Explaining Radiological Emphysema Subtypes with Unsupervised Texture Prototypes: MESA COPD Study. Medical computer vision and Bayesian and graphical models for biomedical imaging : MICCAI 2016

International Workshops, MCV and BAMBI, Athens, Greece, October 21, 2016, Revised selected papers 2017;2017:69-80.

31. Underwood J, Cole JH, Caan M, De Francesco D, Leech R, van Zoest RA, et al. Gray and White Matter Abnormalities in Treated Human Immunodeficiency Virus Disease and Their Relationship to Cognitive Function. Clinical infectious diseases : an official publication of the Infectious Diseases Society of America. 2017;65(3):422-32.

32. Guan WJ, Jiang M, Gao YH, Li HM, Xu G, Zheng JP, et al. Unsupervised learning technique identifies bronchiectasis phenotypes with distinct clinical characteristics. The international journal of tuberculosis and lung disease : the official journal of the International Union against Tuberculosis and Lung Disease. 2016;20(3):402-10.

33. Donges N. The Random Forest Algorithm 2018. Available from: https://towardsdatascience.com/the-random-forest-algorithm-d457d499ffcd.

TABLES

Table C1: Computed tomography imaging features at baseline by subsets of data used for analyses.

	Median (IQR)				
	Total N = 385	Clinically confirmed TB N = 86	Microbiologically confirmed TB N = 299	Unsupervised learning N = 346	Supervised learning N = 293
Time to CT scan (weeks)	4.1 (2.1, 7.0)	4.1 (2.1, 7.3)	4.3 (2.1, 7.0)	4.2 (2.0, 7.3)	4.3 (2.1, 7.0)
Parenchymal pathology					
Atelectasis or banding (%)	7.5 (2.9, 14.2)	4.2 (1.7, 9.0)	8.3 (3.5, 15.0)	6.7 (2.6, 11.7)	7.1 (2.9, 11.3)
Consolidation (%)	0.8 (0.0, 2.1)	0.4 (0.0, 1.7)	0.8 (0.0, 2.5)	0.8 (0.0, 1.7)	0.8 (0, 1.7)
Ground glass opacification (%)	0.0 (0.0, 0.8)	0.0 (0.0, 0.8)	0.0 (0.0, 0.8)	0.0 (0.0, 0.8)	0.0 (0.0, 0.8)
Mosaicism (%)	5.4 (0.8, 14.2)	1.5 (0.0, 9.4)	6.3 (1.3, 15.6)	5.0 (0.8, 12.5)	5.0 (0.8, 13.8)
Emphysema (%)	0.0 (0.0, 0.8)	0.0 (0.0, 0.4)	0.0 (0.0, 0.8)	0.0 (0.0, 0.8)	0.0 (0.0, 0.8)
Emphysematoid destruction (%)	0.0 (0.0, 0.4)	0.0 (0.0, 0.0)	0.0 (0.0, 0.4)	0.0 (0.0, 0.4)	0.0 (0.0, 0.4)
Cavities (%)	0.0 (0.0, 1.7)	0.0 (0.0, 0.4)	0.4 (0.0, 2.1)	0.0 (0.0, 1.7)	0.0 (0.0, 1.7)
Abnormal parenchyma (%)	22.9 (9.2, 39.2)	12.1 (4.3, 34.0)	25.8 (11.3, 40.2)	20.8 (8.3, 36.2)	21.7 (8.3, 36.3)
Abnormal parenchyma excluding mosaicism (%)	12.1 (5.0, 25.0)	8.8 (2.9, 19.9)	12.9 (6.0, 26.0)	10.4 (4.3, 21.7)	10.8 (5.0, 20.4)
Lobes with ≥90% destroyed parenchyma~	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)
Airways pathology					
Total airway plugging severity*	1.0 (0.0, 2.0)	0.5 (0.0, 1.5)	1.0 (0.0, 2.0)	1.0 (0.0, 2.0)	1.0 (0.0, 2.0)
Total tree-in-bud severity*	3.5 (1.5, 6.0)	2.3 (0.6, 4.5)	4.0 (1.5, 6.3)	3.5 (1.0, 6.0)	3.5 (1.5, 6.0)
Total lung bronchiectasis dilatation severity*	2.5 (0.5, 5.0)	1.0 (0.0, 3.0)	2.5 (1.0, 5.0)	2.0 (0.5, 4.5)	2.0 (0.5, 4.5)
	N (%)				
Any cystic bronchiectasis					
Absent	336 (87.3)	79 (91.9)	257 (86.0)	307 (88.7)	261 (89.1)
Present	49 (12.7)	7 (8.1)	42 (14.1)	39 (11.3)	32 (10.9)
Any bronchial wall thickening					
Absent	252 (65.5)	61 (70.9)	191 (63.9)	252 (72.8)	215 (73.4)
Present	94 (24.4)	17 (19.8)	77 (25.8)	94 (27.2)	78 (26.6)
Missing	39 (10.1)	8 (9.3)	31 (10.4)	0 (0.0)	0 (0.0)
Other pathologies					
Any miliary pattern					
Absent	381 (99.0)	84 (97.7)	297 (99.3)	342 (98.8)	290 (99.0)
Present	4 (1.0)	2 (2.3)	2 (0.7)	4 (1.2)	3 (1.0)
Any pleural pathology					
Absent	354 (92.0)	80 (93.0)	274 (91.6)	319 (92.2)	272 (92.8)
Present	31 (8.1)	6 (7.0)	25 (8.4)	27 (7.8)	21 (7.2)

~due to cavities/atelectasis/banding, *severity score 0-18, N: Number of observations, IQR: Interquartile range, CT: Computed tomography

Table C2: Computed tomography imaging features and patient characteristics for the clusters obtained using hierarchical clustering.

	Median (IQR)		
	Cluster 1 N = 175	Cluster 2 N = 69	Cluster 3 N = 102
Time to CT scan (weeks)	4.3 (1.9, 8.4)	4.3 (2.1, 7.6)	4.0 (2.1, 6.5)
Atelectasis or banding (%)	2.9 (1.3, 6.5)	15.0 (10.4, 26.3)	8.8 (6.3, 13.3)
Consolidation (%)	0.0 (0.0, 0.4)	2.5 (0.8, 5.0)	1.3 (0.8, 2.8)
Ground glass opacification (%)	0.0 (0.0, 0.8)	0.4 (0.0, 1.7)	0.0 (0.0, 0.3)
Mosaicism (%)	1.3 (0.0, 7.1)	4.6 (0.8, 14.6)	11.3 (5.6, 21.7)
Emphysema (%)	0.0 (0.0, 0.4)	0.0 (0.0, 5.0)	0.0 (0.0, 0.4)
Emphysematoid destruction (%)	0.0 (0.0, 0.0)	1.7 (0.0, 6.3)	0.0 (0.0, 0.4)
Cavities (%)	0.0 (0.0, 0.0)	3.3 (1.7, 8.3)	0.8 (0.0, 1.7)
Abnormal parenchyma (%)	8.3 (3.8, 15.6)	45.4 (30.0, 64.2)	29.2 (19.9, 37.0)
Abnormal parenchyma excluding mosaicism (%)	4.6 (2.1, 9.4)	36.7 (23.3, 49.2)	15.0 (9.6, 19.9)
Lobes with ≥90% destroyed parenchyma~	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)	0.0 (0.0, 0.0)
Total airway plugging severity*	0.0 (0.0, 0.5)	2.0 (1.0, 3.0)	2.0 (1.0, 3.5)
Total tree-in-bud severity*	1.5 (0.5, 3.0)	6.0 (3.5, 9.0)	5.0 (4.0, 6.5)
Total lung bronchiectasis dilatation severity*	0.5 (0.0, 2.0)	5.5 (3.5, 8.0)	3.5 (2.0, 5.4)
Age (years)	35.0 (28.5, 41.0)	35.0 (30.0, 41.0)	35.0 (27.3, 40.8)
Body mass index (kg/m^2)	20.9 (19.2, 22.8)	19.0 (18.0, 21.1)	20.4 (19.5, 22.3)
Education (years)	10.0 (8.0, 12.0)	10.0 (6.0, 12.0)	10.0 (7.0, 12.0)
	N (%)		
Any cystic bronchiectasis			
Absent	175 (100.0)	30 (43.5)	102 (100.0)
Present	0 (0.0)	39 (56.5)	0 (0.0)
Any bronchial wall thickening			
Absent	164 (93.7)	27 (39.1)	61 (59.8)
Present	11 (6.3)	42 (60.9)	41 (40.2)
Any miliary pattern			
Absent	175 (100.0)	65 (94.2)	102 (100.0)
Present	0 (0.0)	4 (5.8)	0 (0.0)
Any pleural pathology			
Absent	160 (91.4)	57 (82.6)	102 (100.0)
Present	15 (8.6)	12 (17.4)	0 (0.0)
Gender			
Male	118 (67.4)	61 (88.4)	58 (56.9)
Female	57 (32.6)	8 (11.6)	44 (43.1)
Socio-economic status			
Poorest quintile	7 (4.0)	5 (7.2)	6 (5.9)
2nd Poorest Quintile	30 (17.1)	20 (29.0)	26 (25.5)
Middle Quintile	40 (22.9)	17 (24.6)	25 (24.5)
2nd richest Quintile	57 (32.6)	16 (23.2)	28 (27.5)
Richest Quintile	30 (17.1)	7 (10.1)	12 (11.8)
Missing	11 (6.3)	4 (5.8)	5 (4.9)
Smoking status			
Never	138 (78.9)	37 (53.6)	73 (71.6)
Ever	35 (20.0)	30 (43.5)	28 (27.5)
Current	2 (1.1)	2 (2.9)	1 (1.0)
HIV status			
Non-infected	52 (29.7)	33 (47.8)	43 (42.2)
Infected	122 (69.7)	36 (52.2)	59 (57.8)
Missing	1 (0.6)	0 (0.0)	0 (0.0)

N: Number of observations, IQR: Interquartile range, CT: Computed tomography

Table C3: Variable importance scores obtained by principal components regression for abnormal baseline spirometry, 12-month FEV_1 and FVC outcomes.

	Abnormal baseline spirometry	12-month FEV_1	12-month FVC
Principal component 1	100.0	13.1	17.4
HIV status			
Infected	85.8	21.3	23.7
Socio-economic status			
2^{nd} Richest quintile	64.6	0.5	2.1
Richest quintile	63.8	3.1	2.2
Principal component 3	35.7	2.1	1.7
Principal component 5	34.7	0.0	8.3
Age	29.4	10.5	8.8
Principal component 4	25.2	4.8	0.0
Principal component 2	21.1	12.7	1.5
Gender			
Male	19.6	19.5	20.4
Socio-economic status			
Middle quintile	18.2	1.7	2.6
Education years	17.7	8.8	8.5
Smoking status			
Ever	13.0	7.8	3.7
Never	7.7	9.3	5.9
Body mass index	3.9	0.5	6.2
Socio-economic status			
Poorest quintile	0.0	8.2	10.7
Post-bronchodilator baseline spirometry			
FEV_1	-	100.0	14.8
FVC	-	5.0	100.0

Table C4: Variable importance scores obtained by penalized regression for abnormal baseline spirometry, 12-month FEV_1 and FVC outcomes.

	Abnormal baseline spirometry	12-month FEV_1	12-month FVC
Any pleural pathology			
Present	100.0	8.8	15.2
HIV status			
Infected	79.2	8.4	12.9
Lobes with ≥90% destroyed parenchyma	64.1	1.4	9.2
Gender			
Male	45.3	12.6	11.6
Smoking status			
Ever	27.0	0.0	0.0
Socio-economic status			
Richest quintile	24.9	0.9	1.1
Any cystic bronchiectasis			
Present	14.3	6.2	0.0
Socio-economic status			
2nd richest quintile	11.1	0.8	0.0
Smoking status			
Never	10.5	1.3	1.9
Socio-economic status			
Poorest quintile	8.6	6.5	4.2
Emphysematoid destruction (%)	4.5	0.0	1.0
Total lung bronchiectasis dilatation severity	3.7	0.0	0.4
Education years	3.7	0.4	0.6
Cavities (%)	3.4	0.0	0.7
Emphysema (%)	3.3	0.1	0.0
Any bronchial wall thickening			
Present	3.1	0.0	0.0
Atelectasis or banding (%)	2.5	0.0	0.0
Body mass index	1.9	0.3	0.7
Total airway plugging severity	1.6	0.3	0.0
Mosaicism (%)	1.4	0.1	0.4
Abnormal parenchyma (%)	0.9	0.0	0.1
Abnormal parenchyma excluding mosaicism (%)	0.5	0.0	0.0
Total tree-in-bud severity	0.5	0.5	0.9
Consolidation (%)	0.5	0.0	0.0
Time to CT scan (weeks)	0.4	0.2	0.1
Ground glass opacification (%)	0.3	0.4	0.1
Age	0.1	0.3	0.2
Socio-economic status			
Middle quintile	0.0	0.0	3.5
Any miliary pattern			
Present	0.0	0.0	17.5
Post-bronchodilator baseline spirometry			
FEV_1	-	100.0	16.6
FVC	-	0.0	100.0

CT: Computed tomography

Table C5: Variable importance scores obtained by random forest for abnormal baseline spirometry, 12-month FEV_1 and FVC outcomes.

	Abnormal baseline spirometry	12-month FEV_1	12-month FVC
Abnormal parenchyma (%)	100.0	8.7	13.9
Mosaicism (%)	70.1	11.1	10.0
Abnormal parenchyma excluding mosaicism (%)	69.3	8.8	9.0
Atelectasis or banding (%)	64.6	7.0	10.4
Total lung bronchiectasis dilatation severity	52.5	10.6	10.9
Body mass index	47.8	6.5	9.1
Total airway plugging severity	46.1	10.4	7.5
Total tree-in-bud severity	45.5	4.4	11.4
Time to CT scan (weeks)	44.7	3.2	0.0
Education years	44.3	13.2	8.2
Age	44.3	7.6	5.1
Cavities (%)	40.5	2.5	6.6
Consolidation (%)	39.9	6.5	7.1
Ground glass opacification	25.8	3.9	0.7
HIV status			
Infected	21.9	5.5	13.5
Emphysematoid destruction (%)	20.0	2.9	4.7
Emphysema (%)	19.5	0.0	8.7
Any bronchial wall thickening			
Present	18.3	4.3	6.0
Socio-economic status			
Richest quintile	15.5	3.1	5.8
Middle quintile	14.7	5.3	8.1
2^{nd} Richest quintile	12.3	2.1	1.7
Gender			
Male	10.8	13.8	20.8
Any cystic bronchiectasis			
Present	10.6	1.6	6.4
Any pleural pathology			
Present	10.1	4.6	14.9
Smoking status			
Ever	10.0	7.0	7.7
Never	6.8	7.4	11.0
Lobes with ≥90% destroyed parenchyma	3.2	3.1	5.8
Socio-economic status			
Poorest quintile	3.1	5.2	8.2
Any miliary pattern			
Present	0.0	4.2	7.7
Post-bronchodilator baseline			
spirometry	-	100.0	35.2
FEV1	-	12.5	100.0
FVC			

CT: Computed tomography

FIGURES

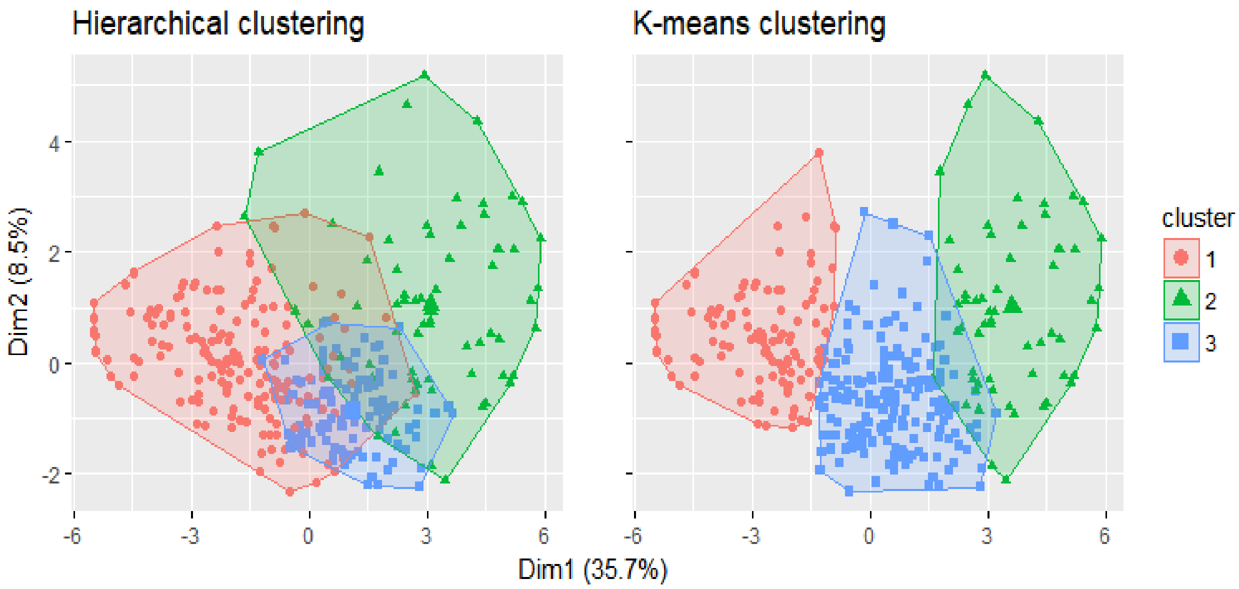

Figure C1: Cluster plots for K-means and hierarchical clustering on the first two principal components.

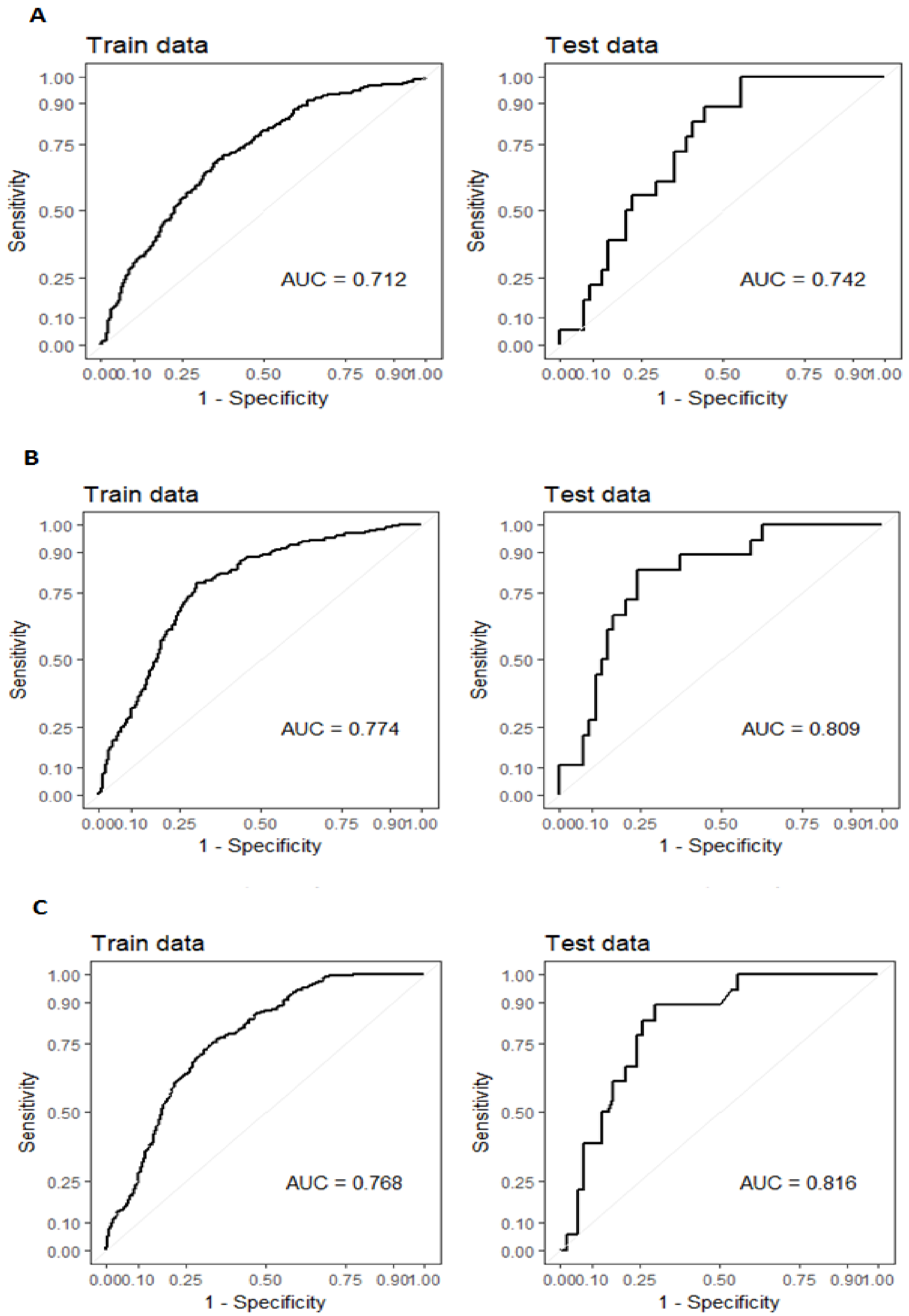

Figure C2: ROC curves for A) principal components regression, B) penalized regression and C) random forest on abnormal baseline spirometry.

PART D: APPENDICES

APPENDIX A: SUPPLEMENTARY TABLES

Table D1: Patient characteristics at baseline by TB diagnosis and subsets of data used for analyses.

	Total N = 405	Clinically confirmed TB N = 92	Microbiologically confirmed TB N = 313	Unsupervised learning N = 346	Supervised learning N = 293
			Median (IQR)		
Age (years)	35.0 (28.0, 41.0)	38.0 (32.6, 43)	34.0 (28.0, 39.0)	35.0 (29.0, 41.0)	35.0 (29.0, 41.0)
Body mass index (kg/m^2)	20.5 (19.0, 22.3)	20.9 (18.9, 22.7)	20.4 (19.0, 22.1)	20.5 (19.0, 22.3)	20.4 (19.0, 22.3)
Education (years)	10.0 (7.0, 12.0)	10.0 (8.0, 12.0)	10.0 (7.0, 12.0)	10.0 (7.3, 12.0)	10.0 (7.0, 12.0)
			N (%)		
Gender					
Males	275 (67.9)	59 (64.1)	216 (69.0)	237 (68.5)	204 (69.6)
Females	130 (32.1)	33 (35.9)	97 (31.0)	109 (31.5)	89 (30.4)
Socio-economic status					
Poorest Quintile	22 (5.4)	2 (2.2)	20 (6.4)	18 (5.2)	15 (5.1)
2nd Poorest Quintile	85 (21.0)	14 (15.2)	71 (22.7)	76 (22.0)	69 (23.5)
Middle Quintile	95 (23.5)	15 (16.3)	80 (25.6)	82 (23.7)	77 (26.3)
2nd richest Quintile	114 (28.1)	35 (38.0)	79 (25.2)	101 (29.2)	89 (30.4)
Richest Quintile	56 (13.8)	19 (20.6)	37 (11.8)	49 (14.2)	43 (14.7)
Missing	33 (8.1)	7 (7.6)	26 (8.3)	20 (5.8)	0 (0.0)
Smoking status					
Never	285 (70.4)	77 (83.7)	208 (66.5)	248 (71.7)	209 (71.3)
Ever	113 (27.9)	14 (15.2)	99 (31.6)	93 (26.9)	79 (27.0)
Current	7 (1.7)	1 (1.1)	6 (1.9)	5 (1.4)	5 (1.7)
HIV status					
Non-infected	159 (39.3)	20 (21.7)	139 (44.4)	128 (37.0)	105 (35.8)
Infected	244 (60.2)	71 (77.2)	173 (55.3)	217 (62.7)	188 (64.2)
Missing	2 (0.5)	1 (1.1)	1 (0.3)	1 (0.3)	0 (0.0)
Spirometry done					
No	40 (9.9)	10 (10.9)	30 (9.6)	35 (10.1)	0 (0.0)
Yes	365 (90.1)	82 (89.1)	283 (90.4)	311 (89.9)	293 (100.0)
CT done					
No	20 (4.9)	6 (6.5)	14 (4.5)	0 (0.0)	0 (0.0)
Yes	385 (95.1)	86 (93.5)	296 (95.5)	346 (100.0)	293 (100.0)
Baseline abnormal spirometry					
No	257 (63.5)	63 (68.5)	194 (62.0)	229 (66.2)	218 (74.4)
Yes	108 (26.7)	19 (20.7)	89 (28.4)	82 (23.7)	75 (25.6)
Missing	40 (9.9)	10 (10.9)	30 (9.6)	35 (10.1)	0.0 (0.0)

N: Number of observations, IQR: Interquartile range, CT: Computed tomography

Table D2: Patient characteristics at 12 months post-treatment.

	Median (IQR)	
	Total N = 405	12-month analyses set N = 264
Age (years)	35.0 (28.0, 41.0)	35.0 (29.0, 42.0)
Body mass index (kg/m^2)	20.5 (19.0, 22.3)	20.4 (19.0, 22.3)
Education (years)	10.0 (7.0, 12.0)	10.0 (7.0, 12.0)
	N (%)	
Gender		
Males	275 (67.9)	188 (71.2)
Females	130 (32.1)	76 (28.8)
Socio-economic status		
Poorest Quintile	22 (5.4)	14 (15.3)
2nd Poorest Quintile	85 (21.0)	56 (21.2)
Middle Quintile	95 (23.5)	72 (27.3)
Richest Quintile	56 (13.8)	4 (15.2)
2nd Richest Quintile	114 (28.1)	82 (31.1)
Missing	33 (8.1)	0 (0.0)
Smoking status		
Never	285 (70.4)	186 (70.5)
Ever	113 (27.9)	73 (27.6)
Current	7 (1.7)	5 (1.9)
HIV status		
Non-infected	159 (39.3)	93 (35.2)
Infected	244 (60.2)	171 (64.8)
Missing	2 (0.5)	0 (0.0)
Spirometry done		
No	40 (9.9)	0 (0.0)
Yes	365 (90.1)	264 (100.0)
CT done		
No	20 (4.9)	0 (0.0)
Yes	385 (95.1)	264 (100.0)

N: Number of observations, IQR: Interquartile range, CT: Computed tomography

Table D3: Spirometry characteristics at baseline by subsets of data used for analyses.

	Mean (SD)				
	Total N = 405	Clinically confirmed TB N = 92	Microbiologically confirmed TB N = 313	Unsupervised learning N = 346	Supervised learning N = 293
Pre-bronchodilator					
FEV_1 (litres)	2.6 (0.7)	2.6 (0.7)	2.6 (0.7)	2.6 (0.7)	2.6 (0.7)
FVC (litres)	3.2 (0.8)	3.2 (0.8)	3.2 (0.8)	3.2 (0.8)	3.2 (0.8)
FEV_1/FVC (%)	80.3 (9.2)	80.6 (8.3)	80.2 (9.4)	80.4 (9.3)	80.4 (9.3)
Missing / N (%)	46 (11.4)	8 (8.8)	36 (11.5)	41 (11.8)	6 (2.0)
Post-bronchodilator					
FEV_1 (litres)	2.6 (0.7)	2.6 (0.7)	2.6 (0.7)	2.6 (0.7)	2.6 (0.7)
FVC (litres)	3.2 (0.8)	3.2 (0.8)	3.2 (0.8)	3.2 (0.8)	3.2 (0.8)
FEV_1/FVC (%)	80.9 (8.9)	81.0 (8.2)	80.9 (9.1)	81.0 (9.1)	81.0 (9.1)
Missing / N (%)	40 (9.9)	10 (10.9)	30 (9.6)	35 (10.1)	0 (0.0)

N: Number of observations, SD: Standard deviation

Table D4: Principal component loadings for the first five principal components.

	Principal component 1	Principal component 2	Principal component 3	Principal component 4	Principal component 5
Time to CT scan (weeks)	0.0058	-0.2305	-0.0246	0.6708	-0.2040
Atelectasis or banding (%)	0.3355	0.0297	-0.0064	-0.0484	0.1310
Consolidation (%)	0.2802	0.0660	-0.2273	-0.1655	-0.0588
Ground glass opacification	0.1070	0.1156	-0.1361	0.6334	0.3555
Mosaicism (%)	0.1906	-0.5061	-0.0249	-0.1098	-0.0703
Emphysema (%)	0.0714	0.4566	-0.3823	-0.0387	-0.0939
Emphysematoid destruction (%)	0.2261	0.2765	-0.0479	0.0070	0.1925
Cavities (%)	0.2878	0.0689	0.2911	-0.0576	-0.1409
Abnormal parenchyma (%)	0.3596	-0.0440	-0.1371	0.0205	0.1026
Abnormal parenchyma excluding mosaicism (%)	0.3563	0.2079	-0.1172	0.0841	0.1352
Lobes with ≥90% destroyed parenchyma	0.1436	0.0402	0.4523	0.2660	-0.2856
Total lung bronchiectasis dilatation severity	0.3296	-0.1179	0.0052	-0.0134	0.0788
Any cystic bronchiectasis	0.1923	0.1805	0.3944	-0.0162	-0.2066
Total airway plugging severity	0.2577	-0.2582	-0.0624	-0.0979	-0.0559
Total tree-in-bud severity	0.2698	-0.2732	-0.1767	-0.0473	-0.1313
Any miliary pattern	-0.0036	0.1902	-0.3678	0.0834	-0.7379
Any pleural pathology	0.0777	0.3342	0.3495	-0.0652	-0.0044
Any bronchial wall thickening	0.2329	-0.0125	0.1006	0.0213	-0.1287

CT: Computed tomography

Table D5: Odds ratios, coefficients and 95% confidence intervals using principal components regression for abnormal baseline spirometry, 12-month FEV_1 and FVC outcomes.

	Abnormal baseline spirometry OR (95% CI)	12-month FEV_1 COEF (95% CI)	12-month FVC COEF (95% CI)
Intercept	26.44 (0.13, 8992.66)	0.32 (-0.04, 0.69)	0.49 (0.15, 0.83)
Principal component 1	1.45 (1.14, 1.91)	0.01 (-0.00, 0.03)	0.02 (0.01, 0.04)
Principal component 2	0.86 (0.53, 1.38)	-0.03 (-0.05, 0.00)	-0.00 (-0.03, 0.02)
Principal component 3	1.36 (0.77, 2.51)	-0.00 (-0.03, 0.02)	0.01 (-0.02, 0.03)
Principal component 4	0.85 (0.55, 1.31)	0.01 (-0.02, 0.04)	-0.00 (-0.03, 0.03)
Principal component 5	1.58 (0.67, 3.99)	-0.00 (-0.04, 0.04)	-0.02 (-0.05, 0.01)
Age (years)	0.98 (0.93, 1.03)	-0.00 (-0.01, 0.00)	-0.00 (-0.01, 0.00)
Body mass index (kg/m^2)	0.99 (0.82, 1.20)	-0.00 (-0.01, 0.01)	-0.01 (-0.02, 0.01)
Education (years)	0.96 (0.82, 1.12)	0.01 (-0.00, 0.02)	-0.01 (-0.02, 0.00)
Gender			
Female	Ref		
Male	0.71 (0.21, 2.30)	0.14 (0.04, 0.23)	0.15 (0.06, 0.25)
Socio-economic status			
2nd poorest quintile	Ref		
Poorest quintile	NA	-0.08 (-0.22, 0.05)	-0.12 (-0.26, 0.02)
Middle quintile	0.68 (0.16, 2.80)	0.01 (-0.07, 0.10)	0.02 (-0.07, 0.11)
2nd richest quintile	0.28 (0.07, 1.03)	0.00 (-0.09, 0.09)	-0.02 (-0.11, 0.07)
Richest quintile	0.19 (0.03, 1.04)	-0.03 (-0.14, 0.09)	-0.03 (-0.14, 0.09)
Smoking status			
Current	Ref		
Never	0.71 (0.02, 11.84)	-0.17 (-0.41, 0.07)	-0.10 (-0.29, 0.09)
Ever	0.55 (0.02, 10.18)	-0.14 (-0.39, 0.10)	-0.07 (-0.26, 0.13)
HIV status			
Non-infected	Ref		
Infected	0.27 (0.09, 0.74)	0.10 (0.04, 0.17)	0.12 (0.06, 0.19)
Post-bronchodilator baseline spirometry			
FEV_1	-	0.96 (0.83, 1.10)	0.15 (0.02, 0.27)
FVC	-	-0.04 (-0.15, 0.07)	0.83 (0.72, 0.94)

OR: Odds ratio, COEF: Coefficient, CI: Confidence interval

Table D6: CT imaging features and patient characteristics for the clusters obtained using K-means clustering.

	Median (IQR)		
	Cluster 1 N = 112	Cluster 2 N = 60	Cluster 3 N = 174
Time to CT scan (weeks)	4.1 (1.6, 6.6)	4.3 (1.9, 7.7)	4.3 (2.5, 9.1)
Atelectasis or banding (%)	2.1 (0.7, 3.3)	17.5 (11.2, 27.1)	7.9 (5.4, 11.3)
Consolidation (%)	1.7 (0.8, 3.5)	3.1 (1.6, 5.0)	0.8 (0.4, 1.7)
Ground glass opacification (%)	0.0 (0.0, 0.4)	0.4 (0.0, 1.7)	0.0 (0.0, 0.8)
Mosaicism (%)	0.4 (0.0, 2.1)	5.8 (1.2, 15.7)	9.6 (3.8, 19.9)
Emphysema (%)	0.0 (0.0, 0.8)	0.0 (0.0, 2.4)	0.0 (0.0, 0.4)
Emphysematoid destruction (%)	0.0 (0.0, 0.0)	3.3 (0.8, 6.7)	0.0 (0.0, 0.0)
Cavities (%)	0.0 (0.0, 0.0)	5.0 (2.5, 9.2)	0.4 (0.0, 1.3)
Abnormal parenchyma (%)	4.8 (1.7, 8.9)	51.9 (37.4, 67.7)	25.8 (16.3, 35.31)
Abnormal parenchyma excluding mosaicism (%)	2.9 (1.3, 5.1)	39.6 (28.7, 52)	12.5 (8.3, 17.5)
Lobes with ≥90% destruction parenchyma~	0.0 (0.0, 0.0)	0.0 (0.0, 0.3)	0.0 (0.0, 0.0)
Total airway plugging severity*	0.0 (0.0, 0.1)	2.0 (1.0, 3.6)	1.5 (0.5, 2.0)
Total tree-in-bud severity*	1.0 (0.0, 1.5)	6.0 (3.0, 8.5)	4.5 (3.0, 6.5)
Total lung bronchiectasis dilatation severity*	0.0 (0.0, 0.6)	6.5 (4.5, 8.5)	2.5 (2.0, 4.0)
Age	34.0 (29.8, 40.3)	36.5 (30.0, 41.3)	35.0 (28.0, 40.8)
Body mass index	21.1 (19.4, 23.2)	18.9 (18.0, 21.1)	20.4 (19.3, 22.2)
Education years	10.0 (8.0, 12.0)	9.5 (5.0, 12.0)	10.0 (7.0, 12.0)
	N (%)		
Any cystic bronchiectasis			
Absent	112 (100.0)	26 (43.3)	169 (97.1)
Present	0 (0.0)	34 (56.7)	5 (2.9)
Any bronchial wall thickening			
Absent	109 (97.3)	16 (26.7)	127 (73.0)
Present	3 (2.7)	44 (73.3)	47 (27.0)
Any miliary pattern			
Absent	110 (98.2)	60 (100.0)	172 (98.9)
Present	2 (1.8)	0 (0.0)	2 (1.1)
Any pleural pathology			
Absent	104 (92.9)	48 (80.0)	167 (96.0)
Present	8 (7.1)	12 (20.0)	7 (4.0)
Gender			
Male	72 (64.3)	52 (86.7)	113 (64.9)
Female	40 (35.7)	8 (13.3)	61 (35.1)
Socio-economic status			
Poorest quintile	4 (3.6)	5 (8.3)	9 (5.2)
2nd Poorest Quintile	15 (13.4)	16 (26.7)	45 (25.9)
Middle Quintile	22 (19.6)	15 (25.0)	45 (25.9)
2nd richest Quintile	39 (34.8)	14 (23.3)	48 (27.6)
Richest Quintile	23 (20.5)	5 (8.3)	21 (12.1)
Missing	9 (8.0)	5 (8.3)	6 (3.4)
Smoking status			
Never	92 (82.1)	32 (53.3)	124 (71.3)
Ever	20 (17.9)	26 (43.3)	47 (27.0)
Current	0 (0.0)	2 (3.3)	3 (1.7)
HIV status			
Negative	25 (22.3)	30 (50.0)	73 (42)
Positive	86 (76.8)	30 (50.0)	101 (58.0)
Missing	1 (0.9)	0 (0.0)	0 (0.0)

N: Number of observations, IQR: Interquartile range, CT: Computed tomography

Table D7: Cross-tabulation of the clusters identified by K-means and hierarchical clustering methods.

		Hierarchical clustering			
		Cluster 1	Cluster 2	Cluster 3	Total
K-means clustering	Cluster 1	109	2	1	112
	Cluster 2	0	56	4	60
	Cluster 3	66	11	97	174
	Total	175	69	102	346

APPENDIX B: LAT ETHICS APPROVAL CERTIFICATE


CERTIFICATE OF ETHICS APPROVAL

This is to certify that the College of Medicine Research and Ethics Committee (COMREC) has reviewed and approved a study entitled:

P.10/15/1813 – Life after pulmonary TB: An evaluation of lung health, physical, and economic well being by Dr. J. Meghji

On 5th January 2016

As you proceed with the implementation of your study, we would like you to adhere to international ethical guidelines, national guidelines and all requirements by COMREC as indicated on the next page

Dr. C. Dzamalala-Chairperson (COMREC)

5th January, 2016

Date

APPENDIX C: CURRENT STUDY ETHICS APPROVAL LETTER

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] 406 6626
Email: shuretta.thomas@uct.ac.za
Website: www.health.uct.ac.za/fhs/research/humanethics/forms

29 June 2018

HREC REF: 177/2018

A/Prof Maia Lesosky
Epidemiology and Biostatistics
Public Health & Family Medicine
Level 5
Falmouth Building

Dear A/Prof Lesosky

PROJECT TITLE: PREDICTION OF POST TUBERCULOSIS LUNG DAMAGE USING CT LUNG IMAGING MEASURES AMONG ADULTS IN MALAWI (MMed-candidate-V P Gunsari)

Thank you for submitting your response to the Faculty of Health Sciences Human Research Ethics Committee.

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

Approval is granted for one year until the 30 June 2019.

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)

Please quote the HREC REF 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.

The HREC acknowledge that the student, Vester Gunsaru will also be involved in this study.

Yours sincerely

PROFESSOR M BLOCKMAN
CHAIRPERSON, FHS HUMAN RESEARCH ETHICS COMMITTEE
Federal Wide Assurance Number: FWA00001637.
Institutional Review Board (IRB) number: IRB00001938

APPENDIX D: IJTLD AUTHOR GUIDELINES

The International Journal of Tuberculosis and Lung Disease

INSTRUCTIONS FOR AUTHORS

THE INTERNATIONAL JOURNAL OF TUBERCULOSIS AND LUNG DISEASE (IJTLD), the Official Journal of the International Union Against Tuberculosis and Lung Disease (The Union), publishes original research articles (including systematic reviews and meta-analyses) of international significance on tuberculosis and lung disease with a focus on clinical, epidemiological and programmatic research. In relation to lung disease, the *Journal's* focus is on research that is relevant to the Union's mission to find health solutions for the poor.

The IJTLD can be accessed in print and electronically via the Union website (http://www.theunion.org). Access to all back issues is free. Access to the current 6 months is available to all paid-up Union members and subscribers using the number provided on their current membership card. Non-members/non-subscribers can access and download individual articles using the 'Pay-per-view' option or contact The Union (membership@theunion.org) for information on how to become a member.

SUBMISSION OF ARTICLES

Articles are submitted online via Manuscript Central http://mc.manuscriptcentral.com/ijtld. Instructions are given on the site.

All other correspondence, such as suggestions for review and perspectives articles, should be sent directly to: **The Editorial Office, The Union, 68 boulevard Saint-Michel, 75006 Paris, FRANCE. e-mail: journal@theunion.org**

Simultaneous submission of a manuscript to more than one journal will automatically result in rejection by the IJTLD.

Each manuscript will be examined by a scientific editor and two referees. Notification of acceptance or rejection will be sent within 2 months from date of receipt. If a revised version is requested, it should be returned to the Editor no later than 1–2 months after notification. A delayed revised article will be treated as a new manuscript.

The Editor reserves the right to make editorial and literary corrections.

Any opinions expressed or policies advocated do not necessarily reflect those of the Union.

AUTHORSHIP

All work must be approved by all co-authors before submission.

Authorship credit should be based on the following criteria: 1) substantial contributions to conception and design, acquisition of data, or analysis and interpretation of data; 2) drafting the article or revising it critically for intellectual content; and 3) final approval of the version to be published.

Multicentre groups: When a multicentre group has conducted the study, all individuals who accept direct responsibility for the manuscript should be identified. When submitting a group author manuscript, the corresponding author should clearly identify all individual authors, as well as the group name.

COPYRIGHT

Copyright forms are submitted electronically. The corresponding author will be requested to log on and submit the form on the website after submission. On acceptance all authors must log on to submit the e-form.

PREPARATION OF MANUSCRIPTS

Manuscripts should conform to the *Uniform Requirements for Manuscripts submitted to Biomedical Journals* (http://www.icmje.org/index.html).

Authors should ensure that they have followed the relevant recommendations for reporting their findings (CONSORT, STARD, MOOSE, STROBE, PRISMA, STREGA). For easier use, here are the direct links to the latest versions:

STARD 15: http://www.equator-network.org/reporting-guidelines/stard/

Articles on clinical research should conform to the standards defined in the Helsinki Declaration, as revised in 2013 (www.wma.net/en/30publications/10policies/b3/index.html). Details of ethics approval (or a statement as to why it was not required) should be provided in the Methods section of all research studies submitted to the *Journal*. All studies involving human subjects should include details of informed consent.

Stigmatising language: Authors are advised to avoid terms that may be perceived to be stigmatising, such as "TB suspect" or "defaulter". Authors can refer to the following publications: Zachariah R, Harries A D, Srinath S, et al., Language in tuberculosis services: can we change to patient-centred terminology and stop the paradigm of blaming the patients? Int J Tuberc Lung Dis 2012; 16: 714–717. http://www.ingentaconnect.com/content/iuatld/ijtld/2012/00000016/00000006/art00003 and the STOP TB Partnership Language Guide: http://www.stoptb.org/assets/documents/resources/publications/acsm/LanguageGuide_ForWeb20131110.pdf

All articles must be submitted in English (US/UK). If the quality of the English requires professional help, authors will be informed. The summaries of all articles are published in French and Spanish.

FORMATTING

Authors should submit a single Microsoft Word document (.doc or .docx) that includes the title page, summary, text, references, tables and figure legends. The article should have 1.5 or double spacing and continuous line numbering.

Title page: This should contain: 1) a concise, informative title of **not more than 110 characters and spaces, without abbreviations**; 2) the names and all affiliations of all contributing authors, clearly indicating who is linked to each institution; 3) a running head of not more than 45 characters and spaces; 4) a word count of the summary, a word count of the text, number of references, tables and figures; 5) 3-5 keywords that do not appear in the title, and 6) the name, full address and contact details of the corresponding author.

Summary: An informative structured abstract of **not more than 200 words** should be provided that can be understood **without reference to the text** (see Ann Intern Med 1990; 113: 69-76). For optimal clarity, the author should use the headings Setting, Objective, Design, Results and Conclusion. Abstracts will be translated into the two other languages on acceptance for publication (authors are welcome to provide translations). Unstructured summaries may be submitted for systematic reviews and/or meta-analyses (250 words) and for short original articles or brief reports (100 words).

Text: Headings should be appropriate to the nature of the article. Normally only two categories of heading are used. Major headings should be typed in capital letters. Minor headings can be typewritten in lower case letters (starting with a capital letter) at the left-hand margin. The subtitles should not be numbered either with figures or alphabetically.

The text should be written as objectively as possible. For word limits, please refer to the section 'Length of text'.

Numerals should be spelt out in full from one to nine (except when referring to a measurement), and when beginning a sentence.

1. Research and experimental papers should follow the usual conventions, as follows:
   - *Introduction:* Setting forth clearly the aim of the study or the main hypothesis, with reference to previous studies and indicating the method used.
   - *Materials or Study population and Methods:* Indicate what measures were taken to assure the quality of the data.
   - *Results:* Presented in logical sequence in the text, with tables and illustrations. All the results of the tables should not be repeated in the text; the most important results should be emphasised.
   - *Discussion:* Refer to the aims, results and importance of the study.
   - *Conclusions*
2. Other papers can be subdivided as the author desires; the use of headings enhances readability.

Acknowledgements: Acknowledge only persons who have made substantial contributions to the study, with their consent. All sources of support in the form of grants, author contributions and all conflicts of interest should be mentioned.

References (Vancouver format): The accuracy of references is the responsibility of the author. They must be numbered in the order in which they are cited in the text, and identified by Arabic numerals in superscript. References that are cited more than once should retain the same number for each citation. The list of references at the end of an article should be arranged in numerical order. **NB: Numbering in tables/figures corresponds to where they are cited in the text.**

The only acceptable references are those of publications that can be consulted.

References to an article in a periodical should include the names of the authors, followed by their initials (list all authors when six or fewer; when there are more, list only the first three and add 'et al.'), the full title of the article in its original language, the name of the journal in its usual abbreviated form (Index Medicus), year of publication, tome or volume number, first and last page numbers in full:

e.g., Gordon J B, Bennett A M. Tuberculosis in reindeer. Scand Rev Respir Dis 1978; 96 (Suppl): 217-219.

References to a piece of work (book/monograph) should include the names of the authors as above, the title of the piece of work in its original language, the number of the publication, the name of the editor, the place and year of publication, the number of the volume and the first and last page numbers.

References to a chapter in a book should include the names of the authors as above, the title of the chapter with the word "In" preceding the reference of the work as above.

e.g., Girling D J. The chemotherapy of tuberculosis. In: Ratledge C, Stanford J, Grange J M, eds. Biology of the mycobacteria. London, UK: Academic Press, 1989: pp 285-323.

Electronic references should be given only when an original citation is unavailable; as much information should be provided as possible, including html address and date of access.

References to an article to be published should give the name of the journal with the mention '(in press)' and *only* appear after having been accepted. Articles under submission can be cited in the text.

Personal communications should be given in the text with the name of the individual cited *and with his/her consent.*

TABLES

Tables should be self-explanatory. They should be referred to consecutively in the text and placed after the references. They should be numbered in Arabic numerals, which are used for reference in the text. A short descriptive title should appear above the table. Each column should have a short or abbreviated title. All abbreviations should be explained in a clear legend below the table.

The number and size of the tables should be kept to a basic minimum to explain the most significant results.

Tables should not have shading or bolding. Explanations of data should be included in the legend and linked to the respective value by a symbol, in the following order: * † ‡ §

NB: A 'moderate-sized' table takes up a quarter of a printed page.

FIGURES

Figures should be referred to consecutively in the text. They can be inserted into the Word document after the tables or uploaded separately as image files (.jpg, .ppt, .gif, .tif or .bmp).

A brief explanatory legend that does not repeat information given in the text should be provided for every Figure.

After acceptance, **all figures must be supplied in editable format** (e.g., .ppt, .xls) to allow editorial modifications.

If editable figures cannot be supplied, they should be provided in high resolution (min 300 dpi) TIFF or JPEG files suitable for reproduction, with all text in black Arial font, 12 pt minimum (or proportional to the size of the figure).

Line drawings (curves, diagrams, histograms) should be in black and white, with solid black lines. For optimal clarity avoid shading.

Flow charts: Flow charts/algorithms should be provided in editable Word format without colour or shadowing.

Lettering The size of the symbols and lettering should be in scale with the figure using black Arial font, of uniform size. All figures should be the same point size.

Half-tone figures should be clear and highly contrasted in black and white. Photo-micrographs should have internal scale markers where appropriate. X-ray films should bring out the detail to be illustrated with the area of importance clearly indicated.

Techniques (staining, magnification, etc) should be defined.

Half tone and colour figures should be supplied at a resolution of a least 300 dpi (preferably 500 dpi).

The cost of reproducing colour illustrations (print or online) will be covered by the authors.

Patient confidentiality Illustrations that show recognisable individuals are strongly discouraged and will be considered for publication only where fully justified. In such cases, consent must be obtained from the individual or legal guardian for publication. A consent form can be obtained on request from the Editorial Office. For portions of the image not essential to the illustration, authors should indicate where it can be cropped.

Permission to reproduce illustrations or tables should be obtained from the original publishers and authors, and submitted with the article. They should be acknowledged as follows: *'Reproduced with the kind permission of (publishers) from (reference)'*.

NB: A 'moderate-sized' figure takes up a quarter of a printed page.

ABBREVIATIONS AND UNITS

Avoid abbreviations in the title or summary. Abbreviations or unusual terms should be described at the first time of use.

Symbols and units of measure must conform to recognised scientific use, i.e., SI units. For more detailed recommendations, authors may consult the Royal Society of Medicine publication *Units, Symbols and Abbreviations: A Guide for Biological and Medical Editors and Authors.*

Designation of diseases must conform to the International Classification of Diseases. Designation of micro-organisms must conform to the norms of biology. Proprietary names of drugs, instruments, etc., should be indicated by the use of initial capital letters. Names of instruments should be accompanied by the manufacturer's name, city, state and country. Any software used should be described in full, including version, manufacturer, city, state and country.

FORMAT OF SUBMITTED ARTICLES

Original research articles:

- Text up to 2500 words, a structured summary of 200 words, up to 7 moderate-sized tables/figures and up to 35 references.
- Brief articles and short reports are welcome. These should be submitted as original articles and will be assessed as such.
- Where it is believed a shorter article is justified, the editors reserve the right to ask authors to shorten their article.

Systematic reviews and meta-analyses are assessed and published as original articles.

- Text up to 4500 words, a structured/ unstructured summary of 250 words, up to 8 moderate-sized tables/figures and 90 references.
- Excess material can be published in an online appendix.
- No pre-submission query required.

Case studies are considered *only* if they contain original and innovative material. Patient consent should be provided (or, in the case of death, the consent of a relative).

- Text up to 1000 words, a summary of 100 words, 2 tables/figures and 10 references.

Scientific correspondence: Text up to 500 words, with 5 references and without tables or figures.

- Correspondence in response to an article published in the IJTLD is welcome, and should be submitted within 3 months of the publication date of the article to which it refers. These are sent to the authors for reply. These exchanges are prioritised for publication in the next available issue of the *Journal.*

Papers that are too long must comply with editorial requirements.

INVITED ARTICLES

The editors may invite submissions in the following formats: Editorials, Perspectives and State of the Art articles. Unsolicited submissions for these sections are not considered. However, prospective authors may submit a one-page proposal for a Perspectives or State of the Art article to ijtld@theunion.org. All articles are subject to peer review.

Editorials are usually commentaries on original research articles published in the same issue or on special events.

- Text up to 500 words and 5 references.

Perspectives articles are focused, expert reviews on cutting-edge issues. One-page proposals will be considered, and will be judged on 1) the scientific importance and novelty of the subject matter, 2) its relevance to the readership of the *Journal*, and 3) the expertise of the proposed authors. Perspectives are expected to draw conclusions and make recommendations that are based on the evidence presented.

- Text up to 2500 words, a structured summary of 200 words, up to 7 moderate-sized tables/figures and 35 references.

State of the Art reviews on the most current research in a specific field relevant to the scope of the IJTLD are regularly invited by the editors. Authors may also suggest relevant topics. A presubmission query is required.

- Text up to 4500 words, a structured/ unstructured summary of 250 words, 8 moderate-sized tables/figures and 90 references.

State of the Art series of at least 6 articles on a specific topic can be proposed by at least two individuals willing to serve as Guest Editors of the series. The interested individuals should submit a proposal listing the topics, potential authors and a timeline.

PUBLICATION CHARGES

Excess page charge: an excess page charge of 200€ per page is applied for all articles that go over length (6 pp for original articles, 10 pp for systematic reviews and SOA articles).

Online appendices: Additional text, tables and figures may be supplied as an Online Appendix, which is appended to the final pdf of the article. A single charge of 100€ will be applied. Online appendices should be submitted with the manuscript for review. Authors may also be requested to move overlarge tables/figures to an online appendix.

Immediate open access: Articles can be made Open Access at the time of online publication, for a charge of 1500€.

CC-BY licence: If for funding reasons an article also requires a Creative Commons Attribution (CC-BY) license and/or deposit onto Europe PMC/PMC, the authors should inform the Editorial Office on acceptance.

PERMISSIONS

Individuals wishing to reproduce material from the IJTLD must request written permission from the Editorial Office.

Likewise, authors of articles published in the IJTLD who wish to include material from other copyrighted sources must seek permission from the copyright holders and provide written evidence of this permission at the time the article is submitted.

PLAGIARISM

The IJTLD uses iTHENTICATE to check for plagiarism.

In cases where plagiarism is suspected, the IJTLD follows the guidelines set out in the flowcharts of the Committee on Publications Ethics (COPE) (http://publicationethics.org/flowcharts). Please see the *Journal* website for details (www.theunion.org).

LINKS FOR SUBMISSION OF ARTICLES

http://mc.manuscriptcentral.com/ijtld, link via www.theunion.org

Full submission instructions are given on the Manuscript Central site (click on "Instructions and Forms", then "Submission instructions"). Please refer to the checklist for the full list of elements.

Please note that if any elements are missing on submission, the article will be returned to the authors for correction.

REVISION OF PROOFS

Proofs are sent to the corresponding author by e-mail in PDF format. Authors should return corrected galley proofs by fax or e-mail directly to the Editorial Office in Paris *within 48 hours* of receipt. Only minor corrections should be made.

OFFPRINTS

An offprint order form will be sent to the corresponding author of each article. The completed form must be returned with the corrected proofs.

Contact: The IJTLD Editorial Office, The Union, Paris, France. e-mail: journal@theunion.org