Investigating the Risk Factors for Contraction and Diagnosis of Human Tuberculosis in Indonesia using Data from the Fifth Wave of RAND's Indonesian Family Life Survey (IFLS 5)

Nathan Adam[1*], Saseendran Pallikadavath[1], Marianna Cerasuolo[2], and Mark Amos[1]

[1]*Portsmouth-Brawijaya Centre for Global Health, Population, and Policy, University of Portsmouth, PO1 2UP, Portsmouth, UK,*

[2] *School of Mathematics and Physics, University of Portsmouth, Portsmouth, UK*

[*]Corresponding author. Email: nathan.adam@myport.ac.uk

Abstract

Tuberculosis (TB) is a globally widespread disease, with approximately a quarter of the world's population currently infected (WHO, 2018). Some risk factors (e.g., HIV status, nutrition and body mass) have already been thoroughly investigated. However, little attention has been given to behavioural and/or psychological risk factors (e.g., stress and education level). In this study, risk factors were investigated via statistical analyses of publicly available data from the most recent wave (2015) of the Indonesian Family Life survey (IFLS 5). For comparison and completeness, variables were divided into levels: individual, household, and community. The most prominent and interesting variables which influence TB diagnosis (on each level) were investigated, and a logistic regression was subsequently developed to understand the extent to which each risk factor acted as a predictor for being diagnosed with TB. Age, health benefit or insurance, stress at work, and living in a rural area all showed significant association with TB diagnosis. The outcomes of this study suggest that suitable control measures, such as BCG vaccinations, schemes for improving mental health/stress reduction, and improved access to healthcare in rural areas should be implemented to address each of the key factors identified.

Keywords: Risk Factors, Logistic Regression, tuberculosis

Introduction

Tuberculosis is an infectious bacterial disease caused primarily (in humans) by *Mycobacterium tuberculosis*. TB typically affects the respiratory system (throat and lungs) and is transmitted via the air when a person with active TB coughs, sneezes or speaks. This 'ancient scourge' has plagued humans and their ancestors throughout known history. Despite the existence of a cure for TB since 1949 and vaccination (administered predominantly to high risk infants shortly after birth) since 1921, TB is still a great public health concern, with approximately 1.7 billion people currently (latently) infected worldwide (Daniel, 2006; WHO, 2018). Around 5-10% of these 1.7 billion people will develop active TB during their lifetime, which means that the disease will become symptomatic and transmissible. Instances of active TB are referred to as TB cases. The majority of worldwide TB cases arise in developing countries such as Indonesia (Bock & Cox, 2017). In the Indonesian context, money for procuring (and access to) healthcare is often lacking. Additionally, Indonesia's Impoverished living conditions are said to contribute to TB's high burden. According to the World Health Organisation, there were approximately 319 new cases of TB per 100,000 people during 2017 in Indonesia alone. To put this into perspective, the incidence rates per 100,000 people in some other high-burden countries are 423 (Central African Republic), 361 (Zambia) and 554 (The Philippines). Two thirds of global TB cases (in 2017) occurred in 8 countries, including Indonesia (and the Philippines) (WHO, 2018). TB causes a great economic burden to the Indonesian economy; approximately US$6.9 billion was lost in 2015, with 86.6% of this accounted for by premature death resulting in lost productivity (Collins et al., 2017). Although this disease is curable, multi-drug resistance caused by improper treatment regimens or incompliance with regimens has made TB difficult to treat effectively. Also, TB was thought to be receding in the past decade, but there is evidence that it has been re-emerging in recent years (Shajahan et al., 2016).

The known risk factors regularly associated with TB are: being young, male, living in a developing country, being a health care worker, poor ventilation, exposure to contaminated air, and having a compromised immune system (e.g., due to smoking). HIV infection and (consequently) AIDS are the most prominent risk factors for contracting TB (Dye et al., 1999; Fogel, 2015). In recent years, the importance of extensive interdisciplinary research into TB has been appreciated as TB's various socio-politico-economic risk/exacerbating factors have come to light, and this has prompted the need to delve beyond the purely biomedical model (Mason et al., 2016; Mason et al., 2017). However, risk factors related to stress or education have not (yet) been examined as extensively within the existing literature; they are rarely included in the analysis, but a lack of education may hinder one's ability to take measures to protect themselves from disease (Baker et al., 2011). There are various stigmatic associations pertaining to TB, such as belonging to a low social class, poverty or indecorous behaviour. Contempt towards TB sufferers can manifest as social exclusion, which heightens allostatic load, resulting in worsened ability to perform daily tasks (Simandan, 2017a), thereby delaying recovery. Anticipation of such contempt can even discourage disclosure of TB sufferers' diagnoses to friends and family, thereby generating avertable risks to loved ones and public health (Simandan, 2017b). Additionally, implementations of politico-economic ideologies such as neoliberalism have led to layoffs and uncertainty regarding employment conditions, inducing fear and anxiety coinciding with (and explaining) the increased incidence of TB in affected countries (Simandan, 2017a). The complicated interplay that stress has with anxiety, substance abuse and depression can lead to incompliance with treatment regimens and aggravate health outcomes (Zvonareva et al., 2019). Worse yet, stress also adversely affects immunity (Morey et al.,2015). Evidently, the detrimental effects that stress, anxiety and allostatic load have on the acquisition, persistence and recrudescence (de Oliveira et al, 2000) of TB warrant further study. Indeed, reciprocity amongst TB researchers from various disciplines is required to understand and ameliorate the burden of this multi-faceted disease (Mason et al., 2016). These causal/exacerbating factors may have various sources or triggers, but this study focuses only on the association between occupational stress (stress at work) and the diagnosis of TB infection.

The widespread burden of TB in Indonesia and its re-emergence prompt an urgent need to cover the aforementioned research gaps, and form the motivation for this study. The aim is to explore a wide variety of risk factors on a variety of levels (individual, household and community) for comparison and completeness, incorporating both existing understanding of the risk factors for TB as well as expanding the understanding of hitherto neglected factors. Bivariate measures of association and logistic regression will be used to quantify the extent to which they act as predictors for contracting and subsequently being diagnosed with TB. With reference to relevant literature, qualitative meaning will be appended to the results and logical control measures will be suggested. The key contribution is the emphasis on behavioural and psychological risk factors (especially stress) as opposed to more commonly understood ones (e.g., HIV). This research is important because the findings may ultimately lead to a reduced burden of TB in Indonesia as well as improved socioeconomic status for Indonesian citizens.

Materials & Methods

Publicly available demographic household survey data from the fifth wave of the Indonesian Family Life Survey (IFLS 5) were used. The primary motivation for using IFLS datasets were that they contain TB markers whereas other common datasets (e.g., DHS) do not. While this is a longitudinal survey, only the most recent wave is used to explore the cross-sectional association between the incidence of TB and relevant risk factors. The survey samples from a frame of approximately 83% of the Indonesian population, and each wave of the survey encapsulates responses from over 30,000 respondents who are living in 13 out of the 27 Indonesian provinces. The names of the surveyed provinces are: North Sumatra, Yogyakarta, West Sumatra, East Java, South Sumatra, Bali, Lampung, West Nusa Tenggara, Jakarta, South Kalimantan, West Java, South Sulawesi, and Central Java. These data were collected by RAND (Research and Development) Corporation staff from anonymous face-to-face household questionnaires (RAND, 2018). The resultant data sets are very rich, with information on various aspects of socio-economic status and health. The data and documentation are publicly available on request at https://www.rand.org/labor/FLS/IFLS/.

Prior to analysis, variables were transformed/modified as appropriate, to ensure that the analyses would be meaningful. Bivariate measures of association were used to assess the relationship between having received a TB diagnosis and each selected variable, based on a chi-square test for categorical variables and ANOVAs for continuous variables. Logistic regression was then used to quantify the relationship between TB diagnosis and selected variables simultaneously. Associations were said to be significant at a 5% level in all cases. Data were analysed using SPSS 20 for Windows and R 3.4.4. for Windows.

Many variables had a very limited number of respondents' answers corresponding to them (such as distance to and knowledge of medical facility). Therefore, the number of variables that were used in this study was restricted in order to preserve the number of cases available for the logistic regression. The number of variables that were non-significant and/or less prolific in the literature compared to the other variables was also limited in order to retain the power of the logistic regression, especially when there was evidence of collinearity in the model. Since the asthma and other lung condition variables have similar symptoms to TB and are both representative of access to healthcare, they have been omitted from the logistic regression due to their redundancy.

Stress and education level are often overshadowed by other, well-known risk factors (e.g., HIV), so less attention has been given to these risk factors. Due to stress and education level being underexplored in the existing literature, these variables were examined with particular attention. In this study, stress is represented by the responses to the survey question "My job involves a lot of stress", with 4 different answers pertaining to level of agreement with this statement. In order to compare high and low levels of stress, a category ('High') was formed for the 2 highest stress levels. Similarly, a category ('Low') was formed for the two lowest stress levels.

Variable Selection

The dependent variable, corresponding to the survey question "Has a doctor/paramedic/nurse/midwife ever told you that you had TB?" (TB Diagnosis Status) is used as the indicator of TB cases for the analysis. Due to TB's non-specific symptoms, misdiagnoses may have occurred, especially in cases where TB infection was not bacteriologically confirmed. Furthermore, TB infection may remain undiagnosed, perhaps due to latent infection (which is asymptomatic), or a lack of access to diagnostic healthcare. Selection of explanatory variables was based on existing literature, which highlights age, smoking, malnutrition, HIV/AIDS, body mass and poverty/impoverished living conditions as the predominant risk factors (Duarte et al., 2018; Keflie et al., 2018). Variables corresponding to age and smoking were readily available, but HIV data were missing. In the absence of clinical nutrition data, a variable pertaining to how satisfied the respondents were with the adequacy of the level of their food consumption was used; this acts as a proxy for both nutrition and body mass. About two thirds of responses for annual salary data were missing, so a variable relating to where the majority of household members went to the toilet was also used — this acts as a proxy for impoverished living conditions/hygiene. Education level was considered because, although it is typically overlooked, there is some evidence that it may have a considerable effect on health (Baker et al., 2011). Stress at work was also used to investigate the apparent association between stress and TB (Morey et al., 2015; Simandan, 2017a). Communicable diseases in general are associated with crowded conditions and urban areas, where people are often in close contact for extended periods of time. Therefore, variables corresponding to the number of people living in each household, and whether the household resides in an urban or a rural area were used. With regards to occupational risk factors, TB is typically associated with working in travel or healthcare industries, so the respondents' sectors of work were considered. Since asthma and other lung conditions are pulmonary diseases, their relationships to TB (which is typically a pulmonary infection) were investigated. The following remaining variables were chosen to be investigated because they relate to health or disease transmission/diagnosis in general: gender (Nhamoyebonde & Leslie, 2014), knowledge of/distance from medical/health facilities, and health benefit/insurance (O'Donnell, 2007).

Results

Cross-tabulation, Chi-Square tests (of association) and ANOVA

To assess the relationships between the explanatory variables and TB, Chi-Square tests (for categorical variables) and ANOVA (for numerical variables) were utilised alongside a cross-tabulation. The results are shown in Table 1 overleaf.

The categorical variables: age, stress, gender, work sector, asthma, other lung conditions, health benefit or insurance, and rural or urban area were found to be significantly associated with TB Diagnosis ($p<0.05$ in Chi-Square tests) in this study.

The non-significant variables in this study were : cigarette consumption, education level, food consumption satisfaction, health facility location knowledge, annual salary, number of people living at household, toilet used by household occupiers, and distance from medical facility (p>0.05 in Chi-Square tests or ANOVA).

Logistic Regression

The results of the logistic regression with 11 variables, 9725 cases, and a Hosmer-Lemeshow p value of 0.679 are shown in Table 2 overleaf. These results should be interpreted with the limitations or assumptions of logistic regression in mind; for example, a uniform relationship between the explanatory variables and the dependent variable is assumed (Ranganathan et al, 2017). The Hosmer-Lemeshow p value of 0.679 (>>0.05) suggests that the model fits the data rather well. However, this metric should be interpreted with caution due to its limitations; for example, it tends to have low power (Hosmer et al., 2013).

From the logistic regression, the significant predictors of being diagnosed with TB are stress at work, living in a rural area, age, and having health benefit or insurance (p < 0.05).

Individual-level Risk Factors

The odds ratio (OR) of 2.76 for the 44+ age group suggests that middle-aged or elderly citizens are 276% as likely to have been diagnosed with TB than very young people (14-30). Cigarette consumption has an OR of 1.00 and is not significant, suggesting that, at least according to these results, it is a poor predictor for being diagnosed with TB. Education level is difficult to interpret here due to the various categories and large standard errors which led to large confidence intervals. At a first glance it looks like those with higher levels of education than the reference category have lower chances of contracting and being diagnosed with TB (each OR < 1: 0.67, 0.29, 0.40, 0.87). However, when considering the ORs CI, it is clear that the association is not significant here. With regards to food consumption (satisfaction), consuming an 'adequate' or more than adequate amount (rather than a less than adequate amount) of food appears to marginally alter the odds of being diagnosed with TB by factors of 1.13 and 0.98 respectively, implying that those with less than adequate or more than adequate food consumption are at a slightly lower risk of contracting TB. However, the authors expected that increased food consumption would lower the risk of TB, since low body mass is a risk factor for TB. The association is also not significant in this case. The OR for gender is just slightly greater than 1 (1.07) and is not significant in this study. Having health benefit or insurance increases the odds of successfully being diagnosed with TB by a factor of 1.53 (1.01, 2.34); this variable probably indicates improved surveillance rather than being a risk factor for contracting TB itself. From this study, there is a clear association between TB diagnosis and (occupational) stress; people with higher levels of stress at work appear to be 2.34 (1.43, 3.83) times as likely to be diagnosed with TB

Household-level Risk Factors

The number of household occupiers variable has an OR of 1.05 and was not found to be associated with TB Diagnosis Status in this study. The waste disposal system appears to have some effect on the chance of being diagnosed with TB; those using toilets without a septic tank are 1.30 times as likely to be diagnosed as those using toilets with septic tanks. Using a toilet with a septic tank appears to have little impact on TB diagnosis compared to the use of 'other' facilities (OR=0.97).

Community-level Risk Factors

In this study, living in a rural area raised the odds of being diagnosed with TB by a factor of 1.80, and this effect is significant. Sector of work appears to have some effect on the chances of being diagnosed with TB, although this effect is not significant according to these results. The odds ratios of 0.74 and 0.63 suggest that those working in agricultural, mining, manufacturing, electrical or construction industries are relatively higher risk occupations compared to the other areas of work considered.

Discussion and Conclusion

The purpose of the study was to investigate several potential risk factors for tuberculosis in Indonesia, especially behavioural or psychological risk factors. The publicly available survey data were investigated using statistical methods alongside appropriate software to investigate the relationships between the explanatory variables and the dependent variable (TB Diagnosis Status). Some of the results agreed with contemporary literature, whereas other results were unexpected.

Individual-level Risk Factors

The results show a significant association between age and TB Diagnosis Status. The average age of those who had not been diagnosed with TB was roughly 5 years younger than those who had been diagnosed. This is expected; the research conducted by Pawelec (2018) states that older people have a higher risk of contracting diseases due to gradual weakening of the immune system with increasing age. The chances of having been diagnosed with TB appeared to be roughly the same for 14-30 and 31-43s year olds (OR=1.10), which is also to be expected; immune function tends to only begin decreasing substantially at about 50-60 years of age. The 44+ age group, predictably, had much higher odds (OR=2.76) of being diagnosed with TB compared to the other age groups.

In this study the relationship between cigarette consumption (smoking) and TB Diagnosis Status contradicted the literature (Duarte et al., 2018) — there did not appear to be any significant relationships between the variables, and the values for cigarette consumption between the TB Diagnosis Status categories were very similar (12.35 and 12.23 for those who had or had not been diagnosed with TB, respectively). This result disagrees with the global consensus that smoking significantly influences contraction of TB (WHO, 2007). However, there is some evidence that in the Indonesian context, the effect of smoking on TB is not significant (Sahiratmadja & Nagelkerke, 2016). It has been shown that smoking causes damage to the throat, lungs, and cilia (which help to remove pathogens from the immune system) (Bates et al., 2007); one would therefore have expected that smoking would significantly impact TB Diagnosis Status. Even when comparing smokers to non-smokers (past and present), there was no significant association; this suggests that smoking may not be so influential as a risk factor, at least in Indonesia. However, it's likely that some people reduced their cigarette consumption upon experiencing symptoms of TB or being diagnosed with TB; smoking exacerbates TB symptoms. Furthermore, since, in Indonesia, smoking is legal in both private and government offices, restaurants, and pubs, many people would be affected by passive smoking, which was not accounted for in the survey. Therefore, it's not possible to understand the relationship between smoking and TB without further analysis.

Since education can have a prominent impact on health (Baker et al., 2011), the highest education level achieved was investigated as a predictor for TB Diagnosis Status. However, no significant association was found. This may be due to the lack of specificity in the data regarding subjects studied and academic attainment. For example, people who studied health or biological sciences at high school were in the same category, regardless of attainment. Additionally, people studying subjects unrelated to health and disease transmission at the same level (e.g., at university) were in the same category as those studying the aforementioned subjects. Therefore, a lot of relevant information, which may have acted as predictors for TB, was not accounted for.

Malnutrition and body mass are said to be prominent risk factors for tuberculosis and diseases in general (Keflie et al., 2018), especially considering the importance of vitamins, minerals, carbohydrates and proteins for maintaining a strong immune system. As a proxy for malnutrition and body mass, a survey question regarding the respondent's satisfaction with their food consumption was used. However, no clear association between this variable and TB Diagnosis Status was found. Data directly relating to nutrition (for example, blood test results for vitamin A and zinc) and body mass (e.g., scale readings) would probably have led to a better investigation into the relationships that nutrition and body mass have with TB.

Gender is also an interesting factor, and it could be argued that it is a behavioural risk factor — in patriarchal societies, like Indonesia, men constitute a considerably larger proportion of the workforce than women (Organisation for Economic Co-operation and Development, 2018), who are likely to stay at home to look after children. Therefore, men generally have considerably more exposure to TB. Despite that, no significant association between gender and TB Diagnosis Status was found in this study. These results disagree with numerous studies, even if the 95% confidence interval contains the odds ratio between men and women, 1.61, estimated in the study by Marçôaa (2018).

Health benefit or insurance is an excellent indicator for access to healthcare. The OR of 1.53 suggests that those with health benefit or insurance are significantly more likely to be successfully diagnosed with TB than those without. It may be noted that the Indonesian government introduced universal health coverage (UHC) in 2014 to improve access to health care.

Stress at work has been a variable of great interest in this study, but it has often been overlooked previously. There is minimal consensus and various schools of thought regarding the classification of events as stressful. One perspective of particular note, arising from occupational stress literature is that a stressful event is one where demands exceed resources. However, many researchers define stressful events differently (e.g., as events consensually seen as harmful or threatening) (Cohen et. al, 2019). In this study, the data is from respondents who individually, subjectively appraised their perceived stress levels (psychological stress) at work. This variable should be considered in the context of allostatic load theory, which aims to delineate the physiological mechanisms which lead from stressors experienced to ill health (Simandan, 2010), all the while offering precise terminology free of the inherent ambiguity of the word 'stress'. The substantial evidence that there is an association, and possibly a causal link between stress (especially allostatic load) and TB, both due to weakened immunity, (Morey et al., 2015) and reduced ability to carry out daily duties (Simandan, 2017a) prompted the need for exploration of pertinent data. According to the results of this study, there is indeed a significant association between (occupational) stress and TB Diagnosis Status, and stress at work is a prominent predictor for contracting TB. The high odds ratio (2.34) suggests that appropriate control measures should, quite urgently, be implemented. In the context of occupational stress, there is probably an inverse relationship between chronicity and intensity, which may be important to consider for future research directions, especially when considering the relationship between chronicity of stress and immune function (Dhabir & McEwen, 1999).

Despite TB's sobriquet of 'poor man's disease', the results do not reveal a significant relationship between annual salary amongst groups with distinct diagnosis statuses. In fact, dissimilarly to the study undertaken by Bhunu et al.(2012), people diagnosed with TB seemed to be wealthier (they had a higher average annual salary) than those who hadn't been diagnosed. This could potentially be explained by difficult circumstances (e.g., due to having to repay debts, or financially assisting family members) resulting in living conditions remaining or becoming impoverished. Additionally, this variable doesn't provide data on total income (just income from the respondent's main job), and the respondent may have been diagnosed with TB prior to getting a higher-paying job.

According to the Chi-Square tests, other lung conditions and asthma were found to be significantly associated with TB Diagnosis Status. However, this may be due to people with asthma or other (known) lung conditions being closely monitored by healthcare professions, so this association may pertain to surveillance.

Household-level Risk Factors

To investigate the effect of overcrowding or population density on TB Diagnosis Status, the number of people living within each household was considered. However, no statistically significant differences between the means for this variable amongst the TB Diagnosis Status categories were found in this study. Additionally, this variable was not a significant predictor in the logistic regression, but logistic regression's assumption of relational uniformity is important to note here — there may have been, for example, more cases of TB at lower values (due to, e.g., struggling single parent families) and at higher values (due to overcrowding), or particular ranges of values with many cases of TB. A lack of ventilation and airflow is also said to contribute to TB transmission in households (Duarte et al., 2018), although there were no data available to investigate this. There may have been more interesting results if there were data on community-level population density.

Due to the importance of investigating sanitation and poverty levels, variables relating to where household occupiers went to the toilet were considered. Unlike annual salary, there were considerably fewer missing data for this variable. However, no significant association between these variables and TB Diagnosis Status were found; this, alongside the investigation of annual salary, suggests that poverty may not be as prominent of a risk factor as the existing literature (Fuady et al., 2018) suggests.

Community-level risk factors

There is no clear consensus in the literature regarding whether living in an urban or rural area has a significant impact on the likelihood of becoming infected with TB. Living in an urban area has often been associated with crowdedness and pollution (Keller, 2018), whereas living in a rural area is associated with having to pay medical expenses out of pocket and having to travel long distances to arrive at medical facilities. Furthermore, rural areas are considerably more likely to have farms, where bovine TB could spread to humans (Davis, 2000; Amo-Adjei, 2016). In principle, either living in an urban area or a rural area could potentially increase one's susceptibility to TB, but from this study, those living in rural areas of Indonesia were significantly more likely to be diagnosed with TB, and this may be due to UHC introduced in 2014 in Indonesia.

Sectors of work were investigated as risk factors. The literature states that people working in travel or healthcare sectors are at a high risk of acquiring TB due to extensive exposure. From the Chi-Square tests, there is a significant association between work sector and TB Diagnosis Status, which aligns with the studies conducted by Rim & Lim (2014). However, inferences regarding specific sectors of work were not made — the sample sizes would have been too small to make these analyses meaningful.

Summary and Conclusion

In summary, several potential risk factors for TB Diagnosis Status were investigated using statistical methods, providing a comprehensive overview of the risk factors for TB diagnosis, expanding on previous analyses by including educational level and stress. Age, stress at work, living in a rural area, and having health benefit or insurance were the most significant predictors of TB Diagnosis Status, and these results are supported by the literature, although stress is rarely explicitly discussed. What's interesting here is that most of these variables relate to immune system deficiencies or vulnerabilities (and surveillance), rather than exposure to the disease itself — this aligns with TB's reputation of being an 'opportunistic' disease (Barnes et al., 1993).People with healthy immune systems will often be able to fight off the infection, even if their exposure to TB is high (Fogel, 2015). Some results were unexpected, with cigarette consumption and annual salary not being significantly different amongst TB Diagnosis Status categories. Although several variables were analysed, it would have been beneficial to have had even more variables to investigate, such as those relating to HIV status and body mass. Survey questions directly related to knowledge of TB would also have been highly desirable. Nonetheless, according to the Hosmer-Lemeshow test, the logistic model fits the data rather well, so the model (as well as some other findings) may potentially be useful in real-world scenarios, especially for investigations regarding the 13 Indonesian provinces in which the IFLS 5 survey was carried out.

The role of stress in these results is a pertinent one. The results may potentially be of interest in the field of behavioural epidemiology and could prompt further research into stress as a risk factor for TB (or other diseases). This research may also be of great interest to employers and companies; one can see from the findings that stress at work is associated with TB, and from the literature review it's clear that TB is detrimental to productivity and, consequently, the Indonesian economy (Collins et al., 2017). It could therefore be highly beneficial for companies to introduce schemes for alleviating the stress levels of their employees. This work may also be of value to medical practitioners, academics, researchers or centres for disease control for determining comparative TB risk in the absence of HIV data and other data. As expected, age was found to be a prominent risk factor for TB, and these findings may act as a catalyst for the implementation of interventions which prioritize care for elderly citizens. This would also be applicable to those with asthma and other lung conditions. The fact that TB was concentrated in people living in rural areas suggest that there is a great need for additional affordable and accessible healthcare institutions to be established in these areas. Indeed, the high basic reproduction number of TB (estimated to be 4.5 by Blower et al. (1995)) suggests that powerful control measure(s) would need to be implemented to eradicate this persistent disease. Further studies should also be conducted to understand the nature of TB risk factors; this could include undertaking multivariate analyses using the IFLS 5 data, or analyses of other variables using alternative data sets, and comparisons of the analyses of similar datasets between different countries or regions. Developing a logistic model tree (or other suitable and sufficiently complex models) would also be an excellent way of expanding upon this work, but this was not within the scope of this study.

Limitations

The questionnaires and resultant secondary data obtained from RAND's IFLS 5 were not tailored specifically to the study, and lacked pertinent information directly relating to HIV, blood pressure, body mass, malnutrition and knowledge of TB. This study was liable to some frame bias; the reliance on household level enumeration excluded traveller or slum dwelling populations which may not have been enumerated and, worse yet, are more likely to be at an elevated risk of developing active TB. Furthermore, only 0.01% of the Indonesian population was surveyed, from 13 out of the 27 Indonesian provinces. People with extreme cases of TB (especially in the case of TB-HIV co-infection) may have been too unwell (and possibly hospitalised) to answer the survey questions. Furthermore, the survey question relating to TB Diagnosis Status was not restricted to any time period, so the interaction between TB and its predictors could not be fully ascertained.

Acknowledgements

The authors would like to thank RAND Corporation for the open-access survey data, and for assistance with handling the data files.

Ethical Approval

The authors assert that all procedures contributing to this work comply with the ethical standards of the relevant national and institutional committees on human experimentation and with the Helsinki Declaration of 1975, as revised in 2008

Declaration of Interest

The authors have no conflicts of interest to declare.

Funding

This research received no specific grant from any funding agency, commercial entity or not-for-profit organization.

References

Amo-Adjei J (2016) Individual, household and community level factors associated with keeping tuberculosis status secret in Ghana. *BMC public health* **16**(1), 1196.

Baker DP, Leon J, Smith-Greenaway EG, Collins J and Movit M (2011) The education effect on population health: a reassessment. *Population and development review* **37**(2), 307–332.

Barnes PF, Le HQ and Davidson PT (1993) Tuberculosis in patients with HIV infection. *The Medical clinics of North America* **77**(6), 1369-1390.

Bates MN, Khalakdina A, Pai M, Chang L, Lessa F and Smith KR (2007) Risk of tuberculosis from exposure to tobacco smoke: a systematic review and meta-analysis. *Archives of internal medicine* **167**(4), 335–342.

Bhunu C, Mushayabasa S and Smith R (2012) Assessing the effects of poverty in tuberculosis transmission dynamics. *Applied Mathematical Modelling* **36**(9), 4173–4185.

Blower SM, Mclean AR, Porco TC, Small PM, Hopewell PC, Sanchez MA and Moss AR (1995) The intrinsic transmission dynamics of tuberculosis epidemics. *Nature medicine* **1**(8), 815-821.

Bock P and Cox H (2017). Acute care–an important component of the continuum of care for HIV and tuberculosis in developing countries. *Anaesthesia* **72**(2), 147–150.

Cohen S, Murphy ML and Prather AA (2019) Ten surprising facts about stressful life events and disease risk. *Annual review of psychology* **70**, 577-597.

Collins D, Hafidz F and Mustikawati D (2017) The economic burden of tuberculosis in indonesia. *The International Journal of Tuberculosis and Lung Disease* **21**(9), 1041–1048.

Daniel TM (2006) The history of tuberculosis. *Respiratory Medicine* **100**(11), 1862 – 1870.

Davis AL (2000) A historical perspective on tuberculosis and its control. In Reichman, LB and Earl, SH (eds) *Tuberculosis: a comprehensive international approach*. CRC Press, pp. 3–54.

de Oliveira HB and Moreira Filho DDC (2000) Tuberculosis recurrence and its risk factors. *Pan American Journal of Public Health* **7**(4), 232-241.

Dhabhar FS and McEwen BS (1999) Enhancing versus suppressive effects of stress hormones on skin immune function. *Proceedings of the National Academy of Sciences*, **96**(3), 1059-1064.

Duarte R, Lönnroth K, Carvalho C, Lima F, Carvalho A, Muñoz-Torrico M and Centis R (2018) Tuberculosis, social determinants and co-morbidities (including HIV). *Pulmonology* **24**(2), 115 – 119.

Dye C, Scheele S, Dolin P, Pathania V, Raviglione MC (1999) Global burden of tuberculosis: estimated incidence, prevalence, and mortality by country. *Jama* **282**(7), 677–686.

Fuady A, Houweling TA, Mansyur M and Richardus JH (2018) Catastrophic total costs in tuberculosis-affected households and their determinants since Indonesia's implementation of universal health coverage. *Infectious diseases of poverty* **7**(1), 3.

Fogel N (2015) Tuberculosis: a disease without boundaries. *Tuberculosis* **95**(5), 527–531.

Hosmer Jr DW, Lemeshow S and Sturdivant RX (2013) *Applied logistic regression* (Vol. 398). John Wiley & Sons.

Keflie TS, Samuel A, Woldegiorgis AZ, Mihret A, Abebe M and Biesalski HK (2018) Vitamin A and zinc deficiencies among tuberculosis patients in Ethiopia. *Journal of clinical tuberculosis and other mycobacterial diseases* **12**, 27-33.

Keller PM (2018) Urban transmission of tuberculosis in china. *The Lancet Infectious Diseases*, **18**(7), 706 – 707.

Marçôaa R (2018) Tuberculosis and gender – factors influencing the risk of tuberculosis among men and women by age group. *Pulmonology* **24**(3), 199-202.

Mason PH, Roy A and Singh P (2017) Reciprocity-building and the importance of interdisciplinary collaboration in tuberculosis research. *Journal of biosocial science* **49**(4), 559-562.

Mason PH, Roy A, Spillane J and Singh P (2016) Social, historical and cultural dimensions of tuberculosis. *Journal of biosocial science* **48**(2), 206-232.

Morey JN, Boggero IA, Scott AB and Segerstrom SC (2015) Current directions in stress and human immune function. *Current opinion in psychology* **5**, 13-17.

O'Donnell O (2007) Access to health care in developing countries: breaking down demand side barriers. *Cadernos de saude publica* **23**(12), 2820-2834.

Organisation for Economic Co-operation and Development (2018) Employment: Labour Force Participation, by Sex and Age Group. URL: https://stats.oecd.org/index.aspx?queryid=54741 (accessed 21st May 2020).

Pawelec G (2018) Age and immunity: what is "immunosenescence"?. *Experimental gerontology* **105**, 4-9.

RAND (2018) RAND IFLS-5 Survey Description. URL: https://www.rand.org/well-being/social-and-behavioral-policy/data/FLS/IFLS/ifls5.html (accessed 21st May 2020).

Ranganathan P, Pramesh CS and Aggarwal R (2017) Common pitfalls in statistical analysis: Logistic regression. *Perspectives in clinical research* **8**(3), 148.

Rim KT and Lim CH (2014) Biologically hazardous agents at work and efforts to protect workers' health: A review of recent reports. Safety and Health at Work **5**(2), 43 – 52.

Sahiratmadja E and Nagelkerke N (2016) Smoking habit as a risk factor in tuberculosis: a case-control study. *Universa Medicina* **30**(3), 189-196

Shajahan R, Navaratnam P, Kasinathan G, Kadirvelu A and Pillai N (2016) Predictors of re-emerging tuberculosis: a review. *Open Access Library Journal* **3**(3), 1-18.

Simandan D (2010) On how much one can take: relocating exploitation and exclusion within the broader framework of allostatic load theory. *Health & Place* **16**(6), 1291-1293.

Simandan D (2017a) Considering neoliberalism, contempt and allostatic load in the social dynamics of tuberculosis. *Journal of biosocial science* **49**(4), 557-558.

Simandan D (2017b) Wisdom and the path-dependent politics of biomedical research. *Journal of biosocial science* **49**(4), 563-565.

Nhamoyebonde S and Leslie A (2014) Biological differences between the sexes and susceptibility to tuberculosis. *The Journal of infectious diseases* **209**(suppl_3), S100-S106.

World Health Organization (WHO) (2007) A WHO/The Union monograph on TB and tobacco control: joining efforts to control two related global epidemics. URL: https://apps.who.int/iris/handle/10665/43812 (accessed 21st May 2020).

World Health Organization (WHO) (2018) Global tuberculosis report 2018. URL: https://www.who.int/tb/publications/global_report/en/ (accessed 21st May 2020).

Zvonareva O, van Bergen W, Kabanets N, Alliluyev A and Filinyuk O (2019) Experiencing syndemic: disentangling the biosocial complexity of tuberculosis through qualitative research. Journal of biosocial science **51**(3), 403-417.

Table 1. A cross-tabulation for each explanatory variable against TB Diagnosis Status.

Explanatory Variables	Categories	Diagnosed with TB, n(%)	Not diagnosed with TB, n(%)	Total, n(%)	P value
Individual-Level Variables:					
Age	14-30	85(0.69)	12229(99.31)	12314(100)	0.000
	31-43	105(0.99)	10525(99.01)	10630(100)	
	44+	138(1.22)	11160(98.78)	11298(100)	
Cigarette Consumption (cigar/ettes per day)[a]		12.35(8.04)	12.23(8.56)	12.23(8.56)	0.876
Diagnosed with Asthma	Yes	45(4.5)	956(95.5)	1001(100)	0.000
	No	283(0.85)	32963(99.15)	33246(100)	
Diagnosed with other lung condition	Yes	43(6.96)	575(93.41)	618(100)	0.000
	No	284(0.84)	33342(99.16)	33626(100)	
Education Level	Elementary/Kindergarten/None	108(1.05)	10148(98.95)	10256(100)	0.108
	High School	149(0.87)	16979(99.13)	17128(100)	
	Adult Education / Pesantren	3(2.33)	126(97.67)	129(100)	
	College/University	6(1.75)	336(98.25)	342(100)	
	Other	49(1.06)	4559(98.94)	4608(100)	
Food Consumption Satisfication	Less than adaquate	39(0.97)	3965(99.03)	4004(100)	0.586
	Just adaquate	161(0.87)	18412(99.13)	18573(100)	
	More than adaquate	87(0.98)	8773(99.02)	8860(100)	
Gender	Male	183(1.12)	16146(98.88)	16329(100)	0.002
	Female	145(0.81)	17769(99.19)	17914(100)	
Health Benefit or Insurance	Yes	202(1.21)	16507(98.79)	16709(100)	0.000
	No	126(0.72)	17348(99.28)	17474(100)	
Health Facility Location Knowledge	Yes	76(1.54)	4860(98.46)	4936(100)	0.407
	No	13(1.71)	746(98.29)	759(100)	
Salary, Annual(rupiah)[a]		25074366.34 (30218003.89)	20473061.51 (30438212.77)	30438024.89 (273838.87)	0.130
Stress at Work	High	29(1.33)	2151(98.67)	2180(100)	0.021
	Low	179(0.87)	20438(99.13)	20617(100)	
Household-Level Variables:					
Number of people at household(#)[a]		4.26(178)	4.24(1.91)	4.24(1.91)	0.862
Toilet used by Householders	Toilet with Septic Tank	245(0.96)	25343(99.04)	25588(100)	0.410
	Toilet without Septic Tank	22(0.77)	2851(99.23)	2873(100)	
	Other	61(1.06)	5677(98.94)	5738(100)	
Community-Level Variables:					
Medical Facility, Distance from(km)[a]		8.11(34.87)	7.32(38.27)	7.33(38.21)	0.858
Urban or Rural area	Urban	223(1.1)	20063(98.9)	20286(100)	0.001
	Rural	105(0.75)	13858(99.25)	13963(100)	
Work Sector	Agriculture/Mining/Manufacture/ Electrical/Construction	80(0.71)	11122(99.29)	11202(100)	0.001
	Wholesale/Transportation/ Finance/Social Services	131(1.02)	12668(98.98)	12799(100)	
	Other	8(2.39)	327(97.61)	335(100)	
[a]Mean(SD)					

With p values for the ANOVA or Chi-Square tests.

Table 2. Results from a binary logistic regression (predicting TB) with odds ratios

Explanatory Variables	OR	95% CI for OR		P value
		Lower	Upper	
Individual-Level Variables:				
Age (ref. category = '14-30')				0.01
Age(31-43)	1.10	0.69	1.74	0.69
Age(44+)	2.76	1.44	5.29	0.00
Cigar/ette Consumption	1.00	0.98	1.02	0.93
Education Level (ref. category= 'Elementary School or Kindergarten or None')				0.38
Education Level (High School)	0.67	0.40	1.14	0.14
Education Level (Adult Education or Pesantren)	0.29	0.04	2.21	0.23
Education Level (College or University)	0.40	0.09	1.74	0.22
Education Level (Other)	0.87	0.40	1.88	0.71
Food Consumption (ref.category = 'Less than Adequate')				0.82
Food Consumption (Adequate)	1.13	0.62	2.06	0.69
Food Consumption (More than Adequate)	0.98	0.49	1.98	0.96
Gender (ref.category = 'Female')	1.07	0.33	3.47	0.91
Health Benefit or Insurance (ref.category = 'Yes')	1.53	1.01	2.34	0.05
Stress at Work (ref.category = 'Low')	2.34	1.43	3.83	0.00
Household-Level Variables:				
Number of Household Occupiers	1.05	0.94	1.18	0.41
Toilet used by Householders (ref.category= 'toilet with septic tank')				0.81
Toilet used by Householders (toilet without septic tank)	1.31	0.56	3.04	0.53
Toilet used by Householders (other)	0.97	0.54	1.74	0.93
C ommunity-Level Variables:				
Urban or Rural Area (ref.category = 'Urban')	1.80	1.10	2.95	0.02
Work Sector (ref. category= 'Agriculture/ Mining/Manufacturing/Electrical/ Construction')				0.38
Work Sector (Wholesale/Transportation/ Finance/Social Services)	0.74	0.47	1.15	0.18
Work Sector (other)	0.63	0.15	2.68	0.53
Constant	14.45			0.01