

1 **Hierarchical true prevalence, risk factors and clinical symptoms of tuberculosis**
2 **among suspects in Bangladesh**

3 **Short title:** Prevalence, risk factors and clinical symptoms of tuberculosis

4 Mohammad Kamruzzaman Khan^{1,2}, Md. Nazimul Islam¹, Jayedul Hassan³, Shaymal

5 Kumar Paul⁴, M. Ariful Islam¹, Konstantinos Pateras⁵, Polychronis Kostoulas⁵, Michael P.

6 Ward⁶, A. K. M. Anisur Rahman¹, Md. Mahbub Alam^{1*}

7 ¹Department of Medicine, Faculty of Veterinary Science, Bangladesh Agricultural
8 University, Mymensingh, Bangladesh

9 ²Department of Community Medicine, Mymensingh Medical College, Bangladesh

10 ³Department of Microbiology and Hygiene, Faculty of Veterinary Science, Bangladesh
11 Agricultural University, Mymensingh, Bangladesh

12 ⁴Department of Microbiology, Netrokona Medical College, Netrokona, Bangladesh

13 ⁵Laboratory of Epidemiology & Artificial Intelligence, Faculty of Public Health, University
14 of Thessaly, Greece

15 ⁶Sydney School of Veterinary Science, The University of Sydney, Camden, New South
16 Wales, Australia

17 *Corresponding author's email: asamahbub2003@yahoo.com

18 Mobile: +8801733792435

19 Address: Department of Medicine, Faculty of Veterinary Science, Bangladesh Agricultural
20 University, Mymensingh-2202, Bangladesh

21

22 **Abstract**

23 **Background**

24 The study was aimed to estimate the true prevalence of human tuberculosis (TB); identify
25 risk factors and clinical symptoms of TB; and detect rifampicin (RIF) sensitivity in
26 Bangladesh.

27 **Methods**

28 The cross-sectional study was conducted in three Bangladesh districts during 2018.
29 Potential risk factors, clinical symptoms, and comorbidities were collected from 684 TB
30 suspects. Sputum specimens were examined by LED microscopy. TB hierarchical true
31 prevalence, risk factors and clinical symptoms were estimated and identified using a
32 Bayesian analysis framework. Rifampicin sensitivity of *M. tuberculosis* (MTB) was
33 detected by GeneXpert MTB/RIF assay.

34 **Results**

35 The median TB true prevalence was 14.2% (3.8; 34.5). Although overall clustering of
36 prevalence was not found, several DOTS centers were identified with high prevalence
37 (22.3% to 43.7%). Risk factors for TB identified (odds ratio) were age (> 25 to 45 years
38 2.67 (1.09; 6.99), > 45 to 60 years 3.43 (1.38; 9.19) and individuals in
39 families/neighborhoods where a TB patient(s) has (ve) already been present (12.31 (6.79;
40 22.60)). Fatigue, night sweat, fever and hemoptysis were identified as important clinical
41 symptoms. Seven of the GeneXpert MTB/RIF positive sputum specimens (65) were
42 resistant to rifampicin.

43 **Conclusions**

44 About one in every seven TB suspects was affected with TB. Around 11% of the TB
45 patients carry multi drug resistant MTB. Hierarchical true prevalence estimation allowed
46 identifying DOTS centers with high TB burden. Insights from this study will enable more
47 efficient use DOTS centers-based TB surveillance to end the TB epidemic in Bangladesh
48 by 2035.

50 **Keywords:** Sputum, Auramine staining, GeneXpert MTB/RIF, Bayesian analysis.

51

52

53

54 Introduction

55 TB has accompanied mankind during its evolution (Jordao and Vieira, 2011) and
56 continues to exert an enormous toll on human health despite the availability of effective
57 anti-tubercular drugs for more than 50 years. There are about 10.4 million new
58 *Mycobacterium tuberculosis* infections and 1.4 million deaths per year and TB is one of the
59 top 10 causes of death worldwide (Koch, 1882; WHO, 2016). Estimated global TB
60 incidence and global TB mortality, in 2017, was 133 and 21 per 100 000 population,
61 respectively (WHO, 2018). The global distribution of TB cases is overrepresented by low-
62 income countries with emerging economies. The highest prevalence of cases is in Asia
63 where China, India, Bangladesh, Indonesia and Pakistan collectively make up over 50% of
64 the global burden (Dye et al. 2005). TB is often described as a barometer of social welfare.
65 Poor quality of life, poor housing, overcrowding, under-nutrition, smoking, lack of
66 education, large families, lack of awareness regarding cause and spread of TB– all these
67 social factors are interrelated and contribute to the occurrence and transmission of TB
68 (Park, 2015).

69 Clinical manifestation varies with TB type and immune status of the patient.
70 Pulmonary complaints are evident in 85% of cases and clinical features significantly
71 associated with active pulmonary TB are cough with expectoration for 2 weeks, low grade
72 evening fever, hemoptysis, loss of appetite, weight loss, fatigue, chest pain and shortness of
73 breath are common (Banu et al. 2013; Hossain et al. 2015; Herchline, 2020). Diagnosis of
74 pulmonary TB is difficult and often requires a combination of tests. Smear microscopy is
75 rapid but with a poor sensitivity which rarely exceeds 68% (WHO, 2009). Culture-based
76 methodology – in solid or liquid media – is considered the gold standard, but it requires
77 several weeks to months for a result (Richter et al. 2009). Emergence of multidrug resistant
78 TB (MDR-TB) is another limitation in TB diagnosis and control (Mathema et al. 2006). To
79 overcome these limitations, rapid screening methods such as GenoTypeMTBDR*plus* line
80 probe assay (LPA) and GeneXpert MTB/RIF were introduced (Ling et al. 2008). These
81 methods allow rapid detection of *M. tuberculosis* along with bacterial resistance against
82 rifampicin and/or isoniazid which are commonly used to treat TB (Ling et al. 2008).

83 Bangladesh is one of the countries with the highest TB burden worldwide (WHO,
84 2019). The country has been implementing a National Tuberculosis Control Program (NTP)
85 since 1965. A project based on the DOTS (Directly Observed Treatment, Short course)
86 strategy was initiated in Bangladesh in 1993 (Kumaresan et al. 1998). Anyone willing to
87 test for TB can attend a DOTS center free of cost. However, health workers visit every
88 household monthly and recommend people having cough for >2 weeks to visit DOTS
89 center for TB testing. The DOTS strategy is now being implemented throughout the whole
90 country via 1131 DOTS centers which have sputum smear microscopy facilities
91 [[https://www.ntp.gov.bd/DOTS list](https://www.ntp.gov.bd/DOTS%20list)]. The DOTS centers are a one stop service where
92 patients receive free diagnostic facilities and supervised treatment with anti-TB drugs. The
93 DOTS strategy helps to reduce treatment failure, lost to follow up and development of
94 MDR TB.

95 Previous studies aiming to elucidate aspects of the human TB epidemiology in
96 Bangladesh were based on the general population or slum dwellers (Banu et al. 2013;

97 Hossain et al. 2015; Zaman et al. 2006). DOTS centers have been widespread throughout
98 the country since 2007 (100% DOTS coverage) and people are also aware about free
99 diagnosis and treatment. DOTS centers have successfully detected and treated 96.4% of
100 bacteriologically confirmed new pulmonary TB cases in 2018 (NTCP, 2020). DOTS
101 centers are dedicated for the surveillance and control of TB in Bangladesh. However, the
102 true prevalence of TB, identification of risk factors and clinical symptoms among TB
103 suspects attending DOTS centers have not been studied previously. Therefore, the aims of
104 this study were to (i) estimate true prevalence, ii) identify risk factors and clinical features
105 of TB and (ii) detect rifampicin sensitivity among TB suspects in Bangladesh.

106 **Methods**

107 **Ethical approval**

108 The study was approved by the Institutional Review Board (IRB) of Mymensingh
109 Medical College, Mymensingh, Bangladesh (Memo no. MMC/IRB/142, date: 18/11/2017).
110 Written informed consent was obtained from all the participants before recruitment.
111 Records were kept strictly confidential.

112 **Design, location and duration of study**

113 The cross-sectional study was conducted from January to December 2018 among
114 the TB suspects attending DOTS centers in the Mymensingh, Sirajganj and Dhaka districts
115 of Bangladesh.

116 **Sample size calculation and sampling techniques**

117 Using the Guilford and Frucher formula ($n = \frac{z^2pq}{d^2}$) (Guilford and Frucher, 1973),
118 assuming a prevalence of 50% [p=0.5] (no previous report) among the subpopulation
119 suspect of TB in Bangladesh, the minimum calculated sample size was 600, at a
120 significance level of 5% and with 4% acceptable margin of error. From 8 administrative
121 divisions of Bangladesh we selected 3 divisions randomly. Again, from each selected
122 division we selected one district randomly namely Mymensingh, Sirajganj and Dhaka.
123 From each district 8 DOTS centers were selected randomly i.e. ultimately, we randomly
124 selected 24 DOTS centers from Mymensingh, Sirajganj and Dhaka districts. Finally, a
125 convenience sample of 684 TB suspects attending those DOTS centers were interviewed
126 using pretested, semi-structured case record forms, which included information on
127 symptoms, socio-demographic characteristics and risk factors for TB.

128 **Collection of specimens**

129 Each TB suspect brought a morning sputum specimen in the plastic container
130 provided by the NTP. After conducting face to face interview another sputum specimen
131 from each TB suspect was collected on the spot.

132 **Auramine staining and LED microscopy**

133 Thin smear slides were made from the purulent part of the sputum. Auramine-
134 rhodamine was used as a fluorescent dye to stain acid-fast bacteria (AFB) and examined

135 under LED microscope on high power (400X). Acid-fast organisms fluoresced bright
136 yellow or orange against a dark background (Tankeshwar, 2015).

137 **GeneXpert MTB/RIF Assay**

138 GeneXpertMTB/RIF assay was performed following the protocol of the
139 manufacturer (Cepheid Inc., Sunnyvale, CA, USA). Sputum specimens were collected in
140 containers provided and treated with sample reagent in a proportion of 2:1 and incubated
141 for 15 minutes at room temperature. Two milliliters reagent treated sample was pipetted
142 into the sample chamber of the Xpert cartridge. The Xpert cartridge was then placed into
143 the GeneXpert instrument system and run. Results were generated after 90 min.

144 **Statistical analysis**

145 An individual was considered to be TB infected if he/she was positive by LED
146 microscopy after auramine staining. TB data were entered into a spreadsheet (Microsoft
147 Excel 2010) and transferred to R 4.0.2 (Team, 2020) for analysis. Age was converted to a
148 categorical variable based on quartiles. The frequency and proportion of TB in each
149 category of independent variables were calculated using the “tabpct” function of the R
150 package “epiDisplay” (Chongsuvivatwong, 2018). The data used for the identification of
151 clinical symptoms and risk factors is available as a Supplementary File 1.
152

153 *True prevalence*

154 The overall and DOTS center level true prevalence of tuberculosis among TB
155 suspects in Bangladesh were estimated using a shiny web-application tPRiors (Pateras and
156 Kostoulas, 2021). The sensitivity (mean=0.627 and lower limit at 95% confidence
157 level=0.50) and specificity (mean=0.987 and lower limit at 95% confidence level=0.90) of
158 the auramine staining and LED microscopy (Gizaw et al. 2020) was used as prior
159 information in the model. This prior information produced Beta (18.34, 10.77) for
160 sensitivity and Beta (95.32, 5.02) for specificity of auramine staining and LED microscopy
161 in the Bayesian model. As there was no previous report on the prevalence of human TB
162 among TB suspects in Bangladesh we used a diffused prior for the prevalence
163 [Beta(0.8292, 0.8246)]. For this Beta distribution we used the following prior information
164 in the tPRiors: 50% mean prevalence, 95% level of confidence that the true value of the
165 mean is greater than the percentile value, the upper limit of the mean as 80.0% at 95% level
166 of confidence, the level of confidence that a certain fraction of the units under study has a
167 prevalence less than the ‘percentile.median’ as 81.1%, the median value of the defined
168 ‘psi.percentile’ as 82% and the value that the ‘percentile.median’ does not exceed 90% with
169 95% confidence. The influence of prior information on the posterior true prevalence was
170 evaluated by providing no prior information in the model [Beta (1,1)]. The model was run
171 using three Markov chains with 100,000 iterations each, half of which were discarded as
172 burn in. The data used for the estimation of hierarchical true prevalence of human TB is
173 provided as a Supplementary file 2.

174 *Mixed-effects Bayesian logistic regression – Univariable screening*

175 To assess the association between the individual TB status and potential risk factors
176 we used mixed-effect univariable logistic regression models with clinic as a random-effects
177 term, within a Bayesian estimation framework. All candidate explanatory variables were
178 initially screened, one-by-one. Variables with a Bayesian p-value <0.25 were then offered
179 to a full model which was, subsequently reduced by backwards elimination, until only
180 significant (P<0.05) variables remained. Collinearity among explanatory variables was
181 assessed by Cramer's phi-prime statistic (R package “vcd,” “assocstats” function). A pair of
182 variables was considered collinear if Cramer's phi-prime statistic was >0.70 (Islam et al.
183 2021).

184 *Multivariable mixed-effects Bayesian logistic regression*

185 Two separate Bayesian logistic regression models with a random-effects term at the
186 DOTS center level were used to identify (i) risk factors for and (ii) clinical symptoms
187 associated with TB according to the previously described method (Goodrich et al. 2018).
188 Confounding was checked by observing the change in the estimated coefficients of the
189 variables that remained in the final model by adding a non-selected variable to the model. If
190 the inclusion of this non-significant variable led to a change of more than 25% of any
191 parameter estimate, that variable was considered a confounder and retained in the model
192 (Dohoo et al. 2009). The two-way interactions of all variables remaining in the final model
193 were assessed for significance based on AIC values, rather than significance of individual
194 interaction coefficients. The spatial autocorrelation of TB prevalence was calculated by the
195 “Moran.I” command of the R package “ape” (Paradis and Schliep, 2019), using sub district
196 centroids. Likewise, spatial autocorrelation of residuals of the final logistic regression
197 model was estimated.

198 **Statistical software**

199 Bayesian logistic regression models were run using the “stan_glmr” function of the
200 R package “stanarm” (Muth et al. 2018). The model was run using five Markov chains with
201 16000 iterations each, half of which were discarded as burn in. The models’ convergence,
202 diagnostics, posterior estimates and posterior predictive checks were observed via a user-
203 friendly graphical user interface of the “launch_shinystan” function of the “shinystan” R
204 package (Gabry, 2018). The R code used for the Bayesian mixed-effects logistic regression
205 analysis with step-by-step explanation is provided in Supplementary File 3. The analysis
206 was performed in R 4.0.2 (Team, 2020) and tPRiors (Pateras and Kostoulas, 2021).

207 **Results**

208 **Descriptive results**

209 Sputum from 684TB suspects attending 24DOTS centers in Mymensingh, Sirajganj and
210 Dhaka districts were examined under LED microscope after auramine staining and 80
211 (11.7%; 95% Confidence Interval: 9.4; 14.4) were found positive for tubercle bacilli (Table
212 1). About 60% (50.3; 69.3) of the tested (108) individuals were GeneXpert MTB/RIF assay
213 positive. Among them 89.2% (78.5; 95.2) and 10.8% (4.8; 21.5) were rifampicin sensitive

214 and resistant, respectively. Child TB was observed only among (3/684) 0.4% (0.1; 1.3) TB
215 suspects.

216 **Table 1. Results of different tests for the diagnosis of human tuberculosis (n=684) in**
217 **selected districts of Bangladesh, 2018.**

Tests	Tested	Negative	Positive (%)	95% CI
Auramine staining and LED microscopy	684	504	80 (11.7)	9.4; 14.4
GeneXpert MTB/RIF	108	43	65 (60.2)	50.3; 69.3
			Sensitive:58 (89.2)	78.5; 95.2
			Resistant: 7 (10.8)	4.8; 21.5

218 CI: Confidence Interval

219 **True prevalence**

220 The overall median true prevalence of tuberculosis among TB suspects was 14.2% (3.8;
221 34.5) (Figure 1, Supplementary file 4). Tuberculosis was predominantly detected in Fulpur
222 (43.7%: 0.2; 87.2) followed by Trishal (43.3%: 16.4; 70.3), Nandail (26.2%: 16.9;69.3),
223 Gauripur (14.7%:9.8; 39.2), Fulbaria (22.3%: 1.2; 43.3), Sirajganj Sadar (18.9%: 6.7; 30.9)
224 and Muktagacha DOTS centers(17.7%:3.8; 31.6). The model without any prior information
225 also yielded the median prevalence of 14.1% (1.4; 32.3) indicating robustness of the model
226 to estimate the posterior TB true prevalence. The spatial distribution of TB prevalence did
227 not show significant clustering (Moran's I = 0.0171, P = 0.3185).

228 **Table 2. Risk factors retained in the final Bayesian mixed-effects multivariable logistic**
229 **regression model for tuberculosis (TB) among TB suspects in Mymensingh, Sirajganj**
230 **and Dhaka, Bangladesh**

231

Risk factors	Category	Estimate	SD	Odds ratio (95% CrI)	n_eff
Age (years)	≤ 25	0.58	0.52	1.79 (0.66, 5.20)	23,750
	>25to 45	0.97	0.47	2.67 (1.09, 6.99)	22,147
	> 45 to 60	1.23	0.48	3.43 (1.38, 9.19)	22,554
	> 60	-	-	Reference	
Tuberculosis patient in the family or neighborhood	Yes	2.51	0.30	12.31(6.79, 22.60)	43,510
	No	-	-	Reference	

232 SD= Standard deviation, CrI: Credible Interval, n_eff=effective sample size, R_{hat} is 1 for
233 all parameters

234

235 **Risk factors for TB**

236 Age and presence of tuberculosis patient in the family or neighborhood were
 237 identified as risk factors for TB. The odds of TB were 2.67 (1.09; 6.99) and 3.43 (1.38;
 238 9.19) times higher among TB suspects aged > 25 to 45 years and > 45 to 60 years,
 239 respectively, than those aged > 60 years. Moreover, the presence of TB patient in the
 240 family or neighborhood increased the risk of TB by 12.31 (6.79; 22.60) times compared to
 241 families or neighborhoods without TB patients (Table 2). Results of the univariable pre-
 242 screening are provided as a Supplementary File 5. Multicollinearity was not detected
 243 among the selected explanatory variables for multivariable regression.

244

245 **Table 3. Clinical symptoms significantly associated with human tuberculosis in the**
 246 **Bayesian mixed-effects multivariable logistic regression model**

247

Clinical symptoms	Category	Estimate	SD	Odds ratio (95% CrI)	n_eff
Fatigue	Yes	1.57	0.32	4.8 (2.6; 9.0)	56,587
	No	-	-	Reference	
Night sweat	Yes	2.17	0.64	8.7 (2.5; 31.9)	54,441
	No	-	-	Reference	
Hemoptysis	Yes	2.26	0.49	9.6 (3.7; 25.8)	50,872
	No	-	-	Reference	
Fever	Yes	2.21	0.72	9.1 (2.6; 43.3)	38,788
	No				

248 SD=Standard deviation, CrI=Credible Interval, n_eff=effective sample size, R_{hat}=1 for all
 249 parameters

250 **Clinical symptoms and comorbidities associated with TB**

251 Individuals with fatigue and experiencing night sweats were 4.8(2.6; 9.0) and 8.7
 252 (2.5; 31.9) times more likely to be TB positive, respectively. Individuals with hemoptysis
 253 and fever had 9.6 (3.7; 25.8) and 9.1 (2.6; 43.3) times higher odds of being TB positive,
 254 respectively (Table 3). Results of the univariable pre-screening are provided as a
 255 Supplementary File 6. No multicollinearity was detected among the selected explanatory
 256 variables for multivariable regression.

257

258 Both risk factor and clinical symptoms models converged well (the R-hat values
 259 were less than 1.1) and the effective sample sizes were more than the total number of
 260 iterations (Tables 2 and 3). The trace plots exhibited good mixing and showed no signs of
 261 convergence problems. The graphical posterior predictive check revealed that our models
 262 adequately fit the data. The residuals of the risk factor model did not show significant
 263 spatial autocorrelation (Moran's I = -0.0025, P = 0.5673).

264

265 **Discussion**

266 This study was conducted on 684 TB suspects visiting DOTS centers in
267 Mymensingh, Sirajganj and Dhaka districts of Bangladesh. Overall, one in every seven TB
268 suspects was found to be affected by TB and some DOTS centers had a high TB burden.
269 We identified risk factors for TB and clinical symptoms associated with TB positivity.
270 About 11% multidrug resistant MTB (MDR TB) was detected. The true prevalence and
271 MDR TB were high among TB suspects. The results of this study will enhance existing TB
272 surveillance by targeting DOTS centers with high prevalence and people with risk factors
273 and clinical symptoms and allocating resources for the management of TB and MDR TB.

274 Age and presence of a TB patient at home or neighborhood were risk factors for TB.
275 The risk of developing TB is high in close family members compared to more distant
276 relatives (Lienhardt et al. 2003). A small number of families with micro-epidemics are
277 responsible for most of the new TB cases which are more infectious. There is an extremely
278 high risk of transmission of TB in these families (Vidal et al. 1997). Socio-demographic
279 characteristics of the patients, such as age, influences the occurrence of tuberculosis
280 (Hossain et al. 2015; Zaman et al. 2006). The economically most active age group (15-54
281 years) is more vulnerable to TB than the inactive or older age group (Herchline, 2020).
282 Evidence of transmission of TB from animals to humans, which has been reported and
283 could be of major concern in Bangladesh, was not found here. Also, a range of potential
284 risk factors (Supplementary file 5) were not associated with TB occurrence. Nevertheless,
285 transmission cannot be excluded because cattle were not tested for TB and potential
286 molecular associations were not investigated. TB transmission for cattle to humans has
287 been reported through close contact or handling of cattle milk or meat and drinking raw
288 milk (Dye et al. 2006).

289 We identified fatigue, night sweats, hemoptysis, and fever (but not the presence of a
290 long lasting cough) as clinical symptoms associated with TB occurrence. Hence, patients
291 with fatigue and/or night sweats and/or hemoptysis and/or fever should be rigorously
292 examined for TB even in the absence of coughing. Fever, cough, fatigue, weight loss, chest
293 pain, hemoptysis, difficult breathing, night sweats, and loss of appetite are the common
294 clinical symptoms associated with TB (Banu et al. 2013; Hossain et al. 2015).

295 After adjusting for the sensitivity and specificity of smear microscopy the median
296 true prevalence of TB among tested people was 14.2%. We identified several DOTS
297 centers with high TB prevalence including Fulpur, Trishal, Nanadail, Muktagacha, Fulbaria
298 and Sirajganj Sadar. Although overall no significant clustering was detected, two
299 subdistricts (Fulpur and Trishal) had higher prevalence (43.7% and 43.3%). The
300 autocorrelation of the residuals of the risk factor model was also non-significant, indicating
301 that this model adequately explained the spatial distribution of risk. The existing TB
302 surveillance should prioritize DOTS centers with high prevalence for more efficient TB
303 control programs. The true prevalence of human TB in different DOTS centers will also
304 enable policy planners to allocate resources for TB treatment. Further investigation of
305 subdistricts with high TB prevalence is warranted. We included DOTS centers in 3
306 divisions out of eight and health service has workforce to reach every household in
307 Bangladesh to convey health related information. So, the true prevalence of tuberculosis
308 among TB suspects likely is representative of the whole country.

309 Almost 11% of the GeneXpert MTB/RIF positive cases were infected with
310 multidrug resistant TB. This finding will also enable policy planners to allocate resources
311 for the treatment of multidrug resistant TB. The WHO-endorsed DOTS Plus strategy that
312 adds components for MDR TB diagnosis, management and treatment, should be
313 implemented in these areas.

314 TB diagnosis in Bangladesh is mostly dependent on microscopy because of its low
315 cost. However, the sensitivity of this method is about 50% (Newage, 2020). High-tech
316 diagnostics such as GeneXpert are very expensive but efficient (sensitivity 93%) to detect
317 TB (Huard et al. 2003). Although GeneXpert would give better estimate of drug resistant
318 *Mycobacterium tuberculosis* in the study population, we could not screen all samples due to
319 financial constraints.

320 **Conclusion**

321 About one in every seven TB suspects was affected with TB and some DOTS
322 centers have high TB burden. We identified age and neighboring infections as risk factors
323 for TB, and fatigue and/or night sweats and/or hemoptysis and/or fever but not coughing as
324 symptoms that should guide TB testing. Around 11% of the TB patients carry multi drug
325 resistant MTB. These results can contribute to the more efficient use of DOTS center-
326 specific surveillance and risk-based NTC program.

327 **Acknowledgements**

328 The authors express thanks to Damien Foundation, Mymensingh for technical assistance in
329 culturing sputum specimens.

330 **Competing interests**

331 None.

332

333 **References**

334 Bangladesh NA. TB kills 129 per day in Bangladesh. 2020;
335 <https://www.newagebd.net/article/98737/tb-kills-129-per-day-in-bangladesh>. (accessed 7
336 July 2021).

337 Banu S, Rahman MT, Uddin MK, Khatun R, Ahmed T, Rahman MM, et al. Epidemiology
338 of tuberculosis in an urban slum of Dhaka City, Bangladesh. PloS One 2013 Oct
339 21;8(10):e77721. <https://doi.org/10.1371/journal.pone.0077721>

340 Chongsuvivatwong V. epiDisplay: Epidemiological data display package. R package
341 version 3.5.0.1. 2018.

342 Dohoo I, Martin W, Stryhn H. Veterinary Epidemiologic Research. 2nd ed. Canada: AVC
343 Inc., Charlottetown; 2009.

344 Dye C, Scheele S, Dolin P, Pathania V, Raviglione M. Consensus statement. Global burden
345 of tuberculosis: estimated incidence, prevalence, and mortality by in an urban community.
346 Int J Tuberc Lung Dis. 2005;9(9):970-6.

- 347 Gabry J. shinystan: Interactive Visual and Numerical Diagnostics and Posterior Analysis
348 for Bayesian Models. 2018.
- 349 Gizaw N, Abera A, Sisay S, Desta K, Kreibich S, Gerwing-Adima L et al. The yield of
350 Auramine O staining using led microscopy with bleach treated sputum samples for
351 detection of pulmonary tuberculosis at St. Peter tuberculosis specialized hospital, Addis
352 Ababa, Ethiopia. *J Clin Tuberc Other Mycobact Dis* 2020;18:100140.
353 [10.1016/j.jctube.2019.100140](https://doi.org/10.1016/j.jctube.2019.100140)
- 354 Goodrich B, Gabry J, Ali I, Brilleman S. rstanarm: Bayesian applied regression modeling
355 via Stan. R package version 2.17.4. 2018.
- 356 Guilford JP, Frucher B. *Fundamental Statistics in Psychology and Education*. New York:
357 MC Graw-Hill; 1973.
- 358 Herchline TE. Tuberculosis (TB) clinical presentation. Medscape. 2020.
359 <https://emedicine.medscape.com/article/230802-clinical>. (accessed 15 July 2020).
- 360 Hossain S, Zaman K, Quaiyum A, Banu S, Husain A, Islam A, et al. Factors associated
361 with poor knowledge among adults on tuberculosis in Bangladesh: results from a
362 nationwide survey. *J Health Popul Nutr* 2015 Dec;34(1):1-7. [10.1186/s41043-015-0002-4](https://doi.org/10.1186/s41043-015-0002-4)
- 363 Huard RC, de Oliveira Lazzarini LC, Butler WR, van Soolingen D, Ho JL. PCR based
364 method to differentiate the subspecies of the Mycobacterium tuberculosis complex on the
365 basis of genomic deletions. *J Clin Microbiol* 2003;41:1637–50. [10.1128/JCM.41.4.1637-
366 1650.2003](https://doi.org/10.1128/JCM.41.4.1637-1650.2003)
- 367 Islam MN, Khan MK, Khan MF, Kostoulas P, Rahman AKMA, Alam MM. Risk factors
368 and true prevalence of bovinetuberculosis in Bangladesh. *PLoS One* 2021;16(2):e0247838.
369 <https://doi.org/10.1371/journal.pone.0247838>.
- 370 Jordao L, Vieira O. Tuberculosis: New aspects of an old disease. *Int J Cell Biol* published
371 online first: 21 Jun 2011. <https://doi.org/10.1155/2011/403623>
- 372 Koch R. Die Aetiologie der Tuberculosis. *Berl Klinische Wochenschr* 1882;19:221–230.
- 373 Kumaresan JA, Ali A, Md AK, Parkkali L M. Tuberculosis control in Bangladesh: success
374 of the DOTS strategy. *Int J Tuberc Lung Dis* 1998 Dec 1;2(12):992-998. PMID: 9869115
- 375 Lienhardt C, Fielding K, Sillah J, Tunkara A, Donkor S, Manneh K, et al. Risk factors for
376 tuberculosis infection in sub-saharan Africa: a contact study in The Gambia. *Am J Respir*
377 *Crit Care Med* 2003; 168:448–55. [10.1164/rccm.200212-1483OC](https://doi.org/10.1164/rccm.200212-1483OC)
- 378 Ling DI, Zwerling AA, Pai M. Rapid diagnosis of drug-resistant TB using line probe
379 assays: from evidence to policy. *Expert review of respiratory medicine*. 2008 Oct
380 1;2(5):583-8. [10.1586/17476348.2.5.583](https://doi.org/10.1586/17476348.2.5.583)
- 381 Mathema B, Kurepina NE, Bifani PJ, Kreiswirth BN. Molecular epidemiology of
382 tuberculosis: current insights. *Clin Microbiol Rev* 2006;19(4):658–85.
383 [10.1128/CMR.00061-05](https://doi.org/10.1128/CMR.00061-05)

- 384 Muth C, Oravecz Z, Gabry J. User-friendly Bayesian regression modeling: A tutorial with
385 rstanarm and shinystan. *Quant Meth Psych* 2018;14(2):99–119.
386 <http://dx.doi.org/10.20982/tqmp.14.2.p099>
- 387 National tuberculosis control program. Available at: [http://www.ntp.gov.bd/wp-](http://www.ntp.gov.bd/wp-content/uploads/2021/09/Annual-Report-2020.pdf)
388 [content/uploads/2021/09/Annual-Report-2020.pdf](http://www.ntp.gov.bd/wp-content/uploads/2021/09/Annual-Report-2020.pdf) (Accessed 20 November 2021).
- 389 Paradis E, Schliep K. ape 5.0: an environment for modern phylogenetics and evolutionary
390 analyses in R. *Bioinformatics* 2019; 35: 526-528. [10.1093/bioinformatics/bty633](https://doi.org/10.1093/bioinformatics/bty633)
- 391 Park K. Park's Textbook of Preventive and Social Medicine. 23rd ed. Jabalpur, India: M/S
392 Banarsidas Bhanol Publishers; 2015. p97-98, p176–202.
- 393 Pateras K, Kostoulas P. tPRiors: A tool for prior elicitation and obtaining posterior
394 distributions of true disease prevalence, 01 November 2021, Preprint (Version 1) available
395 at Research Square. <https://doi.org/10.21203/rs.3.rs-1019762/v1>
- 396 Richter E, Rusch-Gerdes S, Hillemann D. Drug-susceptibility testing in TB: current status
397 and future prospects. *Expert Rev Respir Med* 2009;3:497–510. [10.1586/ers.09.45](https://doi.org/10.1586/ers.09.45)
- 398 Tankeshwar A. Auramine-Rhodamine Fluorochrome Staining.
399 [https://microbeonline.com/auramine-rhodamine-fluorochrome-staining-principle-](https://microbeonline.com/auramine-rhodamine-fluorochrome-staining-principle-procedure-results-limitations/)
400 [procedure-results-limitations/](https://microbeonline.com/auramine-rhodamine-fluorochrome-staining-principle-procedure-results-limitations/) (Accessed on 30 July 2021).
- 401 Team RC. R: A language and environment for statistical computing, version 4.0.2. Vienna,
402 Austria: the R project for Statistical Computing. 2020.
- 403 Vidal R, Miravittles M, Cavla JA, Torrella M, Gracia J de, Morell F. Increased risk of
404 tuberculosis transmission in families with microepidemics. *Eur Respir J*. 1997;10(6):1327-
405 1331. [10.1183/09031936.97.10061327](https://doi.org/10.1183/09031936.97.10061327)
- 406 WHO. Global tuberculosis control 2009. Geneva, Switzerland: World Health Organization
407 Press 2010. PMID: 20210259
- 408 WHO. Global Tuberculosis Report, 2016.
409 <http://apps.who.int/iris/bitstream/10665/250441/1/9789241565394-eng.pdf>. (accessed 27
410 April 2018).
- 411 WHO. Global Tuberculosis Report, 2018.
412 <https://www.who.int/publications/i/item/9789241565646> (accessed 12 June 2019).
- 413 WHO. Global Tuberculosis Report, 2019.
414 <https://www.who.int/publications/i/item/9789241565714> (accessed 3 June 2020).
- 415 Zaman K, Yunus M, Arifeen SE, Baqui AH, Sack DA, Hossain S, et al. Prevalence of
416 sputum smear-positive tuberculosis in a rural area in Bangladesh. *Epidemiol Infect*
417 2006;134(5):1052-59. [10.1017/S0950268806006108](https://doi.org/10.1017/S0950268806006108)

418 **Figure and supplementary files caption**

419 Figure 1. Boxplots showing the DOTS centers level true prevalence of TB among suspects
420 in Mymensingh, Sirajganj and Dhaka districts of Bangladesh.

421 Supplementary File 1 (.xlsx). Data used to identify clinical symptoms and risk factors for
422 human tuberculosis among humans attending at DOTS centers.

423 Supplementary File 2 (.xlsx). Data used to estimate the true prevalence of tuberculosis
424 among humans attending at DOTS centers.

425 Supplementary File3 (.txt). The R code used to identify clinical symptoms and risk factors
426 for human tuberculosis using Bayesian mixed-effects logistic regression models.

427 Supplementary File 4 (.pdf). tPRiors dynamic report of the main Bayesian model to
428 estimate the overall and DOTS centers based true prevalence of human TB.

429 Supplementary File 5 (.docx). Univariable association of demographic and other risk
430 factors with human tuberculosis (TB) in Mymensingh, Sirajganj and Dhaka, Bangladesh
431 (n=684).

432 Supplementary File 6 (.docx). Clinical symptoms and comorbidities associated with human
433 tuberculosis (TB) based on univariable mixed-effects Bayesian logistic regression analyses
434 (n=684).

435

Posterior true prevalence

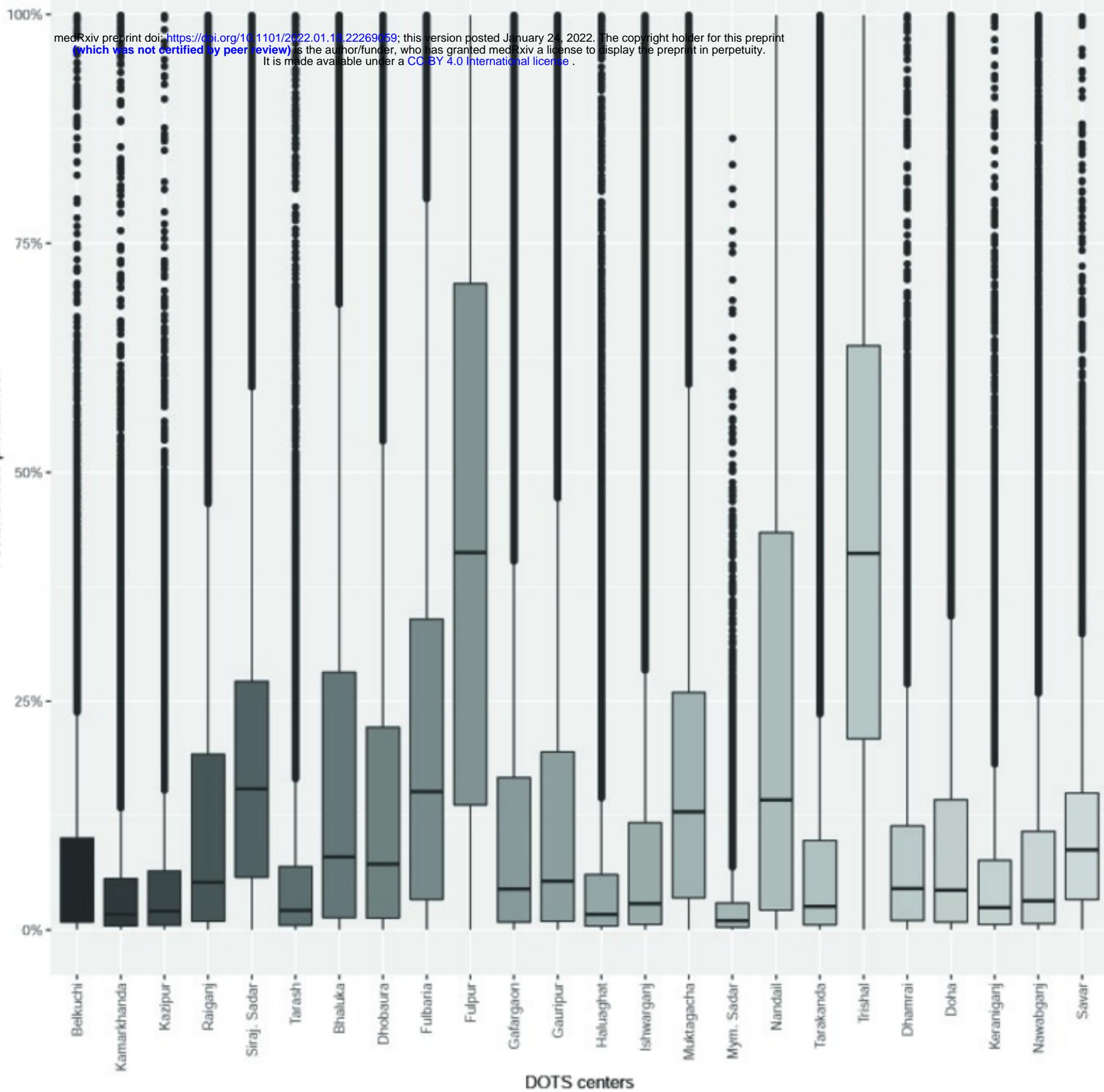


Figure 1