

SUPPLEMENTARY MATERIAL

Sociodemographic, behavioral, and life-course risk factors for tuberculosis in older Indian adults: evidence from a nationally representative survey

Saurav Basu,¹ Shubhanjali Roy,² Puja Samanta,³ Suneela Garg⁴

¹Department of Community Medicine, Employees State Insurance Post-graduate Institute of Medical Science and Research, Joka, Kolkata; ²Indian Institute of Public Health - Delhi, Public Health Foundation of India, New Delhi; ³George Institute of Global Health, New Delhi; ⁴National Academy of Medical Sciences, New Delhi, India

Correspondence: Saurav Basu, Room No. 433, Department of Community Medicine, ESI-PGIMSR & ESIC Medical College, Joka, Diamond Harbor Road, Kolkata, India 700104.

E-mail: saurav.basu1983@gmail.com

Key words: tuberculosis, childhood deprivation, multimorbidity, LASI, India.

Supplementary Table 1. Diagnostic summary of survey-weighted logistic regression models.

Diagnostic	Model I	Model II
Analytical sample, n	30,449	57,765
TB cases, n (%)	243 (0.80)	531 (0.92)
Non-TB, n (%)	30,206 (99.20)	57,234 (99.08)
AUC	0.6755	0.7027
Standard error of AUC	0.0191	0.0114
95% CI for AUC	0.6382–0.7129	0.6803–0.7251
Hosmer–Lemeshow χ^2	2.40	10.08
Hosmer–Lemeshow df	8	8
Hosmer–Lemeshow p-value	0.9664	0.2598
Mean VIF	1.60	1.41
Maximum VIF	3.25	1.91

Abbreviations: AUC, area under the receiver operating characteristic curve; CI, confidence interval; df, degrees of freedom; TB, tuberculosis; VIF, variance inflation factor; χ^2 , chi-square statistic. The Hosmer–Lemeshow test was used to assess model calibration, while VIF values were used to assess multicollinearity.

Supplementary Table 2. Calibration of survey-weighted logistic regression models by deciles of predicted tuberculosis risk.

Decile	Model I observed TB probability	Model I predicted TB probability	Model II observed TB probability	Model II predicted TB probability
1	0.0059	0.0016	0.0017	0.0022
2	0.0030	0.0025	0.0047	0.0033
3	0.0039	0.0032	0.0040	0.0041
4	0.0020	0.0039	0.0052	0.0050
5	0.0076	0.0048	0.0044	0.0060
6	0.0062	0.0060	0.0075	0.0073
7	0.0056	0.0076	0.0102	0.0091
8	0.0089	0.0100	0.0126	0.0117
9	0.0115	0.0148	0.0137	0.0160
10	0.0253	0.0349	0.0279	0.0315

Deciles were created by ranking participants according to model-predicted tuberculosis risk and dividing them into ten approximately equal groups. Calibration was assessed by comparing observed and predicted tuberculosis probabilities within each decile.

Supplementary Table 3. Hosmer-Lemeshow goodness-of-fit assessment by deciles of predicted tuberculosis risk.

Model	Decile	Observed TB	Expected TB	Observed non-TB	Expected non-TB	Total
Model I	1	3	5.5	3042	3039.5	3045
	2	10	8.1	3035	3036.9	3045
	3	12	10.2	3033	3034.8	3045
	4	11	12.2	3034	3032.8	3045
	5	14	14.5	3031	3030.5	3045
	6	16	17.3	3029	3027.7	3045
	7	21	21.1	3024	3023.9	3045
	8	27	27.1	3018	3017.9	3045
	9	42	39.1	3003	3005.9	3045
	10	87	87.9	2957	2956.1	3044
Model II	1	11	12.7	5774	5772.3	5785
	2	15	18.6	5803	5799.4	5818
	3	18	22.8	5714	5709.2	5732
	4	41	28.1	5775	5787.9	5816
	5	30	33.1	5702	5698.9	5732
	6	46	39.9	5734	5740.1	5780
	7	50	49.3	5727	5727.7	5777
	8	57	63.6	5724	5717.4	5781
	9	93	89.4	5679	5682.6	5772
	10	170	173.6	5602	5598.4	5772

Deciles are ten approximately equal groups ranked from lowest to highest predicted tuberculosis risk. Decile 1 indicates the lowest-risk group and Decile 10 the highest-risk group. Calibration compares observed and predicted tuberculosis probability within each decile.

Supplementary Table 4. Diagnostic summary of Firth bias-reduced logistic regression models.

Diagnostic	Model 1A	Model 2B
Number of observations	30,449	57,765
Number of predictors	36	19
Wald χ^2	211.56	367.94
Degrees of freedom	36	19
Model p-value	<0.001	<0.001
Penalized log likelihood	-1259.31	-2798.95
AUC	0.7303	0.7154
Standard error of AUC	0.0168	0.0112
95% CI of AUC	0.6974–0.7633	0.6935–0.7373

Abbreviations: AUC, area under the receiver operating characteristic curve; CI, confidence interval; SE, standard error; χ^2 , chi-square statistic. Model p-values are based on the Wald chi-square test. The penalized log-likelihood was estimated using Firth bias-reduced logistic regression.

Supplementary Table 5. Calibration of Firth bias-reduced logistic regression models by deciles of predicted tuberculosis risk.

Predicted risk decile	Model 1A Frequency	Model 1A observed TB probability	Model 1A predicted TB probability	Model 2B Frequency	Model 2B observed TB probability	Model 2B predicted TB probability
1	3046	0.000985	0.001966	5846	0.001882	0.002246
2	3050	0.003934	0.002873	5751	0.002608	0.003258
3	3039	0.003291	0.0036	5737	0.002963	0.004058
4	3045	0.003284	0.004324	5778	0.007269	0.00491
5	3045	0.005255	0.005126	5771	0.005198	0.005869
6	3045	0.004926	0.006099	5781	0.007784	0.007033
7	3045	0.006568	0.007468	5805	0.008786	0.008694
8	3045	0.009195	0.009574	5744	0.010097	0.011201
9	3045	0.013793	0.013783	5778	0.015576	0.015769
10	3044	0.028581	0.030842	5774	0.029789	0.030595

Supplementary Table 6. Distribution of participant characteristics and bivariate associations with tuberculosis status (n=59910).

Variables	Sample distribution (n=59910) Frequency (Weighted percentage)	Adults not suffering from tuberculosis (n=59360) Frequency (Weighted percentage)	Adults suffering from tuberculosis (n=550) Frequency (Weighted percentage)	P Value
Undernourished				<0.001
No	48,909 (78.61)	48,594 (78.82)	315 (57.79)	
Yes	11,001 (21.39)	10,766 (21.18)	235 (42.21)	
Food insecurity				0.135
No	32,691 (54.54)	32,413 (54.59)	278 (49.30)	
Yes	27,219 (45.46)	26,947 (45.41)	272 (50.70)	
Self-rated childhood health				0.505
Good	53,203 (87.36)	52,717 (87.34)	486 (89.35)	
Fair	5,777 (11.03)	5,727 (11.06)	50 (8.98)	
Poor	902 (1.61)	888 (1.61)	14 (1.67)	
Bedridden for a month or more during childhood				<0.001
No	56,738 (94.12)	56,244 (94.19)	494 (87.28)	
Yes	3,101 (5.88)	3,046 (5.81)	55 (12.72)	
Childhood economic status				<0.001
Pretty well-off	4,300 (7.73)	4,269 (7.75)	31 (5.68)	
Average	31,910 (49.24)	31,664 (49.36)	246 (37.70)	
Poor	23,297 (42.24)	23,027 (42.10)	270 (56.36)	
Varied	354 (0.79)	351 (0.79)	3 (0.26)	
Age (in Years)				0.132
45 to below 60	31,513 (50.07)	31,247 (50.13)	266 (43.90)	
60 to below 70	17,428 (30.01)	17,247 (30.00)	181 (30.57)	
70 and above	10,969 (19.92)	10,866 (19.87)	103 (25.53)	
Gender				0.004
Male	27,793 (45.81)	27,458 (45.70)	335 (55.98)	
Female	32,117 (54.19)	31,902 (54.30)	215 (44.02)	
Marital Status				0.430
Currently married	45,227 (74.36)	44,809 (74.33)	418 (76.91)	
Divorced/Widowed/Separated/Deserted	13,972 (24.51)	13,846 (24.53)	126 (22.52)	
Never Married	710 (1.13)	704 (1.13)	6 (0.57)	
Education				0.838
No education	6,974 (22.63)	6,912 (22.62)	62 (23.51)	
Primary	7,971 (24.69)	7,894 (24.66)	77 (28.04)	
Secondary	13,849 (42.94)	13,752 (42.97)	97 (39.63)	
Higher	2,927 (9.74)	2,908 (9.75)	19 (8.83)	
Wealth index				0.155
Poorest	11,848 (21.08)	11,743 (21.03)	105 (26.44)	
Poorer	12,095 (21.27)	11,979 (21.31)	116 (17.28)	
Middle	12,104 (20.39)	11,997 (20.40)	107 (18.94)	
Richer	12,071 (19.66)	11,962 (19.68)	109 (16.82)	
Richest	11,792 (17.60)	11,679 (17.57)	113 (20.53)	
Religion				0.882
Hindu	43,803 (82.51)	43,418 (82.50)	385 (83.68)	
Muslim	7,008 (11.03)	6,942 (11.04)	66 (10.25)	
Christian	6,012 (2.95)	5,939 (2.95)	73 (2.36)	
Others	3,083 (3.51)	3,057 (3.51)	26 (3.72)	
Caste				0.082
SC/ST	20,693 (28.80)	20,461 (28.73)	232 (35.10)	
OBC	22,715 (46.72)	22,537 (46.79)	178 (39.71)	
Others	14,429 (24.49)	14,307 (24.48)	122 (25.19)	
Working status				0.359
Not working	16,634 (26.14)	16,508 (26.17)	126 (22.98)	
Working	43,271 (73.86)	42,847 (73.83)	424 (77.02)	
Residence				0.031
Rural	39,342 (69.92)	38,932 (69.85)	410 (77.02)	

Urban	20,568 (30.08)	20,428 (30.15)	140 (22.98)	
BMI				<0.001
Underweight	11,001 (21.39)	10,766 (21.18)	235 (42.21)	
Normal	31,338 (51.52)	31,110 (51.62)	228 (41.25)	
Overweight	13,132 (20.28)	13,065 (20.36)	67 (12.04)	
Obese	4,439 (6.81)	4,419 (6.84)	20 (4.51)	
Chronic diseases				<0.001
No diseases	32,088 (53.84)	31,849 (53.97)	239 (41.59)	
1 disease	16,514 (27.56)	16,355 (27.55)	159 (28.55)	
2 or more diseases	11,304 (18.60)	11,152 (18.48)	152 (29.85)	
Tobacco use				0.294
No	37,891 (62.26)	37,595 (62.30)	296 (58.74)	
Yes	21,943 (37.74)	21,689 (37.70)	254 (41.26)	
Alcohol use				0.323
No	49,092 (84.86)	48,680 (84.88)	412 (82.90)	
Yes	10,758 (15.14)	10,620 (15.12)	138 (17.10)	
Region				<0.001
North	10,881 (12.66)	10,785 (12.69)	96 (10.30)	
South	13,988 (23.26)	13,908 (23.37)	80 (12.05)	
East	10,657 (23.81)	10,554 (23.77)	103 (27.55)	
West	7,764 (15.84)	7,660 (15.71)	104 (28.77)	
Central	8,092 (20.97)	8,011 (21.00)	81 (18.18)	
North-East	8,528 (3.47)	8,442 (3.47)	86 (3.16)	

Supplementary Table 7. Determinants of tuberculosis status: logistic regression estimates for sociodemographic, health, and deprivation factors.

Variables	Crude Odds Ratio (95% CI)	P Value	Model I Adjusted Odds Ratio (95% CI)	P Value	Model II Adjusted Odds Ratio (95% CI)	P Value
Undernourished						
No	Reference		Reference		Reference	
Yes	2.72 (2.08–3.55)	<0.001***	3.05 (2.05–4.54)	<0.001***	2.86 (2.04–4.00)	<0.001***
Food insecurity						
No	Reference		Reference		—	
Yes	1.24 (0.94–1.63)	0.135	1.30 (0.86–1.98)	0.212	—	
Self-rated childhood health						
Good	Reference		Reference		—	
Fair	0.79 (0.52–1.21)	0.286	0.53 (0.25–1.15)	0.110	—	
Poor	1.02 (0.46–2.25)	0.966	0.84 (0.28–2.51)	0.761	—	
Bedridden for a month or more during childhood						
No	Reference		Reference		Reference	
Yes	2.36 (1.59–3.51)	<0.001***	2.85 (1.49–5.44)	0.002**	2.23 (1.48–3.37)	<0.001***
Childhood economic status						
Pretty well-off	Reference		Reference		—	
Average	1.04 (0.61–1.78)	0.878	0.96 (0.49–1.89)	0.909	—	
Poor	1.83 (1.09–3.06)	0.022*	1.74 (0.86–3.52)	0.125	—	
Varied	0.45 (0.07–2.92)	0.400	0.03 (0.00–0.23)	0.001**	—	
Age (in Years)						
45 to below 60	Reference		Reference		Reference	
60 to below 70	1.16 (0.89–1.53)	0.273	1.15 (0.75–1.75)	0.521	0.93 (0.70–1.23)	0.612
70 and above	1.47 (0.96–2.23)	0.073	0.80 (0.39–1.63)	0.543	1.05 (0.65–1.69)	0.837
Gender						
Male	Reference		Reference		Reference	
Female	0.66 (0.50–0.88)	0.004**	0.82 (0.44–1.51)	0.515	0.67 (0.51–0.89)	0.006**
Marital Status						
Currently married	Reference		Reference		—	
Divorced/Widowed/Separated/Deserted	0.89 (0.66–1.20)	0.432	0.96 (0.53–1.71)	0.884	—	
Never Married	0.49 (0.14–1.74)	0.269	0.48 (0.09–2.46)	0.377	—	
Education						
No education	Reference		Reference		—	
Primary	1.09 (0.66–1.82)	0.730	1.42 (0.80–2.51)	0.230	—	
Secondary	0.89 (0.55–1.44)	0.626	1.28 (0.73–2.27)	0.391	—	
Higher	0.87 (0.35–2.17)	0.768	1.60 (0.54–4.73)	0.397	—	
Wealth index						
Poorest	Reference		Reference		Reference	
Poorer	0.64 (0.42–0.98)	0.040*	0.75 (0.38–1.49)	0.411	0.61 (0.39–0.95)	0.030*

Middle	0.74 (0.48–1.13)	0.160	0.85 (0.42–1.75)	0.666	0.74 (0.47–1.16)	0.187
Richer	0.68 (0.44–1.05)	0.079	0.76 (0.35–1.65)	0.492	0.71 (0.44–1.13)	0.150
Richest	0.93 (0.57–1.51)	0.766	0.90 (0.42–1.96)	0.797	1.05 (0.59–1.86)	0.873
Religion						
Hindu	Reference		Reference		—	
Muslim	0.92 (0.65–1.29)	0.613	0.79 (0.42–1.47)	0.450	—	
Christian	0.79 (0.40–1.54)	0.484	0.59 (0.26–1.35)	0.211	—	
Others	1.04 (0.51–2.15)	0.906	0.82 (0.26–2.57)	0.739	—	
Caste						
SC/ST	Reference		Reference		Reference	
OBC	0.69 (0.50–0.96)	0.028*	0.48 (0.30–0.77)	0.002**	0.81 (0.57–1.14)	0.228
Others	0.84 (0.60–1.17)	0.312	0.72 (0.42–1.22)	0.221	0.86 (0.60–1.23)	0.420
Working status						
Not working	Reference		Reference		—	
Working	1.19 (0.82–1.72)	0.360	0.77 (0.41–1.43)	0.406	—	
Residence						
Rural	Reference		Reference		Reference	
Urban	0.69 (0.49–0.97)	0.032*	0.89 (0.58–1.37)	0.605	0.77 (0.55–1.08)	0.133
Chronic diseases						
No diseases	Reference		Reference		Reference	
1 disease	1.34 (0.98–1.84)	0.065	1.42 (0.88–2.30)	0.149	1.51 (1.07–2.13)	0.018*
2 or more diseases	2.10 (1.47–2.99)	<0.001***	2.30 (1.36–3.89)	0.002**	2.83 (1.83–4.37)	<0.001***
Tobacco use						
No	Reference		Reference		—	
Yes	1.16 (0.88–1.53)	0.295	0.82 (0.56–1.20)	0.309	—	
Alcohol use						
No	Reference		Reference		—	
Yes	1.16 (0.87–1.55)	0.323	0.98 (0.62–1.54)	0.930	—	
Region						
North	Reference		Reference		Reference	
South	0.63 (0.42–0.97)	0.034*	0.97 (0.48–1.96)	0.925	0.67 (0.42–1.06)	0.086
East	1.43 (0.93–2.18)	0.101	0.84 (0.43–1.65)	0.612	1.27 (0.81–1.99)	0.302
West	2.25 (1.56–3.25)	<0.001***	3.44 (2.00–5.90)	<0.001***	2.47 (1.71–3.58)	<0.001***
Central	1.07 (0.74–1.52)	0.727	1.13 (0.59–2.17)	0.707	1.01 (0.69–1.47)	0.975
North-East	1.12 (0.72–1.75)	0.618	1.45 (0.69–3.06)	0.329	1.21 (0.75–1.95)	0.426

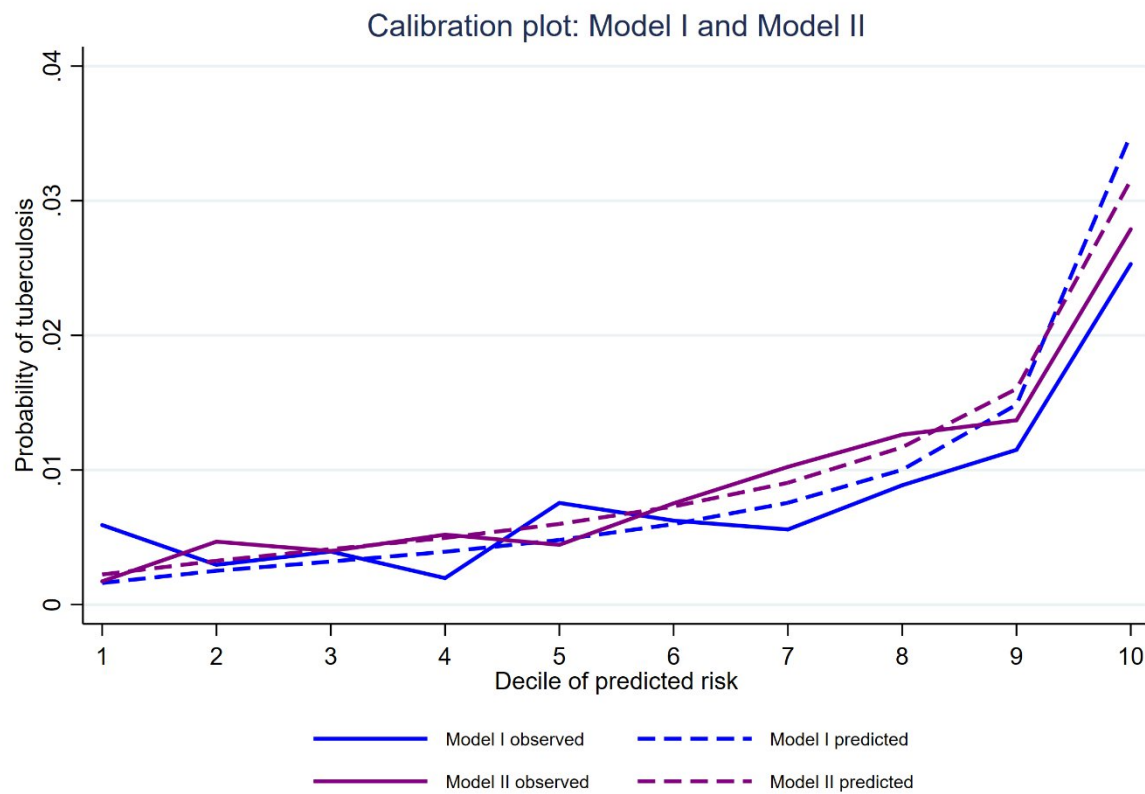
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. **Model I:** Complete model adjusted for undernutrition, food insecurity, childhood health, childhood bedridden status, childhood economic status, age, gender, marital status, education, wealth index, religion, caste, working status, residence, chronic disease status, tobacco use, alcohol use, and region [N=30,449; weighted population=100,975,971; $F(36, 30413) = 6.08$, $p < 0.001$]. **Model II:** Reduced model adjusted for undernutrition, childhood bedridden status, age, gender, residence, wealth index, caste, chronic disease status, and region [N=57,765; weighted population=205,226,846; $F(19, 57746) = 16.05$, $p < 0.001$], CI: Confidence Interval

Supplementary Table 8. Sensitivity analysis of the association between undernutrition and self-reported tuberculosis among older adults in India using Firth bias-reduced logistic regression.

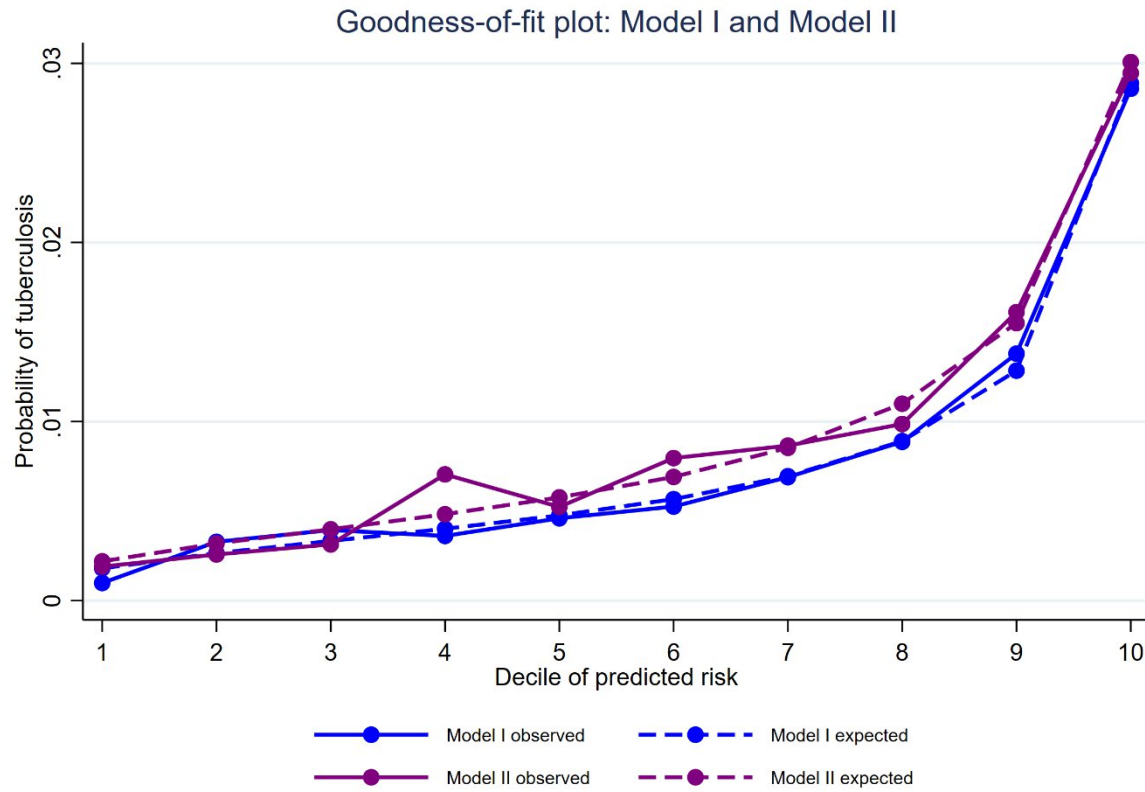
Variables	Model I A Adjusted Odds Ratio (95% CI)	P Value	Model II B Adjusted Odds Ratio (95% CI)	P Value
Undernourished				
No	Reference		Reference	
Yes	4.09 (3.04–5.49)	<0.001***	3.55 (2.94–4.29)	<0.001***
Food insecurity				
No	Reference		—	—
Yes	0.99 (0.76–1.29)	0.940	—	—
Self-rated childhood health				
Good	Reference		—	—
Fair	0.88 (0.55–1.41)	0.586	—	—
Poor	1.59 (0.77–3.30)	0.215	—	—
Bedridden for a month or more during childhood				
No	Reference		Reference	
Yes	1.84 (1.21–2.81)	0.005**	1.84 (1.39–2.45)	<0.001***
Childhood economic status				
Pretty well-off	Reference		—	—
Average	0.98 (0.60–1.62)	0.942	—	—
Poor	1.19 (0.71–1.99)	0.515	—	—
Varied	1.13 (0.21–6.12)	0.884	—	—
Age (in Years)				
45 to below 60	Reference		Reference	
60 to below 70	1.09 (0.82–1.46)	0.550	0.99 (0.81–1.21)	0.937
70 and above	0.64 (0.42–0.98)	0.039*	0.77 (0.61–0.98)	0.035*
Gender				
Male	Reference		Reference	
Female	0.51 (0.34–0.77)	0.001***	0.55 (0.46–0.66)	<0.001***
Marital Status				
Currently married	Reference		—	—
Divorced/Widowed/Separated/Deserted	1.32 (0.92–1.90)	0.131	—	—
Never Married	0.73 (0.21–2.55)	0.617	—	—
Education				
No education	Reference		—	—
Primary	1.30 (0.91–1.86)	0.143	—	—
Secondary	1.09 (0.76–1.55)	0.646	—	—
Higher	1.31 (0.74–2.32)	0.357	—	—
Wealth index				
Poorest	Reference		Reference	

Poorer	1.34 (0.86–2.07)	0.196	1.05 (0.80–1.38)	0.734
Middle	1.04 (0.66–1.65)	0.865	1.03 (0.78–1.36)	0.831
Richer	1.25 (0.80–1.95)	0.327	1.10 (0.83–1.46)	0.490
Richest	1.31 (0.84–2.05)	0.240	1.22 (0.92–1.62)	0.175
Religion				
Hindu	Reference		—	—
Muslim	1.15 (0.72–1.86)	0.554	—	—
Christian	1.50 (0.92–2.45)	0.101	—	—
Others	1.07 (0.59–1.92)	0.824	—	—
Caste				
SC/ST	Reference		Reference	
OBC	0.79 (0.57–1.10)	0.171	0.75 (0.61–0.92)	0.006**
Others	0.79 (0.55–1.13)	0.194	0.80 (0.63–1.01)	0.063
Working status				
Not working	Reference		—	—
Working	0.79 (0.52–1.20)	0.268	—	—
Residence				
Rural	Reference		Reference	
Urban	0.83 (0.62–1.12)	0.222	0.76 (0.62–0.94)	0.012**
Chronic diseases				
No diseases	Reference		Reference	
1 disease	1.38 (1.01–1.89)	0.042*	1.59 (1.29–1.96)	<0.001***
2 or more diseases	2.17 (1.56–3.02)	<0.001***	2.83 (2.26–3.54)	<0.001***
Tobacco use				
No	Reference		—	—
Yes	1.00 (0.74–1.36)	0.988	—	—
Alcohol use				
No	Reference		—	—
Yes	1.16 (0.85–1.60)	0.350	—	—
Region				
North	Reference		Reference	
South	0.65 (0.40–1.07)	0.092	0.66 (0.48–0.90)	0.009**
East	0.86 (0.54–1.38)	0.534	0.98 (0.73–1.31)	0.890
West	1.93 (1.27–2.93)	0.002^^	1.59 (1.19–2.11)	0.002**
Central	1.02 (0.62–1.67)	0.949	1.01 (0.74–1.38)	0.946
North-East	0.97 (0.57–1.66)	0.916	1.18 (0.87–1.62)	0.285

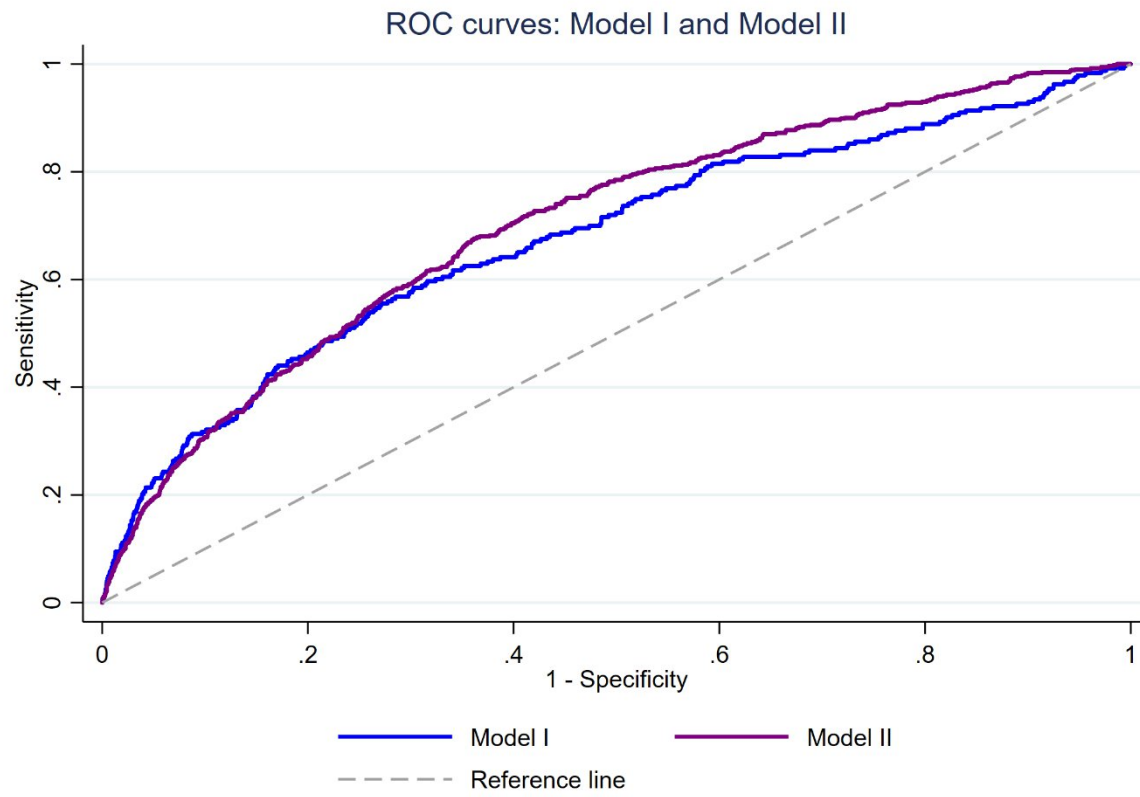
* $p < 0.05$, ** $p < 0.01$, *** $p < 0.001$. **Model I A:** Complete model adjusted for undernutrition, food insecurity, childhood health, childhood bedridden status, childhood economic status, age, gender, marital status, education, wealth index, religion, caste, working status, residence, chronic disease status, tobacco use, alcohol use, and region [N=30,449; Wald $\chi^2(36)$ =211.56, $p < 0.001$; penalized log likelihood = -1259.31; AUC=0.7303 (95% CI: 0.6974–0.7633)]. **Model II B:** Reduced model adjusted for undernutrition, childhood bedridden status, age, gender, residence, wealth index, caste, chronic disease status, and region [N=57,765; Wald $\chi^2(19)$ =367.94, $p < 0.001$; penalized log likelihood = -2798.95; AUC=0.7154 (95% CI: 0.6935–0.7373)]. CI: Confidence Interval



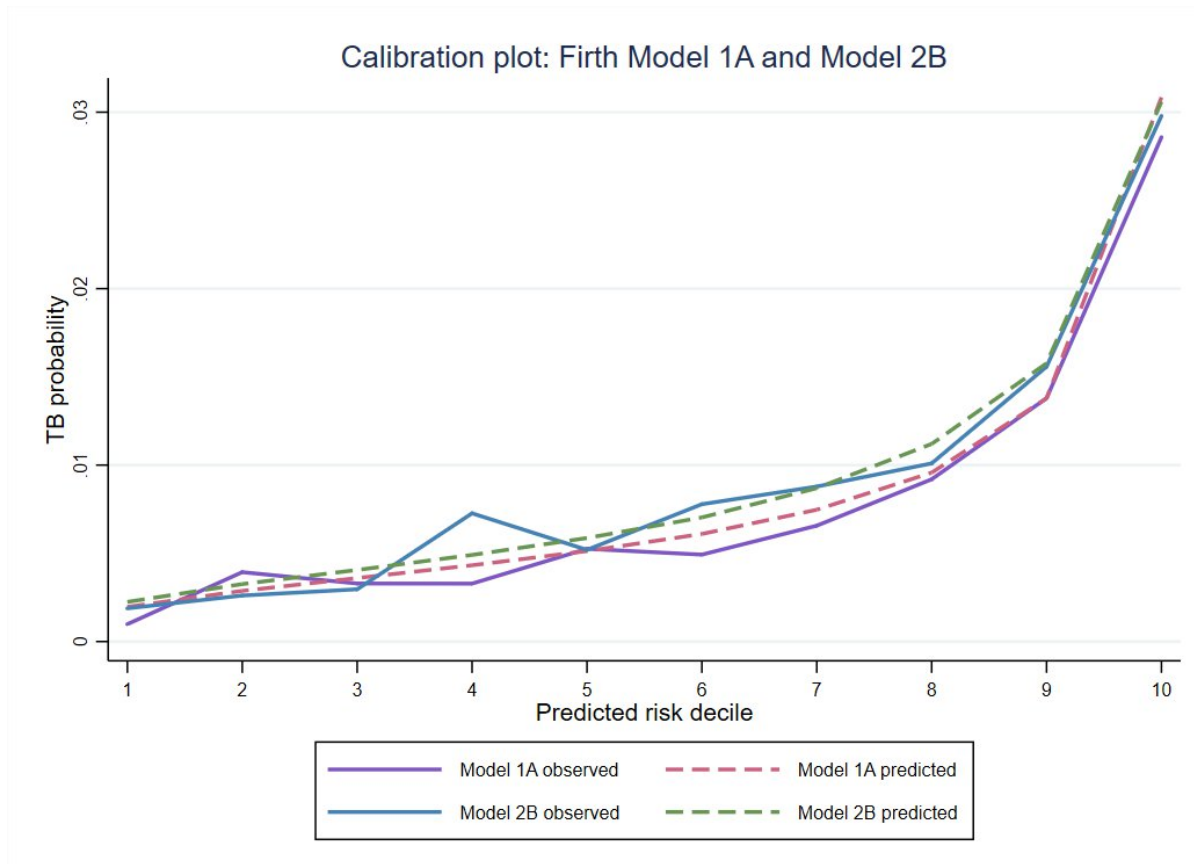
Supplementary Figure 1. Calibration of Model I and Model II for predicting self-reported tuberculosis.



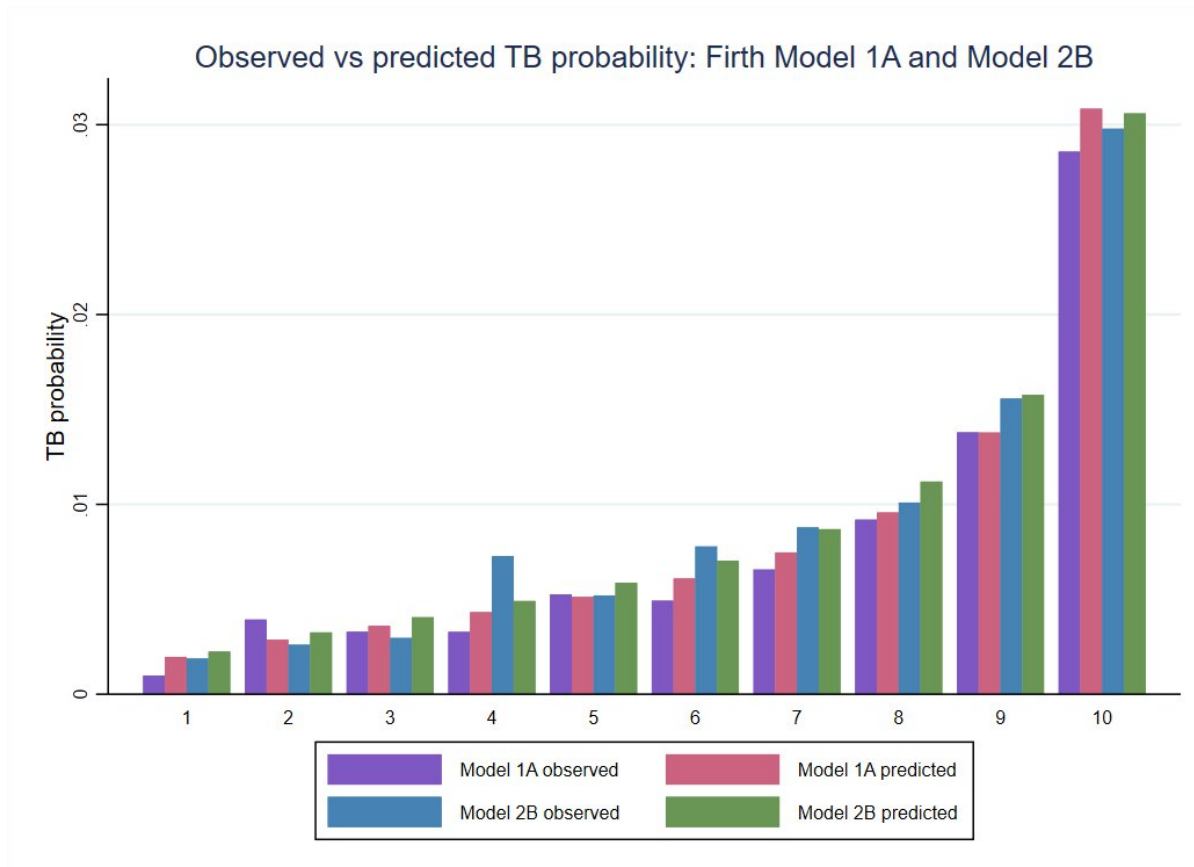
Supplementary Figure 2. Goodness-of-fit plot for Model I and Model II across deciles of predicted tuberculosis risk.



Supplementary Figure 3. Receiver operating characteristic (ROC) curves for Model I and Model II.



Supplementary Figure 4. Calibration plot comparing observed and predicted probabilities of self-reported tuberculosis across risk deciles for Firth Models 1A and 2B.



Supplementary Figure 5. Observed and predicted probabilities of self-reported tuberculosis across risk deciles for Firth Models 1A and 2B.