

Table S1. Characteristics of linked pulmonary TB cases captured by the Enhanced Tuberculosis Surveillance system and cohort review dataset for descriptive analysis, and unlinked cases, South East of England, during the study period[†] May 2011 to December 2015

Characteristic	Stratum	Linked ETS/CR dataset		ETS case without linked cohort review record		% cases unlinked	Odds ratio* of unlinked case (95% CI)	p value
		n	Col %	n	Col %			
Total cases	All	908	100	644	100	41	n/a	n/a
Notification year	2011	47	5	238	37	84	1.00	<0.001
	2012	119	13	283	44	70	0.33 (0.20-0.54)	
	2013	239	26	43	7	15	0.01 (0.01-0.02)	
	2014	261	29	39	6	13	0.01 (0.00-0.01)	
	2015	242	27	41	6	14	0.01 (0.00-0.01)	
Age at notification (years)	Age 65+	184	20	127	20	41	1.00	0.017
	Age 45-64	215	24	136	21	39	1.06 (0.75-1.53)	
	Age 15-44	499	55	359	56	42	1.18 (0.87-1.61)	
	Age 0-14	10	1	22	3	69	4.04 (1.68-9.73)	
Gender [‡] :	Female	363	40	254	40	41	1.00	0.871
	Male	544	60	389	61	42	1.02 (0.81-1.29)	
Sputum Smear [‡]	Negative	334	46	156	44	32	1.00	0.683
	Positive	415	55	201	56	33	1.06 (0.80-1.42)	
Culture	Negative	171	19	188	29	53	1.00	0.704
	Positive	737	81	456	71	38	0.52 (0.73-1.60)	
Chest X Ray ^{††}	NAD	61	9	-	-	-	n/a	n/a
	TB	653	91					
Place of birth ^{**}	Non-UK born	585	65	400	64	41	1.00	

Social risk factors	UK born	316	35	230	37	42	1.21 (0.94-1.56)	0.140
	Absent	787	87	575	89	42	1.00	
	1	75	8	47	7	39	0.89 (0.58-1.38)	
	>1	46	5	22	3	32	0.68 (0.38-1.22)	0.402
Mental Health barriers	Yes	36	4	-	-	n/a	n/a	n/a
	No	872	96					
Language barriers	Yes	807	89	-	-	n/a	n/a	n/a
	No	101	11					
Referral route to TB specialist team ^a	GP	342	44	-	-	n/a	n/a	n/a
	Secondary care	430	55					
	Prison	4	1					
	Private	5	1					
Initiated Treatment ^b	No	10	1	17	3	63	1.00	
	Yes	896	99	620	97	41	0.37 (0.15-0.92)	0.033

CI – Confidence Interval; †Based on notification dates of cohort review cases 04/05/2011 – 24/12/2015; *Taking Local Authority District as a random effect; ‡1 linked case and 1 un-linked cases missing gender data; ±159 linked cases and 287 un-linked cases without sputum smear data; **14 un-linked cases and 7 linked cases missing country of birth data; ††194 linked cases without chest x-ray data, data not available for un-linked cases; ^a127 linked cases missing referral route data, data not available for unlinked cases; ^b7 un-linked and 2 linked cases without treatment initiation data

Table S2. Characteristics of pTB cases with linked cohort review and ETS records by completeness of presentation and health care delay data, and odds ratio for data completeness, South East England 2011-2015

Characteristic	Stratum	Complete presentation delay data		Incomplete presentation delay data		Odds ratio (95% CI)*	p value	Complete health care delay data		Incomplete health care delay data		Odds ratio* (95% CI)	p value
		n	%	n	%			n	%	n	%		

Total cases	All	813	86	128	14	n/a	-	885	96	35	4	n/a	-
Notification year	2011	38	78	11	22	1.00		83	92	4	8	1.00	
	2012	102	84	19	16	0.75 (0.30-1.86)		224	98	2	2	0.20 (0.03-1.18)	
	2013	224	89	28	11	0.55 (0.23-1.28)		506	95	12	5	0.65 (0.19-2.25)	
	2014	232	86	38	14	0.77 (0.33-1.78)		529	97	9	3	0.44 (0.12-1.60)	
	2015	217	87	32	13	0.68 (0.29-1.60)	0.634	465	97	8	3	0.42 (0.12-1.56)	0.373
Age at notification	Age 65+	160	80	41	20	1.00		302	95	10	5	1.00	
	45-64	200	91	20	9	0.40 (0.22-0.74)		409	98	5	2	0.39 (0.13-1.20)	
	15-44	444	87	66	13	0.59 (0.36-0.94)		1073	96	20	2	0.68 (0.31-1.52)	
	0-14	9	90	1	10	0.52 (0.06-0.74)	0.025	23	100	0	0	Empty	0.257
Gender†	Female	327	88	46	12	1.00		356	97	11	3	1.00	0.333
	Male	485	86	82	14	1.18 (0.78-1.78)	0.432	528	96	24	4	1.44 (0.69-2.99)	
Place of birth**	Non-UK born	525	88	74	12	1.00		571	97	20	3	1.00	
	UK born	283	87	43	13	0.91 (0.58-1.43)	0.690	309	97	9	3	0.81 (0.36-1.83)	0.606
SRFs	Absent	704	86	113	14	1.00		768	96	30	4	1.00	
	1	66	86	11	14	0.91 (0.44-1.87)		74	97	2	3	0.64 (0.15-2.80)	
	>1	43	91	4	9	0.46 (0.16-1.39)	0.169	43	93	3	7	1.60 (0.45-5.69)	0.616
Sputum smear‡	Negative	300	86	50	14	1.00		330	96	21	4	1.00	
	Positive	379	90	41	10	0.59 (0.37-0.95)	0.028	406	97	13	3	0.87 (0.39-1.96)	0.741
CXR ^α	NAD	54	87	8	13	1.00		61	100	0	0	1.00	
	pTB	581	86	91	14	1.10 (0.48-2.56)	0.815	642	97	21	3	Empty	-
Culture	Negative	150	83	30	17	1.00		165	97	5	3	1.00	
	Positive	663	87	98	13	0.64 (0.39-1.05)	0.075	720	96	30	4	1.36 (0.51-3.60)	0.539
Referral route to TB specialists†	GP	315	91	32	9	1.00		340	98	7	2	1.00	
	Secondary	391	88	52	12	1.47 (0.89-2.42)		420	97	14	3	1.69 (0.66-4.33)	

	Prison	3	60	2	40	10.44 (1.39-78.81)		4	100	0	0	Empty	
	Private	4	80	1	20	2.29 (0.19-27.11)	0.082	5	100	0	0	Empty	0.278
Initiated	No	10	39	17	61	1.00		0	0	6	100	Empty	
Treatment ^{††}	Yes	801	88	109	12	0.09 (0.04-0.23)	<0.001	885	97	25	3	na	na

**Taking Local Authority District as a random effect; SRF – Social risk factor; ‡1 case missing gender data for presentation delay and health care delay; **16 cases missing place of birth data for presentation delay and 11 for health care delay; †171 cases missing sputum smear data for presentation delay and 159 for health care delay ;^a 207 cases missing CXR data for presentation delay and 196 for health care delay; † 141 cases missing referral route data for presentation delay and 130 for health care delay ;^{††} 4 cases missing treatment initiation data for presentation delay and health care delay*

Table S3. Akaike's Information Criteria (AIC) score for modelling presentation and health care delay in pulmonary TB cases in the South East of England 2011-2015 using accelerated failure time models with various distributions

Model	Distribution	AIC
Presentation delay	Weibull	2811
	Exponential	8038
	Log normal	7874
	Log logistic	7874
	Gamma	7999
Health care delay	Weibull	2162
	Exponential	6413
	Log normal	6341
	Log logistic	6356
	Gamma	6403

Table S4. Crude and adjusted time ratio* of presentation delay by risk factor, for pulmonary tuberculosis cases in the South East of England, 2011-2015 (n=771). Sensitivity analysis after exclusion of 1% of observations with longest presentation delay.

Risk factor	Level	Crude			Adjusted		
		Time Ratio	95% CI	P value	Time ratio	95% CI	P value
Ethnicity	Not white	Ref	-	-	Ref	-	-
	White	1.26	1.04-1.51	0.015	1.06	0.82-1.36	0.611
Mental Health comorbidity	No	Ref	-	-	Ref	-	-
	Yes	1.38	0.86-2.20	0.183	1.14	0.69-1.86	0.611
Rurality	Rural	Ref	-	-	Ref	-	-
	Urban	0.92	0.66-1.27	0.603	0.88	0.63-1.23	0.447
Physical health comorbidity	No	Ref	-	-	Ref	-	-
	Yes	0.90	0.66-1.21	0.487	0.81	0.60-1.11	0.189
Language barrier	No	Ref	-	-	Ref	-	-
	Yes	1.53	1.15-2.03	0.003	1.66	1.23-2.24	0.001
Gender	Female	Ref	-	-	Ref	-	-
	Male	1.11	0.93-1.32	0.261	1.06	0.89-1.27	0.527
Prior TB diagnosis	No	Ref	-	-	Ref	-	-
	Yes	0.98	0.69-1.39	0.904	0.83	0.58-1.18	0.305
Age group ^a	65+	Ref	-	-	Ref	-	-
	45-64	1.23	0.94-1.62	0.128	1.19	0.91-1.57	0.208
	15-44	0.99	0.78-1.25	0.907	1.05	0.80-1.37	0.735
	0-14	0.24	0.10-0.56	0.001	0.23	0.10-0.54	0.001
Social risk factors present ^b	None	Ref	-	-	Ref	-	-
	1	1.57	1.14-2.16	0.005	1.34	0.96-1.88	0.086

	2 or more	1.18	0.78-1.76	0.435	0.96	0.62-1.49	0.845
Time in UK ^c	UK born	Ref	-	-	Ref	-	
	11+ years	0.82	0.64-1.04	0.105	0.84	0.62-1.15	0.281
	2-11 years	0.77	0.62-0.96	0.020	0.76	0.57-1.01	0.057
	<2 years	0.94	0.70-1.26	0.662	0.86	0.61-1.22	0.397
Quintile of Index of Multiple Deprivation ^d	1 (least dep)	Ref	-	-	Ref	-	
	2	1.03	0.76-1.39	0.859	1.05	0.78-1.41	0.751
	3	1.10	0.83-1.46	0.515	1.10	0.83-1.45	0.514
	4	1.21	0.91-1.62	0.193	1.18	0.89-1.57	0.257
	5 (most)	1.29	0.91-1.82	0.149	1.32	0.93-1.85	0.116

**All models include a random effect on 62 Local Authority Districts; ^aWald p value for overall effect of age group categorical variable crude<0.001, adjusted=<0.001; ^bWald p value for overall effect of social risk factor categorical variable crude=0.145, adjusted=0.200; ^cWald p value for overall effect of time in UK categorical variable crude=0.097, adjusted=0.288; ^dWald p value for overall effect of QIMD categorical variable crude=0.510, adjusted=0.550*

Table S5. Crude and adjusted time ratio* of health care delay by risk factor, for pulmonary tuberculosis cases in the South East of England, 2011-2015 (n=614, sensitivity analysis)

Risk factor	Level	Crude			Adjusted		
		Time ratio	95% CI	P value	Time ratio	95% CI	P value
Sputum smear	Negative	Ref	-	-	Ref	-	-
	Positive	0.65	0.64-0.79	<0.001	0.63	0.52-0.78	<0.001
Time in UK ^a	UK born	Ref	-	-	Ref	-	-
	11+ years	1.23	0.94-1.60	0.130	1.12	0.81-1.53	0.514
	2-10 years	0.95	0.76-1.20	0.681	0.90	0.67-1.20	0.468
	<2 years	0.67	0.49-0.92	0.014	0.55	0.39-0.78	0.001
Referred by ^b	Primary Care	Ref	-	-	Ref	-	-
	Secondary	0.68	0.56-0.83	<0.001	0.66	0.55-0.81	<0.001
	Prison	0.55	0.17-1.77	0.318	0.59	0.19-1.90	0.380

Gender	Female	Ref	-	-	Ref	-	
	Male	0.80	0.65-0.97	0.022	0.86	0.70-1.05	0.137
Age group ^c	65+ years	Ref	-	-	Ref	-	
	45-64	0.87	0.65-1.17	0.351	0.79	0.59-1.07	0.136
	15-44	0.76	0.58-0.99	0.040	0.79	0.58-1.06	0.118
	0-14	0.55	0.17-1.77	0.315	0.62	0.20-1.92	0.403
Ethnicity	Non-white	Ref	-	-	Ref	-	
	White	1.00	0.82-1.22	0.995	1.05	0.81-1.37	0.715
Social risk ^d factor	None	Ref	-	-	Ref	-	
	1	0.68	0.50-0.92	0.014	0.88	0.63-1.23	0.450
	2 or more	0.59	0.39-0.89	0.013	0.64	0.42-0.98	0.040
LA incidence	Low/medium	Ref	-	-	Ref	-	
	High	0.86	0.66-1.16	0.251	1.02	0.76-1.36	0.913
Physical co-morbidity	No	Ref	-	-	Ref	-	
	Yes	0.98	0.70-1.36	0.891	1.02	0.74-1.42	0.884
Previous diagnosis	No	Ref	-	-	Ref	-	
	Yes	0.81	0.57-1.16	0.253	0.90	0.61 – 1.27	0.505

**All models include a random effect on 62 Local Authority Districts; ^aWald p value for overall effect of time in UK categorical variable crude=0.009, adjusted=0.001; ^bWald p value for overall effect of referred by categorical variable crude<0.001, adjusted=<0.001; ^cWald p value for overall effect of age group categorical variable crude=0.164, adjusted=0.378; ^dWald p value for overall effect of social risk factor categorical variable crude=0.003, adjusted=0.111*