SUPPLEMENTARY MATERIALS

Self-clearance of *Mycobacterium tuberculosis* infection: implications for lifetime risk and population at-risk of tuberculosis disease

Jon C. Emery[a], Alexandra S. Richards[a], Katie D. Dale[b,c], C. Finn McQuaid[a], Richard G. White[a], Justin T. Denholm[b], Rein M. G. J. Houben[a]

[a] TB Modelling Group, TB Centre and Centre for Mathematical Modelling of Infectious Diseases, Department of Infectious Disease Epidemiology, London School of Hygiene & Tropical Medicine, London WC1E 7HT, UK

[b] Victorian Tuberculosis Program, Melbourne Health, Melbourne, Victoria, Australia

[c] Department of Microbiology and Immunology, The University of Melbourne, Melbourne, Victoria, Australia

TABLE OF CONTENTS

METHODS

Data for self-clearance of *Mycobacterium tuberculosis* infection

Literature review

We reviewed the literature to identify studies from which we could attempt to infer the proportion of an initially infected cohort that have self-cleared over time since infection.

To identify such studies we considered reviews of the spectrum of *Mycobacterium tuberculosis* (*Mtb*) infection and tuberculosis (TB) disease published in the last decade (1–10) before extracting references seemingly pertinent to self-clearance of *Mtb* infection and references therein. We also considered specific references (11–18) known to the authors and references therein.

During reference extraction we considered only studies in humans in which individuals received no chemotherapeutic treatment or BCG-vaccination. We did not consider studies that presented Interferon-Gamma Release Assay (IGRA) reversions only, since the stability of IGRAs and their correlation with underlying pathophysiology remain poorly understood (19). Finally, we did not consider studies that used serial X-ray and/or PET-CT scans alone to infer self-clearance of infection.

With a total of 50 references from which to consider data extraction (11, 20–42, 12, 43–46, 13, 47–49, 14, 50–56, 15, 57–62, 17), each was reviewed for its suitability for inferring the proportion of an initially infected cohort that have self-cleared over time since infection. Studies were excluded if they did not report a known time of infection or for which a time of infection could not be reasonably estimated. Studies were also excluded if relevant details of the study setting, population or design were not sufficiently described.

The four studies that met the above criteria (46, 52, 63, 22) were then further assessed for potential biases that could lead to an overestimate of the extent of self-clearance and, where necessary, the more conservative selection was made. For tuberculin skin test (TST) reversion studies, we opted for the studies with fewest repeat TSTs to avoid biases caused by, for example, TST boosting or persistent non-converters. For autopsy studies, we opted for studies that made the most exhaustive surveys of the lung.

This further assessment resulted in two studies from which data was extracted (52, 63) and used for model fitting. See '*Self-clearance data for model fitting*' below for further details of this data.

This review was not registered in a formal database.
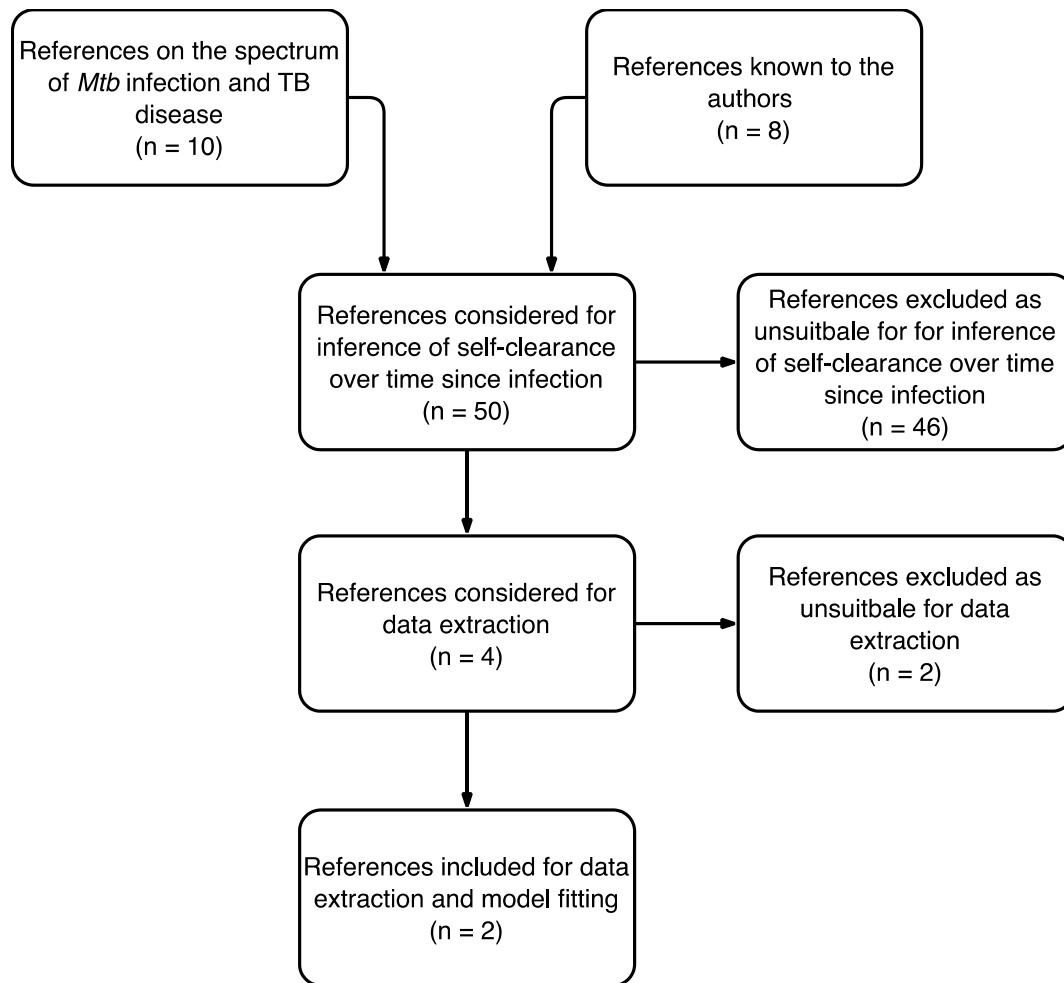
Supplementary figure 1: Flowchart of the literature review for data to infer self-clearance of *Mtb* infection. *Mtb = Mycobacterium tuberculosis.*

Cohort model and lifetime risk of tuberculosis disease

Self-clearance data for model fitting

Tuberculin skin test reversion studies

The data for self-clearance of *Mtb* infection used for model fitting was in part inferred from Ferebee (52), a follow up study to the trial detailed in (64, 65). The trial aimed to investigate the effect of Isoniazid on preventing complications, such as TB meningitis, in TST-positive children. A follow up study then performed a further TST on a subset of the trial group 10-years later. An induration of 5mm or above in response to 5 units of tuberculin was defined as TST-positive in both tests. Supplementary table 1 shows the results from the follow-up study for the placebo group.

Characteristic at study entry	Number TST retested	Number TST negative at retest
5-9mm TST induration	106	53
10+mm TST induration	111	23
0-3 years	144	68
4-6 years	55	6
7+ years	18	2

Supplementary table 1: Results from a follow-up study by Ferebee (52). A subset of initially TST-positive children entering the placebo arm of trial using Isoniazid to prevent complications such as tuberculosis meningitis (64, 65) were retested 10-years later. TST = Tuberculin skin test

The proportion of individuals that cleared their infection were inferred from the proportion that were TST-negative at retest. In order to mitigate against small, chance reversions caused by test instability (57) we include, as a single 0-7 years at study entry cohort, only those individuals that had an induration of 10mm or greater at study entry, leaving only reversions of at least 5mm in magnitude. Supplementary figure 2 shows the final self-clearance of infection data inferred from Ferebee (52), where the error bars show the equal-tailed 95% confidence interval assuming the data are beta distributed.
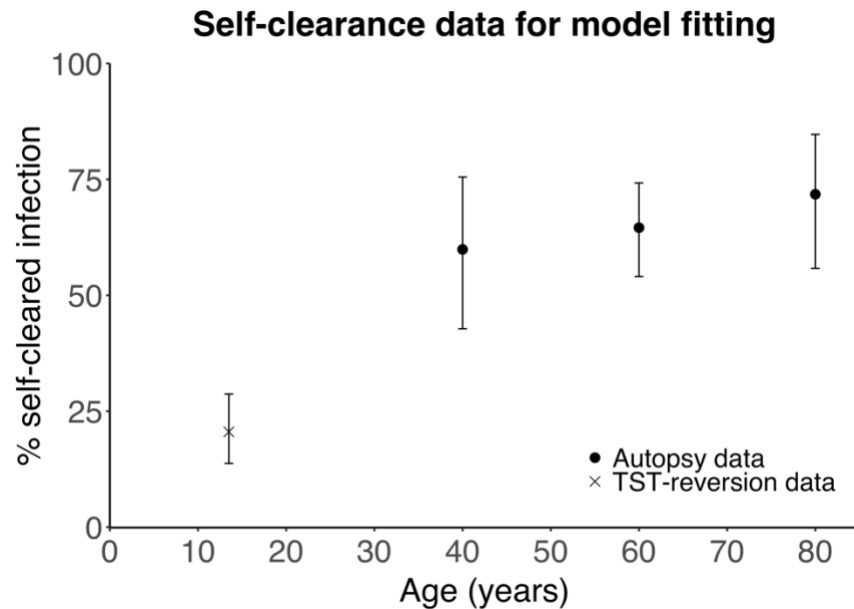
Autopsy studies

The remainder of the data for self-clearance of *Mtb* infection used for model fitting was inferred from a study by Opie and Aronson (63). The authors attempted to identify viable *Mtb* bacilli in the lungs of individuals that had histopathology consistent with a historic *Mtb* infection but died of causes other than TB. Viable bacilli were identified via guinea pig inoculation. The number of autopsies performed in each 10-year cohort and the corresponding number in which viable bacilli could be identified are presented in Supplementary table 2.

Age	Total autopsies performed	Number where viable bacilli were recovered
10-19	4	1
19-29	8	1
30-39	11	3
40-49	23	10
50-59	33	9
60-69	50	21
70-79	26	7
80-89	11	3

Supplementary table 2: Results from the autopsy study by Opie and Aronson (63) attempting to identify viable bacilli in in the lungs of individuals that had histopathology consistent with a historic *Mtb* infection but died of causes other than TB.

The proportion of individuals in each cohort that cleared their infection were inferred from the proportion in which no viable bacilli could be identified.

The data were then aggregated into four 20-year cohorts. The 10-29 years cohort was then discarded as despite aggregation the number of autopsies performed in this group was below the threshold of 30 set out in the methodology. Supplementary figure 2 shows the final self-clearance of infection data inferred from Opie and Aronson (63), where the error bars show the equal-tailed 95% confidence interval assuming the data are beta distributed.

Supplementary figure 2: Self-clearance of *Mtb* infection data used for model fitting, inferred from TST-reversion (crosses) and autopsy studies (points). Error bars show the equal-tailed 95% confidence interval assuming the data is beta distributed. *Mtb* = *Mycobacterium tuberculosis*; TST = Tuberculin skin test.

Progression to tuberculosis disease data for model fitting

First 10-years following infection

To fit progression to TB disease in the first 10-years following infection we follow (66) and generate the cumulative incidence of TB disease using summary data from the British Medical Research Council's Bacillus Calmette–Guérin (BCG) vaccine trials (67–69).

The BCG trials documented 121 cases of pulmonary TB disease in the 10 years following TST conversion amongst 2,550 individuals in the placebo group that converted with an induration greater than 5mm. Adding 41 co-prevalent and 81 extra-pulmonary cases leads to a total of 243 cases in the 10 years following infection. Finally, we distribute this total across the 10-years following infection according to the percentages reported by Styblo (70). The cumulative

number of cases displayed as a percentage of the initial at-risk population for the first 10 years following infection is shown in Supplementary figure 3. The uncertainty displayed is the equal-tailed 95th percentile confidence intervals, assuming the data are beta distributed.

Greater than 10-years following infection

To fit progression to TB disease after 10-years following infection we assume a distal progression risk of 20 per 100,000 per year applied to the BCG vaccine placebo group as if infection is lifelong, which is how distal progression risks are measured and reported. Whilst this value is broadly representative of those found in the literature (71–73), given the variance in estimates for distal progression we explore alternative values in sensitivity analyses.

Supplementary figure 3 shows the cumulative number of cases displayed as a percentage of the initial at-risk population for greater than 10 years following infection assuming a distal progression risk of 20 per 100,000 per year in the BCG vaccine placebo group. Again the uncertainty displayed is the equal-tailed 95th percentile confidence intervals, assuming the data are beta distributed.
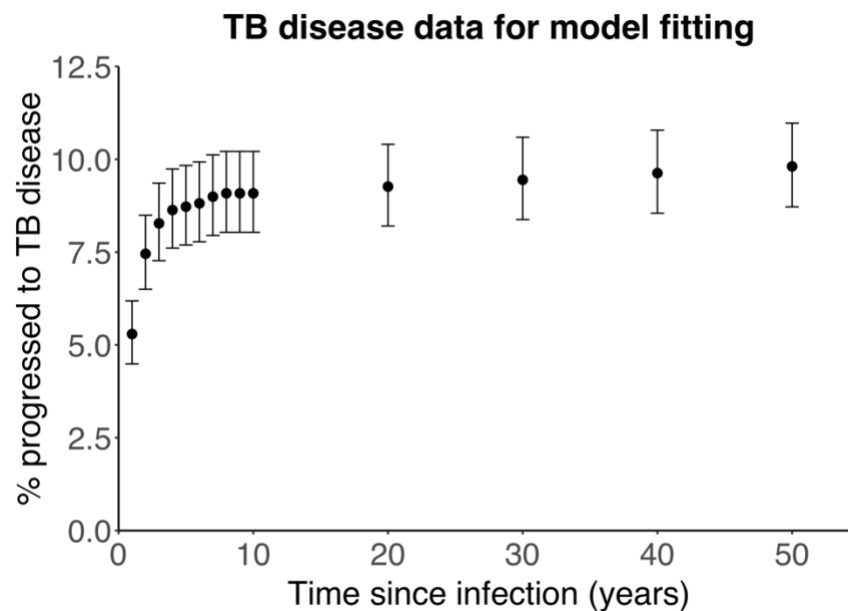
Supplementary figure 3: Progression to TB disease data used for model fitting, presented over time since infection. Error bars show the equal-tailed 95% confidence interval assuming the data is beta distributed.

Age at infection model priors

TST-reversion cohort

Since 92% of the individuals in Ferebee (52) were less than 7 years old at study entry (see Supplementary table 1), we assume a flat prior for the age at infection of the TST-reversion cohort between 0-7 years.

Autopsy cohort

To provide a model prior on the average age at first infection for the autopsy cohorts, we used an average annual risk of infection (ARI) estimated from the earliest available TST surveys in children. For a percentage p of children aged a years old that are TST-positive, the average ARI r at that time can be estimated using $r = 1-(1-p)^{1/a}$ (full details for estimating the ARI from TST survey data can be found in (74)). Supplementary table 3 shows a summary of some of the earliest available TST surveys in children, with estimates for the corresponding average ARI.

Study (Reference)	Survey year	Details	Ages	% TST-positive	Average ARI
Hamburger 1910 (75)	1910	TST survey in adolescents entering a hospital in Vienna for reasons other than TB	8 10 11	70% 80% 90%	14% 15% 19%
Styblo 1969 (76)	1910	ARI in the Netherlands from combined data			11%
Hamburger 1910 (75)	1910	TST survey in adolescents in Prague.	11-14 13-14	70% 80%	8-10% 11-12%
Heynsius 1962 (77)	1927	TST survey in adolescents in Amsterdam	8 10	46% 51%	7% 7%
Grzybowski 1975 (78)	1949	TST survey in Eskimos in Alaska			25%

Supplementary table 3: A summary of some of the earliest available TST surveys in children, with estimates for the corresponding average ARI at the time of the survey. TB = Tuberculosis; TST = Tuberculin skin test; ARI = Annual risk of infection

Although direct TST survey data is not available for the context in Opie and Aronson (63) (i.e. the USA in the late 19th and early 20th century), it is likely that the local ARI would have been at least as high as those presented in Supplementary table 3.

Based on the data in Supplementary table 3, we assumed a normal distribution for the ARI with a median of 10% and an equal-tailed 95% uncertainty interval of 5-15%. To estimate the average age at first infection for each cohort we assumed that the infection process follows a geometric distribution, such that the average age at first infection is given by *1/r*.

This then translates into a prior on the age at first infection of 10 years, with an equal-tailed 95% uncertainty interval of 6.7-20.0 years.

County-level model and population at risk of tuberculosis disease

The age-specific population at risk of developing TB disease in a given country for a particular year was estimated by applying the cohort model described in the main text to individual 5-year birth cohorts before summing over all such cohorts, analogous to the to approach used in (79) to estimate the global burden of latent TB infection.

Births in each cohort were taken from United Nations Population Division estimates of the population of each age group alive in 2019 (80) and background mortality was removed accordingly. The TB natural history components of the model were parametrised using the posterior parameter values derived from the single cohort model. A time-dependent force of infection was applied using ARI estimates.

ARI estimates were provided by re-performing the Gaussian Process Regression with linear trends in (79) with some minor alterations. Firstly, given the revision of the World Health Organisation TB disease prevalence estimates for India following the publication of (79) (see Box 3.3 in (81)), we used TST survey results only in estimating the historical ARI in India. Secondly, we extended the estimates for the ARI in each country from 2014 to 2019. Finally, we used only the median estimated ARI in the country-level model, since the focus in this work is to illustrate the relative difference between the number of individuals at risk of TB disease in the self-clearance/lifelong infection scenarios across different settings, not estimates for the absolute numbers. Uncertainty in the results will then solely be due to that of the natural history parameters.

RESULTS

Cohort model and lifetime risk of tuberculosis disease

Additional results

Cohort model fitted to TB disease data over time since infection
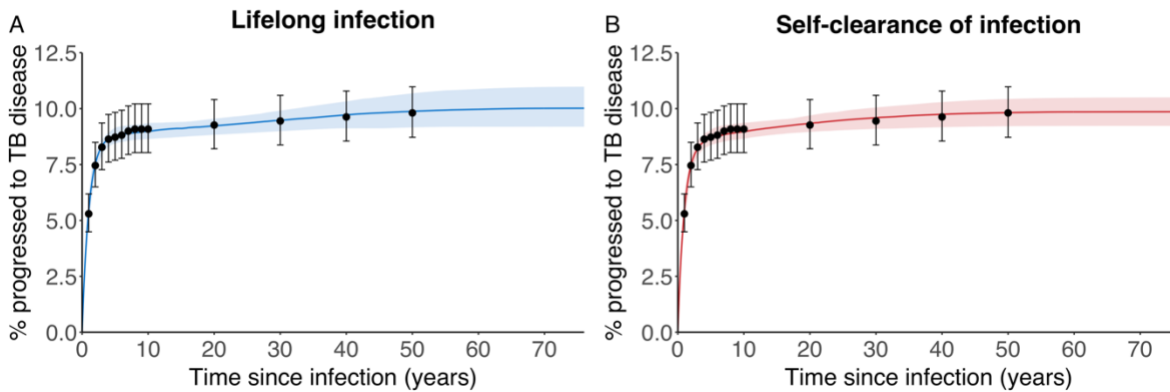
Supplementary figure 4. Cohort model results for progression to TB disease over time since infection for the lifelong infection (A) and self-clearance (B) scenarios. Uncertainty in the data are the equal-tailed 95% confidence intervals. Model fits are median model output and equal-tailed 95% uncertainty intervals. TB = Tuberculosis.

Prior versus posterior parameter values

Lifelong infection		Prior values			Posterior values		
Parameter	Prior distribution	50%	2.5%	97.5%	50%	2.5%	97.5%
f_d	Flat	1.50	0.075	2.93	0.0832	0.0676	0.116
f_s	Flat	1.50	0.075	2.93	0.870	0.674	1.26
s_d	Flat	1.50	0.075	2.93	0.000269	0.0000416	0.000525
Self-clearance of infection		Prior values			Posterior values		
Parameter	Prior distribution	50%	2.5%	97.5%	50%	2.5%	97.5%
f_d	Flat	1.50	0.0750	2.93	0.0859	0.0696	0.111
f_s	Flat	1.50	0.0750	2.93	0.919	0.717	1.25
s_d	Flat	1.50	0.0750	2.93	0.000622	0.000135	0.00113
s_c	Flat	1.50	0.0750	2.93	0.0330	0.0254	0.0467
f_c	Flat	1.50	0.0750	2.93	1.40E-6	2.90E-7	2.61E-6
$1/a_{AUT}$	Normal	1/10.0	1/20.0	1/6.67	1/14.1	1/18.3	1/8.97
a_{TST}	Flat	3.50	0.175	6.83	4.92	0.971	6.88

Supplementary table 4: Comparison of prior and posterior model parameters for the self-clearance and lifelong infection scenarios. Parameters are illustrated in the model diagram in Figure 1 of the main text. Note that the inverse of the age of first infection for the autopsy cohort is shown, since the prior distribution is on the contextual annual risk of infection from which the average age at first infection is derived. f_d = Rate of progression to TB disease from infected (fast progression) (*per year*); f_s = Rate of transition from infected (fast progression) to infected (slow progression) (*per year*); s_d = Rate of progression to TB disease from infected (slow progression) (*per year*); s_c = Rate of self-clearance from infected (slow progression) (*per year*); f_c = Rate of self-clearance from infected (fast progression) (*per year*); a_{AUT} = Age at which autopsy cohort is first infected (years); a_{TST} = Age at which TST-reversion cohort is first infected (years); μ = Mean background life expectancy (*years*); v = Standard deviation of background life expectancy (*years*).

Sensitivity analyses – Life expectancy

The primary analysis presented in the main paper used a mean background life expectancy of 60 years for fitting the cohort model. For sensitivity analysis we re-fitted the cohort model using values of 50 and 70 years.

Posterior model parameters

Lifelong infection									
	50 years			60 years			70 years		
Parameter	50%	2.5%	97.5%	50%	2.5%	97.5%	50%	2.5%	97.5%
f_d	0.0831	0.0670	0.109	0.0832	0.0676	0.116	0.0832	0.0664	0.108
f_s	0.867	0.681	1.17	0.870	0.674	1.26	0.865	0.669	1.16
s_d	0.000289	0.0000452	0.000589	0.000269	0.0000416	0.000525	0.000244	0.0000383	0.000502
Self-clearance of infection									
	50 years			60 years			70 years		
Parameter	50%	2.5%	97.5%	50%	2.5%	97.5%	50%	2.5%	97.5%
f_d	0.0826	0.0671	0.104	0.0859	0.0696	0.111	0.0847	0.0685	0.115
f_s	0.882	0.686	1.13	0.919	0.717	1.25	0.894	0.705	1.25
s_d	0.000598	0.0000930	0.00138	0.000622	0.000135	0.00113	0.000504	0.000101	0.000968
s_c	0.0362	0.0277	0.0518	0.0330	0.0254	0.0467	0.0311	0.0241	0.0409
f_c	3.71E-7	1.60E-8	1.09E-6	1.40E-6	2.90E-7	2.61E-6	5.18E-7	3.30E-8	1.27E-6
a_{AUT}	15.9	9.46	19.2	14.1	8.97	18.3	12.7	8.03	16.9
a_{TST}	4.42	0.578	6.73	4.92	0.971	6.88	4.18	0.77	6.67

Supplementary table 5: Estimated median and equal-tailed 95% uncertainty intervals for the posterior model parameters in the lifelong infection and self-clearance scenarios. Mean background life expectancies of 50, 60, and 70 years are assumed. Parameter symbols are described in Table 1 and illustrated in the model diagram in Figure 1 of the main text.

Self-cleared over time since infection (%)

| Time since infection	50 years			60 years			70 years		
	50%	2.5%	97.5%	50%	2.5%	97.5%	50%	2.5%	97.5%
10 years	25.5	19.1	35.4	24.4	17.8	32.6	21.8	17.0	28.3
Lifetime	69.2	58.4	79.7	73.1	64.6	81.7	75.6	67.1	82.1

Supplementary table 6: Estimated median and equal-tailed 95% uncertainty intervals for the percentage self-cleared over time since infection. Mean background life expectancies of 50, 60, and 70 years are assumed.

Lifetime risk of TB disease (%)

Scenario	50 years			60 years			70 years		
	50%	2.5%	97.5%	50%	2.5%	97.5%	50%	2.5%	97.5%
Lifelong infection	13.0	10.4	15.7	12.6	10.1	15.0	12.4	10.1	15.0
Self-clearance	17.1	11.6	28.0	17.0	10.9	22.5	15.1	10.9	21.3

Supplementary table 7: Estimated median and equal-tailed 95% uncertainty intervals for the lifetime risk of progressing to TB disease in the lifelong infection and self-clearance scenarios. Mean background life expectancies of 50, 60, and 70 years are assumed.

Cohort model fitting results
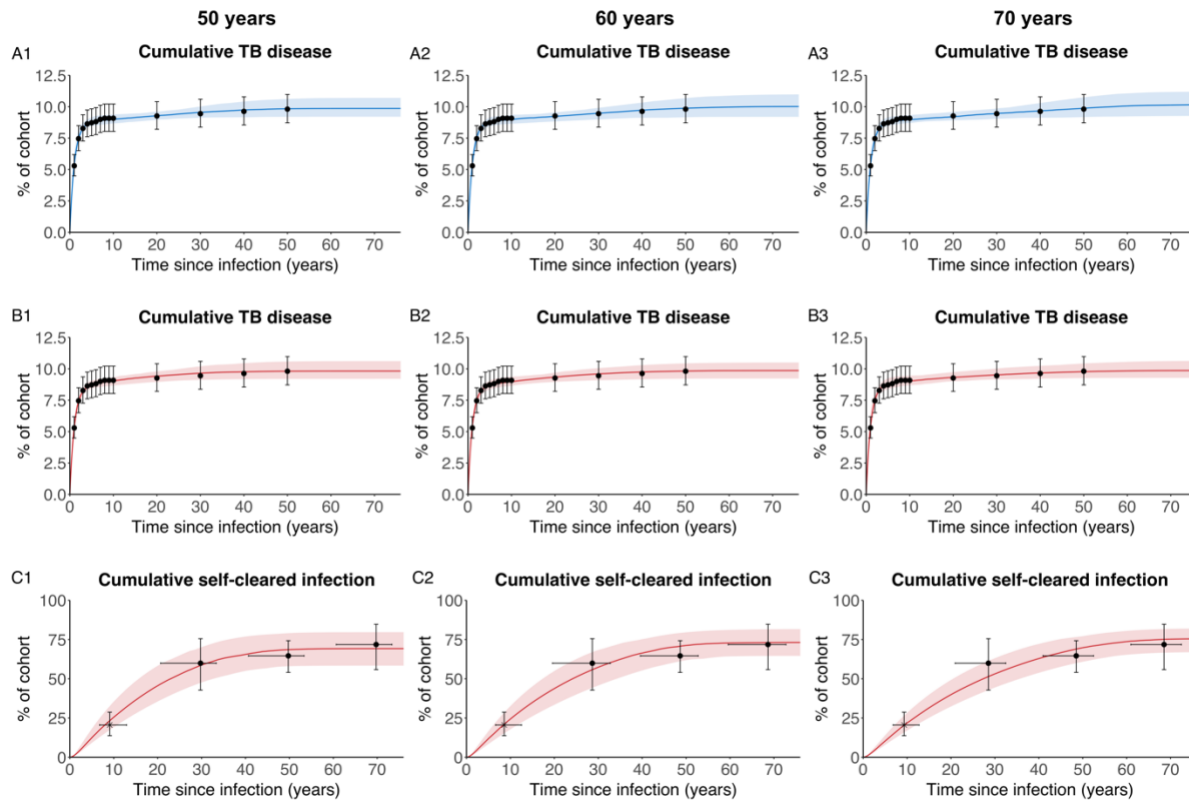
Supplementary figure 5: Cohort model assuming lifelong infection fitted to data on progression to TB disease (row A) and cohort model assuming self-clearance of *Mtb* infection fitted to data on progression to TB disease (row B) and self-clearance (row C), all over time since infection. Mean background life expectancies of 50, 60, and 70 years are assumed (columns 1, 2 and 3, respectively). Uncertainty in the data in the vertical axes are the equal-tailed 95% confidence intervals. Uncertainty in the horizontal axis in row C reflects the 95% uncertainty interval for the estimate for the age of infection. Model fits are median model output and equal-tailed 95% uncertainty intervals. *Mtb = Mycobacterium tuberculosis;* TB = Tuberculosis.

Sensitivity analyses – Distal progression rate

The primary analysis presented in the main paper assumed a distal progression to TB disease risk of 20/100,000 per year for fitting the cohort model. For sensitivity analysis we re-fitted the cohort model using values of 5/100,000 and 35/100,000 per year.

Posterior model parameters

Lifelong infection								
	5/100,000 per year			**20/100,000 per year**			**35/100,000 per year**		
Parameter	**50%**	**2.5%**	**97.5%**	**50%**	**2.5%**	**97.5%**	**50%**	**2.5%**	**97.5%**
f_d	0.0835	0.0671	0.111	0.0832	0.0676	0.116	0.0843	0.0680	0.109
f_s	0.865	0.676	1.18	0.870	0.674	1.26	0.890	0.688	1.18
s_d	0.000148	0.00000981	0.000403	0.000269	0.0000416	0.000525	0.000390	0.000149	0.000674
Self-clearance of infection									
	5/100,000 per year			**20/100,000 per year**			**35/100,000 per year**		
Parameter	**50%**	**2.5%**	**97.5%**	**50%**	**2.5%**	**97.5%**	**50%**	**2.5%**	**97.5%**
f_d	0.0867	0.0680	0.110	0.0859	0.0696	0.111	0.0866	0.0690	0.118
f_s	0.905	0.705	1.19	0.919	0.717	1.25	0.940	0.711	1.34
s_d	0.000328	0.0000348	0.000753	0.000622	0.000135	0.00113	0.000863	0.000291	0.00144
s_c	0.0325	0.0253	0.0437	0.0330	0.0254	0.0467	0.0335	0.0254	0.0431
f_c	2.83E-7	1.80E-8	7.81E-7	1.40E-6	2.90E-7	2.61E-6	4.91E-7	1.50E-8	1.30E-6
a_{AUT}	11.0	7.37	19.7	11.4	7.25	20.4	11.8	7.54	19.2
a_{TST}	4.66	0.658	6.86	4.92	0.971	6.88	4.38	0.645	6.80

Supplementary table 8: Estimated median and equal-tailed 95% uncertainty intervals for the posterior model parameters in the lifelong infection and self-clearance scenarios. A distal progression to TB disease risk of 5, 20 and 35/100,000 per year are assumed for model fitting. Parameter symbols are described in Table 1 and illustrated in the model diagram in Figure 1 of the main text. TB = Tuberculosis.

Self-cleared over time since infection (%)

|	**5/100,000 per year**			**20/100,000 per year**			**35/100,000 per year**		
Time since infection	**50%**	**2.5%**	**97.5%**	**50%**	**2.5%**	**97.5%**	**50%**	**2.5%**	**97.5%**
10 years	23.2	19.0	30.3	24.4	17.8	32.6	24.4	19.4	29.9
Lifetime	72.2	64.7	80.2	73.1	64.6	81.7	72.6	64.8	80.2

Supplementary table 9: Estimated median and equal-tailed 95% uncertainty intervals for the percentage self-cleared over time since infection. A distal progression to TB disease risk of 5, 20 and 35/100,000 per year are assumed for model fitting. TB = Tuberculosis

Lifetime risk of TB disease (%)

	5/100,000 per year			20/100,000 per year			35/100,000 per year		
Scenario	50%	2.5%	97.5%	50%	2.5%	97.5%	50%	2.5%	97.5%
Lifelong infection	10.8	9.50	14.0	12.6	10.1	15.0	14.2	11.5	17.1
Self-clearance	13.8	9.90	17.6	17.0	10.9	22.5	19.3	12.9	26.5

Supplementary table 10: Estimated median and equal-tailed 95% uncertainty intervals for the lifetime risk of progressing to TB disease in the lifelong infection and self-clearance scenarios. A distal progression to TB disease risk of 5, 20 and 35/100,000 per year are assumed for model fitting. TB = Tuberculosis

Cohort model fitting results
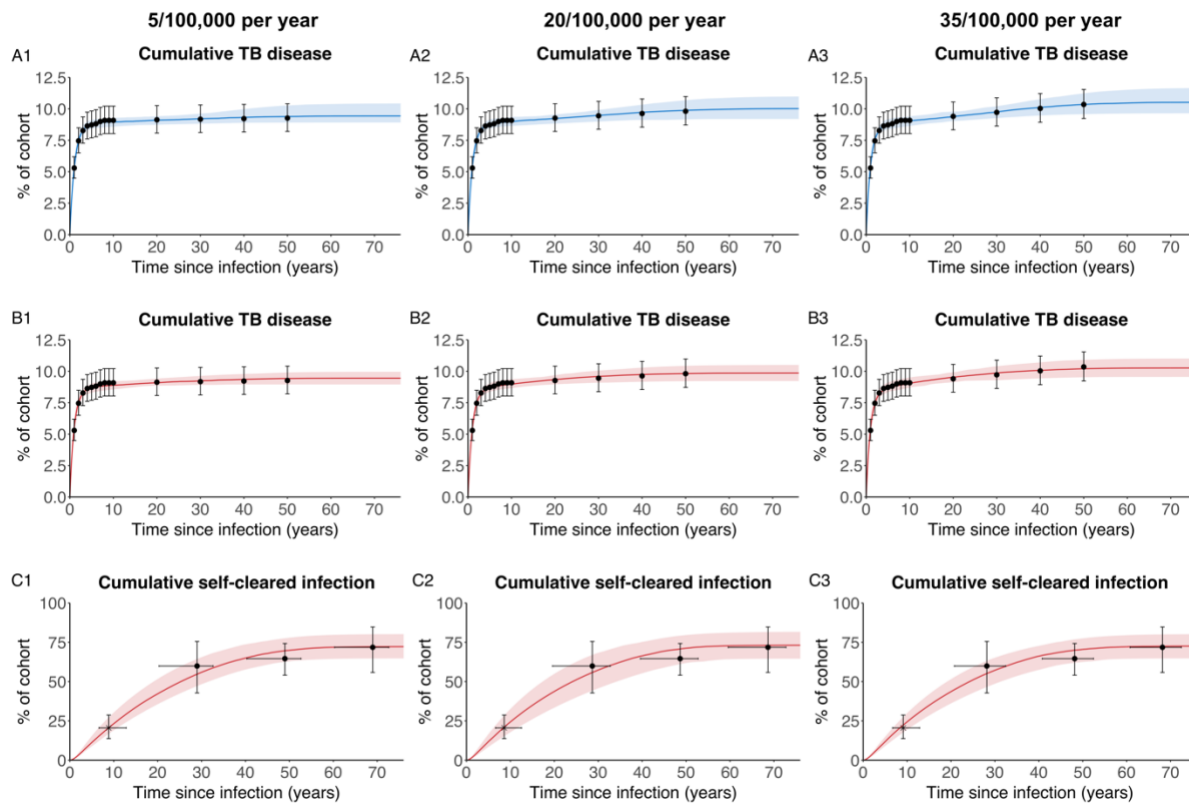
Supplementary figure 6: Cohort model assuming lifelong infection fitted to data on progression to TB disease (row A) and cohort model assuming self-clearance of *Mtb* infection fitted to data on progression to TB disease (row B) and self-clearance (row C), all over time since infection. A distal progression to TB disease risk of 5, 20 and 35/100,000 per year are assumed for model fitting (columns 1, 2 and 3, respectively). Uncertainty in the data in the vertical axes are the equal-tailed 95% confidence intervals. Uncertainty in the horizontal axis in row C reflects the 95% uncertainty interval for the estimate for the age of infection. Model fits are median model

output and equal-tailed 95% uncertainty intervals. *Mtb = Mycobacterium tuberculosis*; TB = Tuberculosis.

County-level model and population at risk of tuberculosis disease

Sensitivity analyses – Life expectancy

The primary analysis presented in the main paper used a mean background life expectancy of 60 years in the cohort model to infer the TB natural history parameters used in the country-level model. For sensitivity analysis we used values of 50 and 70 years in the cohort model before re-running the country-level models with the updated TB natural history parameters.

Estimated population with a viable Mtb infection in 2019
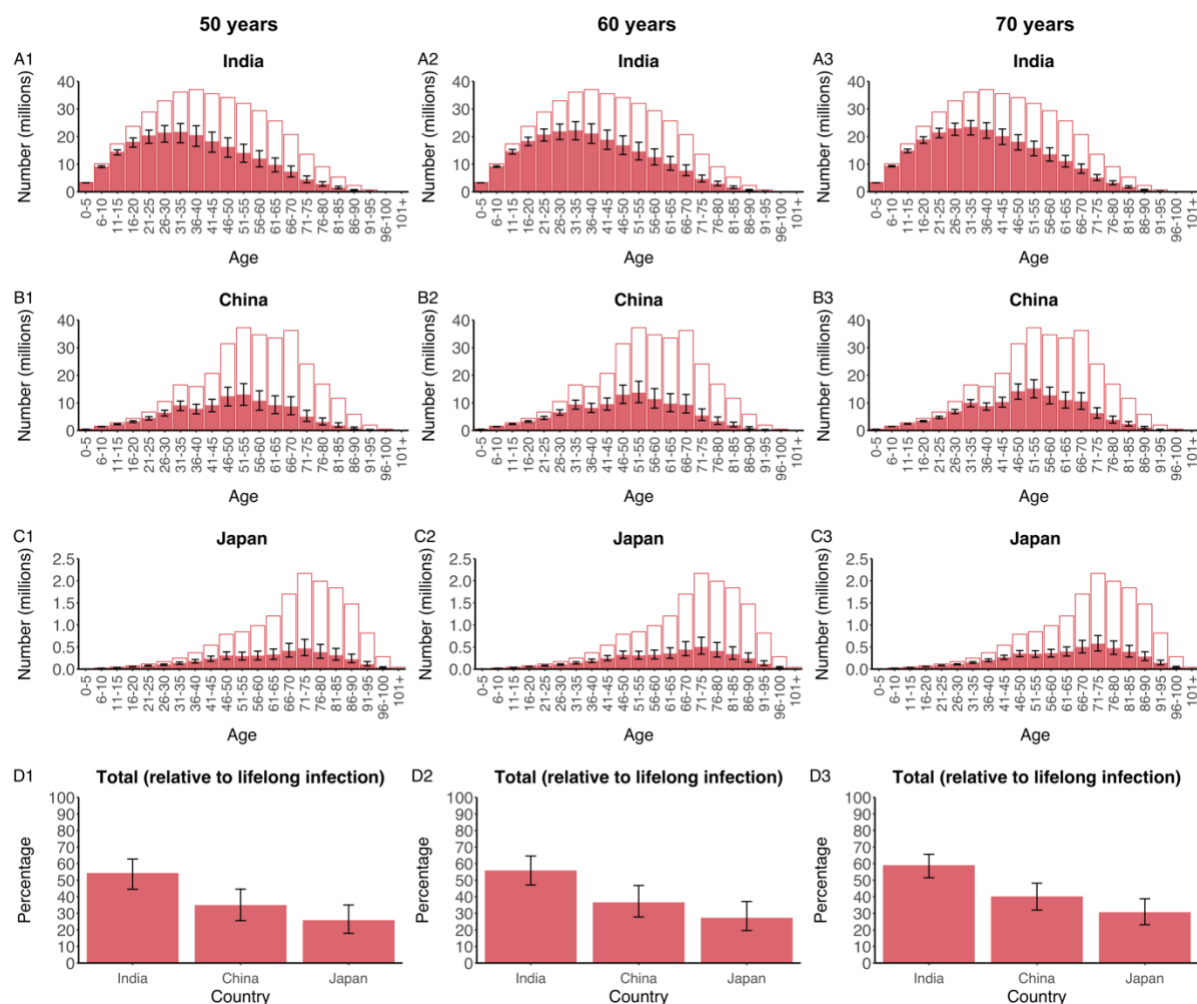
Supplementary figure 7: Rows A-C show the estimated population with a viable *Mtb* infection in 2019 disaggregated by age in three epidemiologically distinct settings for the cases of lifelong infection (red outlined) and self-clearance (red filled). Row D shows the total population with a

viable infection in each setting, assuming self-clearance, expressed as a percentage of the total population with a viable infection assuming lifelong infection. Mean background life expectancies of 50, 60, and 70 years (columns 1, 2 and 3, respectively) are assumed in the cohort model from which the country-level model parameters are estimated. Point values and error bars show median and equal-tailed 95% uncertainty intervals from the model outputs. Note that median estimates for the historical ARI have been used, such that uncertainty presented in the results is associated with uncertainty in the natural history model parameters only. For clarity, uncertainty in the lifelong infection results are not shown since they are much smaller than in the self-clearance scenario. *Mtb = Mycobacterium tuberculosis*

*Proportion that retain a viable infection compared to the lifelong infection scenario (%)*

	50 years			60 years			70 years		
Country	50%	2.5%	97.5%	50%	2.5%	97.5%	50%	2.5%	97.5%
India	54.4	44.6	62.8	55.9	47.1	64.6	59.1	51.5	65.7
China	35.0	25.6	44.6	36.6	27.8	46.8	40.2	32.0	48.2
Japan	25.8	17.9	35.0	27.3	19.6	37.2	30.7	23.1	38.8

Supplementary table 11: Estimated median and equal-tailed 95% uncertainty intervals for the proportion of people in each country that retain a viable infection in the self-clearance scenario compared to the lifelong infection scenario. Mean background life expectancies of 50, 60, and 70 years are assumed in the cohort model from which the country-level model parameters are estimated.

Sensitivity analyses – Distal progression rate

The primary analysis presented in the main paper used a distal progression to TB disease risk of 20/100,000 per year in the cohort model to infer the TB natural history parameters used in the country-level model. For sensitivity analysis we used values of 5/100,000 and 35/100,000 per year in the cohort model before re-running the country-level models with the updated TB natural history parameters.

Estimated population with a viable Mtb infection in 2019
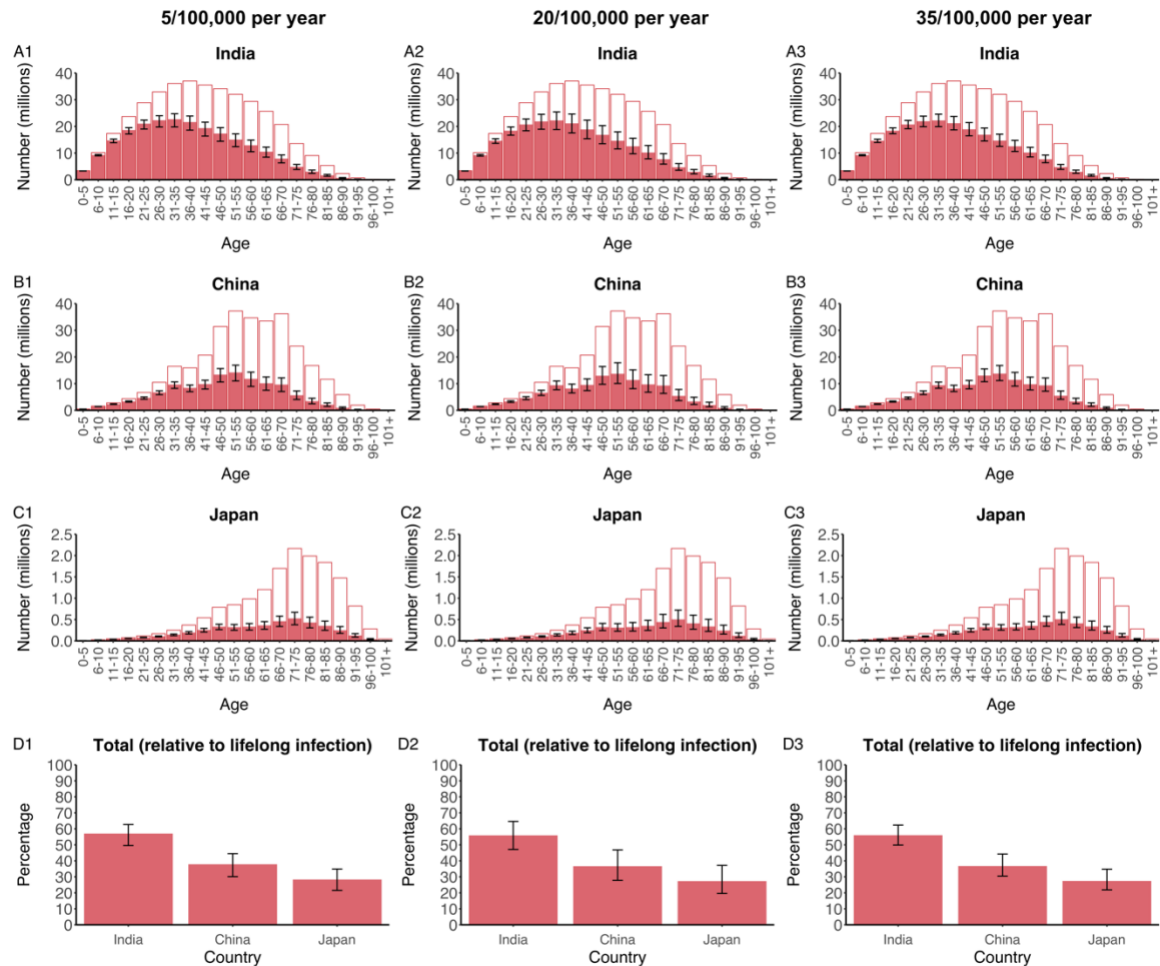
Supplementary figure 8: Rows A-C show the estimated population with a viable *Mtb* infection in 2019 disaggregated by age in three epidemiologically distinct settings for the cases of lifelong infection (red outlined) and self-clearance (red filled). Row D shows the total population with a viable infection in each setting, assuming self-clearance, expressed as a percentage of the total population with a viable infection assuming lifelong infection. A distal progression to TB disease risk of 5, 20 and 35/100,000 per year (columns 1, 2 and 3, respectively) are assumed for fitting the cohort model from which the country-level model parameters are estimated. Point values and error bars show median and equal-tailed 95% uncertainty intervals from the model outputs. Note that median estimates for the historical ARI have been used, such that uncertainty presented in the results is associated with uncertainty in the natural history model parameters only. For clarity, uncertainty in the lifelong infection results are not shown since they are much smaller than in the self-clearance scenario. TB = Tuberculosis; *Mtb* = *Mycobacterium tuberculosis*.

Proportion that retain a viable infection compared to the lifelong infection scenario (%)

Country	5/100,000 per year			20/100,000 per year			35/100,000 per year		
	50%	2.5%	97.5%	50%	2.5%	97.5%	50%	2.5%	97.5%
India	57.1	49.6	62.7	55.9	47.1	64.6	56.0	49.8	62.3
China	37.8	30.1	44.5	36.6	27.8	46.8	36.7	30.4	44.2
Japan	28.3	21.5	34.8	27.3	19.6	37.2	27.4	21.8	34.8

Supplementary table 12: Estimated median and equal-tailed 95% uncertainty intervals for the proportion of people in each country that retain a viable infection in the self-clearance scenario compared to the lifelong infection scenario. A distal progression to TB disease risk of 5, 20 and 35/100,000 per year are assumed for fitting the cohort model from which the country-level model parameters are estimated.

REFERENCES

1. C. E. Barry, *et al.*, The spectrum of latent tuberculosis: rethinking the biology and intervention strategies. *Nat. Rev. Microbiol.* **7**, 845–855 (2009).

2. D. B. Young, H. P. Gideon, R. J. Wilkinson, Eliminating latent tuberculosis. *Trends Microbiol.* **17**, 183–188 (2009).

3. P. L. Lin, J. L. Flynn, Understanding latent tuberculosis: a moving target. *J. Immunol.* **185**, 15–22 (2010).

4. H. Esmail, C. E. Barry, R. J. Wilkinson, Understanding latent tuberculosis: the key to improved diagnostic and novel treatment strategies. *Drug Discov. Today* **17**, 514–521 (2012).

5. H. Esmail, C. E. Barry, D. B. Young, R. J. Wilkinson, The ongoing challenge of latent tuberculosis. *Philos. Trans. R. Soc. Lond., B, Biol. Sci.* **369**, 20130437 (2014).

6. K. Dheda, C. E. Barry, G. Maartens, Tuberculosis. *Lancet* **387**, 1211–1226 (2016).

7. M. Pai, *et al.*, Tuberculosis. *Nat Rev Dis Primers* **2**, 16076 (2016).

8. M. A. Behr, P. H. Edelstein, L. Ramakrishnan, Revisiting the timetable of tuberculosis. *BMJ* **362**, k2738 (2018).

9. P. K. Drain, *et al.*, Incipient and Subclinical Tuberculosis: a Clinical Review of Early Stages and Progression of Infection. *Clin. Microbiol. Rev.* **31** (2018).

10. P. L. Lin, J. L. Flynn, The End of the Binary Era: Revisiting the Spectrum of Tuberculosis. *The Journal of Immunology* **201**, 2541–2548 (2018).

11. A. E. Keller, R. H. Kampmeier, Tuberculin Survey. Observations on Medical Students and Undergraduate Nurses at Vanderbilt University. *am rev tuberc* **39**, 657–665 (1939).

12. A. Maclouf, The Yearly Persistence of Cutaneous Sensibility to Tuberculin. *Rev. Path. Comparee et Hyg. Gen.* **46**, 241–57 (1946).

13. R. N. Johnston, R. T. Ritchie, I. H. F. Murray, Declining Tuberculin Sensitivity with Advancing Age. *Br Med J* **2**, 720–724 (1963).

14. H. Mysakowska, *et al.*, An appraisal of the chemoprophylaxis of tuberculosis in the student population of Lublin, Poland, following thirty months' observation. *Am. Rev. Respir. Dis.* **93**, 628–629 (1966).

15. K. A. Sepkowitz, Tuberculin skin testing and the health care worker: lessons of the Prophit Survey. *Tuber. Lung Dis.* **77**, 81–85 (1996).

16. T. R. Hawn, *et al.*, Tuberculosis vaccines and prevention of infection. *Microbiol. Mol. Biol. Rev.* **78**, 650–671 (2014).

17. V. Nduba, A. H. Van't Hoog, E. M. H. Mitchell, M. Borgdorff, K. F. Laserson, Incidence of Active Tuberculosis and Cohort Retention Among Adolescents in Western Kenya. *Pediatr. Infect. Dis. J.* **37**, 10–15 (2018).

18. M. A. Behr, P. H. Edelstein, L. Ramakrishnan, Is Mycobacterium tuberculosis infection life long? *BMJ* **367** (2019).

19. M. Pai, *et al.*, Gamma interferon release assays for detection of Mycobacterium tuberculosis infection. *Clinical microbiology reviews* **27**, 3–20 (2014).

20. A. S. Griffith, Types of tubercle bacilli in human tuberculosis. *British Journal of Tuberculosis* **11**, 160–165 (1917).

21. E. L. Opie, J. D. Aronson, Tubercule bacilli in latent tuberculosis lesions and in lung tissue without tuberculosis lesions. *Arch. Pathol.* **4**, 1–21 (1927).

22. W. H. Feldman, A. H. Baggenstoss, The residual infectivity of the primary complex of tuberculosis. *Am. J. Pathol.* **14**, 473-490.3 (1938).

23. G. Canetti, I. Sutherland, E. Svandova, Endogenous reactivation and exogenous reinfection: their relative importance with regard to the development of non-primary tuberculosis. *Bull Int Union Tuberc* **47**, 116–34 (1972).

24. J. Keane, *et al.*, Tuberculosis associated with infliximab, a tumor necrosis factor alpha-neutralizing agent. *N. Engl. J. Med.* **345**, 1098–1104 (2001).

25. J. J. Gómez-Reino, *et al.*, Treatment of rheumatoid arthritis with tumor necrosis factor inhibitors may predispose to significant increase in tuberculosis risk: a multicenter active-surveillance report. *Arthritis Rheum.* **48**, 2122–2127 (2003).

26. F. Wolfe, K. Michaud, J. Anderson, K. Urbansky, Tuberculosis infection in patients with rheumatoid arthritis and the effect of infliximab therapy. *Arthritis Rheum.* **50**, 372–379 (2004).

27. M. M. Klote, L. Y. Agodoa, K. Abbott, Mycobacterium tuberculosis infection incidence in hospitalized renal transplant patients in the United States, 1998-2000. *Am. J. Transplant.* **4**, 1523–1528 (2004).

28. B. Baldin, A. Dozol, A. Spreux, R.-M. Chichmanian, Tuberculoses lors de traitements par l'infliximab. *La Presse Médicale* **34**, 353–357 (2005).

29. A. Atasever, *et al.*, Tuberculosis in renal transplant recipients on various immunosuppressive regimens. *Nephrol. Dial. Transplant.* **20**, 797–802 (2005).

30. J. J. Gómez-Reino, L. Carmona, M. Á. Descalzo, Risk of tuberculosis in patients treated with tumor necrosis factor antagonists due to incomplete prevention of reactivation of latent infection. *Arthritis Care & Research* **57**, 756–761 (2007).

31. J. Torre-Cisneros, *et al.*, Tuberculosis after solid-organ transplant: incidence, risk factors, and clinical characteristics in the RESITRA (Spanish Network of Infection in Transplantation) cohort. *Clin. Infect. Dis.* **48**, 1657–1665 (2009).

32. C. R. Horsburgh, *et al.*, Revisiting rates of reactivation tuberculosis: a population-based approach. *Am. J. Respir. Crit. Care Med.* **182**, 420–425 (2010).

33. J.-W. Ai, *et al.*, The Risk of Tuberculosis in Patients with Rheumatoid Arthritis Treated with Tumor Necrosis Factor-α Antagonist: A Metaanalysis of Both Randomized Controlled Trials and Registry/Cohort Studies. *J. Rheumatol.* **42**, 2229–2237 (2015).

34. W.-C. Fan, *et al.*, Long-term risk of tuberculosis in haematopoietic stem cell transplant recipients: a 10-year nationwide study. *Int. J. Tuberc. Lung Dis.* **19**, 58–64 (2015).

35. H.-J. Lee, *et al.*, The demanding attention of tuberculosis in allogeneic hematopoietic stem cell transplantation recipients: High incidence compared with general population. *PLoS ONE* **12**, e0173250 (2017).

36. N. Agrawal, *et al.*, Incidence and clinical profile of tuberculosis after allogeneic stem cell transplantation. *Transpl Infect Dis* **20** (2018).

37. M. P. Cheng, *et al.*, Risk of Latent Tuberculosis Reactivation After Hematopoietic cell Transplantation. *Clin. Infect. Dis.* **69**, 869–872 (2019).

38. G. J. Hamman L., The subsequent history of 1000 patients who recieved tuberculin tests. *Bull. Johns Hopkins Hosp.* **24**, 180–186 (1913).

39. M. Paretzky, The disappearance of specific skin hypersensitiveness in tuberculosis, a report based on 80 cases. *Am. Rev. Tuberc.* **33**, 370 (1936).

40. A. W. Dahlstrom, The Instability of the Tuberculin Reaction. Observations on Dispensary Patients with Special Reference to the Existence of Demonstrable Tuberculous Lesions and the Degree of Exposure to Tubercle Bacilli. *American Review of Tuberculosis and Pulmonary Diseases* **42**, 471–87 (1940).

41. D. Zacks, P. E. Sartwell, Development of Tuberculosis and Changes in Sensitivity to Tuberculin in an Institution for the Feebleminded—A Ten Years' Study. *Am J Public Health Nations Health* **32**, 732–738 (1942).

42. R. R. Puffer, H. C. Stewart, R. S. Gass, E. F. Harrison, Serial Tuberculin Tests and Stability of the Tuberculin Reaction,,. *am rev tuberc* **54**, 541–552 (1946).

43. J. D. Aronson, The Fluctuation of the Tuberculin Reaction in Different Geographic Areas and its Relationship to Resistance. *American Review of Tuberculosis and Pulmonary Diseases* **63**, 121–39 (1951).

44. R. H. Marks, G. Tokuyama, A. Peterson, A Five-Year Pilot Study of Tuberculin Testing in Honolulu Schools. Progress Report of the First Two Years of Observation. *American Review of Tuberculosis and Pulmonary Diseases* **78**, 871–83 (1958).

45. K. J. Kunofsky S, The Stability of the Tuberculin Reaction. *American Review of Tuberculosis and Pulmonary Diseases* **78**, 862–70 (1958).

46. J. M. Adams, *et al.*, Reversal of tuberculin reaction in early tuberculosis. *Dis Chest* **35**, 348–356 (1959).

47. S. Grzybowski, E. A. Allen, THE CHALLENGE OF TUBERCULOSIS IN DECLINE. A STUDY BASED ON THE EPIDEMIOLOGY OF TUBERCULOSIS IN ONTARIO, CANADA. *Am. Rev. Respir. Dis.* **90**, 707–720 (1964).

48. R. C. Dickson, The Tuberculin Test. *Canadian Medical Association Journal* **92**, 25–9 (1965).

49. S. Grzybowski, Ontario Studies on Tuberculin Sensitivity. Section III. Instability of Tuberculin Reactions. *Canadian Journal of Public Health* **56**, 190–92 (1965).

50. V. N. Houk, D. C. Kent, K. Sorensen, J. H. Baker, The eradication of tuberculosis infection by isoniazid chemoprophylaxis. *Arch. Environ. Health* **16**, 46–50 (1968).

51. I. Sutherland, "The ten-year incidence of clinical tuberculosis following 'conversion' in 2,550 individuals aged 14 to 19 at time of conversion" (Tuberculosis Surveillance Research Unit, 1968).

52. S. H. Ferebee, Controlled chemoprophylaxis trials in tuberculosis. A general review. *Bibliotheca tuberculosea* **26**, 28–106 (1970).

53. N. O. Atuk, Serial Tuberculin Testing and Isoniazid Therapy in General Hospital Employees. *JAMA: The Journal of the American Medical Association* **218**, 1795 (1971).

54. E. J. Perez-Stable, D. Flaherty, G. Schecter, G. Slutkin, P. C. Hopewell, Conversion and reversion of tuberculin reactions in nursing home residents. *Am. Rev. Respir. Dis.* **137**, 801–804 (1988).

55. F. M. Gordin, *et al.*, Stability of positive tuberculin tests: are boosted reactions valid? *Am. Rev. Respir. Dis.* **144**, 560–563 (1991).

56. R. L. Sepulveda, D. Araya, X. Ferrer, R. U. Sorensen, Repeated Tuberculin Testing in Patients With Active Pulmonary Tuberculosis. *Chest* **103**, 359–363 (1993).

57. P. E. Fine, *et al.*, Tuberculin sensitivity: conversions and reversions in a rural African population. *Int. J. Tuberc. Lung Dis.* **3**, 962–975 (1999).

58. D. Menzies, Interpretation of repeated tuberculin tests. Boosting, conversion, and reversion. *Am. J. Respir. Crit. Care Med.* **159**, 15–21 (1999).

59. K. Bjartveit, Olaf Scheel and Johannes Heimbeck: their contribution to understanding the pathogenesis and prevention of tuberculosis. *Int. J. Tuberc. Lung Dis.* **7**, 306–311 (2003).

60. P. C. Hill, *et al.*, Longitudinal assessment of an ELISPOT test for Mycobacterium tuberculosis infection. *PLoS Med.* **4**, e192 (2007).

61. H. G. Wiker, T. Mustafa, G. A. Bjune, M. Harboe, Evidence for waning of latency in a cohort study of tuberculosis. *BMC Infectious Diseases* **10** (2010).

62. C. M. Stein, *et al.*, Long-term stability of resistance to latent M. tuberculosis infection in highly exposed TB household contacts in Kampala, Uganda. *Clin. Infect. Dis.* (2018) https:/doi.org/10.1093/cid/ciy751.

63. E. L. Opie, J. D. Aronson, Tubercule bacilli in latent tuberculosis lesions and in lung tissue without tuberculosis lesions. *Arch. Pathol.* **4**, 1–21 (1927).

64. S. Ferebee, F. W. Mount, A. Anastasiades, Prophylactic effects of isoniazid on primary tuberculosis in children; a preliminary report. *Am Rev Tuberc* **76**, 942–963 (1957).

65. F. W. Mount, S. H. Ferebee, Preventive effects of isoniazid in the treatment of primary tuberculosis in children. *N. Engl. J. Med.* **265**, 713–721 (1961).

66. N. A. Menzies, *et al.*, Progression from latent infection to active disease in dynamic tuberculosis transmission models: a systematic review of the validity of modelling assumptions. *Lancet Infect Dis* (2018) https:/doi.org/10.1016/S1473-3099(18)30134-8.

67. I. Sutherland, "The ten-year incidence of clinical tuberculosis following 'conversion' in 2,550 individuals aged 14 to 19 at time of conversion" (Tuberculosis Surveillance Research Unit, 1968).

68. I. Sutherland, Recent studies in the epidemiology of tuberculosis, based on the risk of being infected with tubercle bacilli. *Adv Tuberc Res* **19**, 1–63 (1976).

69. I. Sutherland, E. Svandova, S. Radhakrishna, The development of clinical tuberculosis following infection with tubercle bacilli. *Tubercle* **63**, 255–68 (1982).

70. K. Styblo, Styblo 1991 - Selected papers - Chapter 7. Development of bacillary pulmonary tuberculosis capable of transmitting tubercle bacilli following infection with M.tuberculosis (the disease ratio).pdf. *KNCV Tuberculosis Foundation - Selected papers* (1991).

71. K. D. Dale, J. M. Trauer, P. J. Dodd, R. M. G. J. Houben, J. T. Denholm, Estimating long-term tuberculosis reactivation rates in Australian migrants. *Clin. Infect. Dis.* (2019) https:/doi.org/10.1093/cid/ciz569.

72. L. A. Ronald, *et al.*, Demographic predictors of active tuberculosis in people migrating to British Columbia, Canada: a retrospective cohort study. *Canadian Medical Association Journal* **190**, E209–E216 (2018).

73. K. Styblo, Selected papers (Vol 24). *KNCV Tuberculosis Foundation - Selected papers* (1991).

74. G. M. Cauthen, A. Pio, H. G. ten Dam, Annual risk of tuberculous infection. 1988. *Bull. World Health Organ.* **80**, 503–511; discussion 501-502 (2002).

75. F. Hamburger, R. Monti, Allgemeine Pathologie und Diagnostik der Kindertuberkulose. *Franz Deuticke*, (pp. 56-58) (1910).

76. K. Stýblo, J. Meijer, I. Sutherland, [The transmission of tubercle bacilli: its trend in a human population]. *Bull. World Health Organ.* **41**, 137–178 (1969).

77. M. R. Heynsius van den Berg, Leerboek der tuberculosebestrijding. *The Hague, KNCV*, (p. 149) (1962).

78. S. Grzybowski, K. Styblo, E. Dorken, Tuberculosis in Eskimos. *Tubercle* **57**, S1-58 (1976).

79. R. M. G. J. Houben, P. J. Dodd, The Global Burden of Latent Tuberculosis Infection: A Re-estimation Using Mathematical Modelling. *PLoS Med* **13**, e1002152 (2016).

80. United Nations, Department of Economic and Social Affairs, Population Division (2019)., World Population Prospects 2019, Online Edition. (December 18, 2019).

81. WHO, Global Tuberculosis Report 2016. Geneva: World Health Organization; 2016 (2016).