

1 **Survival of people with untreated tuberculosis: effects of time, geography and setting**

2

3 Carly A Rodriguez^{1*}, Sarah V Leavitt^{2*}, Tara C. Bouton^{3,4}, C Robert Horsburgh^{1,2}, Pia
4 Abel zur Wiesch^{5,6}, Brooke Nichols⁷, Helen E Jenkins^{2**}, Laura F White^{2**}

5 ¹Department of Epidemiology, Boston University School of Public Health, Boston, MA,
6 USA

7 ²Department of Biostatistics, Boston University School of Public Health, Boston, MA,
8 USA

9 ³Section of Infectious Diseases, Boston Medical Center, Boston, MA, USA

10 ⁴Boston University School of Medicine, Boston, MA, USA

11 ⁵Division for Infection Control and Environmental Health, Norwegian Institute of Public
12 Health, Norway

13 ⁶Center of Infectious Disease Dynamics, Pennsylvania State University, USA

14 ⁷Department of Global Health, Boston University School of Public Health, Boston, MA
15 USA

16 *Joint first author (equal contribution)

17 **Joint senior author (equal contribution)

18

19 **ABSTRACT**

20 **Background** An estimated 40% of people who developed tuberculosis in 2021 were not
21 diagnosed or treated. Pre-chemotherapy era data are a rich resource on survival for
22 people with untreated TB. We aimed to identify heterogeneities in these data to inform
23 more precise use of them.

24 **Methods** We extracted survival data from pre-chemotherapy era papers reporting TB
25 specific mortality and/or natural recovery data. We used Bayesian parametric survival
26 analysis to model the survival distribution, stratifying by geography (North America
27 versus Europe), time (pre-1930 versus post-1930), and setting (sanitoria versus non-
28 sanitoria).

29 **Results** We found 12 studies with TB-specific mortality data. Ten-year survival was
30 69% in North America (95 CI: 54%-81%) and 36% in Europe (95% CI: 10%-71%). Only
31 38% (95% CI: 18%-63%) of non-sanitorium individuals survived to 10 years compared
32 to 69% (95% CI: 41%-87%) of sanitoria/hospitalized patients. There were no significant
33 differences between people diagnosed pre-1930 and post-1930 (five-year survival pre-
34 1930: 65%; 95% CI: 44%-88% versus post-1930: 72%; 95% CI: 41%-94%).

35 **Conclusions** Mortality and natural recovery risks vary substantially by location and
36 setting. These heterogeneities need to be considered when using pre-chemotherapy
37 data to make inferences about expected survival of people with undiagnosed TB.

38

39

40 INTRODUCTION

41 Tuberculosis (TB) is a major infectious cause of global mortality, with case detection (an
42 estimated 40% were undiagnosed in 2021) and failure to cure those treated being major
43 drivers of mortality and transmission ¹. Clinical outcomes of individuals with untreated
44 TB are important from a public health perspective and in modelling studies.

45 Several studies have mined the rich pre-chemotherapy era TB literature to understand
46 outcomes for people with untreated TB²⁻⁴. As we continue to use these important pre-
47 chemotherapy era data, it is important to remember that these studies are not
48 homogeneous. They span decades during which tumultuous events occurred (e.g.
49 World Wars, the Great Depression), sanatoria were for treatment to varying degrees,
50 studies were carried out in differing geographic regions (North America and Europe),
51 and the quality and availability of medical care increased over time. TB outcomes
52 among untreated individuals could vary significantly by these factors.

53 Here, we use studies from the pre-chemotherapy era that recorded mortality and natural
54 recovery outcomes over time. We aim to estimate TB mortality rates during this era
55 overall and by geographic and temporal context hypothesizing that individuals from
56 countries with more resources during times of great stability would have better
57 outcomes than individuals in studies from more turbulent times. We also aim to explore
58 outcome differences between sanatoria patients versus those not in sanatoria and also
59 hypothesize that individuals in sanatoria had less severe disease initially than those who
60 were not.

61

62 **METHODS**

63 *Search strategy*

64 Since the development of anti-TB chemotherapy in the mid-twentieth century, studies
65 reporting outcomes of untreated populations are unethical and pre-chemotherapy era
66 studies are often not indexed in databases (e.g. PubMed, EMBASE). Therefore, we
67 reviewed our personal libraries, which include publications of pre-chemotherapy era
68 cohorts from a previous study of untreated children with TB ⁴, and publications used by ²
69 in a similar review to this current study (Table S1). We additionally reviewed the
70 reference lists of these publications for further sources, and reference lists of those
71 papers, where possible, to capture additional data on adults to supplement the original
72 review results.

73

74 *Review of studies*

75 We included cohort studies comprised of TB patients who did not receive anti-TB
76 chemotherapy, and that reported the proportion of the cohort within specified time
77 interval(s) with the following four outcomes: died from TB, died from a non-TB cause,
78 naturally recovered, and lost to follow-up (Figure S1). Outcome definitions would
79 depend on specific publications and are included in the results section and Supplement
80 (Table S2). “Natural recovery” in this literature was not clinically well-defined and our
81 definition would need to be driven by publications that we identify. However, natural
82 recovery would likely include patients who no longer had detectable *Mycobacterium*
83 *tuberculosis* (*Mtb*) in sputum smears, were able to work, or had symptom resolution ⁵.

84

85 We excluded publications that were not about TB, did not contain data on natural
86 recovery or TB-specific mortality, were restricted to patients with specific forms of TB
87 (e.g. TB meningitis) or narrow age groups (e.g. children <2 years old), case reports,
88 autopsy studies, or were primarily patients receiving anti-TB drugs. We excluded
89 publications presenting population-level outcome rates, or that did not specify follow-up
90 time or loss to follow-up. Additionally, we excluded reviews or editorials but we did
91 search their reference lists for additional studies of interest.

92

93 Two reviewers (two of TCB, HEJ, CAR, LFW) independently read each publication to
94 determine inclusion eligibility. Non-English publications were translated by authors
95 fluent in the relevant language (PAzW, BN) and independently reviewed by HEJ and
96 LFW, along with the relevant translator, to determine eligibility. Discrepancies were
97 resolved by group discussion.

98

99 *Data extraction*

100 Two reviewers (two of TCB, HEJ, CAR, LFW, PAzW, BN) independently extracted data
101 from each publication using a standardized electronic form. We extracted: study type
102 (e.g. population-based cohort, sanatorium-based cohort), cohort characteristics
103 (geographic location, time period, forms of TB disease, age range), determination of
104 death, outcome definitions, treatment (e.g. sanatorium, surgery, anti-TB drugs), and the
105 start of follow-up time (i.e. notification/diagnosis, sanatorium entry or exit). For
106 consistency among sanatorium studies, if the follow-up start time was the sanatorium
107 exit and the average sanatorium stay was reported, the follow-up times were shifted

108 ahead by this average length of stay so that the follow-up start was approximately entry
109 into the sanatorium. If the average length of stay was not reported, then the follow-up
110 times were shifted by 165 days, the average length of sanatoria stay in the United
111 States between 1934-1938 ⁶. We extracted data into life tables and stratified by disease
112 severity when possible.

113

114 Publications often reported on multiple cohorts and occasionally, the same study
115 reported data over multiple publications. Therefore, we enumerated findings using the
116 following terminology: “publications” are individual published reports in the literature;
117 “studies” are a population of patients recruited, diagnosed, and treated in a similar
118 fashion during the same time period, which may be reported over multiple publications;
119 and “cohorts” are a patient population that may have been recruited, diagnosed, and/or
120 treated differently over varying time periods but were in the same publication.

121

122 Discrepancies between reviewers were resolved by group discussion and reconciled
123 extractions were entered into REDCap electronic data capture tools hosted at Boston
124 University Medical Campus, Boston, MA, USA. Ethics approval was not required
125 because all data were from published reports.

126

127 *Mortality meta-analysis*

128 We transformed the extracted life-table structured data and created individual-level data
129 with the interval from entry/diagnosis to death or time of censoring for each individual in
130 each cohort. We used parametric survival analysis with interval censoring to estimate

131 the survival times for TB-specific mortality. TB deaths were considered events and
132 deaths from other causes were considered censored observations.

133
134 We used a Bayesian framework to estimate the survival distribution which included a
135 frailty term for each study (see Supplement). We analyzed all of the studies together in
136 one model and then we compared mortality outcomes by time (pre-1930 to post-1930),
137 geography (North America versus Europe) and setting (sanatorium/hospital studies and
138 non-sanatorium studies) due to the difference in care pattern and start of follow-up time.
139 We choose 1930 as a cut-off as the approximate mid-point of the total time period
140 covered by our included studies. For all analyses, we plotted the study-level survival
141 curves overlaid with the overall survival curves (the mean of the frailty distribution) with
142 95% credible intervals. We estimated study-specific one, five, and ten-year survival
143 probabilities and median survival times with 95% credible intervals.

144
145 We compare the baseline reported severity in individuals treated in sanatoria/hospitals to
146 those not treated in sanatoria, where disease severity reporting is available. We provide
147 summary statistics of this data and test for associations using a chi square test.

148
149 We used R v4.0.2 with Bayesian models fit using JAGS with R2jags v0.6-1. Model
150 diagnostic results are in the supplement. All data and code are available on GitHub at
151 https://github.com/sarahleavitt/TB_mortality.

152

153 **RESULTS**

154 *Literature review*

155 We identified 153 publications (93 from personal libraries, 60 from a manual reference
156 search of personal library publications) during screening and obtained 142 (93%) for
157 full-text review (Figure 1). We excluded 122 publications, most due to not reporting an
158 outcome of interest (n=48), unclear follow-up time (n=23), or reporting population-level
159 rates (n=17). Data from 20 publications were included; three publications included data
160 from the same cohort but were stratified by disease severity across three publications.
161 For two studies ^{7,8}, data required modification to conform to a life table format (see
162 Supplement).

163

164 *Characteristics of cohorts*

165 Of the 20 included publications, 10 (50%) were from North America, 5 (25%) from the
166 United Kingdom, 3 (15%) from Germany (one from an area that is now part of Poland),
167 1 (5%) from the Netherlands, and 1 (5%) from Norway (Table 1). Twelve studies (60%)
168 were conducted primarily before 1930. Most publications enrolled patients between the
169 1910s and 1940s. Ten (50%) publications included patients treated at a sanatorium or
170 an inpatient hospital. Six (30%) publications indicated a subset of patients received
171 surgical treatment; anti-TB chemotherapy was provided to a minority in two publications
172 (Table 1).

173

174 Outcome definitions were not standardized during the pre-chemotherapy era and
175 definition reporting was minimal in some publications (Table S2). Only five (25%)
176 studies described determination of death. These included clinician-determined cause of

177 death, via questionnaire mailed to households, or through multiple sources such as
178 family report or vital record review. Natural recovery definitions included patients with no
179 longer detectable *Mtb*, able to work, and the more stringent, standardized definitions of
180 the United States' National Tuberculosis Association's (USNTA) *Diagnostic Standards*,
181 which required symptom resolution, negative culture, and improved chest x-ray for \geq two
182 years to establish recovery (Table S2) ⁵.

183
184 The 20 publications represented 18 studies. The 18 studies followed 84 cohorts with
185 some studies following one cohort and others identifying different cohorts at each time
186 point for a total of 35,900 patients. Of the 18 studies, 12 (67%) representing 53 (63%) of
187 the cohorts reported TB-specific death. Of the 18 studies, 12 (67%) representing 63
188 (72%) of the cohorts classified alive patients as either having chronic TB or having
189 naturally recovered by each time point.

190
191 *Overall mortality results*
192 Using all data, there were 12 studies, 53 cohorts and 17,166 patients with follow-up
193 data on TB-specific mortality. Assuming a lognormal distribution, we estimated TB-
194 specific mortality survival (meanlog=2.51, sdlog=1.72). Overall, 93% of individuals
195 survived one year (95% CI: 0.84-0.98) while 55% of individuals survived 10 years (95%
196 CI: 0.36-0.73) with a median survival time of 12.33 years (95% CI: 5.47-28.93) (Table 2,
197 Figure S3). Model fit diagnostic plots for all models are in the supplement (Figures S11-
198 S17).

199

200 *Sanatoria results*

201 Survival was higher in the sanatorium studies than in the non-sanatorium studies (Table
202 2, Figure 2, Figure S4). Only 38% (95% CI: 18%-63%) of non-sanatorium individuals
203 survived to ten years compared to 69% (95% CI: 41%-87%) of sanatoria/hospitalized
204 patients. There were more individuals with far advanced disease at baseline among
205 those treated outside of sanatoria (43.0% versus 22.0%) and overall, there was a
206 significant difference in disease severity between those treated in sanatoria compared to
207 those treated elsewhere ($p < 0.001$) (Table S7).

208

209 *Mortality by time*

210 In studies primarily enrolling individuals prior to 1930, ten-year survival was 50% (95%
211 CI: 26%-75%), which was similar to survival estimated in studies primarily enrolling after
212 1930 [60% (95% CI: 29%-89%)]. Median TB survival time was nearly twice as long for
213 studies conducted after 1930 (17.23 years, 95% CI: 2.98-129.05 versus 9.90 years,
214 95% CI: 4.10-25.24) (Table 2, Figures 2 and 3, Figure S5).

215

216 *Mortality by geography*

217 Overall mortality outcomes in North American studies compared to European studies
218 were much better. Ten-year survival was estimated to be 69% in North America (95 CI:
219 54%-81%) compared to 36% in European studies (95% CI: 10%-71%). Median survival
220 times were four-fold higher in North American studies: 21.33 years (95% CI: 11.54,
221 37.95) versus 5.02 (95% CI: 0.86, 28.73) (Table 2, Figure 2, Figure S6).

222

223 *Natural recovery data*

224 The raw natural recovery data are summarized in the Supplement (Figures S6-S9).

225 Consistent with findings from mortality data, overall trends in reported recovery rates
226 were higher among individuals treated in sanatoria, post-1930 and from North America.

227

228 **DISCUSSION**

229 Data from TB patients during the pre-chemotherapy era are a rich resource that can
230 inform current TB management. However, our study illustrates that researchers should
231 be mindful of the context when using these data. There are substantial differences in
232 outcomes based on time, geography, and whether or not patients were in a sanatorium.
233 Inferences from these data should take account of these potential biases and not view
234 pre-chemotherapy era data as homogeneous.

235

236 Geographically, people in North American studies experienced better outcomes than
237 those in Europe. This may be because North America was more shielded from the worst
238 effects of the two world wars in the early 20th century than Europe. Also, an increased
239 proportion of the North American studies used data from sanatoria (6/8 studies versus
240 2/10 studies). Our results indicate that outcomes from sanatoria were better than in
241 studies using community data. However, for studies that reported data on severity at
242 presentation, we found that patients with less severe disease were more likely to spend
243 time in a sanatorium, consistent with the hypothesis that sanatoria were less likely to
244 admit moribund patients as they could not help them. Thus, providing an additional
245 reason why data from North American studies could be biased towards improved

246 outcomes, and illustrating a bias intrinsic to sanatoria data. Stadler et al.⁹ also
247 suggested that sanatoria preferentially treated people of higher socioeconomic status;
248 even if their stay was paid, some people declined a sanatorium stay because they had
249 to work to support family; and some people were excluded as “unsuited” for sanatoria.
250 All of these factors are still seen today regarding the inequities of who does and does
251 not receive TB care. We note that we did not observe significant differences in survival
252 among studies enrolling individuals pre-1930 compared to those enrolling after 1930.

253
254 It is well documented that people with fewer resources, poor access to care, or live in
255 political or social turmoil, such as war-torn areas, have poorer TB outcomes^{10–18}. Our
256 results are consistent with this in that people without access to sanatoria, or living
257 through wartime experienced poorer TB outcomes. Aggregating the pre-chemotherapy
258 era data obscures this important heterogeneity. When making inferences from these
259 data for modern day TB management, it might be necessary to identify specific relevant
260 studies for analysis, depending on the research question.

261
262 A study strength is our focus on TB-specific mortality. Previous studies have looked at
263 all-cause mortality, which would not only over-estimate TB-specific mortality, but also
264 might suffer from different types of heterogeneities, thus obscuring differences in TB-
265 specific mortality between groups^{2,3}. Also, we have used parametric survival modeling
266 to estimate parameters that can be used to generate survival probabilities and
267 parameterize modeling studies.

268

269 Our study has some limitations. The sanatorium discharge start time was used for the
270 sanatoria studies with the average stay added (two of seven sanatoria studies) to this
271 whereas it was notification/diagnosis for the non-sanatoria studies. Therefore, only
272 individuals who survived their sanatorium stay were included. This may have
273 contributed to improved survival in sanatorium studies. Asymptomatic patients were not
274 included in our cohorts since they would not have presented to medical services.
275 Nonetheless, our analyses will have included nearly all symptomatic smear-positive
276 patients. “A few” patients were reported to receive anti-TB drugs and therefore our
277 estimated survivals might be overestimates. However, the availability of anti-TB drugs
278 and types of regimens provided during our study period was limited and inferior to those
279 used nowadays, so it is unlikely that this affected our results substantially. Lastly, our
280 data do not account for the impact of HIV co-infection on TB survival and recovery,
281 since they are from the pre-HIV era.

282

283 Since around 40% of people estimated to develop TB disease in 2021 were not
284 diagnosed ¹, results from the pre-chemotherapy era provide important insights that are
285 relevant nowadays. These data are also critical for modeling studies. However, the
286 context must be considered when using these data to ensure that we make suitable
287 inferences for 21st century TB management.

288

289

290 REFERENCES

- 291 1. Global tuberculosis report 2022. Accessed November 6, 2022.
292 <https://www.who.int/publications/i/item/9789240061729>
- 293 2. Tiemersma EW, van der Werf MJ, Borgdorff MW, Williams BG, Nagelkerke NJD. Natural history of
294 tuberculosis: Duration and fatality of untreated pulmonary tuberculosis in HIV negative patients:
295 A systematic review. *PLoS One*. 2011;6(4). doi:10.1371/journal.pone.0017601
- 296 3. Ragonnet R, Flegg JA, Brilleman SL, et al. Revisiting the Natural History of Pulmonary
297 Tuberculosis: A Bayesian Estimation of Natural Recovery and Mortality Rates. *Clinical Infectious
298 Diseases*. 2021;73(1):e88-e96. doi:10.1093/cid/ciaa602
- 299 4. Jenkins HE, Yuen CM, Rodriguez CA, et al. Mortality in children diagnosed with tuberculosis: a
300 systematic review and meta-analysis. *Lancet Infect Dis*. 2017;17(3). doi:10.1016/S1473-
301 3099(16)30474-1
- 302 5. *Diagnostic Standards and Classification of Tuberculosis*.; 1940.
- 303 6. TUBERCULOSIS FACILITIES IN THE UNITED STATES. *J Am Med Assoc*. 1940;114(9):765-804.
304 doi:10.1001/jama.1940.02810090043014
- 305 7. Mitchell RS. Mortality and relapse of uncomplicated advanced pulmonary tuberculosis before
306 chemotherapy: 1,504 consecutive admissions followed for fifteen to twenty-five years. *Am Rev
307 Tuberc*. 1955;72:502-512.
- 308 8. Stephens MG. Follow-up of 1,041 Tuberculous Patients. *Am Rev Tuberc*. 1941;44(4):451-462.
309 doi:10.1164/art.1941.44.4.451
- 310 9. Stadler E. Der Einfluss der Lungentuberkulose of Lebensdauer und Erwerbsfähigkeit und der Werth
311 der Volksheilstättenbehandlung. *Deutsch Arch Klin Med*. 1903;95:412-440.
- 312 10. Houk V, Baker J, Sorensen K, Kent D. The epidemiology of tuberculosis infection in a closed
313 environment. *Arch Environ Health*. 1968;16:26-35.
- 314 11. Ochs C. The epidemiology of tuberculosis. *JAMA*. 1962;179:247-252.
- 315 12. Enarson D, Wang JS, Dirks J. The incidence of active tuberculosis in a large urban area. *Am J
316 Epidemiol*. 1989;129:1268-1276.
- 317 13. Bhaatti N, Law M, Morris J. Increasing incidence of tuberculosis in England and Wales: a study of
318 the likely causes. *Br Med J*. 1995;310:967-969.
- 319 14. Spence D, Hotchkiss J, Davies P. Tuberculosis and poverty. *Br Med J*. 1993;307:759-761.
- 320 15. McKeown T, Record R. Reasons for the decline in mortality in England and Wales during the
321 nineteenth century. *Popul Studies*. 1962;16:94-122.
- 322 16. Terris M. Relation of Economic Status to Tuberculosis Mortality by Age and Sex. *Am J Public
323 Health*. 1948;38:1061-1070.

- 324 17. Ragan EJ, Kleinman MB, Sweigart B, et al. The impact of alcohol use on tuberculosis treatment
325 outcomes: A systematic review and meta-analysis. *International Journal of Tuberculosis and Lung*
326 *Disease*. 2020;24(1):73-82. doi:10.5588/ijtld.19.0080
- 327 18. Vasankari T, Holmström P, Ollgren J, Liippo K, Kokki M, Ruutu P. Risk factors for poor tuberculosis
328 treatment outcome in Finland: A cohort study. *BMC Public Health*. 2007;7. doi:10.1186/1471-
329 2458-7-291
- 330 19. Heise FH, Hennigar WAP. The condition in 1931 of patients discharged from the Trudeau
331 Sanatorium from 1916-1930. *Tubercle*. 1933;15:120-123.
- 332 20. Rutledge JA, Crouch JB. The Ultimate Results in 1654 Cases of Tuberculosis Treated at the
333 Modern Woodmen of America Sanatorium. *Am Rev Tuberc*. 1919;2(12):755-763.
334 doi:10.1164/art.1919.2.12.755
- 335 21. Ferguson RG. Follow up information on 1747 tuberculosis patients one to seven years after
336 discharge from a sanatorium. *Am Rev Tuberc*. 1926;14(6):625-645.
- 337 22. Wherrett GJ. Follow-up information on 2031 tuberculosis patients one to thirteen years after
338 discharge from sanatoria. *Am Rev Tuberc*. 1935;31(1):62-73.
- 339 23. Zacks D. Pulmonary Tuberculosis in the Second Decade of Life. *Am Rev Tuberc*. 1939;39(6):683-
340 702. doi:10.1164/art.1939.39.6.683
- 341 24. Munchbach W. Das Schicksal des lungentuberkulösen Erwachsenen. Ergebnisse der
342 Heilstättenbehandlung von annähernd 10000 Männern und Frauen. *Tuberkulose-Bibliothek*.
343 1933;49(64).
- 344 25. Baart de la Faille R. *Onderzoek Naar de Resultaten Der Tuberculosebehandeling in Het*
345 *Sanatorium "Berg En Bosch."* 1939.
- 346 26. Mitchell RS. Mortality and relapse of uncomplicated advanced pulmonary tuberculosis before
347 chemotherapy: 1,504 consecutive admissions followed for fifteen to twenty-five years. *Am Rev*
348 *Tuberc*. 1955;72:502-512.
- 349 27. Alling DW, Lincoln S, Bosworth EB. The after history of pulmonary tuberculosis. *Am Rev Tuberc*.
350 1954;70(6):995-1008.
- 351 28. Alling DW, Bosworth EB, Lincoln NS. The after-history of pulmonary tuberculosis. V. Moderately
352 advanced tuberculosis. *Am Rev Tuberc*. 1955;71(4):519-528. doi:10.1164/artpd.1955.71.4.519
- 353 29. Lincoln NS, Bosworth EB, Alling DW. The after-history of pulmonary tuberculosis. III. Minimal
354 tuberculosis. *Am Rev Tuberc*. 1954;70(1):15-31. doi:10.1164/art.1954.70.1.15
- 355 30. Braeuning H, Neissen A. Die Prognose der offenen Lungentuberkulose. *Tuberkulose-Bibliothek*.
356 1933;52(40).
- 357 31. Tattersall WH. The survival of sputum-positive consumptives: A study of 1,192 cases in a County
358 Borough between 1914 and 1940. *Tubercle*. 1947;28(6):107-123.
359 doi:[https://doi.org/10.1016/S0041-3879\(47\)80088-1](https://doi.org/10.1016/S0041-3879(47)80088-1)

- 360 32. Holst P, Nicolaysen L, Ustvedt Y. Untersuchungen über die Lebensdauer der Schwindsüchtigen in
361 Norwegen. *Deutsch Arch Klin Med.* 1906;88(325-350).
- 362 33. Cox L. Sanatorium treatment contrasted with home treatment; after-histories of 4,067 cases.
363 *British Journal of Tuberculosis.* 1923;17:27-30.
- 364 34. Bentley F J, Grzybowski S, Benjamin B. Tuberculosis in childhood and adolescence. The National
365 Association for the Prevention of Tuberculosis. London, England: Waterlow and Sons Ltd, 1954.
- 366 35. Lowe CR. Recent trends in survival of patients with respiratory tuberculosis. *Br J Prev Soc Med.*
367 1954;8:91-98.
- 368 36. Springett VH. Ten-year results during the introduction of chemotherapy for tuberculosis.
369 *Tubercle.* 1971;52(2):73-87. doi:10.1016/0041-3879(71)90015-8
- 370
- 371

372 **ACKNOWLEDGEMENTS** We acknowledge Margaret Zimmer for research assistance
373 and Annelies Mesman for translation support. The results reported herein correspond to
374 specific aims of grants R01GM122876 and R35GM141821 to investigator LFW, HEJ,
375 and CRH, and R01GM122876-04S1 to CAR from the National Institute of General
376 Medical Sciences of the National Institutes of Health. This work was also supported by
377 grant K23 AI152930 from the National Institutes of Health to TCB and a Burroughs
378 Wellcome Fund/American Society for Tropical Medicine and Hygiene Postdoctoral
379 Fellowship in Tropical Infectious Diseases to TCB. The content is solely the
380 responsibility of the authors and does not necessarily represent the official views of the
381 National Institutes of Health.

382
383 **DATA AND CODING:** We used R v4.0.2 with Bayesian models fit using JAGS with
384 R2jags v0.6-1. All data and code are available on GitHub at
385 https://github.com/sarahleavitt/TB_mortality.

386

387

388 **TABLES**

389 **Table 1. Descriptive characteristics of pre-treatment era cohorts with data on tuberculosis-specific mortality**
 390 **and/or, natural recovery from tuberculosis (N=20 publications, 18 studies)**

Reference	Enrollment period (years)	Setting	Population	Diagnostic method	Treatment	Stratification
Heise ^{19b}	1917-1931	New York, USA	Sanatorium cohort of patients with pulmonary tuberculosis treated at Trudeau Sanatorium	Chest X-ray	Sanatoria (100%)	By disease severity
Rutledge ^{20b}	1909-1914	Colorado, USA	Sanatorium cohort treated at the Modern Woodmen of America sanatorium	Sputum smear and clinical findings	Sanatoria (100%)	By disease severity

Stephens ^{8a,b}	1919-1938	New York, USA	Employment-based sanatorium cohort of employees ages 17-64 of the Metropolitan Life Insurance Company	Signs and symptoms (pre-1924) and chest X-ray (1925+)	Sanatoria (100%)	By disease severity
Ferguson ^{21a,b}	1917-1924	Saskatchewan, Canada	Sanatorium survivor cohort of patients ages 0-69 treated at Fort Qu'Appelle Sanatorium	Not reported	Sanatoria (100%), surgery (1%)	None
Wherrett ^{22a}	1917-1924	Saskatchewan, Canada	Sanatorium survivor cohort of patients treated at Fort Qu'Appelle Sanatorium	Not reported	Sanatoria (100%)	None
Zacks ^{23b}	Approximately 1927-	Massachusetts	Cohort of children with pulmonary TB; unclear	Not reported	Surgery and/or	None

	1932	tts, USA	whether cohort were inpatient, outpatient, or population-representative		sanatoria (55%)	
Munchbach ^{24a}	1920-1927	Black Forest, south-west Germany	Sanatorium cohort of patients with pulmonary TB	Not reported	Sanatoria (100%)	None
Baart de la Faille ^{25a}	1922-1935	Bilthoven, Netherlands	Sanatorium-based cohort of adults with pulmonary tuberculosis at the Berg en Bosch sanatorium	Chest x-ray	Sanatoria (100% and collapse therapy (34%)	None
Mitchell ^{26a,b}	1930-1939	New York, USA	Sanatorium cohort of patients aged 15+ with pulmonary TB treated at the Trudeau	Chest X-ray and symptoms	Sanatoria (100%)	By disease severity

Sanatorium

Alling, 1954 ²⁷ ,	1938-1948	New York,	Outpatient clinical	Sputum smear	Among far-	By disease
Alling, 1955 ²⁸ ,	(Alling,	USA	cohort of smear+	and chest X-	advanced	severity
Lincoln ^{29a,b}	1954 and		patients ages 10+ with	ray	patients: Bed	
	1955)		far advanced (Alling,		rest (27%),	
	1937-1947		1954), moderately		collapse therapy	
	(Lincoln)		advanced (Alling,		(27%)	
			1955), and minimal		Among	
			(Lincoln) TB treated at		moderately-	
			Hermann M Biggs		advanced	
			Memorial Hospital		patients: Anti-	
					TB drugs (“a	
					few”), collapse	
					therapy (37%),	
					bed rest (63%)	
					Among minimal	

					patients: NR	
Braeuning ^{30a,b,c}	1920/21, 1927/28, 1930, 1933	Stettin, Germany (at the time of the study, currently Poland)	Population-based cohort of pulmonary TB patients attending a public TB clinic	Chest x-ray	Surgery or sanatoria (% receiving each unclear)	By disease severity
Stadler ^{9a}	1893-1902	Marburg, Germany	Population based cohort of pulmonary TB patients attending public outpatient clinics	Not reported	Not reported	None
Tattersall ^{31b}	1914-1940	Reading, UK	Population-based cohort of smear+ patients from the Tuberculosis Service of Reading County	Smear positive	Sanatoria (38%)	By disease severity

			Borough			
Holst ^{32a}	1891-1900	Norway	People with TB as reported through forms filled out by doctors	Clinical signs and symptoms	Not reported	None
Lissant Cox ^{33a,b}	Approximately 1934-1935	Lancaster, UK	Cohort of smear+ patients while in residential treatment at sanatoria and pulmonary hospitals	Sputum or other specimen smear, signs and symptoms	Sanatoria (100%)	By disease severity
Bentley ^{34a,b}	1942-1946	Brentwood, Essex, UK	Hospital-based cohort of pediatric patients ages 0-16 years treated at High Wood Hospital	Chest X-ray and clinical findings	Sanatoria (100%)	None
Lowe ^{35b}	1930-1951	Birmingham,	Population-based cohort of all patients	Not reported	Not reported	None

UK
with respiratory TB
identified from
registers

Springett ^{36a}	1947	Birmingham, UK	Outpatient clinical cohort of sputum+, respiratory TB patients ages 15+ seeking care the Birmingham Chest Service	Sputum smear	Surgery (30%), Anti-TB drugs (0.2% ^d)	None
--------------------------	------	-------------------	--	--------------	---	------

391 ^aData available on natural recovery

392 ^bData available on TB-specific mortality

393 ^cIncludes TB-specific mortality data that were not stratified by severity and for a subset of this cohort, TB-specific mortality
394 and natural recovery data stratified by severity

395

Table 2. Survival analysis results TB-specific mortality with the combined and stratified models

Setting	Survival Distribution	1-Year Survival (95% CI)	5-Year Survival (95% CI)	10-Year Survival (95% CI)	Median Survival Time (95% CI)	Number of Studies	Number of Cohorts	Number of Individuals
Combined	lognormal(2.51, 1.72)	0.93 (0.84, 0.98)	0.7 (0.52, 0.85)	0.55 (0.36, 0.73)	12.33 (5.47, 28.93)	12	53	17,166
Time Stratified								
Pre-1930	lognormal(2.29, 1.40)	0.95 (0.84, 0.99)	0.69 (0.44, 0.88)	0.50 (0.26, 0.75)	9.90 (4.10, 25.24)	7	30	10,732
Post-1930	lognormal(2.85, 2.13)	0.91 (0.70, 0.99)	0.72 (0.41, 0.94)	0.60 (0.29, 0.89)	17.23 (2.98, 129.05)	5	23	6,434
Geography Stratified								
North America	lognormal(3.06, 1.56)	0.98 (0.94, 0.99)	0.83 (0.71, 0.90)	0.69 (0.54, 0.81)	21.33 (11.54, 37.95)	7	28	10,619
Europe	lognormal(1.85, 2.13)	0.8 (0.47, 0.99)	0.5 (0.18, 0.89)	0.36 (0.10, 0.89)	5.02 (0.86, 28.93)	5	25	6,547

	61, 1.92)	0.96)	0.82)	0.71)	28.73)			
Setting Stratified								
Sanatorium/hospital	lognormal(3.08, 1.56)	0.98 (0.90, 1.00)	0.83 (0.59, 0.94)	0.69 (0.41, 0.87)	21.77 (7.09, 59.58)	7	41	12,330
Non-Sanatorium	lognormal(1.67, 2.03)	0.80 (0.58, 0.93)	0.51 (0.28, 0.75)	0.38 (0.18, 0.63)	5.29 (1.53, 19.92)	5	12	4,836

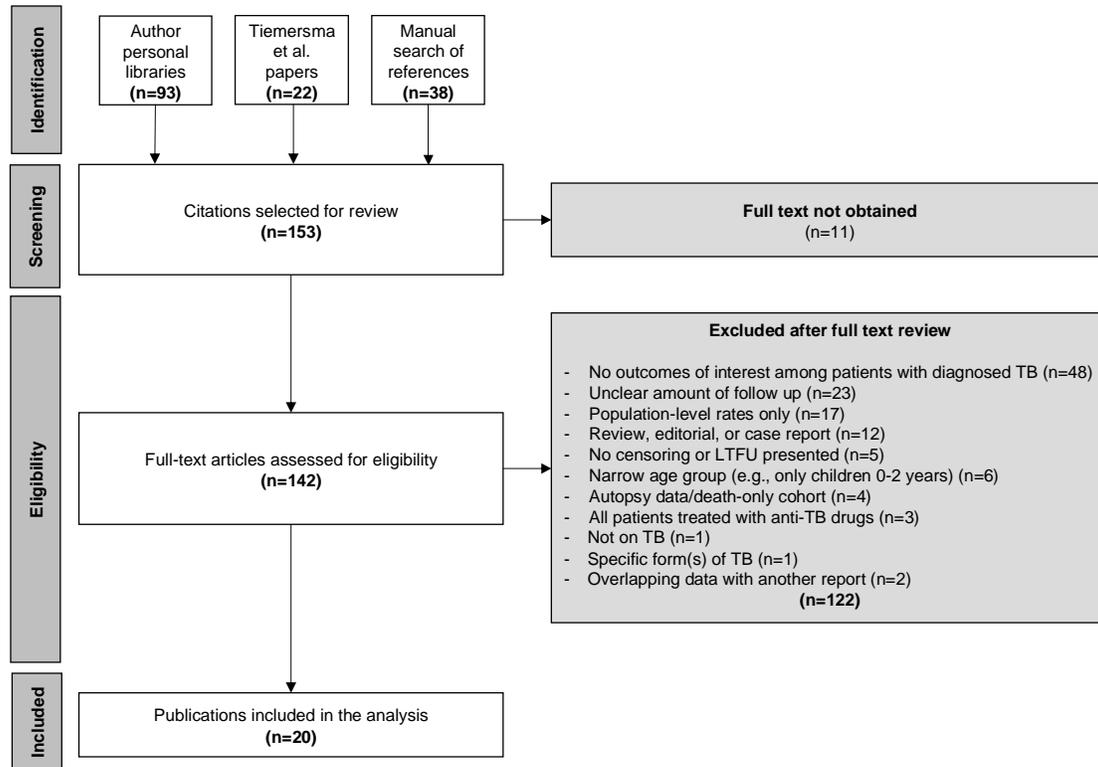
397

398

399

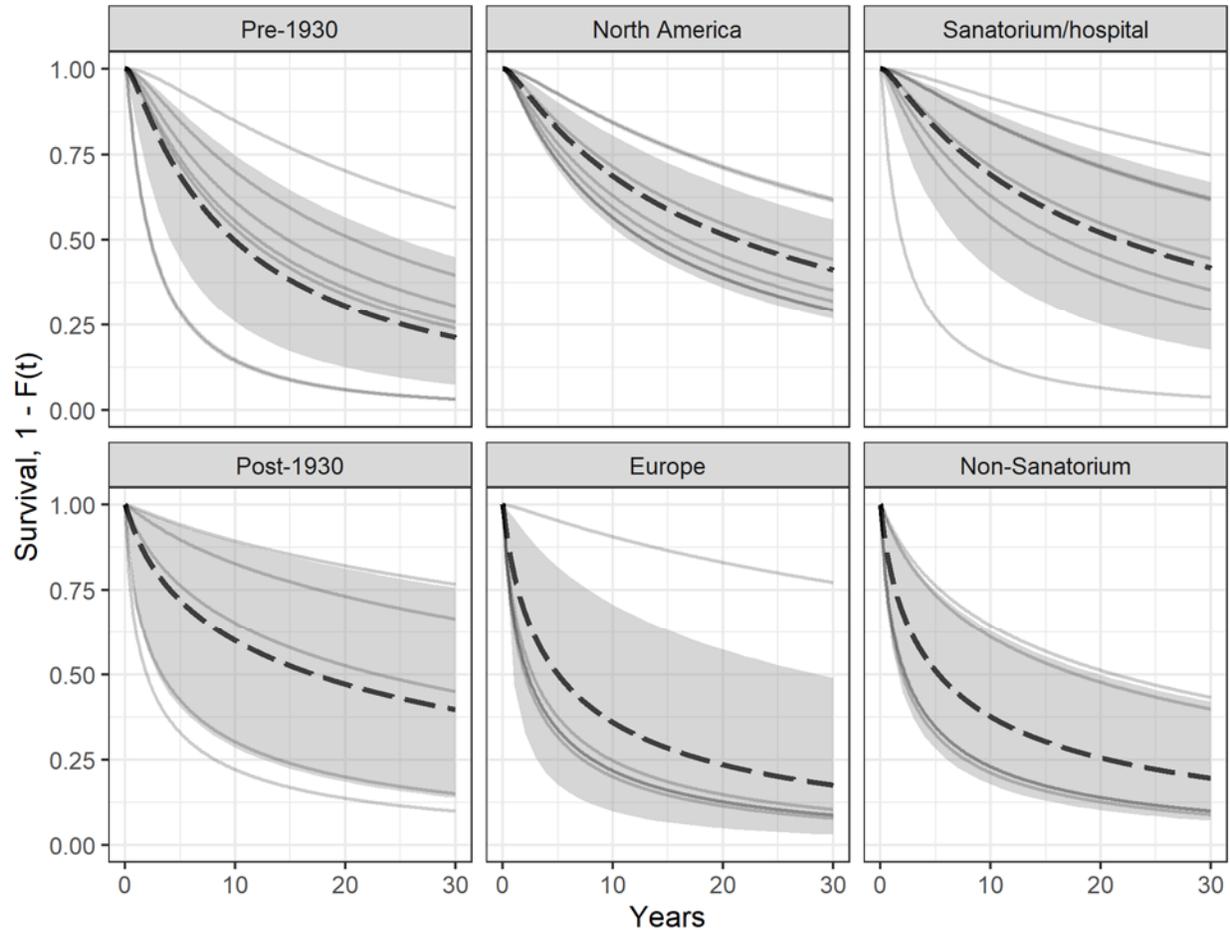
400

401 **FIGURES**



402

403 **Figure 1 Flow chart describing the review process and exclusions**



404

405 **Figure 2. Survival curves for TB-specific mortality for the time period, study**

406 **location, and treatment location stratified analyses.** The thin, grey lines represent

407 the specific survival curves for each study. The black dotted line represents the overall

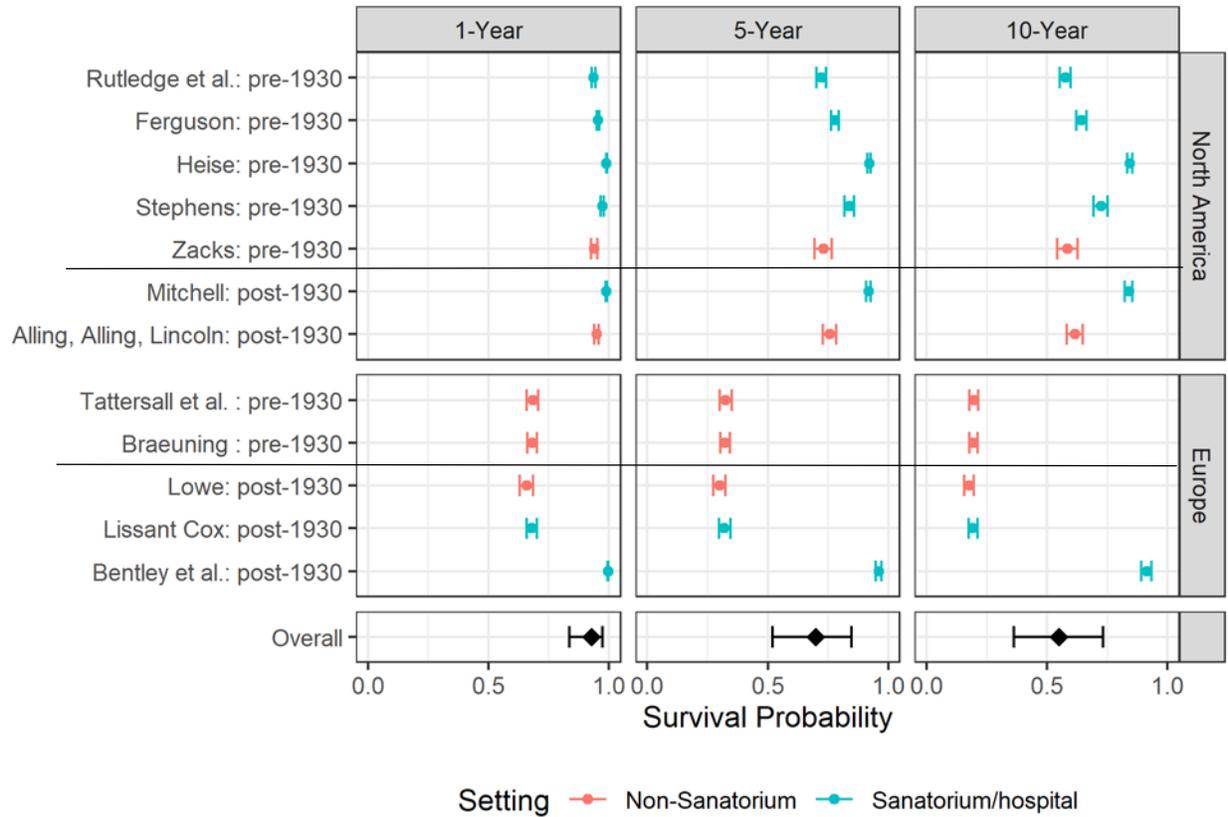
408 survival curve (the mean of the frailty distribution). The grey shaded area is the 95%

409 credible interval for the overall survival curve.

410

411

412



413
414 **Figure 3. Forest plot of one, five, and ten-year survival probabilities for TB-**
415 **specific mortality for the combined model.** The studies are sorted by first year of
416 enrollment (earliest to latest) with a horizontal line splitting the pre-1930 and post-1930
417 studies. The two panels for stratify the studies by geography and the dots and bars are
418 colored by setting. The rows are the survival probabilities with 95% credible intervals for
419 the different studies. The overall estimate of the survival probabilities is located at the
420 bottom of the plot and represented with a diamond

421

422