The Temporal Dynamics of Relapse and Reinfection Tuberculosis After Successful Treatment: A Retrospective Cohort Study

Florian M. Marx,1,4 Rory Dunbar,1 Donald A. Enarson,1,5 Brian G. Williams,2 Robin M. Warren,2 Gian D. van der Spuy,2 Paul D. van Helden,2,a and Nulda Beyers1,a

1Desmond Tutu TB Centre, Department of Paediatrics and Child Health, Stellenbosch University, and 2DST/NRF Centre of Excellence for Biomedical Tuberculosis Research, US/MRC Centre for Molecular and Cell Biology, Division of Molecular Biology and Human Genetics, Stellenbosch University, Tygerberg Campus, Cape Town, and 3DST/NRF Centre of Excellence in Epidemiological Modelling and Analysis (SACEMA), Stellenbosch University, Stellenbosch, South Africa; 4Department for Pediatric Pneumology and Immunology, Charité–Universitätsmedizin, Berlin, Germany; and 5International Union Against Tuberculosis and Lung Disease, Paris, France

Background. There is increasing evidence from tuberculosis high-burden settings that exogenous reinfection contributes considerably to recurrent disease. However, large longitudinal studies of endogenous reactivation (relapse) and reinfection tuberculosis are lacking. We hypothesize a relationship between relapse vs reinfection and the time between treatment completion and recurrent disease.

Methods. Population-based retrospective cohort study on all smear-positive tuberculosis cases successfully treated between 1996 and 2008 in a suburban setting in Cape Town, South Africa. Inverse gaussian distributions were fitted to observed annual rates of relapse and reinfection, distinguished by DNA fingerprinting of Mycobacterium tuberculosis strains recultured from diagnostic samples.

Results. Paired DNA fingerprint data were available for 130 (64%) of 203 recurrent smear-positive tuberculosis cases in the 13-year study period. Reinfection accounted for 66 (51%) of 130 recurrent cases overall, 9 (20%) of 44 recurrent cases within the first year, and 57 (66%) of 86 thereafter (P<.001). The relapse rate peaked at 3.93% (95% confidence interval [CI], 2.35–5.96%) per annum 0.35 (95% CI, .15–.45) years after treatment completion. The reinfection tuberculosis rate peaked at 1.58% (95% CI, .94–2.46%) per annum 1.20 (95% CI, .55–1.70) years after completion.

Conclusions. To our knowledge, this is the first study of sufficient size and duration using DNA fingerprinting to investigate tuberculosis relapse and reinfection over a lengthy period. Relapse occurred early after treatment completion, whereas reinfection dominated after 1 year and accounted for at least half of recurrent disease. This temporal relationship may explain the high variability in reinfection observed across smaller studies. We speculate that follow-up time in antituberculosis drug trials should take reinfection into account.

Keywords. Mycobacterium tuberculosis; recurrence; reinfection; DNA fingerprinting; South Africa.

Recurrent tuberculosis after successful treatment constitutes a challenge to tuberculosis control, particularly in populations with a high prevalence and with high rates of human immunodeficiency virus (HIV) coinfection. The question whether recurrent tuberculosis is due predominantly to endogenous reactivation (relapse) or exogenous reinfection has been a subject of debate for decades [1]. Models have suggested that the contribution of reinfection increases with the incidence of tuberculosis and risk of infection in the particular setting [2–4].

DNA fingerprint technologies [5] have been used to differentiate between relapse and exogenous reinfection. Reinfection seems to be an important [6, 7] and even major [8, 9] underlying cause of recurrent tuberculosis in high-burden settings. However, studies have yielded conflicting evidence, with the proportion of reinfection tuberculosis...
ranging from 0% to 100% in studies conducted in earlier years (ie, 1993–2001) [10]. At least 2 studies in more recent years concluded that reinfection was not a common cause of disease recurrence in a high-burden setting [11, 12]. Thus, the actual extent to which reinfection contributes to recurrent tuberculosis after successful treatment remains uncertain. There is more consistent evidence that reinfection rather than relapse may explain high rates of disease recurrence observed among individuals coinfected with HIV [13–15].

Longitudinal studies of relapse and reinfection tuberculosis in high-burden settings have been limited, because they depend on lengthy follow-up and sample availability. Given that small studies may be subject to sampling bias [16], large longitudinal studies are needed to better clarify the contribution and timing of relapse and reinfection to recurrent tuberculosis [10, 17]. Better knowledge about the timing of relapse and reinfection may be of value for the design of antituberculosis drug trials, which usually do not take reinfection into account. At least 2 of the current large phase 3 antituberculosis drug trials rely on 24-month posttreatment follow-up for their primary outcome [18, 19].

We conducted a large longitudinal study in a tuberculosis high-burden setting with known high rates of recurrent disease [20]. Recurrent tuberculosis contributes significantly to the overall burden of tuberculosis in the area, irrespective of its association with HIV [21]. The purpose of our study was to investigate, among successfully treated smear-positive tuberculosis cases, the relationship between the type of recurrence—that is, endogenous reactivation (relapse) or exogenous reinfection—and the time to recurrent smear-positive tuberculosis.

METHODS

Study Setting
The study was conducted in a suburban area in Cape Town, South Africa, covering 3.4 km², with approximately 39 000 inhabitants of low socioeconomic status [22]. The directly observed treatment, short-course (DOTS) strategy [23] was introduced in 2 primary health care clinics in 1996. Treatment was supervised daily by health care staff in the clinics or in the community by trained health care workers. Treatment outcomes were documented according to standard definitions [24]: cure, for patients who were smear negative at or 1 month before the completion of treatment and on ≥1 previous occasion; or treatment completed, for patients who had completed treatment but without proof of cure because smear results were unavailable. The term treatment success includes both cure and treatment completed.

Study Design
The study was nested in a retrospective cohort study [20] conducted among all patients with smear-positive tuberculosis receiving standard first-line treatment documented in the local treatment registers who either successfully completed or defaulted from treatment between 1996 and 2008. The underlying study used routinely collected treatment register data to investigate the frequency of retreatment for smear-positive tuberculosis by previous treatment outcome [20].

The present study included all patients with smear-positive tuberculosis who successfully completed their index treatment episode (documented cure or treatment completed) and were subsequently treated again for sputum smear-positive tuberculosis. Any recurrent smear-positive treatment episode recorded from the first day after the index treatment episode had been completed was considered, as recommended by Lambert et al [10]. The principal determinant was the type of recurrence (ie, relapse or reinfection), and the main outcome was the time to retreatment for smear-positive tuberculosis.

Ascertainment of Relapse and Reinfection
Sputum samples used for this study were routinely collected from all patients, processed in the routine laboratory, and subsequently transported to the research laboratory for culture and genotypic analysis. Decontaminated samples were cultured in BACTEC460, MGIT 960 or Löwenstein-Jensen medium for DNA extraction [25]. Isolates were classified using the internationally standardized IS6110 DNA fingerprinting method [5]. Each DNA fingerprint was entered into a GelCompar database and analyzed using GelCompar software (version 6.5). For each case included in the study, we identified diagnostic samples collected for both, the index treatment and the recurrent episode. A diagnostic sample was defined as the first chronological sputum sample with a sample date between 2 months before and 2 months after the onset date of the corresponding treatment episode.

Paired DNA fingerprint patterns (ie, from the index and recurrent treatment episodes for the same individual) were compared on the basis of their restriction fragment length polymorphism patterns using UPGMA (Unweighted-pair group method with arithmetic mean) and the Dice coefficient. Relapse was inferred when both patterns either were identical or differed by only 1 or 2 bands, provided that an identical strain pattern was not identified in a sample of another case before the corresponding treatment episode of the patient (implying an intrapatient evolutionary event) [26]. Reinfection was inferred if paired samples did not meet the definition for relapse. Previous studies from this laboratory have shown that laboratory error and cross-contamination were on the order of 2.1% [8, 27]. To investigate the possibility of rifampicin drug resistance at baseline and recurrence, DNA sequencing for rpoB genotype mutation was conducted for all diagnostic samples.

Time to Recurrence
The time to recurrence was defined as the time between the documented end date of the index treatment episode and the date when the recurrent treatment episode was recorded. In a few cases with missing dates, the end date was identified via a patient folder search.
**Data Analysis**

Survival analysis was used to compare the distributions of relapse and reinfection tuberculosis cases over time since treatment completion. The Kolmogorov-Smirnov test was used to test for a difference in the observed distributions. Annual rates of relapse and reinfection tuberculosis were calculated by dividing the number of relapse and reinfection cases by the population at risk at quarterly (3-month) intervals, assuming that the proportion of relapse and reinfection cases observed at each time interval would not differ between recurrent cases with and without ascertained type of recurrence. Inverse gaussian functions [28] were fitted to the observed annual rates to estimate the trends in relapse and reinfection tuberculosis over time. Sensitivity analysis was conducted by including smear-negative/culture-positive recurrent episodes in the analysis.

**Ethics Statement**

The study was approved by the Committee for Human Research, Faculty of Medicine and Health Sciences, Stellenbosch University (N09/05/144 and amendments 1 and 3).

**RESULTS**

A total of 2359 cases with an index episode of treatment for smear-positive tuberculosis were recorded in the treatment registers, of whom 1869 (79.2%) were successfully treated (1743 were cured, 126 completed treatment). A recurrent episode of treatment for smear-positive tuberculosis after the index episode was recorded for 203 (10.9%) of 1869 successfully treated cases. Figure 1 illustrates cases included in and excluded from the study. The median time to smear-positive retreatment was 21 months (interquartile range [IQR], 8–48 months). Two of the 203 recurrent cases had been classified as cured/treatment completed at the index treatment episode despite having a positive sputum smear result documented at the end of treatment, but in both cases the smear result was classified as “scanty.”

**Completeness of Paired DNA Fingerprint Results**

For 130 (64%) of the 203 recurrent smear-positive cases, a DNA fingerprint result was available for both the diagnostic sample collected at the index episode and that collected at the recurrent episode (Table 1). Availability of both results was associated

---

**Figure 1.** Overview of tuberculosis cases included in and excluded from the study.

*Not including 17 treatment outcomes that were retrieved via patient file search*
with the calendar year of treatment completion ($P = .03$). There was no association between the time to recurrence and availability of both DNA fingerprint results ($P = .63$). The median time to recurrence was 20.4 months (IQR, 9.0–45.5 months) in recurrent cases with both results available vs 24.8 months (IQR, 6.1–53.7 months) in recurrent cases with $\geq 1$ result missing ($P = .92$).

### Relapse and Reinfection Tuberculosis

Of the 130 recurrent smear-positive tuberculosis cases with both DNA fingerprint results available, 64 (49%) were relapse cases and 66 (51%) reinfection tuberculosis cases (Table 2). Reinfection accounted for 9 (20%) of 44 recurrent cases within the first year, and 57 (66%) of 86 recurrent cases thereafter. After 2 years 43 (72%) of 60 recurrent cases were due to reinfection (Table 2). There was a strong association between relapse and earlier recurrence (Kolmogorov-Smirnov test for difference in the observed distributions, $P < .001$; Figure 2).

The association between relapse, reinfection, and the time to recurrence remained after the analysis was restricted to cases after cure (66% relapse within 2 years vs 27% later than 2 years; $P < .001$) and after it was restricted to 82 recurrent cases with a documented HIV-negative test result (59% relapse within 2 years vs 21% later than 2 years; $P = .02$). Ten of 100 patients with documented HIV-positive tuberculosis had an episode of recurrent smear-positive tuberculosis. DNA fingerprint results were available in 7, of whom 6 had relapse cases.

Table 3 compares characteristics between patients with relapse or reinfection tuberculosis and those without recurrent treatment. None of the index episode diagnostic samples and 2 of the 130 recurrent episode samples (1.5%), 1 relapse and 1 reinfection, were found to be positive for $rpoB$ mutation.

### Annual Rates of Recurrent Tuberculosis Due to Reinfection and Relapse

The relapse rate peaked at 3.93% (95% confidence interval [CI], 2.35%–5.96%) per annum 0.35 (95% CI, .15–.5) years after completion of treatment, followed by a steady decline (Figure 3). The reinfection tuberculosis rate peaked at 1.58% (95% CI, .94%–2.46%) per annum 1.20 (95% CI, .55–1.70) years after completion of treatment (Figure 3). Table 2 summarizes numbers and annual rates of recurrent, relapse and reinfection tuberculosis cases.

### Sensitivity Analysis: Smear-Negative, Culture Positive Recurrent Episodes

Including smear-negative/culture-positive recurrent episodes in the analysis did not change the observed association between relapse, reinfection and the time to recurrence. Of 276 recurrent cases confirmed by either smear or culture, 159 (58%) had DNA fingerprint results available for both the index and the recurrent treatment episode. Of these, 82 (52%) were relapse cases, of which 54 (66%) were treated again within 2 years, compared with only 26 (34%) of 77 reinfection cases ($P < .001$).
Mixed Strain Infection at the Index Episode

Two different strains of *M. tuberculosis* were detected at the index episode in 2 of the 130 recurrent cases. At recurrence, both were classified as reinfection according to the study definition. However, for both, the single strain found at the recurrent episode was equal to the underlying strain at the index episode (all rpoB wild-type infections).

**DISCUSSION**

This is the first longitudinal study of sufficient duration and size to investigate relapse and reinfection in relation to the time to recurrent smear-positive tuberculosis. Relapse occurs predominantly in the first year, whereas reinfection tuberculosis predominates over the subsequent years after successful treatment. The delayed appearance of reinfection tuberculosis is plausible, given that the risk of reinfection is related to the disease burden in the community, and there is a lag-time period between infection and progression to disease.

Our results are consistent with findings from the earlier British Medical Research Council trials, which showed that the majority of relapse cases would occur within 6–12 months of treatment [29, 30], notably without being able to distinguish between relapse and reinfection. Our results are also consistent with those of studies from South Africa [13, 15] and China [31] indicating that relapse predominated in patients with early recurrence, whereas reinfection was associated with longer intervals between tuberculosis episodes.

Our results have implications for studies of relapse and reinfection tuberculosis in high-burden populations. Time-of-observation bias is likely for retrospective analyses of clinical trials, which usually rely on relatively short follow-up periods, favoring higher proportions of relapse, as observed recently in a large retrospective analysis in Uganda [12]. Furthermore, sampling bias may occur in retrospective studies if the sample of patients was small and not representative of all recurrent cases. Variation in follow-up or observation time and, consecutively, in the time to recurrent tuberculosis may thus serve to explain the high variability in the contribution of relapse and reinfection observed across studies [10].

Our results have implications for clinical trials. We show that reinfection was the dominant cause of recurrent tuberculosis soon after the first year since treatment completion. Phase 3 antituberculosis drug trial designs may thus benefit from shortening of follow-up time in order to avoid the undesired “contaminating” effect of reinfection tuberculosis. This is in accordance with previous recommendations based on recurrence rates in the earlier British Medical Research Council trials [30, 32].

We show that, in a high-burden setting, recurrence of smear-positive tuberculosis after treatment success may be due to reinfection in at least half of all recurrent cases. Two prior studies conducted in the area concluded that exogenous reinfection was the major cause of tuberculosis after cure and that the rate of

<table>
<thead>
<tr>
<th>Category</th>
<th>Total</th>
<th>Year 1</th>
<th>Year 2</th>
<th>Year 3</th>
<th>Year 4</th>
<th>Year 5+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Recurrent cases, No. (%)</td>
<td>203</td>
<td>67 (33)</td>
<td>39 (19)</td>
<td>28 (14)</td>
<td>20 (10)</td>
<td>51 (25)</td>
</tr>
<tr>
<td>Recurrent cases with DNA fingerprint data, No. (%)</td>
<td>130</td>
<td>44 (34)</td>
<td>26 (20)</td>
<td>17 (13)</td>
<td>16 (12)</td>
<td>27 (21)</td>
</tr>
<tr>
<td>Due to relapse</td>
<td>64</td>
<td>35 (55)</td>
<td>12 (19)</td>
<td>4 (6)</td>
<td>5 (8)</td>
<td>8 (13)</td>
</tr>
<tr>
<td>Due to reinfection</td>
<td>66</td>
<td>9 (14)</td>
<td>14 (21)</td>
<td>13 (20)</td>
<td>11 (17)</td>
<td>19 (29)</td>
</tr>
<tr>
<td>Incidence risk of recurrencea</td>
<td>. . .</td>
<td>3584</td>
<td>2164</td>
<td>1578</td>
<td>1140</td>
<td>. . .</td>
</tr>
<tr>
<td>Incidence risk of relapsea (95% CI)</td>
<td>. . .</td>
<td>2836 (2131–3693)</td>
<td>998 (593–1574)</td>
<td>371 (158–811)</td>
<td>342 (126–743)</td>
<td>. . .</td>
</tr>
<tr>
<td>Incidence risk of reinfection tuberculosisa (95% CI)</td>
<td>. . .</td>
<td>733 (410–1253)</td>
<td>1165 (723–1776)</td>
<td>1184 (734–1804)</td>
<td>798 (437–1336)</td>
<td>. . .</td>
</tr>
</tbody>
</table>

Abbreviation: CI, confidence interval.

* Per 100 000 successfully treated cases.

![Proportion Recurrence Free](https://academic.oup.com/cid/article-abstract/58/12/1676/2895361)

**Figure 2.** Kaplan-Meier survival estimates for relapse (n = 64), reinfection tuberculosis (n = 66), and unknown type of recurrence (n = 73) (Kolmogorov-Smirnov test for difference in relapse vs reinfection, *P* < .001).
reinfection tuberculosis exceeded that of new tuberculosis cases [8, 9]. Both studies were limited by the small number of patients studied, and neither took into account the time to recurrence.

Although the overall risk of relapse may be reduced by provision of standardized and supervised treatment, the increased risk of reinfection in high-burden settings exacerbates the burden of recurrent tuberculosis. Thus, successfully treated cases, as well as treatment defaulters [20], represent an important risk group in tuberculosis high-burden settings, a point worth considering in screening and active case finding strategies.

It remains uncertain to what extent previously treated cases contribute to tuberculosis transmission within communities. den Boon et al [33] reported that previously treated individuals represented more than half of patients with prevalent smear-positive tuberculosis. Both their and our study took place in a community where tuberculosis rates remain constantly high, >10 years after the introduction of the internationally recommended control strategy. The strength of our study is the large sample of recurrent tuberculosis cases, sufficient to study relapse and reinfection tuberculosis in a high-burden setting over a lengthy period.

Our study has limitations. We used a programmatic definition of treatment success on the basis of routinely documented information. The source and availability of data did not enable us to describe clinical and biological mechanisms underlying relapse and reinfection tuberculosis. In particular, we were unable to exclude the possibility of drug resistance beyond rpoB.
genotype results at treatment initiation or to describe acquisition of drug resistance during treatment as a potential cause of relapse.

Both DNA fingerprint results were available in 64% of all recurrent cases—a higher proportion than in earlier studies. Availability was lower in patients treated in earlier years and with lower sputum smear grades, which may relate to the probability of bacterial survival and successful reculturing. We did not observe an association between availability and time between both treatment episodes; it is therefore unlikely that selection bias explains the observed temporal distributions of relapse and reinfection.

Using retreatment probably underestimated true rates of disease recurrence. Some patients with recurrent disease might not have returned for treatment or sought treatment elsewhere, and the population at risk in this study includes those who moved or died. Time to retreatment involves not only the time to recurrent disease but also time to diagnosis and treatment.

The definition of relapse and reinfection was based on comparison of *M. tuberculosis* DNA fingerprint patterns using restriction fragment length polymorphism analysis, the state of the art method when this study was conducted. More advanced methods, such as whole genome sequencing, may be more powerful for discriminating different strains, but they were beyond the scope of our study. Lower discriminative power may have led to an underestimation of reinfection but is unlikely to explain the observed temporal distributions.

Although strain diversity is known to be high in this area [34], individuals may have been reinfected within clusters of the same strain, for example, after being repeatedly exposed in the household or in social networks. Conversely, mixed-strain infections [35, 36] may have led to misclassification: Rather than reinfection, a selection process of the underlying strain may have led to recurrent disease in individuals with mixed-strain infection, particularly if the underlying strain was drug resistant.

We conducted our study in a setting where HIV coinfection is less common than in other high-burden settings. Low numbers of coinfected patients with recurrent tuberculosis precluded any rate estimates for this subgroup. Settings with more frequent HIV coinfection may have higher overall rates of reinfection [15, 37, 38] and shorter time to relapse and reinfection. Furthermore, the frequency and timing of reinfection tuberculosis may vary across settings with a high vs low annual risk of infection.

Future research should serve to improve our understanding of programmatic and individual risk factors for relapse and reinfection tuberculosis. A more comprehensive model of strain persistence and repeated infections that takes into account mechanisms such as treatment adherence, mixed infections, and acquisition of drug resistance would be valuable. Reducing relapse and reinfection tuberculosis may be an important means of reducing the overall tuberculosis burden in high-prevalence settings.

**Notes**

**Acknowledgments.** We are grateful to our laboratory staff, in particular Marianna de Kock, Michael Stead, Elizabeth Streicher, and Louise Vos, our study nurses, and the data management team. We express our appreciation to the local staff working in the clinics. We thank the City of Cape Town Directorate of Health for giving permission to access the program data (research ID 10142).

**Financial support.** This work was supported by a United States Agency for International Development cooperative agreement (Technology, Research, Education and Technical Assistance for Tuberculosis [TREAT TB] agreement GHN-A-00-08-00004-00) managed by the International Union Against Tuberculosis and Lung Disease (The Union), and by the Oskar Helene Heim Foundation (research grant to F. M. M.). The contents of this article are the sole responsibility of the authors and can under no circumstances be regarded as reflecting the positions of The Union or of the donors.

**Potential conflicts of interest.** All authors: No reported conflicts.

All authors have submitted the ICMJE Form for Disclosure of Potential Conflicts of Interest. Conflicts that the editors consider relevant to the content of the manuscript have been disclosed.

**References**


