Pathogens and Disease, 75, 2017, ftx020

doi: 10.1093/femspd/ftx020
Advance Access Publication Date: 18 February 2017
Minireview

MINIREVIEW

Relapse, re-infection and mixed infections in tuberculosis disease

Amanda McIvor[1], Hendrik Koornhof[2] and Bavesh Davandra Kana[1,3,*]

[1]DST/NRF Centre of Excellence for Biomedical TB Research, School of Pathology, Faculty of Health Sciences, University of the Witwatersrand and the National Health Laboratory Service, Johannesburg 2000, South Africa, [2]Centre for Tuberculosis, National Institute for Communicable Diseases and National Health Laboratory Service, Johannesburg, 2000, South Africa and [3]CAPRISA, Centre for the AIDS Programme of Research in South Africa, Durban, 4001, South Africa

*Corresponding author: DST/NRF Centre of Excellence for Biomedical TB Research, School of Pathology, Faculty of Health Sciences, University of the Witwatersrand and the National Health Laboratory Service, PO Box 1038, Johannesburg 2000, South Africa. Tel: +27-11-489 9030; Fax: +27-11-489-9397; E-mail: bavesh.kana@nhls.ac.za

One sentence summary: This review provides a synthesis on the prevalence of mixed tuberculosis infection, which can provide insight on transmission rates, strain diversity and drug tolerance and resistance within select high and low burden settings.

Editor: Patrick Brennan

ABSTRACT

Tuberculosis (TB) disease can be characterized by genotypic and phenotypic complexity in *Mycobacterium tuberculosis* bacilli within a single patient. This microbiological heterogeneity has become an area of intense study due its perceived importance in drug tolerance, drug resistance and as a surrogate measure of transmission rates. This review presents a descriptive analysis of research describing the prevalence of mixed-strain TB infections in geographically distinct locations. Despite significant variation in disease burden and a rampant human immunodeficiency virus (HIV)-TB co-epidemic, there was no difference in the prevalence range of mixed infections reported in African countries when compared to the rest of the world. The occurrence of recurrent TB was associated with a higher prevalence of mixed-strain infections, but this difference was not reported as statistically significant. These interpretations were limited by differences in the design and overall size of the studies assessed. Factors such as sputum quality, culture media, number of repeated culture steps, molecular typing methods and HIV-infection status can affect the detection of mixed-strain infection. It is recommended that future clinical studies should focus on settings with varying TB burdens, with a common sample processing protocol to gain further insight into these phenomena and develop novel transmission blocking strategies.

Keywords: tuberculosis; recurrent disease; relapse; re-infection; mixed infections; heteroresistance

INTRODUCTION

Efforts to contain tuberculosis (TB) with combination chemotherapy for the past 50 years have met with mixed success as this disease remains a global public health concern and is now the leading cause of death due to an infectious bacterial agent (WHO 2016a). Current TB treatment is characterized by a variety of disease outcomes including clinical cure, treatment failure and the development of latent infection, with the latter two carrying the risk of subsequent relapse disease. Chemotherapy of TB is complex and while ultimately aimed at eliminating tubercle bacilli, treatment can also result in the selection of various forms of antimicrobial resistance. This can lead to emergence of one dominant resistant bacterial genotype or the appearance of heteroresistance, comprising mixtures of drug-susceptible and drug-resistant organisms.

Received: 7 December 2016; **Accepted:** 16 February 2017

This complexity in bacterial populations, which can occur in a single patient, presents unique challenges and hampers the rapid clearance of organisms during chemotherapy. In addition, differences in bacterial growth states can further contribute to the establishment of drug-tolerant populations as chemotherapeutic agents are generally targeted at metabolic pathways that are a hallmark of active bacterial replication (Zumla, Nahid and Cole 2013; Kana *et al.* 2014). This is evidenced by the demonstration that sputum from TB patients contains a mixture of actively replicating and differentially culturable tubercle bacteria (DCTB), the latter displaying tolerance to first-line TB chemotherapeutics (Mukamolova *et al.* 2010; Chengalroyen *et al.* 2016; Loraine *et al.* 2016). It is proposed that these drug-tolerant populations provide the framework for the selection and emergence of genetically stable drug-resistant strains (Loraine *et al.* 2016). In view of this, the genomic plasticity and resulting strain complexity of *Mycobacterium tuberculosis* populations during infection is a key feature of the success of this deadly pathogen and merits further consideration. As a result, the complexity in *M. tuberculosis* population dynamics in human populations has become the focus of intense study. The advent of molecular typing methods and next-generation sequencing technologies has allowed for the genetic characterization of *M. tuberculosis* strains at a resolution that was previously not possible (Gan *et al.* 2016; Dheda *et al.* 2017).

DNA fingerprinting technologies, developed over the past two decades, have been widely used to study the diversity of *M. tuberculosis* strains and to investigate episodes of exogenous re-infection and relapse (van Rie *et al.* 1999; Bandera *et al.* 2001; Caminero *et al.* 2001), as well as the presence of heteroresistance in sputum isolates (Kaplan *et al.* 2003; Post *et al.* 2004; Shamputa *et al.* 2004), and the occurrence of mixed *M. tuberculosis* strain infections (Warren *et al.* 2004; van Rie *et al.* 2005; Shamputa *et al.* 2006; Cohen *et al.* 2011). Prior to these advances, TB disease was assumed to occur as a consequence of infection with a single *M. tuberculosis* strain, this event was thought to confer a measure of protection against infection with a secondary strain (Stead 1967). Recurrence of the disease was understood to be the result of endogenous re-activation of a non-replicating, immune-subversive variant of the strain that was responsible for the original disease episode (Stead 1967). In this context, the importance of re-infection remained relatively unexplored. However, molecular-based genotyping has subsequently demonstrated that exogenous re-infection with different *M. tuberculosis* strains does indeed occur in both high- and low-incidence settings, thus confirming that previous infection does not provide protection against subsequent infection (Small *et al.* 1993; van Rie *et al.* 1999; Sonnenberg *et al.* 2001; Kruuner *et al.* 2002; Andrews *et al.* 2008; Charalambous *et al.* 2008). Furthermore, studies have shown that many TB cases occur as a result of recent transmission and depending on setting; these transmission events are likely to take place outside of the household within the community (Verver *et al.* 2004; Dheda *et al.* 2017). High rates of re-infection have important implications for TB control strategies and highlight the need to reduce TB transmission in the community, together with the importance of early diagnosis, coupled with treatment initiation. TB disease due to recent transmission and re-infection has become an increasingly relevant topic of study in the era of human immunodeficiency virus (HIV) infection, as the resulting reduced immunity may result in vulnerable populations becoming increasingly susceptible to infection or re-infection with *M. tuberculosis*. Moreover, increased transmission in communities with a diversity of circulating *M. tuberculosis* strains can lead to mixed-strain infections, which have been documented in many distinct geographical locations, emphasizing the importance of interventions that limit or interrupt transmission. The complex environment of the lung as a site for TB disease, combined with sustained exposure to tubercle bacilli in high-TB-endemic settings, enhances the occurrence of mixed-strain infections, which have been associated with unfavorable treatment outcomes (Theisen *et al.* 1995; Niemann *et al.* 2000; Baldeviano-Vidalon *et al.* 2005; van Rie *et al.* 2005; Kamakoli *et al.* 2017). Given the emerging importance of phenotypic diversity in bacterial populations during TB pathogenesis, this review takes a retrospective view of published studies that report the prevalence of mixed-strain infections within select populations. The prevalence of these phenomena was interrogated amongst studies conducted within and outside the African continent to identify patient and population-level risk factors associated with mixed-strain TB infections, as well as possible avenues for future research. Recurrent TB and factors associated with relapse versus re-infection are discussed, together with any association with mixed-strain infections. For clarity, the terminology and methods used to describe and study these clinically complex presentations of TB disease are detailed below.

DEFINITIONS AND TERMINOLOGY

Definitions of terms and descriptions of features relating to TB disease, pathogenesis and epidemiology that are commonly encountered in the literature and feature in this review are given below. These are further clarified in Fig. 1.

Primary TB disease

Primary TB, Fig. 1, is defined as disease in a patient who has never been treated for TB or has taken anti-TB medication for less than a month (WHO 2014).

Endogenous re-activation

Endogenous reactivation refers to incident cases of TB (both new and recurrent) that occur as a result of re-activation/re-emergence of a previous infection that was contained by the host immune response, in the case of individuals with latent TB infection, and/or the application of chemotherapy in patients with active TB disease, Fig. 1. This is in contrast to TB disease (both new and recurrent) that is a result of a recent transmission event (Mathema *et al.* 2006).

Recurrent TB disease

Recurrent TB refers to a repeat occurrence (second, third or subsequent episode) of TB disease in a patient that occurs as a result of either relapse or re-infection. Recurrent TB occurs after the previous/initial episode has been classified as clinically cured according to WHO guidelines [smear or culture-negative sputum specimens in the last month of treatment and on at least one previous occasion, (WHO 2016a)].

Re-infection

Recurrent TB disease also occurs as a result of re-infection, whereby a patient is exogenously infected with a *Mycobacterium tuberculosis* strain that is distinct from the organism that caused the original infection, Fig. 1B (Mathema *et al.* 2006). A caveat here is that in high-incidence settings, patients, on rare occasions,

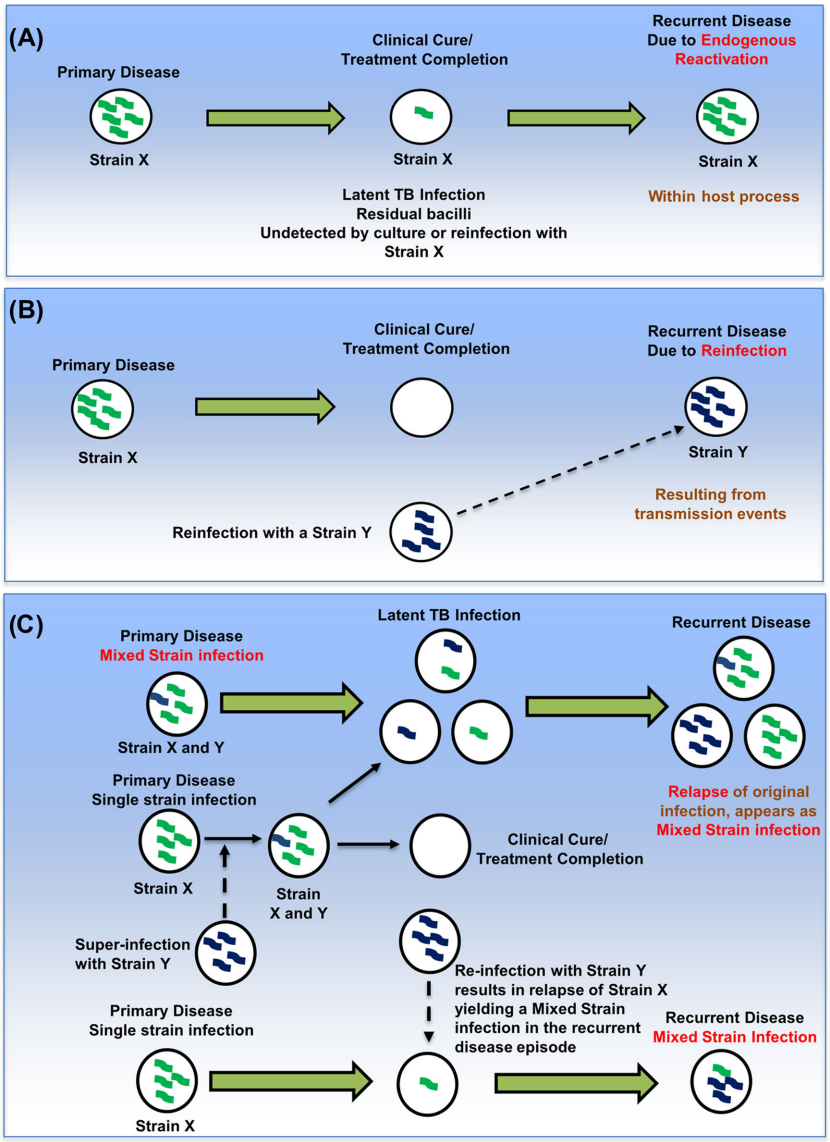

Figure 1. Recurrent disease and mixed-strain *M. tuberculosis* infections. Shown are the various scenarios that give rise to recurrent TB disease and mixed infections. (**A**) Relapse disease occurs due to the re-emergence of an *M. tuberculosis* strain that caused the original infection, suggesting that complete eradication of tubercle bacteria was not achieved during the primary disease episode. In this case, the second disease episode results from the presence of persisting organisms that emerge when treatment is stopped and environmental factors are favorable for bacillary survival. (**B**) Re-infection occurs when a patient is infected with an *M. tuberculosis* strain that is distinct from the strain that caused the original infection. High rates of re-infection can be attributed to both environmental and host factors. Alternatively, an individual can be functionally cured, with no surviving bacteria but can then be re-infected with the same strain that caused the primary disease episode. In the absence of molecular tools to distinguish these events, this would result in the incorrect classification of re-activation, see (**A**). (**C**) Mixed infections can occur in both primary TB disease and recurrent TB disease. During primary disease, mixed infections can occur during a single infectious episode, whereby two genetically distinct strains are present at the same time during the initial infectious episode. Alternatively, infection with a second distinct strain (super infection) can occur during latent infection or active disease. Re-infection with a second strain may also result in re-activation of an underlying strain.

may be exposed to or be infected by a very similar or the same strain as in the primary infection, which makes differentiation between relapse and re-infection in these particular cases difficult.

Relapse

Relapse disease is defined as a second (or third) episode of active TB disease due to re-emergence of the original infection, as determined by genotypic analysis of the prevailing tubercle bacilli (Mathema *et al.* 2006). As indicated for re-infection, substantive homogeneity in *M. tuberculosis* strains in any given setting, combined with high transmission rates, will make the classification relapse versus re-infection difficult. In these cases, whole-genome sequencing (WGS) to identify minor differences will provide the greatest insight (Gan *et al.* 2016).

Mixed-strain TB infection

This refers to TB disease caused by more than one clonally distinct *M. tuberculosis* strain, either through a single transmission event involving more than one distinct strain or through multiple transmission events (super-infection) during a single-

disease episode, Fig. 1C. The simultaneous transmission of multiple strains resulting in mixed infection may occur in populations of vulnerable individuals, whereby both strains are able to bypass the host's defense system and resist killing (Warren *et al.* 2004). Super-infection can also occur when the severity of a current disease episode is such that it compromises the host innate immune response to a point that leads to increased susceptibility to infection with a secondary strain (Warren *et al.* 2004). Alternatively, mixed-strain infections can arise if a subsequent infectious episode, which was caused by a distinct strain, results in relapse of the original infection, yielding disease with two unique *M. tuberculosis* strains that may have the same or different drug susceptibility profiles.

Heteroresistance

The term heteroresistance refers to the occurrence of populations of both drug-susceptible and drug-resistant isolates within the same clinical sample (Rinder 2001). This phenomenon may arise in a single infection as tubercle bacilli undergo genetic change through mutation of genes associated with drug resistance. It may also arise in mixed-strain infections, whereby one strain is resistant to a particular TB drug, while the other strain is susceptible. When heteroresistance occurs within the same strain type, it involves sub-populations of *M. tuberculosis*, for example resistant mutants co-existing with susceptible wild-type genotypes of specific resistance-related genes such as *katG or rpoB* in the same sputum specimen. This phenomenon is known as clonal heteroresistance, which is dependent solely on within-host processes affecting a single causative strain, in comparison to a mixed infection that is reliant on the presence of more than one strain.

METHODOLOGY USED TO STUDY RECURRENT TB AND STRAIN HETEROGENEITY

Numerous approaches have been employed to describe the microbiological complexity associated with TB disease, further detailed below. While some methods rely on variations in the biochemical properties of tubercle bacilli, the most popular methods used for the identification of mixed-strain infections include IS*6110*-based restriction fragment length polymorphism (IS*6110* RFLP) analysis, spacer oligonucleotide genotyping (spoligotyping) and the mycobacterial interspersed repetitive-unit-variable-number of tandem repeats (MIRU-VNTR) typing. In more recent work, metagenomics (Kay *et al.* 2015) and WGS approaches have been used (Ford *et al.* 2012; Gan *et al.* 2016).

Phage typing of *Mycobacterium tuberculosis* strains

Mixed-strain infections were first described in the 1970s using phage typing of *M. tuberculosis* cultures (Mankiewicz and Liivak 1975; Bates, Stead and Rado 1976). This method demonstrated that distinct strains of *M. tuberculosis* could be isolated from a single sputum sample by the typing of multiple colonies on solid culture media (Mankiewicz and Liivak 1975).

IS6110 RFLP typing

IS*6110* RFLP typing is an attractive tool to study transmission dynamics and relatedness of *M. tuberculosis* isolates (Cave *et al.* 1994; Warren *et al.* 2004). Strain identification using this method relies on the number and position of the transposable insertion sequence element, IS*6110*, in *M. tuberculosis* strains after restriction digestion, yielding a unique fingerprint pattern (van Embden *et al.* 1993). In studies investigating recurrent TB disease, strains exhibiting an identical banding pattern to the primary infectious episode point to relapse of TB disease, whereas re-infection and recent transmission is implied when distinct banding patterns are obtained. IS*6110* RFLP technology has also been utilized to identify mixed infections either by the identification of low-intensity banding patterns (LIBs) present in the same sputum sample (de Boer *et al.* 2000) or by the identification of unique fingerprints from the selection of multiple single colonies from the original sample (de Viedma *et al.* 2003; Das *et al.* 2004; Shamputa *et al.* 2004). One limitation of this method is that the identification of LIBs requires that at least 10% of the extracted DNA originates from the minority strain in order to be detected (Cohen *et al.* 2012). An additional limitation is that it fails to reliably discriminate between strains with less than six IS*6110* copies i.e. low-copy-number isolates (Gutacker *et al.* 2006). Some of these strains, displaying identical banding patterns on IS*6110* typing, have been shown to have unique genetic identities when using a secondary genotyping method (Bauer *et al.* 1999; Rhee *et al.* 2000).

Spoligotyping

Spoligotyping is the most commonly used PCR-based method for genotyping and differentiating between *M. tuberculosis* complex (MTBC) strains and is based on the visualization of 43 interspersed spacer sequences in a genomic direct repeat locus of MTBC strains (Kamerbeek *et al.* 1997). Although spoligotyping is a popular and simple molecular tool used in the study of molecular epidemiology, it has lower discriminatory power when compared to IS*6110* RFLP typing and is generally not recommended for use on its own in the study of mixed infections or transmission patterns (Kremer *et al.* 1999).

MIRU-VNTR typing

This method uses PCR to categorize the number and sizes of repeats in a minimum of 12 independent loci, each of which have a unique repeat sequence (Barnes and Cave 2003). As this method examines heterogeneity within a limited set of loci, its discriminatory power is linked to the number of loci that are evaluated. An advantage of MIRU-VNTR typing is that it can identify strains that are under-represented in a mixed-strain infection. Thus, a minority strain representing as little as 1% of the entire bacterial population can be identified using this technique (de Viedma *et al.* 2005). One limitation of this method is that it may present difficultly in differentiating between mixed infection and microevolution of strains (Cohen *et al.* 2012).

Whole-genome sequencing

WGS has become an established method in the field of molecular epidemiology. The application of this technology provides the capability of in-depth analysis into the diversity of *M. tuberculosis* strains that cannot be achieved by other currently used molecular methods. However, its use has been hampered by high cost and the complexity of data analysis but it is likely that the application of WGS will increase as new-generation sequencing becomes more affordable and simpler data analysis tools become available.

APPROACH FOR RETROSPECTIVE ANALYSIS

To retrospectively assess the prevalence of mixed-strain infections, the published literature was interrogated for reports that described these and related occurrences. Publications selected for analysis were obtained through the Pubmed database (www.ncbi.nlm.gov/pubmed) using the following keywords and Boolean operators: (mixed OR multiple) AND tuberculosis AND (strain OR strains), which yielded 1749 hits. Relevant English articles published from the beginning of records till December 2016 that reported the occurrence of mixed-strain infections using molecular genotyping techniques in adult TB patients were chosen. These articles were selected using several criteria which included, geographic location, patient population, recurrent TB and where possible, congruent processing protocols. We also endeavored to differentiate our analysis to a previous review published 4 years ago (Cohen *et al.* 2012), which affected our inclusion/exclusion of studies. Studies with a starting sample size of less than 10 patients were excluded from analysis. In addition, relevant articles that were referenced by papers obtained through the original search were also included for analysis. The analysis of mixed infections was further stratified by primary and recurrent disease. In some instances, where not reported, the proportion of mixed infections was calculated in each cohort based on the information available. The resulting synthesis of research on mixed *Mycobacterium tuberculosis* infections worldwide, spanning over the last 15–17 years is provided in this review.

MIXED *MYCOBACTERIUM TUBERCULOSIS* INFECTIONS

Studies of TB cases in deceased patients from the 18th century, while limited, revealed that in five out of eight bodies investigated using metagenomics, more than one *Mycobacterium tuberculosis* genotype was identified, suggesting that mixed-strain infection may have been a common occurrence during a time of rampant TB in Europe (Kay *et al.* 2015). Current studies investigating the frequency of present day mixed *M. tuberculosis* infections confirm that they can occur regularly but the frequency of detection varies according to the design of the studies, the sample size and methodology used.

African versus non-African studies

For the greatest coverage possible, geographically distinct regions covering Africa and countries outside Africa, respectively, were included in this review and the findings summarized in Tables 1 and 2. The rationale for this stratification was based on the observation that countries in Africa (especially southern Africa) carry a high burden of HIV-associated TB (WHO 2016a,b). We hypothesized that the compromised immune response in individuals associated with these dual epidemics, together with a high incidence of TB and elevated transmission rates, would give rise to a greater prevalence of mixed infections on the African continent. The reported TB incidence figures per 100 000 population in the select countries used for this analysis were South Africa, 834; Rwanda, 63; Zambia, 406; Malawi, 164; Uganda, 161. In similar studies from outside of Africa, the TB incidence figures per 100 000 population were Spain, 12; India, 167; Bangladesh, 227; Georgia, 106; China, 68; Kyrgyzstan, 142; Vietnam, 140; Guyana, 93; Surinam, 33 and Pakistan, 270 (WHO 2016b). Amongst studies from African countries, 15 that investigated the occurrence of mixed infections in patients with pulmonary TB were included for analysis. These studies were selected on the basis of the methodology employed as well as the number of patient samples analyzed. Of the selected African studies, nine were conducted in South Africa, and the remaining six were carried out in Rwanda, Zambia, Malawi (one study from each country) and Uganda (three studies). The frequency of mixed infections observed in sputum samples from African studies varied between 2.3% and 19% (Table 1). Frequencies of mixed infections reported in 19 studies performed in countries outside of Africa were similar to those in Africa and varied from <0.4% to 15.7% (Table 2). Given the substantive variation in overall study design, it is unclear whether the difference in the upper limit of mixed infection prevalence, 14%–19%, can be considered as significant. Possible reasons for the large variation in the occurrence of mixed infection in any particular setting or between the select studies are outlined below.

Effect of quality, volume and processing of specimens

A major contribution to variations between studies in similar settings is likely to be sputum collection and processing. The collection of a single, low-volume sputum sample may greatly decrease the odds of finding mixed infection episodes. For example, a poor-quality sputum sample containing a significant amount of upper airway secretions is unlikely to be representative of the entire *M. tuberculosis* population in the bronchial tree, while the collection of multiple sputum samples may increase the odds of sampling from different cavities within the lung. Several studies reviewed here, relied on the collection of sputum cultures from multiple pre-treatment specimens to increase the sensitivity of their assay (Richardson *et al.* 2002a; Shamputa *et al.* 2006; Peng *et al.* 2013). Similarly, for the demonstration of mixed infections, typing of cultures from multiple sputum samples from smear-positive individuals, collected serially over time, was used (Braden *et al.* 2001; Baldeviano-Vidalon *et al.* 2005; van Rie *et al.* 2005; Dickman *et al.* 2010; Mallard *et al.* 2010; Mulenga *et al.* 2010; Cohen *et al.* 2011). Considering this, an in-depth comparison of findings between these studies was not possible due to differences in methodology and study design. However, intra-study comparisons of the frequency of mixed infections have concluded that the analysis of a single sputum sample, as opposed to multiple samples, decreases the likelihood of detecting mixed infections (Shamputa *et al.* 2006; Peng *et al.* 2013).

Effect of specimen processing

We hypothesize that additional factors that could affect the detection of mixed-strain infections include delays in transportation, breaks in cold-chain transport and delays in sputum processing. Furthermore, it is not known whether the decontamination procedure may result in a loss of bacterial populations. To our knowledge, no research has focused on the impact of these procedures on detecting mixed strains, and further research is required to corroborate these hypotheses. Findings of studies on mixed infections are also not directly comparable as the population structure of mixed infection is strain genotype and culture medium dependent (Hanekom *et al.* 2013). Moreover, the clonal composition of cultures from clinical specimens could be modified by the culture techniques used (Martín *et al.* 2010). Hanekom *et al.* (2013) used a PCR-based molecular typing approach to identify strain genotypes in both Löwenstein-Jensen (LJ) and mycobacterial growth indicator tube (MGIT) cultures. The investigators noted that Beijing and Haarlem strain

Table 1. Studies conducted on mixed tuberculosis infection in Africa.

Year	2001	2002a	2004	2005	2007	2008	2009	2010	2010	2010	2011	2013	2013	2014	2015
Author	du Plessis *et al.*	Richardson *et al.*	Warren *et al.*	Van Rie *et al.*	Umubyeyi *et al.*	Andrews *et al.*	Stavrum *et al.*	Dickman *et al.*	Mallard *et al.*	Mulenga *et al.*	Cohen *et al.*	Hanekom *et al.*	Muwonge *et al.*	Shin *et al.*	Ssengooba *et al.*
Location	Western Cape, RSA	Cape Town, RSA	Cape Town, RSA	Cape Town, RSA	Rwanda	KwaZulu-Natal, RSA	Eight provinces in RSA	Kampala, Uganda	Karonga District, Malawi	Ndola, Zambia	KwaZulu-Natal, RSA	Khayelitsha, RSA	Uganda	Botswana	Kampala, Uganda
Strain typing method used	IS*6110* RFLP DR Sequence	IS*6110* RFLP DR Sequence	PCR (Beijing and non-Beijing) Spoligo	IS*6110* RFLP PCR	Spoligo MIRU-VNTR	Spoligo	Spoligo MIRU-VNTR IS*6110* RFLP	MIRU-VNTR Spoligo	PCR (LAM and Beijing)	Spoligo MIRU-VNTR	MIRU-VNTR Spoligo	PCR (LAM, Haarlem, S, Beijing, LCC)	MIRU-VNTR Spoligo	MIRU-VNTR	Spoligo MIRU-VNTR
Number of enrolled	17	210	407	768	710 (69 MDR)	23[c]	252 isolates	113	72	361	240	535	74	370	66
Number of patients/ samples analyzed	13[a]	131	186	48 MDR-TB	22 MDR 1 Non-MDR	17	54 isolates	113	72	273	56	206 cultures[f]	72	370	51
First incidence of TB	No	210	200	23	10	14	141	NS	62	Yes	NA	NA	49	142	47
Retreatment records	Yes	NA	207	25	12	3	102	3	10	Yes	NA	NA	23	228	4
HIV status	All negative	All negative	10% HIV+	2% HIV+, 47% HIV– 51% NA	54.5% HIV +	15 + 2 NA	44% HIV + 22% HIV– 34% NA	14.4% HIV+	62.7% HIV+	NA	96% HIV+	NA	24.7% HIV+ 17% HIV– 59% NA	75.4% HIV+	All positive
Samples															
Single	No	No	Yes	No	47	No	NS	No	25	Yes	No	Yes	Yes	Yes	No
Multiple	Yes[b]	Yes	No	Yes	22	Yes	NS	Yes	47	Yes	Pooled[e]	No	No	No	Yes[g]
AFB +ve	Yes	Yes	Yes	NA	Yes	NA	NS	Yes	Yes	Yes	No	NS	NS	NA	Yes/No
Sputum	Yes/No	Yes	Yes	Yes	Yes	Yes	NS	Yes	Yes	Yes	No	Yes	Yes/No	Yes	Yes/No
Culture step	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	Yes	Yes	Yes	Yes	Yes	Yes
Collected before treatment	Yes (4); No (9)	Yes	Yes	Yes/No	Yes/No	Yes/No	Yes/No	Yes/No	NS	Yes/No	Yes (42%) No (58%)	NA	Yes/No	Yes/No	Yes/No

Table 1. (*Continued*).

Year	2001	2002a	2004	2005	2007	2008	2009	2010	2010	2010	2011	2013	2013	2014	2015
Percentage of mixed infections in:															
Primary TB disease	NA	2.3	17	NS	NS	NS	NS	NS	2.8	NS	NA	NA	12	5.6	53
Recurrent TB disease	15.4	NA	23	NS	NS	NS	NS	NS	0	NS	NA	NA	4.3	12.7	25
Total	IS*6110*: **15.4** DR: **7.7**	**2.3**	PCR:**19** Spoligo: **4.8**	PCR: **10.4** IS*6110*: **2**	**4.3**	**11.7 (29.4)**[d]	**18.5**	**7.1** (HIV +: **37.5**)	**2.8**	**3.2**	**9**	**15**	**11.1**	**10**	**4** (blood and sputum: **51**)

[a]Twelve autopsied cases and one pneumonectomy, both lung and extra-pulmonary samples were collected.
[b]Sampled from 6 to 15 lesions from anatomically distinct sites.
[c]Patients who developed multi drug-resistant (MDR) or extremely drug-resistant (XDR-TB) after treated for less resistant form.
[d]Two patients infected with multiple genotypes at initiation (11.7%). During follow-up, an additional three patients had two strains with different DST patterns in a single episode (29.4%).
[e]Pooled biopsy specimens from lung, liver and spleen.
[f]Convenience sample consisting of paired MGIT and LJ cultures.
[g]Two sputum and one blood sample collected at baseline for culture.
AFB: acid fast bacilli
DR: direct repeat sequence
IS*6110* restriction fragment length polymorphism analysis
LCC: low copy clade
MDR: multidrug-resistant tuberculosis
MIRU-VNTR: mycobacterial interspersed repetitive-unit-variable-number of tandem repeats
NA: not available
NS: not specified
RSA: Republic of South Africa
Spoligo: spacer oligonucleotide genotyping
The boldface is the total percentage of mixed-strain infections reported in each study.

Table 2. Large studies conducted on mixed tuberculosis infections in countries outside of Africa.

Year	1999	1999	2003	2004	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2015	2015	2015	2015	2016
Author	Chaves *et al.*	Yeh, Hopewell and Daley	de Viedma *et al.*	Das *et al.*	Shamputa *et al.*	de Viedma *et al.*	Sham-puta *et al.*	Cox *et al.*	Fang *et al.*	Mokrousov *et al.*	Huang *et al.*	Wang *et al.*	Huyen *et al.*	Peng *et al.*	Wang *et al.*	Pang *et al.*	Streit, Millet and Rastogi	Zheng *et al.*	Mustafa *et al.*
Location	Madrid, Spain (TB prison hospital)	United States	Madrid, Spain	Thirivallur, India	Bangladesh (Hospitals in Mymensingh)	Madrid, Spain	Georgia (TB prison hospital)	Central Asia	Shanghai, China	Kyrgyzstan	Eastern Taiwan	Taiwan	Southern Vietnam	Heilon-gjiang, China	Inner Mongolia, China	China	Guyana and Suriname	Sichuan Province, China	Punjab District Pakistan
Typing method	IS6110 RFLP	IS6110 RFLP	Spoligo DRE-PCR IS6110 RFLP	IS6110 RFLP DR-RFLP	IS6110 RFLP Spoligo MIRU-VNTR	IS6110 RFLP Spoligo MIRU-VNTR	IS6110 RFLP MIRU-VNTR	IS6110 RFLP Spoligo	MIRU-VNTR	IS6110 inverse PCR MIRU-VNTR, spoligo	PCR (Beijing and non-Beijing)	PCR (Beijing and non-Beijing)	IS6110 RFLP Spoligo MIRU-VNTR	MIRU-VNTR	MIRU-VNTR	MIRU-VNTR	Spoligo MIRU-VNTR	Spoligo MIRU-VNTR	PCR (Beijing and non-Beijing)
Number of patients enrolled	226	49	123	543	132	NS	385	416 strains	561	56	185	868	1890	174	384	3929 strains	161 isolate[c]	5090	102[d]
Number of patients/ samples analyzed	37	49	50	543	97	115 Cultures	199	397 isolates	249	56	185	466	1248	89	384	3248 strains	154	499 isolates	95 MDR-TB
First incidence of TB	25	NS	NA	NS	1	NA	134	198	217	56	144	436	1107	NA	325	2519	NA	NA	44
Retreat-ment records	12	NS	NA	NS	131	NA	65	184	249	NA	41	30	139	NA	59	729	NA	NA	51
HIV status	All positive	NS	84% positive	NA	NA	NA	NA	NA	NA	All negative	1 HIV + patient identified	NA	NA	NA	NA	NA	26.6% positive (not all tested)	NA	NA
Samples																			
Single	No	No	No	No	Yes	NS	No	NS	Yes	Yes	Yes	Yes	Yes	No	Yes	No	NS	Yes	Yes
Multiple	Yes	Yes	Yes[a]	Yes	No	NS	Yes	NS	Yes	No	No	No	No	Yes	No	Yes	NS	No	No
AFB +ve	NS	NS	NA	NA	Yes	Yes	Yes	NS	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	NS	NA	Yes
Sputum	Yes/No	Yes	Yes/No	Yes	Yes	Yes	Yes	Yes	isolates	Yes	Yes	Yes (292)	Yes	Yes	NS	Yes	NS	NS	Yes
Culture step	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	No	No	Yes (174)	Yes	Yes	Yes	Yes	Yes	Yes	Yes

Table 2. (*Continued*).

Year	1999	1999	2003	2004	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2015	2015	2015	2015	2016
Collected before TX	NS	Yes/No	NS	Yes/No	Yes	NS	Yes	Yes/No	Yes (217)	NS	Yes (144)	Yes	NS	NS	NS	Yes	NS	NS	NS
DST	S,I,R,E	NA	S,I,R,E	NA	Yes	NA	S,I,R,E	S,I,R,E, PZA	Yes	I, R	S,I,R,E	Yes	S,I,R,E	Yes	S,I,R,E,O,K	NA	Yes	S,I,R,E	S,I,R,E
# Colonies selected	0	0	10	15–20[b]	10	Yes	0	0	30	0	0	0	0	0	30[b]	0	0	0	0
Percentage of mixed infections in:																			
Primary disease	NA	NA	NA	NA	NA	NA	15.7	NA	4.1	13.7	10.4	2.8	3.4	NA	2.8	3.3	NA	NA	27.2
Recurrent disease	NA	NA	NA	NA	2.1	NA	7.7	NA	15.6	NA	14.6	6.6	0.7	NA	5.1	4.4	NA	NA	5.9
Total	**3.6**	**2**	**6**	**<0.4**	**2.1**	VNTR: **2.6**	IS6110: **9** VNTR:**13.1**	**5**	**5.6**	**13.7**	**11.3**	**3%**	RFLP/ spol: **3.1** VNTR: **4.8**	**11.2%**	**3.13**	**3.5**	**0.6**	**2.4**	**15.7**

[a]One respiratory and one extra-respiratory sample analyzed less than 30 days apart.
[b]Monoclones were subsequently isolated and tested after mixed infections were identified using initial typing methods.
[c]Seventy-four isolates from Guyana and 80 from Suriname.
[d]Patients suspected of MDR-TB.
AFB: acid fast bacilli
DST: drug susceptibility testing
E: Ethambutol
IS6110 RFLP: IS6110 restriction fragment length polymorphism analysis
I: isoniazid
K: kanamycin
MDR: multidrug-resistant tuberculosis
MIRU-VNTR: mycobacterial interspersed repetitive-unit-variable-number of tandem repeats
NA: not available or not applicable
NS: not specified
O: Ofloxacin
PZA: Pyrazinamide
R: Rifampicin
Spoligo: spacer oligonucleotide genotyping
S: streptomycin
The boldface is the total percentage of mixed-strain infections reported in each study.

families were more likely to be associated with mixed infections and showed a predilection for some culture media (Beijing MGIT, $P = 0.02$, LJ $P < 0.01$; Haarlem MGIT $P < 0.01$, LJ $P = 0.01$). Other strains investigated included the Latin American Mediterranean (LAM) and S-family genotypes (Hanekom *et al.* 2013). In addition, mixed infections were identified in 23 of 206 MGIT cultures and 28 of 206 LJ cultures, indicating that LJ media may be marginally preferable to liquid media in identifying mixed infections (Hanekom *et al.* 2013). In a separate study, the clonal complexity of *M. tuberculosis* populations from clinical specimens was investigated by mixing strains in different volumetric proportions before culturing was performed. Using MIRU-VNTR typing, marked changes in the clonal composition were observed, including strains with different resistance profiles. The authors concluded that clonal complexity can be lost after culture, therefore genotyping techniques should be optimized in such a way that they can be performed directly on clinical samples (Martín *et al.* 2010). Four studies included herein were performed directly on sputum samples of which three used a PCR amplification method to detect different lineages (Mokrousov *et al.* 2009; Huang *et al.* 2010; Mallard *et al.* 2010; Wang *et al.* 2011). PCR methods to detect Beijing and non-Beijing lineages identified mixed infection frequencies to occur at 3% and 11% (Huang *et al.* 2010; Wang *et al.* 2011) while PCR to detect LAM and Beijing lineages reported mixed infection frequencies of 2.8% (Mallard *et al.* 2010). In the remaining study, 13.7% was identified by 12 locus MIRU-VNTR typing (Mokrousov *et al.* 2009). Another study investigated the proportion of mixed infections in both sputum and liquid culture-positive medium and reported a frequency of 2.1% and 4.6%, respectively (Wang *et al.* 2011). In a subsequent study, Wang *et al.* (2015) further sub-cultured the mixed strains initially observed and selected 30 clones from the sub-culture for each mixed infection. The authors reported that after the addition of a culture step, 8 of 12 (66.7%) strains with mixed infections had converted into a single genotype (Wang *et al.* 2015). Mixed-strain infections involving heteroresistance may also be missed if sputum samples are collected from a patient who has started treatment at a time when resistance is being amplified by first-line therapy (Hingley-Wilson *et al.* 2013). Pre-treatment samples may therefore be preferable for studies trying to identify mixed infections/clonal heteroresistance within an infecting bacterial population.

Techniques used for demonstrating mixed infections

The prevalence of mixed-strain infections varied between <0.4% using IS*6110* genotyping on colonies from 543 patients in India (Das *et al.* 2004), to 15.4% using the same technique on 12 HIV-negative autopsy cases in the Western Cape (du Plessis *et al.* 2001). In studies employing MIRU-VNTR typing, mixed infections were detected at an average frequency of 5.6% in Shanghai, China, with re-treatment cases having a higher rate of mixed infections than new TB cases (15.6% versus 4.1%) (Fang *et al.* 2008). Compared to these figures, mixed infections were encountered in 10% of all TB cases in Botswana (Shin *et al.* 2014). Only one study used spoligotyping as its primary method of genotyping and reported mixed infections to occur in 18% of samples analyzed, which is high in comparison to rates reported in other studies using other molecular methods (Andrews *et al.* 2008).

Frequencies of certain *Mycobacterium tuberculosis* genotypes involved in mixed infections

In the African studies, the incidence of mixed infections varied from 2.3% in diagnostic sputum cultures in patients from Cape Town, South Africa (Richardson *et al.* 2002a) to 19% in sputum cultures obtained from other patients suspected of suffering from TB from the same region (Warren *et al.* 2004). The highest frequency of mixed infections in sputum samples was obtained using a novel PCR technique demonstrating that patients analyzed in the study were simultaneously infected with a Beijing and non-Beijing *M. tuberculosis* strains (Warren *et al.* 2004). These results probably underestimate the frequency of mixed infections in this population as the PCR was conducted with primers specific to the Beijing lineage, and mixed infections may have also been present within the non-Beijing lineages. However, this method was still superior to spoligotyping, which identified only 4.5% of mixed cases in the same population (Warren *et al.* 2004). In another study, a PCR-based approach was used to determine the occurrence of mixed infections in sputum samples taken from patients in Malawi (Mallard *et al.* 2010). The investigators applied two lineage-specific PCR assays to sputum samples, targeting the LAM and non-LAM genotypes, as LAM strains are the most common genotypes present in northern Malawi. This yielded mixed infections in 2.8% of sputum samples, significantly lower than the 19% mixed infection rates observed in the Cape Town study that investigated the occurrence of Beijing and non-Beijing strains (Warren *et al.* 2004; Mallard *et al.* 2010). It was proposed that the differences found between the two studies could be explained by differences in the annual risk of infection [3% in Cape Town (Kritzinger *et al.* 2009) versus 1% in the rural Karongo district in Malawi (Crampin, Glynn and Fine 2009)]. Alternatively, inherent properties such as transmissibility of the Beijing versus LAM genotypes could explain the discrepancy. In addition, there were differences in the methods employed in the two studies.

Middelkoop *et al.* (2014) investigated factors associated with *M. tuberculosis* strain success in a high-TB-burdened community, with high strain diversity. Using IS*6110* RFLP, the authors noted that only 4% of strain types were persistently successful over a 10-year study period. These were strains from the W-Beijing family (W451 W724, W181, W330) and four from the CC-related lineage (CC27, CC61, CC73, CC77), as well as NG24 and AI265 strains. As host factors examined, such as age and gender, were not associated with strain success, it was concluded that pathogen rather than host characteristics appear to play a greater role in strain success (Middelkoop *et al.* 2014).

Effect of disease burden on prevalence of mixed infections

A high number of incident cases should logically influence the frequency of mixed infections, as was the case with the increased prevalence of 19% of mixed infections identified in the Cape Town study by Warren *et al.* (2004). However, Richardson *et al.* (2002a) who conducted their study in the same setting but found the frequency of mixed infections to be markedly lower, at 2.3%. This result was similar to the mixed infection rate seen in Bangladesh (2.1%) and Spain (2.6%). As mentioned previously, detailed comparisons between these studies are not scientifically valid due to differences in study design and methodology; however, the findings do suggest that overall disease burden is not the only arbiter of strain diversity or mixed infections. Alternatively, the differences in methods used in these studies may have resulted in an underestimate of mixed infections, as studies that used IS*6110*-based RFLP as their primary typing technique tended to report a low incidence of mixed infections, Tables 1 and 2. In contrast, when used in a congregate setting,

IS*6110*-based RFLP on multiple pre-treatment sputum samples identified mixed infections at a frequency of 9% in adult inmates in a prison TB hospital in Georgia (Shamputa *et al.* 2006). When MIRU-VNTR was used as a secondary typing tool, mixed infections were observed in 13.1% of the patients. The high mixed infection rates seen in this study occurred in settings of overcrowding coupled with a high incidence of TB.

Mixed infections in extra-pulmonary sites

While the majority of analyses reported herein was focused on sputum, the occurrence of mixed infections has also been demonstrated in extra-pulmonary samples taken from culture-positive biopsy specimens in two studies, both reporting a frequency of 9% (Cohen *et al.* 2011; Lieberman *et al.* 2016). Sampling of different anatomical sites in another study yielded a frequency of 3.6% (Chaves *et al.* 1999) while a frequency of 6% was recorded in another study where both extra-pulmonary and pulmonary specimens were collected (de Viedma *et al.* 2003). In an autopsy study, using lung and extra-pulmonary samples, a mixed infection frequency of 15.4% was reported (du Plessis *et al.* 2001). The highest frequency of mixed infections involving extra-pulmonary samples was reported by Ssengooba *et al.* (2015) who found that 51% of HIV-infected individuals with CD4 counts $>$ 200 cells/mm^3 displayed discordant *M. tuberculosis* genotypes between isolates from blood and sputum samples, respectively, providing evidence of substantive mixed infection frequencies in different body compartments (Ssengooba *et al.* 2015). In agreement with this, genomic analysis of postmortem lung and extra-pulmonary biopsies from 44 HIV-positive patients in KwaZulu-Natal reported that *M. tuberculosis* sub-lineages, including mixed infections, were distributed differentially throughout the lung, pointing towards temporary barriers to pathogen migration in the lung (Lieberman *et al.* 2016). In addition, multiple genotypes were identified between *M. tuberculosis* organisms in the lungs and extra-pulmonary sites, which may have occurred through a single transmission event involving multiple genotypes or by means of super-infection episodes (Lieberman *et al.* 2016). The genomic diversity identified by the occurrence of mixed-strain infections as well diversified sub-lineages provide an interesting record of evolution and dissemination across the body in HIV-infected individuals.

RECURRENT TB DISEASE

A select few studies analyzed here investigated whether mixed infections were more likely to occur during a primary versus a recurrent disease episode. From these, the frequency of mixed infections were reported as being higher in the recurrent disease episode in seven out of 13 studies, although most of these studies found this difference not to be significant (Warren *et al.* 2004; Fang *et al.* 2008; Huang *et al.* 2010; Wang *et al.* 2011, 2015; Shin *et al.* 2014; Pang *et al.* 2015). The six remaining studies reported lower frequencies of mixed infections in the recurrent episode, Tables 1 and 2. However, in three of these studies, less than 20% of the patients analyzed were recurrent cases, making it difficult to draw any definitive conclusions on the likelihood of previous TB exposure being a risk factor for mixed-strain infections.

A review of earlier studies yielded little consensus regarding the proportions of recurrent disease due to relapse and re-infection (Lambert *et al.* 2003). Although re-infection is most likely to occur in geographical locations where the incidence of TB disease is high (Romeyn 1970; Vynnycky and Fine 1997; Godfrey-Faussett *et al.* 2000; Richardson *et al.* 2002b; Middelkoop *et al.* 2015), studies performed in low- to moderate-incidence settings have also reported high rates of re-infection (Caminero *et al.* 2001). A poor association between re-infection and recurrent TB disease in high-burden settings has also been reported, suggesting that a consistent trend in this regard is lacking (Shamputa *et al.* 2007; Luzze *et al.* 2013). Annual rates of relapse versus re-infection in patients with recurrent disease in Cape Town, South Africa, indicate that relapse generally occurred within a year following treatment, while re-infection, responsible for at least half of all cases of recurrent disease in this study, predominated after the first year following treatment (Marx *et al.* 2014; Guerra-Assunção *et al.* 2015). Hence, it is difficult to make direct comparisons between studies where the length of follow-up after treatment differs. Co-infection with HIV and antiretroviral treatment has been shown to be associated with increased risk for recurrent disease due to re-infection as opposed to relapse (Godfrey-Faussett *et al.* 1994; Sonnenberg *et al.* 2001; Middelkoop *et al.* 2012, 2015; Guerra-Assunção *et al.* 2015). It was also demonstrated that there is a 2.4 times higher hazard ratio for recurrent disease in HIV-1-positive individuals, compared to HIV-1-negative counterparts (Sonnenberg *et al.* 2001). The risk of re-infection has also been linked to previous TB infection, where disease attributable to re-infection after successful treatment completion was four times higher than that attributable to new TB disease (Verver *et al.* 2005). A major limitation of this study is that the HIV status of most patients was unknown; however, the authors concluded that re-infections observed could not be explained by HIV status alone, as the HIV prevalence in this study group was relatively low (1.2% to 5.2% in antenatal clinics and 11% amongst newly diagnosed TB patients). These findings suggest that a subset of individuals within a population may be predisposed to TB disease, with previous TB exposure itself being a risk factor for recurrent disease. We hypothesize that mixed infection prevalence will be higher in these individuals.

INTERPRETATION AND RECOMMENDATIONS

Interruption of TB transmission is pivotal to eliminating this disease from human society (Pai *et al.* 2016). The science surrounding transmission of the tubercle bacillus is an evolving discipline and much work is still needed to elucidate the biophysical, environmental, genetic and immunologic determinants in those individuals that transmit disease and those that get infected. We propose that the analysis of mixed infection rates provides an interesting and useful proxy for the level of transmission in communities within TB-endemic regions. However, retrospective analyses of studies conducted in various settings to assess this are difficult due to numerous factors which include differences in study design, sample processing, bacterial culture and molecular typing tool employed. Although molecular approaches for identifying mixed-strain TB infections are limited in their sensitivity, studies relying on sputum samples alone have reported that up to 19% of individuals are infected with more than one *Mycobacterium tuberculosis* strain. Furthermore, studies utilizing a combination of both pulmonary and extra-pulmonary samples have reported mixed infections to occur at a frequency of up to 51%. As these results are likely to be a conservative estimate, this suggests that mixed-strain infections occur at reasonably high frequencies. Whether this is a result of high transmission rates in endemic settings remains unclear as mixed infection rates seems similar in high- and low-

burden settings, based on the studies analyzed herein. Various factors need to be considered when assessing the prevalence or risk of mixed infections. These are detailed below.

What is the role of HIV?

From our collective analysis, factors associated with mixed-strain infection among HIV-infected patients include prior TB treatment [adjusted prevalence ratio, 2.11] and CD4+ T-cell count < 100 cells/μl (adjusted prevalence ratio, 10.18), suggesting that patients with advanced immune suppression are at increased risk of contracting a mixed-strain infection (Shin *et al.* 2014). Consistent with this, a study in Uganda on 113 smear and culture-positive patients established that mixed infections are more likely to occur in HIV-seropositive individuals, 37.5% versus 12.6% in HIV-positive versus HIV-negative patients with TB, respectively (Dickman *et al.* 2010). These results may still underestimate the proportion of mixed infections as TB/HIV co-infection often results in disseminated/extra-pulmonary disease, therefore reducing the chances of detecting mixed infections from pulmonary specimens alone, which may be paucibacillary in nature (de Viedma *et al.* 2003; Dickman *et al.* 2010). Various studies have now shown that multiple genotypes can be detected by sampling from both respiratory and extra-pulmonary sites in HIV-positive individuals, illustrating the presence of migration routes within and between organs. To our knowledge, no studies have investigated this sampling strategy in HIV-negative individuals. Despite the demonstrated increased risk of mixed infections in HIV-infected individuals, we found no substantive differences in mixed infection incidence when collectively comparing studies within Africa, an HIV-endemic region, and the rest of the world. This could reflect the lack of studies in the correct populations or the above-mentioned limitations in our comparative analysis.

How do mixed infections affect disease/treatment outcome?

The presence of mixed-strain infections involving heteroresistance results in poor treatment outcome (Theisen *et al.* 1995; Niemann *et al.* 2000; Baldeviano-Vidalon *et al.* 2005; van Rie *et al.* 2005; Kamakoli *et al.* 2017). For this reason, patients with various forms of drug-resistant TB may require a combination of first- and second-line antibiotics to target both drug-sensitive and -resistant organisms, respectively (Post *et al.* 2004; van Rie *et al.* 2005). Diagnostically, if only the sensitive strain is cultured from a mixed-strain infection, the administration of a drug-sensitive regimen will result in the selection of the resistant strain. This highlights the importance of early detection of mixed-strain infections with different resistance profiles to tailor the treatment regimen accordingly.

Mixed infections and the immune response

It has been demonstrated that different *M. tuberculosis* strains elicit distinct immune responses in the lung, ultimately leading to differences in disease pathology and mortality (Post *et al.* 2004). Consistent with this, immune protection conferred by one *M. tuberculosis* strain may differ when compared to distantly related strains due to differences in antigenic properties. Indeed, studies assessing the effectiveness of Bacille Calmette-Guerin (BCG) have varied worldwide and are possibly complicated by the heterogeneity in *M. tuberculosis* strains circulating within vaccinated communities (Fine 1995). Direct intra-tracheal injection of BALB/c mice indicated that BCG vaccination was less effective against infection with Beijing strains when compared to H37Rv (Lopez *et al.* 2003). Furthermore, infection with Beijing strains was characterized by excessive pneumonia, early but short-lived tumor necrosis factor-α (TNF-α) and nitric oxide synthetase (INOS) expression and an increase in mortality, while infection with *M. canetti* resulted in limited pneumonia, sustained TNF-α and INOS expression in the lung, with 100% survival (Lopez *et al.* 2003). This indicates that different *M. tuberculosis* strains elicit varying immune-pathological events. In this regard, the simultaneous interaction between multiple strains and the host immune system in a mixed infection scenario remains to be explored. In addition, it is not known whether all resident strains in an infected individual are transmitted simultaneously or whether one strain is more transmissible than the other. These remain as critical knowledge gaps that limit the development of novel interventions to reduce transmission.

Phenotypic heterogeneity

An important aspect of sputum microbiology, for which there has been growing appreciation, is the failure to detect non-replicating organisms that could influence the probability of diagnosing mixed-strain infections (Dartois *et al.* 2016). It is predicted that a sub-population of organisms withstands treatment and can enter into a drug-tolerant, quiescent state (Kell and Young 2000). Consistent with this, *M. tuberculosis* grown under hypoxic conditions is able to enter into a drug-tolerant, non-replicating state and accumulate lipid inclusion bodies (Garton *et al.* 2008). Bacteria with a similar physiology have been isolated directly from patient sputum, suggesting that these cells reflect those studied *in vitro* (Garton *et al.* 2008). Subsequent work has also demonstrated that pre-treatment sputum samples are dominated by a DCTB population that can only be grown in liquid media (Mukamolova *et al.* 2010; Chengalroyen *et al.* 2016; Loraine *et al.* 2016). Considering this, we hypothesize that it is possible that specific strains of *M. tuberculosis* adopt this differentially culturable state, and thus are not identifiable by routine culture. Currently, there is no data to corroborate this, but the upregulation of the DosR regulon in W-Beijing strains, and concomitant accumulation of triacylglycerides, suggests that this family of strains may have a greater propensity to adopt these differential growth states (Reed *et al.* 2007). Hence, the association of Beijing lineage 2 strains with relapse (Lan *et al.* 2003; Bryant *et al.* 2013) may be linked to a proportion of this strain assuming a differentially culturable state, thereby providing the organism with an adaptive advantage against immune assault and drug treatment. Furthermore, the occurrence of DCTB could result in an underestimation of mixed infections as some strains may not emerge on solid media.

Limitations in molecular typing tools

Current molecular methods do not provide adequate information on the mechanism of transmission and order in which each individual infection was acquired, for the latter continuous and very expensive molecular surveillance would be required. This information would be invaluable in the interpretation of epidemiological data. For example, in a scenario where re-infection resulted in the re-activation of an underlying strain, IS*6110* may only identify the secondary strain which could be interpreted as the strain causing the original infection and result in inaccurate identification of the source of infection in contact tracing. High

rates of mixed infections in a recurrent disease episode highlight the importance of re-infection in disease transmission, which has previously been questioned (Stead 1967). These limitations call for wider use of high-resolution techniques such as WGS.

Diagnosing mixed infections

In high-burden settings, where the majority of TB disease is likely (but not necessarily) due to re-infection /recent infection, TB control strategies will be highly reliant on the early identification and early treatment of infectious individuals to prevent transmission of both drug-susceptible and -resistant strains within the community. High rates of disease recurrences due to re-infection pose a serious problem to vulnerable populations, such as HIV-infected mining communities where post-treatment prophylaxis was proposed as an important public health strategy (Sonnenberg *et al*. 2001). Exogenous re-infection has also been implicated in the acquisition of drug resistance in patients (Andrews *et al*. 2008). Several studies in South Africa investigating extremely drug-resistant (XDR)-TB have been associated with primary transmission (Ghandi *et al*. 2006; Klopper *et al*. 2013). These cases have been attributed to 'super spreaders'—a small percentage of infected individuals who have the propensity to transmit TB disease as identified by a cough aerosol sampling system (Jones-López *et al*. 2013). The long-term treatment outcome in patients with XDR-TB is very poor, and the growing infectious pool of drug-resistant TB disease is of major concern as it threatens global TB control efforts (Dheda *et al*. 2010). Where drug resistance is often the result of primary transmission of resistant strains, the rapid identification and treatment of patients harboring these strains should be a high priority. Furthermore, if mutations are more likely to emerge from certain strains associated with re-infection, strain differentiation may need to be included in the diagnostic process for early identification and the development of tailored treatment regimens. With regard to implementation, while the addition of strain typing to current diagnostic algorithms in high-income, low-burden nations would seem plausible, the introduction of this technology in developing nations with a high disease burden would be unrealistic due to the high costs and the laboratory expertise required. These challenges highlight the need for the development of simpler and more affordable technologies for the characterization of TB strains (e.g. point-of-care multiplex tool to differentiate strain type). Successful interventions to decrease transmission will reduce the levels of recurrent TB disease (due to re-infection) as well as reduce the opportunities for mixed-strain TB infection.

CONCLUDING REMARKS

Molecular fingerprinting techniques have allowed for the differentiation of recurrent TB disease into relapse versus re-infection, as well as the identification of TB infection with multiple strains. The development of improved techniques for the identification of relapse, re-infection and mixed infections has important implications for the understanding of TB epidemiology and the implementation of related technologies in TB-control programs, patient management and outcome analysis of clinical trials. Ongoing research using WGS to provide more detailed insight into these phenomena is imperative to design effective transmission intervention strategies. However, a limitation with the substantive research effort in these areas over the last two decades is the lack of a common sample collection and processing protocol for specimens to allow for a direct comparison of mixed infections and recurrent disease in geographically distinct settings with varying burdens of disease. Such studies should take high priority on the global TB research agenda.

ACKNOWLEDGEMENTS

We would like to thank CBTBR members for insightful comment and discussion.

FUNDING

This work was supported by funding from an International Early Career Scientist Award from the Howard Hughes Medical Institute (to BDK), the South African National Research Foundation (to BDK), the South African Medical Research Council (to BDK), the South African Medical Research Council in terms of the National Health Scholars Programme (to AM) and the National Health Laboratory Services Research Trust (to BDK). AM was also supported by a grant from the Soweto Matlosana Collaborating Centre for HIV/Aids and TB.

Conflict of interest. None declared.

REFERENCES

Andrews JR, Gandhi NR, Moodley P *et al*. Exogenous reinfection as a cause of multidrug-resistant and extensively drug-resistant tuberculosis in rural South Africa. *J Infect Dis* 2008;**198**:1582–9.

Baldeviano-Vidalon G, Quispe-Torres N, Bonilla-Asalde C *et al*. Multiple infection with resistant and sensitive *M. tuberculosis* strains during treatment of pulmonary tuberculosis patients. *Int J Tuberc Lung D* 2005;**9**:1155–60.

Bandera A, Gori A, Catozzi L *et al*. Molecular epidemiology study of exogenous reinfection in an area with a low incidence of tuberculosis. *J Clin Microbiol* 2001;**39**:2213–8.

Barnes PF, Cave MD. Molecular epidemiology of tuberculosis. *New Engl J Med* 2003;**349**:1149–56.

Bates JH, Stead WW, Rado TA. Phage type of tubercle bacilli isolated from patients with two or more sites of organ involvement. *Am Rev Respir Dis* 1976;**114**:353–8.

Bauer J, Andersen AB, Kremer K *et al*. Usefulness of spoligotyping to discriminate IS6110 low-copy-number *Mycobacterium tuberculosis* complex strains cultured in Denmark. *J Clin Microbiol* 1999;**37**:2602–6.

Braden CR, Morlock GP, Woodley CL *et al*. Simultaneous infection with multiple strains of *Mycobacterium tuberculosis*. *Clin Infect Dis* 2001;**33**:e42–7.

Bryant JM, Harris SR, Parkhill J *et al*. Whole-genome sequencing to establish relapse or re-infection with *Mycobacterium tuberculosis*: a retrospective observational study. *Lancet Respir Med* 2013;**1**:786–92.

Caminero JA, Pena MJ, Campos-Herrero MI *et al*. Exogenous reinfection with tuberculosis on a European island with a moderate incidence of disease. *Am J Respir Crit Care* 2001;**163**:717–20.

Cave M, Eisenach KD, Templeton G *et al*. Stability of DNA fingerprint pattern produced with IS6110 in strains of *Mycobacterium tuberculosis*. *J Clin Microbiol* 1994;**32**:262–6.

Charalambous S, Grant AD, Moloi V *et al*. Contribution of reinfection to recurrent tuberculosis in South African gold miners. *Int J Tuberc Lung D* 2008;**12**:942–8.

Chaves F, Dronda F, Alonso-Sanz M *et al*. Evidence of exogenous reinfection and mixed infection with more than one strain of *Mycobacterium tuberculosis* among Spanish HIV-infected inmates. *AIDS* 1999;**13**:615–20.

Chengalroyen MD, Beukes GM, Gordhan BG *et al*. Detection and quantification of differentially culturable tubercle bacteria in sputum from tuberculosis patients. *Am J Respir Crit Care* 2016;**194**:1532–40.

Cohen T, van Helden PD, Wilson D *et al*. Mixed-strain *Mycobacterium tuberculosis* infections and the implications for tuberculosis treatment and control. *Clin Microbiol Rev* 2012;**25**:708–19.

Cohen T, Wilson D, Wallengren K *et al*. Mixed-strain *Mycobacterium tuberculosis* infections among patients dying in a hospital in KwaZulu-Natal, South Africa. *J Clin Microbiol* 2011;**49**:385–8.

Cox HS, Kalon S, Allamuratova S *et al*. Multidrug-resistant tuberculosis treatment outcomes in Karakalpakstan, Uzbekistan: treatment complexity and XDR-TB among treatment failures. *PLoS One* 2007;**2**:e1126.

Crampin AC, Glynn JR, Fine PE. What has Karonga taught us? Tuberculosis studied over three decades. *Int J Tuberc Lung D* 2009;**13**:153–64.

Dartois V, Saito K, Warrier T *et al*. New evidence for the complexity of the population structure of *Mycobacterium tuberculosis* increases the diagnostic and biologic challenges. *Am J Respir Crit Care* 2016;**194**:1448–51.

Das S, Narayanan S, Hari L *et al*. Simultaneous infection with multiple strains of *Mycobacterium tuberculosis* identified by restriction fragment length polymorphism analysis. *Int J Tuberc Lung D* 2004;**8**:267–70.

de Boer AS, Kremer K, Borgdorff MW *et al*. Genetic heterogeneity in *Mycobacterium tuberculosis* isolates reflected in IS6110 restriction fragment length polymorphism patterns as low-intensity bands. *J Clin Microbiol* 2000;**38**:4478–84.

de Viedma DG, Marín M, Serrano MJR *et al*. Polyclonal and compartmentalized infection by *Mycobacterium tuberculosis* in patients with both respiratory and extrarespiratory involvement. *J Infect Dis* 2003;**187**:695–9.

de Viedma DG, Rodriguez NA, Andrés S *et al*. Characterization of clonal complexity in tuberculosis by mycobacterial interspersed repetitive unit-variable-number tandem repeat typing. *J Clin Microbiol* 2005;**43**:5660–4.

Dheda K, Limberis J, Pietersen E *et al*. Outcomes, infectiousness, and transmission dynamics of patients with extensively drug resistant tuberculosis and home-discharged patients with programmatically incurable tuberculosis: a prospective cohort study. *Lancet Respir Med* 2017, in press. http://dx.doi.org/10.1016/52212-2600(16)30433-7.

Dheda K, Shean K, Zumla A *et al*. Early treatment outcomes and HIV status of patients with extensively drug-resistant tuberculosis in South Africa: a retrospective cohort study. *Lancet* 2010;**375**:1798–807.

Dickman KR, Nabyonga L, Kateete DP *et al*. Detection of multiple strains of *Mycobacterium tuberculosis* using MIRU-VNTR in patients with pulmonary tuberculosis in Kampala, Uganda. *BMC Infect Dis* 2010;**10**:349.

du Plessis DG, Warren R, Richardson M *et al*. Demonstration of reinfection and reactivation in HIV-negative autopsied cases of secondary tuberculosis: multilesional genotyping of *Mycobacterium tuberculosis* utilizing IS 6110 and other repetitive element-based DNA fingerprinting. *Tuberculosis* 2001;**81**:211–20.

Fang R, Li X, Li J *et al*. Mixed infections of *Mycobacterium tuberculosis* in tuberculosis patients in Shanghai, China. *Tuberculosis* 2008;**88**:469–73.

Fine PE. Bacille Calmette-Guerin vaccines: a rough guide. *Clin Infect Dis* 1995;**20**:11–4.

Ford C, Yusim K, Ioerger T *et al*. *Mycobacterium tuberculosis* heterogeneity revealed through whole genome sequencing. *Tuberculosis* 2012;**92**:194–201.

Gan M, Liu Q, Yang C *et al*. Deep whole-genome sequencing to detect mixed infection of *Mycobacterium tuberculosis*. *PLoS One* 2016;**11**:e0159029.

Garton NJ, Waddell SJ, Sherratt AL *et al*. Cytological and transcript analyses reveal fat and lazy persister-like bacilli in tuberculous sputum. *PLoS Med* 2008;**5**:e75.

Ghandi N, Moll A, Sturm A *et al*. Extensively drug-resistant tuberculosis as a cause of death in patients co-infected with tuberculosis and HIV in a rural area of South Africa. *Lancet* 2006;**368**:1575–80.

Godfrey-Faussett P, Githui W, Batchelor B *et al*. Recurrence of HIV-related tuberculosis in an endemic area may be due to relapse or reinfection. *Tuber Lung Dis* 1994;**75**:199–202.

Godfrey-Faussett P, Sonnenberg P, Shearer SC *et al*. Tuberculosis control and molecular epidemiology in a South African gold-mining community. *Lancet* 2000;**356**:1066–71.

Guerra-Assunção JA, Houben RM, Crampin AC *et al*. Recurrence due to relapse or reinfection with *Mycobacterium tuberculosis*: a whole-genome sequencing approach in a large, population-based cohort with a high HIV infection prevalence and active follow-up. *J Infect Dis* 2015;**211**:1154–63.

Gutacker MM, Mathema B, Soini H *et al*. Single-nucleotide polymorphism-based population genetic analysis of *Mycobacterium tuberculosis* strains from 4 geographic sites. *J Infect Dis* 2006;**193**:121–8.

Hanekom M, Streicher E, Van de Berg D *et al*. Population structure of mixed *Mycobacterium tuberculosis* infection is strain genotype and culture medium dependent. *PLoS One* 2013;**8**:e70178.

Hingley-Wilson SM, Casey R, Connell D *et al*. Undetected multidrug-resistant tuberculosis amplified by first-line therapy in mixed infection. *Emerg Infect Dis* 2013;**19**:1138–41.

Huang H-Y, Tsai Y-S, Lee J-J *et al*. Mixed infection with Beijing and non-Beijing strains and drug resistance pattern of *Mycobacterium tuberculosis*. *J Clin Microbiol* 2010;**48**:4474–80.

Huyen MN, Kremer K, Lan NT *et al*. Mixed tuberculosis infections in rural South Vietnam. *J Clin Microbiol* 2012;**50**:1586–92.

Jones-López EC, Namugga O, Mumbowa F *et al*. Cough aerosols of *Mycobacterium tuberculosis* predict new infection. A household contact study. *Am J Respir Crit Care* 2013;**187**:1007–15.

Kamakoli M, Sadegh H, Farmanfarmaei M *et al*. Evaluation of the impact of polyclonal infection and heteroresistance on treatment of tuberculosis patients. *Sci Rep* 2017;**7**:41410, DOI: 10.1038/srep41410.

Kamerbeek J, Schouls L, Kolk A *et al*. Simultaneous detection and strain differentiation of *Mycobacterium tuberculosis* for diagnosis and epidemiology. *J Clin Microbiol* 1997;**35**:907–14.

Kana BD, Karakousis P, Parish T *et al*. Future target-based drug discovery for tuberculosis? *Tuberculosis* 2014;**94**:551–6.

Kaplan G, Post FA, Moreira AL *et al*. *Mycobacterium tuberculosis* growth at the cavity surface: a microenvironment with failed immunity. *Infect Immun* 2003;**71**:7099–108.

Kay GL, Sergeant MJ, Zhou Z *et al*. Eighteenth-century genomes show that mixed infections were common at time of peak tuberculosis in Europe. *Nat Commun* 2015;**6**:6717, DOI: 10.1038/ncomms7717.

Kell DB, Young M. Bacterial dormancy and culturability: the role of autocrine growth factors: commentary. *Curr Opin Microbiol* 2000;**3**:238–43.

Klopper M, Warren RM, Hayes C *et al.* Emergence and spread of extensively and totally drug-resistant tuberculosis, South Africa. *Emerg Infect Dis* 2013;**19**:449–55.

Kremer K, van Soolingen D, Frothingham R *et al.* Comparison of methods based on different molecular epidemiological markers for typing of *Mycobacterium tuberculosis* complex strains: interlaboratory study of discriminatory power and reproducibility. *J Clin Microbiol* 1999;**37**:2607–18.

Kritzinger FE, den Boon S, Verver S *et al.* No decrease in annual risk of tuberculosis infection in endemic area in Cape Town, South Africa. *Trop Med Int Health* 2009;**14**:136–42.

Kruuner A, Pehme L, Ghebremichael S *et al.* Use of molecular techniques to distinguish between treatment failure and exogenous reinfection with *Mycobacterium tuberculosis*. *Clin Infect Dis* 2002;**35**:146–55.

Lambert ML, Hasker E, Van Deun A *et al.* Recurrence in tuberculosis: relapse or reinfection? *Lancet Infect Dis* 2003;**3**:282–7.

Lan NT, Lien HT, Tung le B *et al. Mycobacterium tuberculosis* Beijing genotype and risk for treatment failure and relapse, Vietnam. *Emerg Infect Dis* 2003;**9**:1633–5.

Lieberman TD, Wilson D, Misra R *et al.* Genomic diversity in autopsy samples reveals within-host dissemination of HIV-associated *Mycobacterium tuberculosis*. *Nat Med* 2016;**22**:1470–6.

Lopez B, Aguilar D, Orozco H *et al.* A marked difference in pathogenesis and immune response induced by different *Mycobacterium tuberculosis* genotypes. *Clin Exp Immunol* 2003;**133**:30–7.

Loraine J, Pu F, Turapov O *et al.* Development of an in vitro assay for detection of drug-induced resuscitation-promoting-factor-dependent mycobacteria. *Antimicrob Agents Ch* 2016;**60**:6227–33.

Luzze H, Johnson DF, Dickman K *et al.* Relapse more common than reinfection in recurrent tuberculosis 1-2 years post treatment in urban Uganda. *Int J Tuberc Lung D* 2013;**17**:361–7.

Mallard K, McNerney R, Crampin AC *et al.* Molecular detection of mixed infections of *Mycobacterium tuberculosis* strains in sputum samples from patients in Karonga district, Malawi. *J Clin Microbiol* 2010;**48**:4512–8.

Mankiewicz E, Liivak M. Phage types of *Mycobacterium tuberculosis* in cultures isolated from Eskimo patients. *Am Rev Respir Dis* 1975;**111**:307–12.

Martín A, Herranz M, Serrano MJR *et al.* The clonal composition of *Mycobacterium tuberculosis* in clinical specimens could be modified by culture. *Tuberculosis* 2010;**90**:201–7.

Marx FM, Dunbar R, Enarson DA *et al.* The temporal dynamics of relapse and reinfection tuberculosis after successful treatment: a retrospective cohort study. *Clin Infect Dis* 2014;**58**:1676–83.

Mathema B, Kurepina NE, Bifani PJ *et al.* Molecular epidemiology of tuberculosis: current insights. *Clin Microbiol Rev* 2006;**19**:658–85.

Middelkoop K, Bekker LG, Mathema B *et al.* Factors affecting tuberculosis strain success over 10 years in a high TB- and HIV-burdened community. *Int J Epidemiol* 2014;**43**:1114–22.

Middelkoop K, Bekker LG, Shashkina E *et al.* Retreatment tuberculosis in a South African community: the role of re-infection, HIV and antiretroviral treatment. *Int J Tuberc Lung D* 2012;**16**:1510–6.

Middelkoop K, Mathema B, Myer L *et al.* Transmission of tuberculosis in a South African community with a high prevalence of HIV infection. *J Infect Dis* 2015;**211**:53–61.

Mokrousov I, Valcheva V, Sovhozova N *et al.* Penitentiary population of *Mycobacterium tuberculosis* in Kyrgyzstan: exceptionally high prevalence of the Beijing genotype and its Russia-specific subtype. *Infect Genet Evol* 2009;**9**:1400–5.

Mukamolova GV, Turapov O, Malkin J *et al.* Resuscitation-promoting factors reveal an occult population of tubercle bacilli in sputum. *Am J Respir Crit Care* 2010;**181**:174–80.

Mulenga C, Shamputa IC, Mwakazanga D *et al.* Diversity of *Mycobacterium tuberculosis* genotypes circulating in Ndola, Zambia. *BMC Infect Dis* 2010;**10**:1.

Mustafa S, Hasnain J, Hashmi J *et al.* Emergence of mixed infection of Beijing/non-Beijing strains among multi-drug resistant *Mycobacterium tuberculosis* in Pakistan. *Biotech* 2016;**6**:108.

Muwonge A, Kankya C, Olea-Popelka F *et al.* Molecular investigation of multiple strain infections in patients with tuberculosis in Mubende district, Uganda. *Infect Genet Evol* 2013;**17**:16–22.

Niemann S, Richter E, Rusch-Gerdes S *et al.* Double infection with a resistant and a multidrug-resistant strain of *Mycobacterium tuberculosis*. *Emerg Infect Dis* 2000;**6**:548–51.

Pai M, Behr M, Dowdy D *et al.* Tuberculosis. *Nat Rev Dis Primers* 2016;**2**:16706.

Pang Y, Zhou Y, Wang S *et al.* Prevalence and risk factors of mixed *Mycobacterium tuberculosis* complex infections in China. *J Infect* 2015;**71**:231–7.

Peng Y, Yang C, Li X *et al.* Multiple samples improve the sensitivity for detection of mixed Mycobacterium infections. *Tuberculosis* 2013;**93**:548–50.

Post FA, Willcox PA, Mathema B *et al.* Genetic polymorphism in *Mycobacterium tuberculosis* isolates from patients with chronic multidrug-resistant tuberculosis. *J Infect Dis* 2004;**190**:99–106.

Reed MB, Gagneux S, DeRiemer K *et al.* The W-Beijing lineage of *Mycobacterium tuberculosis* overproduces triglycerides and has the DosR dormancy regulon constitutively upregulated. *J Bacteriol* 2007;**189**:2583–9.

Rhee JT, Tanaka MM, Behr MA *et al.* Use of multiple markers in population-based molecular epidemiologic studies of tuberculosis. *Int J Tuberc Lung D* 2000;**4**:1111–9.

Richardson M, Carroll NM, Engelke E *et al.* Multiple *Mycobacterium tuberculosis* strains in early cultures from patients in a high-incidence community setting. *J Clin Microbiol* 2002a;**40**:2750–4.

Richardson M, van Lill SW, van der Spuy GD *et al.* Historic and recent events contribute to the disease dynamics of Beijing-like *Mycobacterium tuberculosis* isolates in a high incidence region. *Int J Tuberc Lung D* 2002b;**6**:1001–11.

Rinder H. Hetero-resistance: an under-recognised confounder in diagnosis and therapy? *J Med Microbiol* 2001;**50**:1018–20.

Romeyn JA. Exogenous reinfection in tuberculosis. *Am Rev Respir Dis* 1970;**101**:923–7.

Shamputa IC, Jugheli L, Sadradze N *et al.* Mixed infection and clonal representativeness of a single sputum sample in tuberculosis patients from a penitentiary hospital in Georgia. *Respir Res* 2006;**7**:99.

Shamputa IC, Rigouts L, Eyongeta LA *et al.* Genotypic and phenotypic heterogeneity among *Mycobacterium tuberculosis* isolates from pulmonary tuberculosis patients. *J Clin Microbiol* 2004;**42**:5528–36.

Shamputa IC, Van Deun A, Salim MA *et al.* Endogenous reactivation and true treatment failure as causes of recurrent tuberculosis in a high incidence setting with a low HIV infection. *Trop Med Int Health* 2007;**12**:700–8.

Shin SS, Modongo C, Ncube R *et al.* Advanced immune suppression is associated with increased prevalence of mixed-strain

Mycobacterium tuberculosis infections among persons at high risk for drug-resistant tuberculosis in Botswana. *J Infect Dis* 2014;**211**:347–51.

Small PM, McClenny NB, Singh SP *et al*. Molecular strain typing of *Mycobacterium tuberculosis* to confirm cross-contamination in the mycobacteriology laboratory and modification of procedures to minimize occurrence of false-positive cultures. *J Clin Microbiol* 1993;**31**:1677–82.

Sonnenberg P, Murray J, Glynn JR *et al*. HIV-1 and recurrence, relapse, and reinfection of tuberculosis after cure: a cohort study in South African mineworkers. *Lancet* 2001;**358**:1687–93.

Ssengooba W, Cobelens FG, Nakiyingi L *et al*. High genotypic discordance of concurrent *Mycobacterium tuberculosis* isolates from sputum and blood of HIV-infected individuals. *PLoS One* 2015;**10**:e0132581.

Stavrum R, Mphahlele M, Øvreås K *et al*. High diversity of *Mycobacterium tuberculosis* genotypes in South Africa and preponderance of mixed infections among ST53 isolates. *J Clin Microbiol* 2009;**47**:1848–56.

Stead WW. Pathogenesis of a first episode of chronic pulmonary tuberculosis in man: recrudescence of residuals of the primary infection or exogenous reinfection? *Am Rev Respir Dis* 1967;**95**:729–45.

Streit E, Millet J, Rastogi N. *Mycobacterium tuberculosis* polyclonal infections and microevolution identified by MIRU-VNTRs in an epidemiological study. *Int J Mycobacteriol* 2015;**4**:222–7.

Theisen A, Reichel C, Rüsch-Gerdes S *et al*. Mixed-strain infection with a drug-sensitive and multidrug-resistant strain of *Mycobacterium tuberculosis*. *Lancet* 1995;**345**:1512–3.

Umubyeyi AN, Shamputa IC, Rigouts L *et al*. Molecular investigation of recurrent tuberculosis in patients from Rwanda. *Int J Tuberc Lung D* 2007;**11**:860–7.

van Embden JD, Cave MD, Crawford JT *et al*. Strain identification of *Mycobacterium tuberculosis* by DNA fingerprinting: recommendations for a standardized methodology. *J Clin Microbiol* 1993;**31**:406–9.

van Rie A, Warren R, Richardson M *et al*. Exogenous reinfection as a cause of recurrent tuberculosis after curative treatment. *New Engl J Med* 1999;**341**:1174–9.

van Rie A, Victor TC, Richardson M *et al*. Reinfection and mixed infection cause changing *Mycobacterium tuberculosis* drug-resistance patterns. *Am J Respir Crit Care* 2005;**172**:636–42.

Verver S, Warren RM, Beyers N *et al*. Rate of reinfection tuberculosis after successful treatment is higher than rate of new tuberculosis. *Am J Respir Crit Care* 2005;**171**:1430–5.

Verver S, Warren RM, Munch Z *et al*. Proportion of tuberculosis transmission that takes place in households in a high-incidence area. *Lancet* 2004;**363**:212–4.

Vynnycky E, Fine PE. The natural history of tuberculosis: the implications of age-dependent risks of disease and the role of reinfection. *Epidemiol Infect* 1997;**119**:183–201.

Wang JY, Hsu HL, Yu MC *et al*. Mixed infection with Beijing and non-Beijing strains in pulmonary tuberculosis in Taiwan: prevalence, risk factors, and dominant strain. *Clin Microbiol Infect* 2011;**17**:1239–45.

Wang X, Liu H, Wei J *et al*. An investigation on the population structure of mixed infections of *Mycobacterium tuberculosis* in Inner Mongolia, China. *Tuberculosis* 2015;**95**:695–700.

Warren RM, Victor TC, Streicher EM *et al*. Patients with active tuberculosis often have different strains in the same sputum specimen. *Am J Respir Crit Care* 2004;**169**:610–4.

WHO. *Definitions and Reporting Framework for Tuberculosis – 2013 revision*. Geneva, Switzerland: WHO Press, World Health Organization, 2014.

WHO. *Global Tuberculosis Report*. Geneva, Switzerland: WHO Press, World Health Organization, 2016a.

WHO. *Tuberculosis Country Profiles*. 2016b, http://www.who.int/tb/country/data/profiles/en/ (26 October 2016, date last accessed).

Yeh RW, Hopewell PC, Daley CL. Simultaneous infection with two strains of *Mycobacterium tuberculosis* identified by restriction fragment length polymorphism analysis. *Int J Tuberc Lung D* 1999;**3**:537–9.

Zheng C, Li S, Luo Z *et al*. Mixed infections and rifampin heteroresistance among *Mycobacterium tuberculosis* clinical isolates. *J Clin Microbiol* 2015;**53**:2138–47.

Zumla A, Nahid P, Cole S. Advances in the development of new tuberculosis drugs and treatment regimens. *Nat Rev Drug Dis* 2013;**12**:388–404.